Vedic Research International
Bioinformatics & Proteomics
eISSN 2330-0272
JOURNAL HOME PAGE AT WWW.VEDICJOURNALS.COM
RESEARCH ARTICLE
DOI: http://dx.doi.org/10.14259/bp.v2i1.92

Comparison of ADAM and ADAMTS Protein Families in Human, Frog, Fly and Worm Genomes

KRISHNA K INAMPUDI*, GR HEMALATHA AND L GURUPRASAD
School of Chemistry, University of Hyderabad, Hyderabad 500 046, India

Article Info: Received: *December 7th, 2013*; **Accepted:** *December 13th, 2013*

Abstract
ADAM and ADAMTS family members play a crucial role in modulating the turnover of extracellular matrix, mediate cell-cell, cell matrix interactions, and are involved in several disease states. The completion of the human, fly, frog and worm genomes have now provided an opportunity to study the representative genomes in which these closely related proteins are present. In this work, we have identified and analyzed ADAMs and ADAMTS to understand the distribution of the members and domain architecture. We report that, the human genome is encoded by 90 ADAMs and 92 ADAMTS genes. 11 ADAMs and 2 ADAMTS genes encode the frog genome, 19 ADAM and 6 ADAMTS genes encode the fly genome and 7 ADAMs and 8 ADAMTS genes encode the worm genome. The phylogenetic tree of ADAM and ADAMTS is organized into 6 clades and the phylogenetic tree of the corresponding metalloproteinase domain is organized into 9 clades. We identified a different domain architecture pattern in ADAMTS protein family which is not as similar as to the previous report.

Keywords: *ADAM, ADAMTS, Human, Frog, Fly, Worm, Genome, Tsp*

Introduction
ADAM (A Disintegrin And Metalloproteinase) family proteins are characterized by the presence of both disintegrin and metalloproteinase domains. Another closely related protein family is the ADAMTS (A Disintegrin And Metalloproteinase with Thrombospondin motifs). Several members of these two protein families are known in humans. In the recent years, a number of publications explain the importance of ADAMs and ADAMTSs in diseases such as prostate, breast and non-small-cell lung cancers, arthritis and Alzheimer. These proteins are also of great physiological importance in regulating cell-cell and cell matrix interactions. Keeping in view the variety of functions regulated by these enzymes in humans, and the completed genome sequencing of several organisms, we intended to analyze the ADAM and ADAMTS protein families in the completed genomes of mammals- *Homo sapiens* (human), amphibian- *Xenopus laevis* (frog), arthropoda *Drosophila melanogaster* (fly) and nematode- *Caenorhabitis elegans* (worm). These four genomes represent all the organisms according to phylogeny, in which the ADAM and ADAMTS protein families are present.

ADAMs and ADAMTSs are multi-domain protein families and play multiple roles in cell signaling, cell fusion, and cell-cell interactions. Members of ADAM family comprise a C- terminal transmembrane segment and are therefore cell surface proteins. At the N- terminus, these proteins comprise a propeptide domain that contains a sequence motif similar to the "cysteine switch" of the matrixins. A zinc dependent proteinase domain follows this region, the proteolytic activity of ADAMs is due to the zinc proteinase domain and its activity is directed towards the extracellular domains of the transmembrane proteins [1]. The mechanism of activation of these proteins involves the cleavage of the pro-part followed by conformational changes in the protein. Activation of metalloproteinases is an additional important mechanism for regulating activity of these enzymes. The adjacent disintegrin domain is responsible for the adhesive properties of the protein, thus mediating cell-cell and cell matrix interactions. This is followed by the cysteine rich domain that supports cell adhesion. In addition to these four

*Corresponding Author
Krishna K Inampudi, PhD
School of Chemistry, University of Hyderabad, Hyderabad 500 046 India.
Phone: +33 0781037608
Email: kishorehcu@gmail.com

ISSN 2330-0272 (Online); CODEN: VRIBCN, VRI Bioinformatics & Proteomics, VRI Press, Vedic Research Inc. USA

domains, at the C-terminus, ADAMs also comprise a EGF domain, transmembrane segment and cytoplasmic tail responsible for signaling.

ADAMTS protein family members are soluble, multi-domain proteins. The first two domains in ADAMTS are similar to that in ADAMs, i.e. they comprise a pro-domain, zinc dependent proteinase. In addition they also comprise the thrombospondin repeats. Thrombospondin repeats are 50 amino acid residues long and are responsible for binding to extracellular matrix ligands including fibrinogen, fibronectin, some collagens, latent and active transforming growth factor-beta-1, TSG6, heparin, plasmin, cathepsin G, neutrophil elastase, some MMPs, tissue factor pathway inhibitor, and heparan sulfate proteoglycans.

The zinc proteinase domain is also common to matrix metalloproteinases (MMPs) superfamily, known as the "metzincins" [2]. The name is derived from consensus sequence and structural features involving a methionine residue, which forms a conserved structure known as the "metturn" and "zinc binding motif" (zincin). The essential components necessary for the catalytic proteinase mechanism are three histidines, a glutamic acid and a water molecule [3,4]. The key amino acid residues are arranged in a highly characteristic sequence; HExxHxxGxxH forming a zinc-binding motif. The triad of histidine residues coordinates the zinc ion, which in combination with the glutamic acid forms the critical sequence components of the catalytic mechanism in the proteins of MMPs, ADAM and ADAMTS families. The first two histidines are separated by a single turn of a helix, and allow the side chain imidazole ring to be positioned toward the catalytic zinc. The carboxylate group of glutamic acid residue acts as a neutrophile, with which an associated water molecule promotes cleavage of the substrate peptide scissile bond. The conserved glycine residue in the sequence motif allows a sharp turn, permitting the third histidine in the triad to associate with the zinc ion. On the C-terminal to the zinc ion binding motif is the conserved methionine residue which is responsible for the "metturn" in metzincin structures and provides a hydrophobic base for the histidine triad [2]. In addition to zinc, these enzymes also require calcium for stabilization of the protein tertiary structure.

