academicJournals

Vol. 15(8), pp. 252-263, 24 February, 2016
DOI: 10.5897/AJB2015.14818
Article Number: 7A1E93B57331
ISSN 1684-5315

http://www.academicjournals.org/AJB

African Journal of Biotechnology

Full Length Research Paper

Identification of accelerated evolution in the metalloproteinase domain of snake venom metalloproteinase sequences (SVMPs) through comparative analysis

Mahmudul Islam[1]*, Zahida Yesmin Roly[1], Mirza Ruhullah[1], Zohorul Islam[1], Alsan Sharia[1], Rafsan Zani Tanvir[2] and Khaja Foyasal[2]

[1]Department of Genetic Engineering and Biotechnology, Faculty of Life and Earth Sciences, University of Rajshahi, Rajshahi-6205, Bangladesh.
[2]Department of Biotechnology and Genetic Engineering, Islamic University, kushtia-7003, Bangladesh.

Received 25 June, 2015, Accepted 31 August, 2015

Computational protein sequence analysis is one of the most important tools used for understanding the evolution of closely related proteins sequences including snake venom metalloproteinase sequences (SVMPs) which give valuable information regarding genetic variations. The fundamental objective of the present study is to screen the evolution distributed in metalloproteinase domain regions of protein sequences among different SVMPs in snake species which are involved in a range of pathological disorders such as arthritis, atherosclerosis, liver fibrosis, cardiovascular, cancer, liver and neurodegenerative disorders. In fact, SVMPS are responsible for hemorrhage and may also interfere with the hemostatic system. A comparative characterization of the metalloproteinase sequences has been carried out to analyze their multiple sequence alignment, phylogenic tree, homology, physico-chemical, secondary structural and functional properties. DNAMAN software was used for multiple sequence alignment, phylogenic tree and homology and Expasy's Prot-param server was used for amino acid composition, physico-chemical and functional characterization of these SVMPs sequences. Studies of secondary structure of these SVMPs were carried out by computational program. Based on the observed patterns of occurrence of atypical features, we hypothesize that amino acids of metalloproteinase domain region (66.63% identity) of protein sequences are highly changeable; whereas, signal peptide region (93.98% identity) is the lowest changeable protein sequence and the remaining other three domains such as propeptide region (87.36% identity), desintegrin domain region (78.63% identity) and cysteine-rich domain region (75.70% identity) show moderate changeable protein sequence. SVMPs might be an accelerated evolution, which is a key player in causing diseases. From the data, it can be suggested that over -changed metalloproteinase domain regions in snake venom metalloproteinase might be responsible for the generation of functional variation of proteins expressed, which in turn may lead to different disorders in humans after snake bite. The results of this study would be an effective tool for the study of mutation, drugs resistance mechanisms and development of new drugs for different diseases.

Key words: SVMPs, evolution, multiple sequence alignment, phylogenic tree, secondary structure, homology.

INTRODUCTION

Metalloproteinase is a ubiquitous enzyme that exists in nearly all organisms from animal to plants. However, apart from its different expression sites in different plants and animals for performing distinct physiological roles, metalloproteinase also exists in the toxin/venom of several venomous creatures (snake, caterpillar, scorpion etc.) to cause agony, suffering and even death of the prey/victim. Among them, snake venom is a very rich source of metalloproteinase and they are termed as snake venom metalloproteinases (SVMPs). Several diseases are shown to be associated with metalloproteinase. For example, genetic polymorphisms in matrix metalloproteinase genes MMP1, MMP9 and MMP12 are shown to be important in the development of chronic obstructive pulmonary disease (COPD) (Wallace and Sandford, 2002). Metalloproteinases also play role in the development of renal cyst (Obermüller et al., 2001), uterine cervical carcinoma (Libra et al., 2009), angiogenesis (Pepper, 2001) and various inflammatory diseases of the central nervous system such as bacterial meningitis.