The identification of these enzymes under normal physiological and disease states has led to the functional characterization. For example, ADAM12-S stimulates bone growth by modulating chondrocyte proliferation and maturation through mechanisms probably involving both metalloproteinase and adhesion activities [5]. ADAM33 has been identified as a susceptible gene for asthama [6]. ADAM19 has a constitutive alpha-secretase activity for amyloid precursor protein [7] and it has been suggested that ADAM19 may have a modulatory role in the dysfunctional renal allograft state [8]. ADAM12 and ADAMTS1 are implicated in non-small-cell lung cancer [9]. ADAM23 is frequently silenced in gastric cancers by homozygous deletion or aberrant promoter methylation [10]. ADAM15 generally overexpressed in adenocarcinoma is associated with metastatic progression of prostate and breast cancers [11] and ADAM28 is overexpressed in an activated form in breast carcinoma cells. ADAM29 expression ratio is a novel prognosis indicator in chronic lymphocytic leukemia [12]. ADAM9 is induced in human prostate cancer cells [13]. ADAMTS8 and ADAMTS15 have emerged as novel predictors of survival in patients with breast carcinoma [14]. One of the factors responsible for asthama is ADAM8. ADAM20 and ADAM21 are human testis-specific membrane metalloproteinases and ADAM30 show testis-specific gene expression [15]. ADAM18 is a sperm surface protein for oocyte recognition [16]. ADAM2 is required for sperm egg fusion. ADAM22, a brain-specific cell surface protein, mediates growth inhibition using an integrin dependent pathway. It is expressed in normal brain but not in high grade gliomas [17]. ADAMTS10 plays a major role in growth, skin, lens, and heart development in humans. ADAM17 is TNF-alpha convertase enzyme (TACE) from human arthritis affected cartilage [18]. ADAM10 represents an important molecular modulator of FasL-mediated cell death. It has been shown that there is genetic association between polymorphisms in the ADAMTS14 gene and multiple sclerosis [19]. ADAMTS1 expression is associated with decidualization of the endometrial stroma *in vivo* [20]. ADAMTS8 has a role in brain tumorigenesis. ADAMTS12 is important for the initiation and progression of arthritis. The L1565 variant of von Willebrand factor has a role in susceptibility to proteolysis by ADAMTS13 [21]. ADAMTS5 mainly contributes to ECM (extra cellular matrix) metabolism in growth plate and condylar cartilage during growth. ADAMTS1 and ADAMTS4 may be involved in ECM turn over in articular cartilage [22]. The importance of the ADAMs and ADAMTSs in these disease states makes them important drug targets and comparative genome studies of these protein families will help validate the drug targets. These comparative studies will help estimate the divergence of these protein families during evolution as required for their adaptation.

Methods

Search for ADAM and ADAMTS in the human, frog, fly and worm genomes

Sequences encoded by the ADAM and ADAMTS were obtained from protein database at NCBI (www.ncbi.nlm.nih.gov/). Preliminary searches for ADAM and ADAMTS were performed individually using BLAST [23]. Sequences belonging to these families with greater than >30% identities among themselves were considered as queries. Reciprocal searches were carried out with each of the protein sequence using PSI-BLAST [24] till no new proteins were identified. Non redundant (nr) protein sequence database was chosen. BLOSUM62 matrix, with existence 11 and extension 1 as gap penalties, Expect threshold 10 and PSI-BLAST threshold 0.005 was chosen for all the PSI-BLAST searches. To identify human proteins, the Organism option in PSI-BLAST was chosen as *homo sapiens* (taxid: 9606). Similarly, for proteins from frog, fly and worm, the corresponding Organism options were; *Xenopus laevis* (taxid: 8355), *Drosophila melanogaster* (taxid: 7227) and *Caenorhabditis elegans* (taxid:6239). Thus, the

mammalian genome, *homo sapiens*, the amphibian genome *Xenopus laevis*, the fly genome *Drosphila melanogaster* and the nematode genome *Caenorhabditis elegans* were analysed using PSI-BLAST searches to identify all ADAMs and ADAMTSs. The PSI BLAST hits were further scanned using the online SMART database [25,26] in the batch mode and INTERPRO database [27]. The E values from the BLAST output, and the SMART or INTERPRO annotation of each BLAST hit were verified before including a protein into the superfamily. The hits obtained from these methods were merged together after removing the redundant proteins. The proteins present as fragments or identical to larger proteins were also discarded from the dataset. The hits sharing 100% sequence identity with other proteins were considered as redundant and hence only the representative sequence that had the longest amino acid sequence length was included in further steps of analysis.

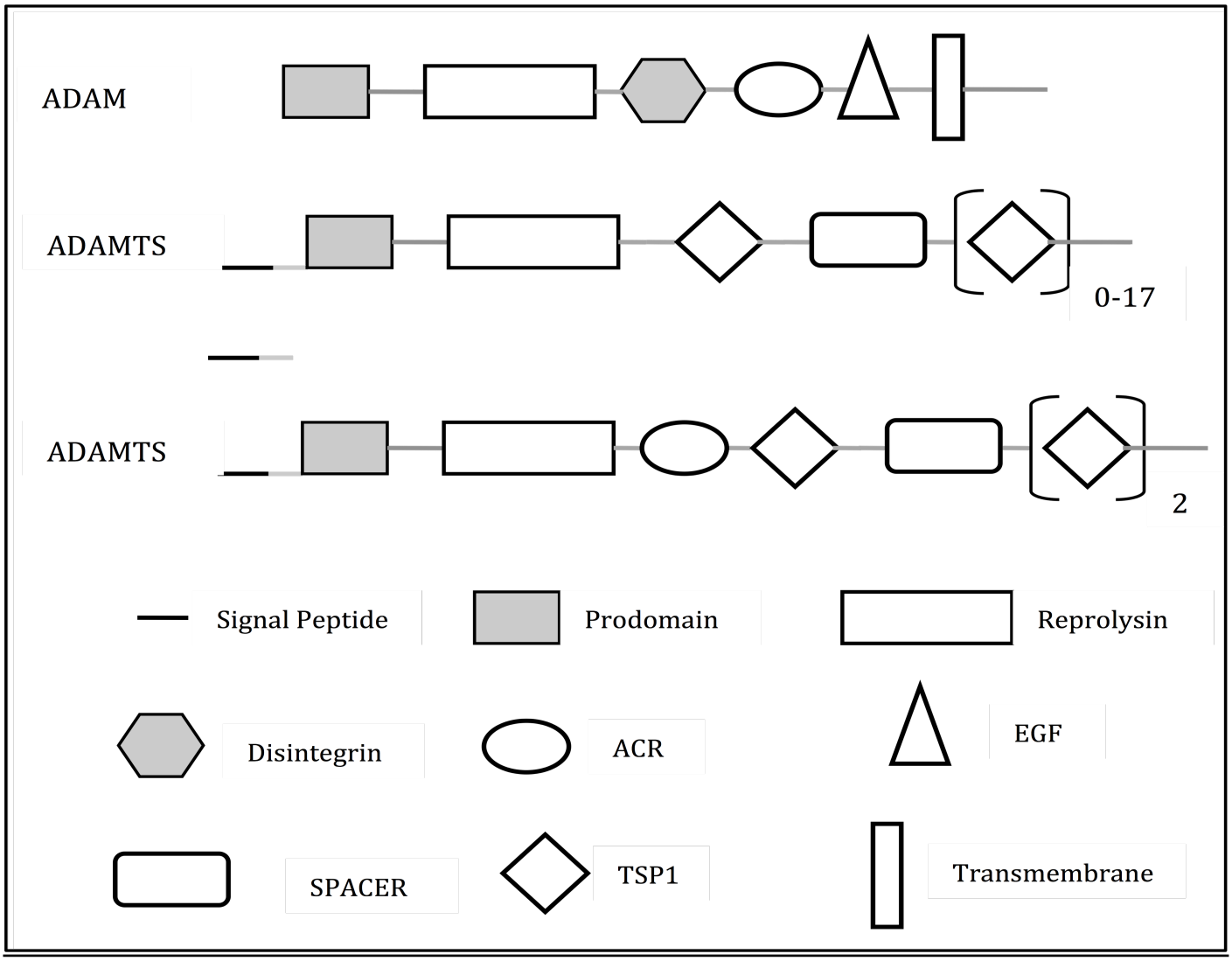
Sequence alignment and evolutionary relationship analysis

Multiple sequence alignment of the ADAM and ADAMTS proteins from human, frog, fly and worm was performed using CLUSTALW [28]. BLOSUM62 matrix, an open gap penalty of 10 and an extension penalty of 0.05 were the parameters employed for multiple sequence alignment. For generating the phylogenetic trees, Bootstrapping was performed 1000 times to obtain support values for each internal branch. Pairwise distances were determined with protpars protein parsimony method [29]. Representations of the calculated trees were constructed using TreeView [30]. The multiple sequence alignments and the phylogenetic trees were constructed separately for the full-length proteins as well as the amino acid sequence region corresponding to the zinc proteinase domain alone.

Results and Discussion

The genome sizes human, fly, and worm correspond to 3100 Mb, 180 Mb and 100 Mb respectively and are encoded by corresponding number of genes (28920, 19778, and >20,000 genes [www.ncbi.nlm.nih.gov/]). The frog genome, *Xenopus laevis* is an ongoing project at DDBJ (www.ddbj.nig.ac.jp) with a 3100 Mb genome size. PSI BLAST searches identified 235 genes belonging to the ADAM and ADAMTS protein families in human genome, and after removing redundancy there were

Figure 1: Domain organization of ADAM and ADAMTS proteins.