SVMPs are more abundantly found in viper snake venom; however, they are also from few elapid families (Birrell et al., 2007; Fry et al., 2003). They are synthesized as zymogens in the venom gland and contain a pro-peptide which is cleaved off during maturation. They have a common zinc-binding motif with a consensus sequence of HEXXHXXGXXH (Bode et al., 1993). They are classified into different types (PI to P-IV) on the basis of the other domains that are present in these complexes (Hite et al., 1994). These families of enzymes are responsible for haemorrhagic, local myonecrotic, antiplatelet, edema-inducing and other inflammatory effects. Recently, it has been shown that SVMPs are potential tools in the development of drugs for the prevention and treatment of several illnesses. These enzymes are extensively used in the treatment and prevention of thrombotic disorders, since they serve as defibrinogenating agents (Costa et al., 2010; Bjarnason and Fox, 1994). Animal models of septic shock have also delivered proof-of-concept that MMPs can be of therapeutic interest (Vanlaere and Libert, 2009).

Evolution and diversification of snake venom is a very interesting phenomenon. Snake venom glands are believed to have evolved by the modification of the salivary glands, and various body proteins have been recruited in the venom gland and adapted to attack and damage various physiological system of the prey (Reza et al., 2006). Therefore, study of the expressed venom protein among and within a particular family enables us to understand the mode and direction of evolution of that gene family. A lot of variation is evident in the SVMPs among the species even within the same species with indication of accelerated evolution of this particular venom component. Therefore, this study was undertaken to perform a detailed bioinformatics analysis of the different domains of snake venom metalloproteinase sequences in order to understand their pattern of accelerated evolution.

MATERIALS AND METHODS

Sequence retrieval

Twelve (12) SVMPs sequences from different venomous snake species including *Agkistrodon contortrix laticinctus, Agistrodon piscivorus leucostoma, Deinagkistrodon acutus, Gloydius halys, Sistrurus catenatus edwardsi, Naja naja atra, Bothrops jararaca, Bothrops insularis, Protobothrops flavoviridis, Bungarus multicinctus, Crotalus viridis* and *Crotalus atrox* were obtained from National Center for Biotechnology Information (NCBI) with the following accession numbers: O42138, C9E1S0, Q9W6M5, Q8AWI5, ABG26979, A8QL59, O93523, Q8QG88, Q90ZI3, ABN72537, C9E1R8 and Q9DGB9, respectively.

Multiple sequence alignment

Twelve (12) SVMPs sequences from different venomous snake species were used for multiple sequence alignment of the species, with the aid of DNAMAN software. After multiple sequence alignment of SVMPs sequences of the different snake species, black color indicates positions with fully conserved residue; pink color indicates that one of the following high scoring groups is fully conserved; cyan color indicates that one of the following moderate scoring groups is fully conserved; white color indicates that one of the following 'weaker' scoring groups is fully conserved.

Classification of SVMPs sequences

SVMPs sequences were divided into five domains: signal peptide, propeptide, metalloprotease, disintegrin and cysteine-rich domain based on their domain organization after multiple sequence alignment. Signal peptide sequence (about 18 residues long), propeptide (about 176 residues long), metalloproteinase (about 205 residues long), desintegrin (about 95 residues long) and cyestine-rich (about 194 residues long) domains are aligned separately. Cys-switch sites (PKMCGV) and Zn^{2+} binding motifs (HEXXHXXGXXH) are marked in the box of black color.

Phylogenetic tree and homology construction

Phylogenetic tree of 12 SVMPs sequences was done using

*Corresponding author. E-mail: mahmudulgebru@gmail.com.

Table 1. The name of twelve SVMPs sequences of different venomous snake species with number of accession and amino acids

Snake species	Accession number	Number of amino acids
Agkistrodon contortrix laticinctus	O42138	620
Agkistrodon piscivorus leucostoma	C9E1S0	613
Deinagkistrodon acutus	Q9W6M5	610
Gloydius halys	Q8AWI5	610
Sistrurus catenatus edwardsi	ABG26979	612
Naja atra	A8QL59	621
Bothrops jararaca	O93523	610
Bothrops insularis	Q8QG88	610
Protobothrops flavoviridis	Q90ZI3	612
Bungarus multicinctus	ABN72537	614
Crotalus viridis viridis	C9E1R8	609
Crotalus atrox	Q9DGB9	610

molecular evolutionary genetic analysis (MEGA) software (version 4.0.02) (Tamura et al., 2007), with UPGMA method. Each node was tested using the bootstrap approach by taking 1,000 replicates; the bootstrap analysis indicates strong support. Homology of 12 SVMPs sequences was done using DNAMAN software.