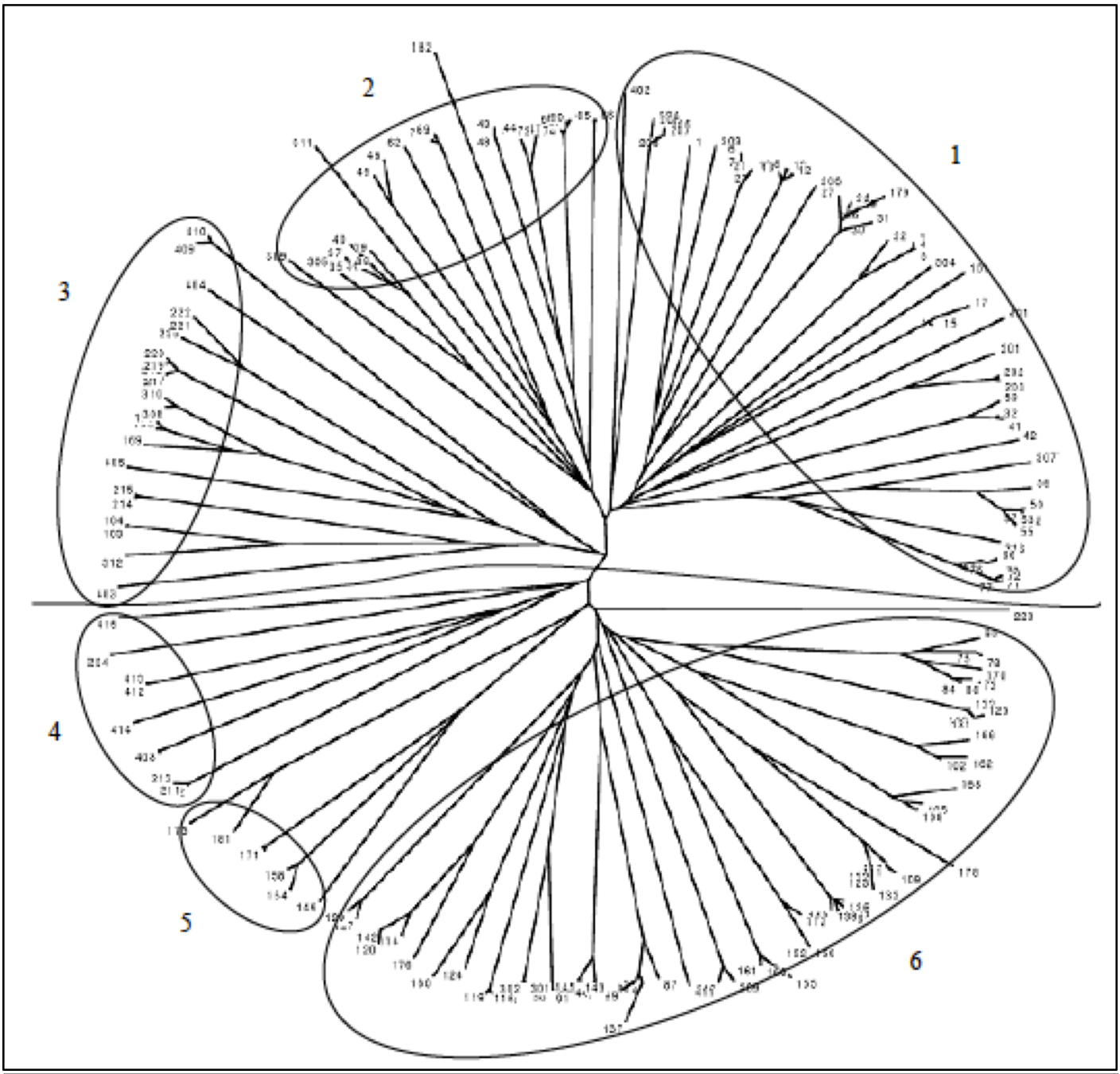


182 proteins. Among these, 90 are ADAMs and 92 are ADAMTS genes. Similarly, the frog genome is encoded by 11 ADAMs and 2 ADAMTS genes. The fly genome is encoded by 19 ADAM and 6 ADAMTS genes and the worm genome is encode by 7 ADAMs and 8 ADAMTS genes. These figures indicate that the relative ratio of the numbers ADAM and ADAMTS are comparable in human and worm, while the relative ratio of the numbers of ADAMTS are far fewer compared to ADAMs in frog and fly. A list of these proteins are provided as appendix the above mentioned proteins are unique and distinct and when present as isoforms share high sequence homology but are not identical. We discuss below the occurrence, domain organization and phylogenetic analysis of these two protein families in four distinctly different representative organisms.

Domain organization of ADAM and ADAMTS:

The ADAMs comprises variable sequence length, between 375 and 1538 amino acid residues and the ADAMTSs comprises a variable sequence length, between 344 and 2165 amino acid residues. The domain organization of these two protein families has been analyzed using the SMART and INTERPRO database. Most ADAMs follows a similar domain architecture pattern as shown in Fig 1. A typical ADAM comprises pro-domain, zinc dependent metalloproteinase, Disintegrin, ACR region followed by the EGF domain and a transmembrane segment at

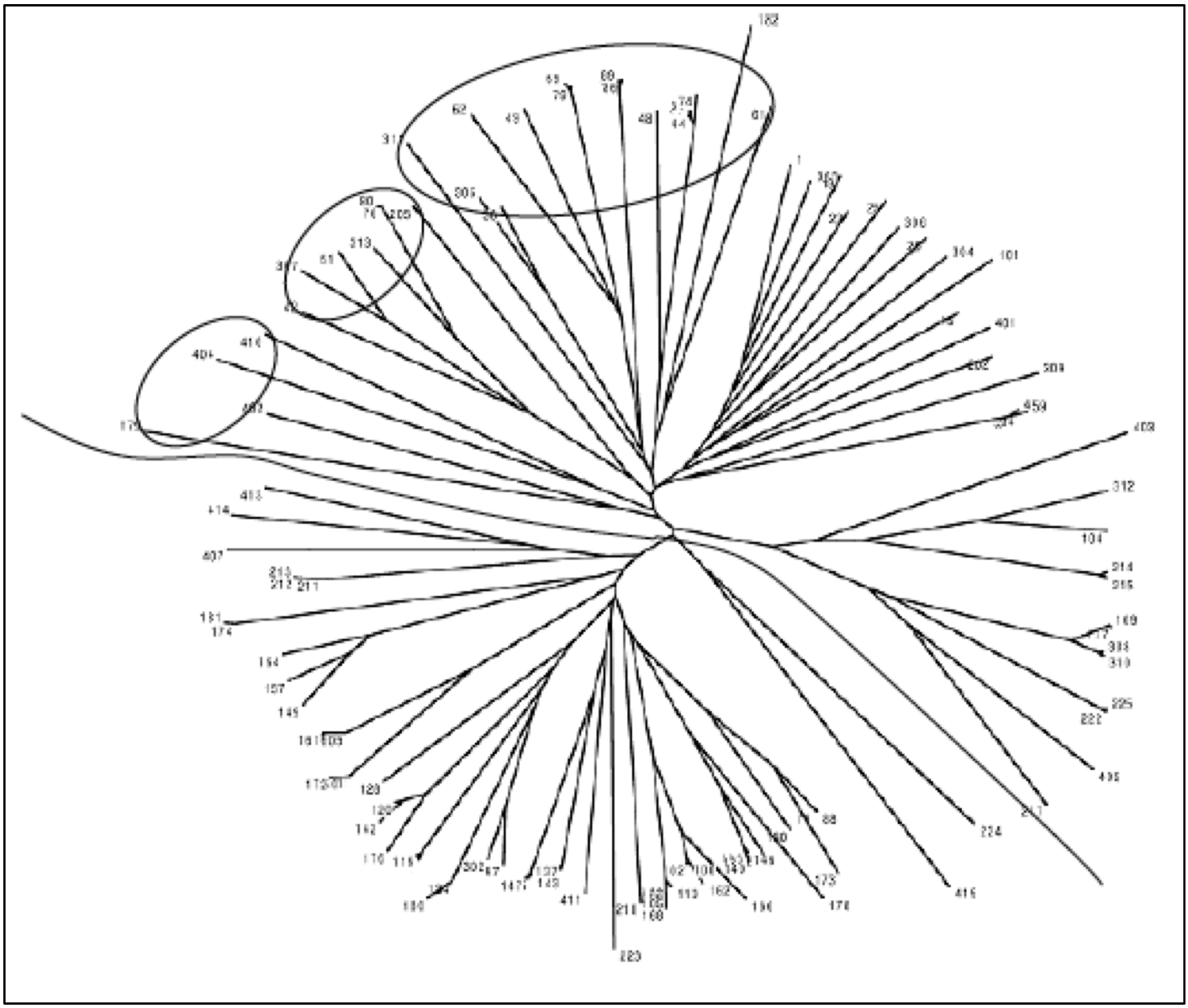
Figure 2a: the tree generated for the full length ADAM and ADAMTS members.



the C-terminus that makes ADAMs membrane bound. These domains are followed by a C-terminal cytoplasmic tail. While the order of domains remains unchanged in all the proteins, depending upon their amino acid sequence length, some of the domains either from N- or C- terminus are absent. In some instances, SMART did not identify some domains corresponding to a region, and we attribute this to the low sequence homology shared in these regions.

At the N-terminus, the soluble ADAMTS proteins comprise pro-domain and zinc dependent metalloproteinase. These two domains are followed by region that varies in length between 90-100 amino acid residues. This region identifies the ACR region of ADAMs with low sequence homology (from the second iteration of PSI-BLAST searches). This region is followed by a single TSP1 repeat, which is followed by amino acid sequence region that comprises 100-110 amino acid residues. PSI-BLAST searches identified this region to be unique to ADAMTSs and only in few very instances; ACR domain of ADAM is identified with very low sequence homology (from the third iteration of PSI-BLAST searches. This region is followed by ADAM_spacer1 domain comprising 110-120 amino acid residues and TSP1 repeats. Typically, the number of TSP1 repeats varies from 0 to 17 according to SMART. The number of TSP1 repeats depends on the length of the protein. Few exceptions to the standard domain architecture of ADAMs and ADAMTSs have been noticed. Some notable exceptions are in the drosophila proteins. For example, the protein NP_996218 has an N-terminal 165 amino acid residue insertion. Also proteins, AAC47275, ABV53679, ABV53680, AAQ22412, NP_001014481, NP_651716 and AAS48649 have specific inserts that vary in length between 155 and 264 amino acid residues. In both the cases, the insert regions are restricted to proteins from drosophila alone and may mediate specific

Figure 2b: The tree generated for the region corresponding to the zinc dependent metalloproteinase domain region for ADAMs and ADAMTSs.



interactions required by the organism.

Multiple sequence alignment:

The multiple sequence alignment of entire ADAMs and ADAMTSs was generated using CLUSTALW. According to CLUSTALW, the sequence homology between ADAMs and ADAMTSs across the four genomes analyzed can be as low as 1%. We explain that, this is partly due to the low sequence homology shared between similar domains in the proteins and mainly due to the significant variation in the lengths of protein sequences. Therefore we propose that in spite of low sequence homology, the proteins that belong to either ADAM or ADAMTS protein family are highly similar in terms of 3-D structure and function. We observe that the members of ADAM and ADAMTS are distinctly different and share similar domains only at the N-terminus (pro-domain and metalloproteinase domains).

The zinc binding catalytic region represented by REPRO domain mostly varies in length between 210 and 225 amino acid residues, and is responsible for proteolytic activity. The zinc-binding motif HExxHxxGxxH is present in most ADAMs and ADAMTS with the exception of few ADAMs in which mutations have been observed.

Phylogenetic analyses:

The phylogenetic trees were constructed for proteins belonging to the four representative organisms. To observe the differences between the ADAM and ADAMTS full length proteins, and their corresponding metalloproteinase domain alone we have built two phylogenetic trees, one using full length proteins and the second corresponding to their domain sequences alone. Figure 2a indicates the tree generated for the full length ADAM and ADAMTS members, Figure 2b indicates the tree generated for the region corresponding to the zinc dependent metalloproteinase domain region for ADAM and ADAMTS members. From the examination of the phylogenetic trees shown in Figure 2a and 2b, it is obvious that the members of ADAM and ADAMTS fall into two distinct clusters. We also observe that there are some distinct differences between the trees built for full length ADAMs and ADAMTSs and the corresponding metalloproteinase domain alone. The proteins can be organized into 6 clades according to the phylogenetic tree of full length ADAM and ADAMTS proteins. These are termed as Clades 1 to 6. Clades 1, 2 and 3 comprise ADAMs and clades 4, 5 and 6 comprise ADAMTS. The phylogenetic tree corresponding to the metalloproteinase domain is organized into 9 Clades and these are termed as Clades I to IX. The clades I to V comprise ADAMs and clades VI to IX comprise ADAMTSs. The increased number of clades in this phylogenetic tree implies that, the degree of divergence during evolution is greater within the metalloproteinase domain than compared to the full length ADAMs and ADAMTSs.

Clade1: The ADAMs that belong to this clade are 11, 22, 23, 7, neu3, 28, 8, 15, 19, 12, 13, 33, mind meld and ADAM decysin. This clade constitutes proteins from all four organisms. ADAM decysin is a secreted protein belonging to the disintegrin metalloproteinase family and its expression is up regulated during dendritic cells maturation [31]. ADAMs in drosophila is encoded by 3 isoforms of neu3 genes. Using whole-genome microarray assays, Stathopoulos colleagues [32] have showed that neu3 is expressed in broad lateral stripes in wild-type embryos, but is expressed throughout the dorsal-ventral axis of mutant embryos derived from *Tollrm9/Tollrm10* females. Another ADAM gene in drosophila, mind meld, is neuronally expressed. Neu3 and mind meld are restricted to drosophila alone. ADAM13 homolog is restricted to frog alone, while ADAM11, 22 and 15 homologs are present in both humans and frog. Two members of worm ADAM11 and ADAM22 belong to this clade. ADAM19, 15, 8 and 12 are closely related members. ADAM28 types of proteins are lymphocyte-expressed proteins and their alternative splicing results in two transcript variants. The shorter version encodes a secreted isoform, while the longer version encodes a transmembrane isoform. ADAM33 proteins are implicated in asthma and bronchial hyper-responsiveness. ADAM19 proteins serve as a marker for dendritic cell differentiation. ADAM8 proteins may be involved in cell adhesion during neurodegeneration. ADAM15 family members are type I transmembrane glycoproteins known to be involved in cell adhesion and proteolytic ectodomain processing of cytokines and adhesion molecules. Through its disintegrin-like domain, these proteins specifically interact with the integrin beta chain, beta 3. It also interacts with Src family protein-tyrosine kinases in a phosphorylation-dependent manner, suggesting that these proteins may function in cell-cell adhesion as well as in cellular signaling. ADAM7 family is composed of zinc-binding proteins that can function as adhesion proteins and/or endopeptidases. They are involved in a number of biological processes, including fertilization, neurogenesis, muscle development, and immune response. ADAM11 genes represent candidate tumor supressor genes for human breast cancer based on its location within a minimal region of chromosome 17q21 defined by tumor deletion mapping. In frog, this acts as a probable ligand for integrin in the brain and it can be detected in testis and barely expressed in heart and muscle, In developing embryos, this expression is restricted to neural crest derivatives. ADAM12 has an important role to play in myoblast differentiation. ADAM22 and 23 family members are highly expressed in the brain and may function as an integrin ligand in the brain.

The members of Clade1 are divided into Clades I, III and some members of Clade IV. The members of Clade I are present in four organisms and are active with the exception of ADAM7. The members of ADAM7 have the mutation HQ (HE) in the zinc binding sequence motif and it is unlikely that the activity is retained. The members of clade III are present in human, frog and drosophila and are all inactive. The members of this clade, ADAM11 and 23 families have all three histidines mutated and members of ADAM22 have the histidines, H1 and H3 mutated, and these proteins therefore lack the zinc proteinase activity. Further, the members of mind meld proteins from drosophila have the mutation HM (HE) in the zinc binding sequence motif and it is unlikely that the activity is retained. One ADAM15 member of clade 1 now belongs to clade IV.

This protein has the mutation HL (HE) in the zinc binding sequence motif and it is unlikely that its activity is retained.

Clade 2: The ADAMs that belongs to this clade is ADAM30, 29, 21, 20, 18, 2, 4, 9, 32. These proteins correspond to the members of human and frog. ADAM9 proteins interact with SH3 domain-containing proteins, bind mitotic arrest deficient 2 beta protein, and is also involved in TPA-induced ectodomain shedding of membrane-anchored heparin-binding EGF-like growth factor. The expression of ADAM20 and 21 is testis-specific. ADAM29 proteins encoded genes are highly expressed in testis and may be involved in human spermatogenesis. ADAM24 and 25 are expressed exclusively in testis and more specifically on the surface of mature sperm. ADAM30 is testis-specific and contains a polymorphic region, resulting in isoforms with varying numbers of C-terminal repeats. ADAM3 and 5 germ-cell specific metalloproteinase-disintegrin that may play a role in cell-cell and cell-matrix interactions during spermatogenesis. ADAM18 is a sperm surface protein. ADAM32 is similar to ADAM2. ADAM18 expressed on cell-cell and cell-matrix interactions, including fertilization, muscle development, and neurogenesis. ADAM1 and 2 are expressed on the plasma membrane of developing spermatogenic cells and sperm (This member is a subunit of an integral sperm membrane glycoprotein called fertilin, which plays an important role in sperm-egg interactions). ADAM4 and 6 have high similarity to human metargidin and it may participate in dual proteolysis and integrin-mediated cell-cell, cell-matrix interaction. We therefore conclude that all members of this clade are important for mediating cell-cell and cell-matrix interactions thus play a major role in fertilization.

The members of ADAM2, 18 and 32 have all three histidines mutated and members of ADAM21 have the histidine, H2 mutated, and these proteins therefore lack the zinc proteinase activity. The members of ADAM29 and a member of ADAM4 have the mutation HH/N (HE) in the zinc binding sequence motif respectively and this makes them unlikely to function as a proteinase. All members of clade 2 are present in clade II.

Clade 3: The ADAMs that belongs to this clade is ADAM10, 17. Members of this clade are present in all four organisms. Members of ADAM10 cleave many proteins including TNF-alpha, E-cadherin and myelin basic protein. The homolog of ADAM10 in drosophila is Kuz and is required for proper development of peripheral and central nervous system. Kuz is also required for axon extension, vein formation and wing margin formation and is involved in NOTCH mediated lateral inhibition in drosophila. Any mutations in Kuz result in production of excessive number of neurons and bristles in the central nervous system. Kuz homolog in worm is SUP-17, and is also known to modify Notch-mediated cell fare decisions in the organism. Recently it has been shown that ADAM10 is required for the formation of optic projection by xenopus retinal ganglion cell (RGC) axons and its mRNA is expressed in the dorsal neuroepithelium through which RGC axons extend [33]. In Humans, a similar role for ADAM10 in the vertebrate development is speculated. All members of ADAM10 are active.

ADAM17 proteins function as a tumor necrosis factor-alpha converting enzyme (TACE) and normal release of soluble TNF in humans. Their homologues are also present drosophila, frog and worm and are active with the exception of two members from *C. elegans*, that have the mutation HQ (HE) in the zinc binding sequence motif and it is unlikely that the zinc metalloproteinase activity is retained. Members of clade 3 are present in clade V and some members in clade IV. The members present in clade IV are four ADAMs from worm, among which only one represents an active metalloproteinase. The members of clade V are present in all organisms and are active proteins.

Clade 4: This clade constitutes ADAMTS proteins from drosophila and worm. These proteins do not belong to any known ADAMTS types of proteins, but are described as ADAMTS like subgroup members. These proteins belong to angiogenesis inhibitor homologs, ADT-2, MIG-17 and T19D2. The MIG-17 protein is secreted from muscle cells of the body wall and localizes in the basement membranes of gonad. MIG-17 is essential for its function in controlling distal tip cells (DTC) migration. All members of this clade are active. Members of clade 4 are distributed in clades VI and VII. The MIG-17 members in drosophila and worm belong to clade VII.

Clade 5: The ADAMTSs that belong to this clade are ADAMTS2, 3, 13, 14. This clade constitutes proteins only from the human genome. ADAMTS14 mainly expressed in lung, and is highly similar to ADAMTS2 and 3 and possess the aminoprocollagen peptidase activity. The ADAMTS2 gene provides instructions for making an enzyme that processes several types of procollagen molecules. Procollagens are the precursors of collagens, which are complex molecules that add strength, support, and elasticity (the ability to stretch) to many body tissues. Specifically, the ADAMTS2 enzyme clips a short chain of protein building blocks off one end of procollagens. This clipping step is necessary for the resulting collagen molecules to assemble into strong, slender fibrils outside cells. ADAMTS13 is a plasma metalloproteinase that cleaves von Willebrand factor to smaller, less thrombogenic forms. This protein mainly expressed in liver. All members of this clade are active. Members of clade 5 are present in clade VIII.

Clade 6: The ADAMTSs that belong to this clade are ADAMTS7, 9, 1, 12, 6, 10, 16, 18, 19, 17, 20, 4, 15, 8, and 5. Members of this clade are present in all four organisms. ADAMTS12 is an important enzyme that causes cartilage degradation in arthritic disorders and play important role in the development and progression of inflammatory and tumor processes. ADAMTS7 exhibited higher expression in musculoskeletal tissues, and its concentration was found to be up-regulated in the cartilage and synovium of patients with rheumatoid arthritis. ADAMTS10 and 6 have a role in inflammatory eye disease. ADAMTS10 plays a vital role in growth and in skin, lens, and heart development. ADAMTS16

is specific to human and highly expressed both in the kidney and in the ovary, where they are predominantly expressed in the parietal granulose cells of pre-ovulatory follicles but only slightly expressed in cells of the cumulus oophorus. ADAMTS16 is capable of cleaving a2-macroglobulin MG, a common substrate for proteinases, which is present at high concentrations in the follicular fluid of ovarian follicles. ADAMTS16 plays a physiological role of ovarian follicles, during the pre-ovulatory phase. ADAMTS18 functions as a tumor suppressor. ADAMTS1 has anti-angiogenic activity. The expression of these proteins is associated with various inflammatory processes as well as development of cancer cachexia. ADAMTS20 protein is also overexpressed in some human malignant tumors, including brain, colon, and breast carcinomas. ADAMTS15 is very similar to ADAMTS1 and 8. ADAMTS4 and 8 are inflammatory regulated enzymes expressed in macrophage-rich areas of atherosclerotic plaques and the ADAMTS5 proteins cleave the aggrecan interglobular domain. ADAMTS9 is a secreted, cell-surface-binding metalloproteinase that cleaves the proteoglycans versican and aggrecan. ADAMTS17 is mainly expressed in fetal tissues, especially in lung, ADAMTS19 is in kidney. ADAMTS19 is virtually undetectable in adult tissues, suggesting that this functional role is specifically restricted to processes occurring during human fetal development. The members of this clade appear to participate in vital roles in the developmental biology of various organs in the four representative organs. One member of ADAMTS16 (EAX08118) has the histidine, H3 mutated and this protein will therefore not have the zinc proteinase activity. We report that this is the only inactive ADAMTS in the dataset analysed. Members of clade 6 are present in clade IX.

Conclusions

In this work we have present a comprehensive study and an early bioinformatic overview of the ADAM and ADAMTS families in the human, *drosophila*, frog and worm genomes. Classification of the zinc proteinase proteins on the basis of domain architecture and cellular localization helps us to associate the proteins to the various biochemical pathways and the different cellular niches. Shuffling of domains or modules among the various ADAM and ADAMTS seems to be each one adapting to diverse biological roles. Reports of several novel domain combinations indicate an increase in the functional repertoire. We identified a different domain architecture pattern in ADAMTS protein family which is not as similar as to the previous report. Comparison of domain architectures in human ADAM and ADAMTS with frog, *drosophila* and worm reveals non-linear inheritance of several domain architectures or deselection of specific architecture across the genomes. The presence of several unexpected domains provides insights into the unknown regions of several known protein families. Such whole genome surveys and cross-genome comparisons using computations should be useful to design rational experiments and enhance our understanding of the specific biological roles of the ADAM and ADAMTS family proteins.

Acknowledgement

KKI and Hemalatha thanks CSIR, New Delhi, India for senior research fellowship. LGP thanks UPE and CMSD, University of Hyderabad for providing computational facilities.

References

1. Wolfsberg TG, White JM: **ADAMs in fertilization and development**. *Developmental biology* 1996, **180**:389-401.
2. Bode W, Gomis-Rüth F-X, Stöckler W: **Astacins, serralysins, snake venom and matrix metalloproteinases exhibit identical zinc-binding environments (HEXXHXXGXXH and Met-turn) and topologies and should be grouped into a common family, the 'metzincins'**. *FEBS letters* 1993, **331**:134-140.
3. Skiles JW, Gonnella NC, Jeng AY: **The design, structure, and therapeutic application of matrix metalloproteinase inhibitors**. *Current medicinal chemistry* 2001, **8**:425-474.
4. Lovejoy B, Hassell AM, Luther MA, Weigl D, Jordan SR: **Crystal structures of recombinant 19-kDa human fibroblast collagenase complexed to itself**. *Biochemistry* 1994, **33**:8207-8217.
5. Kveiborg M, Albrechtsen R, Rudkjaer L, Wen G, Damgaard-Pedersen K, Wewer UM: **ADAM12-S Stimulates Bone Growth in Transgenic Mice by Modulating Chondrocyte Proliferation and Maturation**. *Journal of Bone and Mineral Research* 2006, **21**:1288-1296.
6. Holgate ST, Yang Y, Haitchi H-M, Powell RM, Holloway JW, Yoshisue H, Pang YY, Cakebread J, Davies DE: **The genetics of asthma: ADAM33 as an example of a susceptibility gene**. *Proceedings of the American Thoracic Society* 2006, **3**:440-443.
7. Tanabe C, Hotoda N, Sasagawa N, Sehara-Fujisawa A, Maruyama K, Ishiura S: **ADAM19 is tightly associated with constitutive Alzheimer's disease APP α-secretase in A172 cells**. *Biochemical and biophysical research communications* 2007, **352**:111-117.
8. Melenhorst W, Van Den Heuvel M, Stegeman C, Van Der Leij J, Huitema S, Van Den Berg A, Van Goor H: **Upregulation of ADAM19 in chronic allograft nephropathy**. *American journal of transplantation* 2006, **6**:1673-1681.
9. Rocks N, Paulissen G, Calvo FQ, Polette M, Gueders M, Munaut C, Foidart J-M, Noel A, Birembaut P, Cataldo D: **Expression of a disintegrin and metalloprotease (ADAM and ADAMTS) enzymes in human non-small-cell lung carcinomas (NSCLC)**. *British journal of cancer* 2006, **94**:724-730.
10. Takada H, Imoto I, Tsuda H, Nakanishi Y, Ichikura T, Mochizuki H, Mitsufuji S, Hosoda F, Hirohashi S, Ohki M: **ADAM23, a possible tumor suppressor gene, is frequently silenced in gastric cancers by homozygous deletion or aberrant promoter hypermethylation**. *Oncogene* 2005, **24**:8051-8060.
11. Kuefer R, Day KC, Kleer CG, Sabel MS, Hofer MD, Varambally S, Zorn CS, Chinnaiyan AM, Rubin MA, Day ML: **ADAM15 disintegrin is associated with aggressive prostate and breast cancer disease**. *Neoplasia (New York, NY)* 2006, **8**:319.
12. Oppezzo P, Vasconcelos Y, Settegrana C, Jeannel D, Vuillier F, Legarff-Tavernier M, Kimura EY, Bechet S, Dumas G, Brissard M: **The LPL/ADAM29 expression ratio is a novel prognosis indicator in chronic lymphocytic leukemia**.

Blood 2005, **106**:650-657.

13. Shigemura K, Sung SY, Kubo H, Arnold RS, Fujisawa M, Gotoh A, Zhau HE, Chung LW: **Reactive oxygen species mediate androgen receptor-and serum starvation-elicited downstream signaling of ADAM9 expression in human prostate cancer cells**. *The Prostate* 2007, **67**:722-731.
14. Porter S, Span PN, Sweep FC, Tjan-Heijnen VC, Pennington CJ, Pedersen TX, Johnsen M, Lund LR, Rømer J, Edwards DR: **ADAMTS8 and ADAMTS15 expression predicts survival in human breast carcinoma**. *International journal of cancer* 2006, **118**:1241-1247.
15. Cerretti DP, DuBose RF, Black RA, Nelson N: **Isolation of two novel metalloproteinase-disintegrin (ADAM) cDNAs that show testis-specific gene expression**. *Biochemical and biophysical research communications* 1999, **263**:810-815.
16. Frayne J, Hurd EA, Hall L: **Human tMDC III: a sperm protein with a potential role in oocyte recognition**. *Molecular human reproduction* 2002, **8**:817-822.
17. Giovanna M, Ng K, Paradiso L, Godde NJ, Kaye A, Novak U: **ADAM22, expressed in normal brain but not in high-grade gliomas, inhibits cellular proliferation via the disintegrin domain**. *Neurosurgery* 2006, **58**:179-186.
18. Solomon DH, Stedman M, Licari A, Weinblatt ME, Maher N: **Agreement between patient report and medical record review for medications used for rheumatoid arthritis: The accuracy of self-reported medication information in patient registries**. *Arthritis Care & Research* 2007, **57**:234-239.
19. Goertsches R, Comabella M, Navarro A, Perkal H, Montalban X: **Genetic association between polymorphisms in the< i> ADAMTS14</i> gene and multiple sclerosis**. *Journal of neuroimmunology* 2005, **164**:140-147.
20. Ng YH, Zhu H, Pallen CJ, Leung PC, MacCalman CD: **Differential effects of interleukin-1β and transforming growth factor-β1 on the expression of the inflammation-associated protein, ADAMTS-1, in human decidual stromal cells in vitro**. *Human Reproduction* 2006, **21**:1990-1999.
21. Davies JA, Bowen DJ: **An association between the L1565 variant of von Willebrand factor and susceptibility to proteolysis by ADAMTS13**. *haematologica* 2007, **92**:240-243.
22. Mitani H, Takahashi I, Onodera K, Bae J-W, Sato T, Takahashi N, Sasano Y, Igarashi K, Mitani H: **Comparison of age-dependent expression of aggrecan and ADAMTSs in mandibular condylar cartilage, tibial growth plate, and articular cartilage in rats**. *Histochemistry and cell biology* 2006, **126**:371-380.
23. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool**. *Journal of molecular biology* 1990, **215**:403-410.
24. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs**. *Nucleic acids research* 1997, **25**:3389-3402.
25. Schultz J, Milpetz F, Bork P, Ponting CP: **SMART, a simple modular architecture research tool: identification of signaling domains**. *Proceedings of the National Academy of Sciences* 1998, **95**:5857-5864.
26. Letunic I, Copley RR, Pils B, Pinkert S, Schultz J, Bork P: **SMART 5: domains in the context of genomes and networks**. *Nucleic acids research* 2006, **34**:D257-D260.
27. Mulder N, Apweiler R: **InterPro and InterProScan**. In *Comparative genomics*. Edited by: Springer; 2007:59-70.
28. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice**. *Nucleic acids research* 1994, **22**:4673-4680.
29. Felsenstein J: **Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods**. *Methods in enzymology* 1996, **266**:418-427.
30. Page RD: **TreeView: an application to display phylogenetic trees on personal computers**. *Comput Appl Biosci* 1996, **12**:2.
31. Mueller CG, Rissoan M-C, Salinas B, Ait-Yahia S, Ravel O, Bridon J-M, Briere F, Lebecque S, Liu Y-J: **Polymerase chain reaction selects a novel disintegrin proteinase from CD40-activated germinal center dendritic cells**. *The Journal of experimental medicine* 1997, **186**:655-663.
32. Stathopoulos A, Levine M: **Whole-Genome Expression Profiles Identify Gene Batteries in< i> Drosophila</i>**. *Developmental cell* 2002, **3**:464-465.
33. Chen Y, Hehr CL, Atkinson-Leadbeater K, Hocking JC, McFarlane S: **Targeting of retinal axons requires the metalloproteinase ADAM10**. *The Journal of neuroscience* 2007, **27**:8448-8456.

Visit us @ www.vedicjournals.com : DOI: http://dx.doi.org/10.14259/bp.v2i1.92

Authors Column