Analysis of physico-chemical properties

The SVMPs sequences were utilized as the input data type to compute the percentage of amino acid composition (%) (Islam et al., 2013), molecular weight, theoretical isoelectric point (pI), number of positively and negatively charged residues, extinction coefficient, instability and aliphatic index, Grand Average of Hydropathy (GRAVY), using Expasy Protparam tool (http://web.expasy.org/protparam).

Analysis of secondary structure

SOPMA tool (Self-Optimized Prediction Method with Alignment) of NPS@ (Network Protein Sequence Analysis) server was used to characterize the secondary structural features of the proteins such as, alpha helix, 310 helix, Pi helix, beta bridge, extended strand, beta turn, bend region, random coil, ambiguous and other states (Geourjon and Deleage, 1995; Roly et al., 2014a, Islam et al., 2015).

Analysis of functional properties

The analysis of the selected 12 SVMPs sequences was done with the help of Motif scan (http://myhits.isb-sib.ch/cgibin/motif_scan) tool (Roly et al., 2014b). The input data type was in FASTA format and scanned against 'PROSITE Patterns' which is a selected protein profile.

RESULTS AND DISCUSSION

In our present investigation, the NCBI database was used as source to collect SVMPs sequences from different venomous snake species with accession number (Table 1). A total of 12 SVMPs sequences (after removing the duplicates and partial sequences) were obtained from different venomous snake species. SVMPs sequences were reckoned into five domains: signal peptide, propeptide, metalloproteinase, desintegrin and cyestine-rich domains based on their domain organization. Some researchers reported same result (Brust et al., 2013; Casewell, 2012; Ryan et al., 2003). Signal peptides of all the sequences are highly conserved and they are nearly identical. There are 18 residues in signal peptide which show 93.98% identity (Figure 1). However, the 13th residues in five sequences (*A. c. laticinctus, D. acutus, B. jararaca, B. insularis, C. v. viridis*) are Alanine while the remaining two (*A. p. leucostoma, G. halys, S. c. edwardsi, N. n. atra, P. flavoviridis, B. multicinctus, C. atrox*) is valin. However, as the properties of these two amino acid residues are almost the same we do not expect any change in signaling the secretion of the protein or in the removal of the signal peptide after secretion of the protein. Propeptide sequences of all the sequences are highly conserved and they are nearly identical. They are about 176 residues long showing 87.36% (Figure 2). The Cys-switch site (PKMCGV) within the propeptide is in the position of 165th residues (Figure 2). Cys-switch site (PKMCGV) is a short peptide of prodomain and is blocking the active site of metalloproteinase. When this peptide is removed, metalloproteinase is active. Metalloproteinase domains are 205 residues long and they have 66.63% identity (Figure 3). Desintegrin domains are approximately 95 residues long and have 78.63% identity. Same sort of grouping like metalloproteinase is also evident in the Desintegrin domain. The cystine-rich domains are 194 residues long and show 75.70% identity. In this study we showed that amino acids of metalloproteinase domain

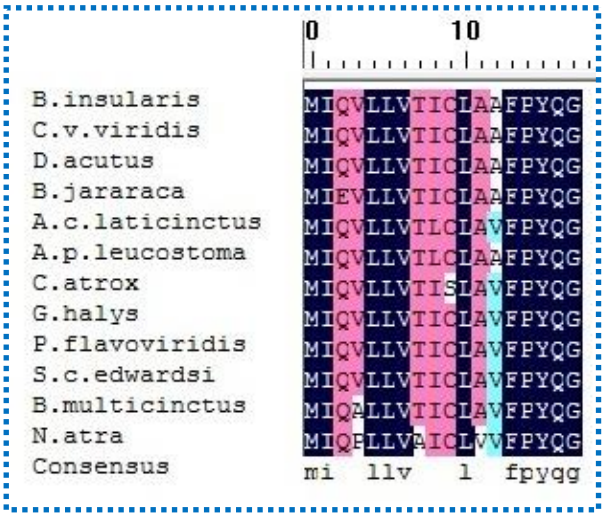

Figure 1. Multiple sequence alignment of signal peptide (93.98% identity) of SVMPs sequences: black color indicates positions with fully conserved residue; pink color indicates that one of the following high scoring groups is fully conserved; cyan color indicates that one of the following moderate scoring groups is fully conserved; white color indicates that one of the following 'weaker' scoring groups is fully conserved.

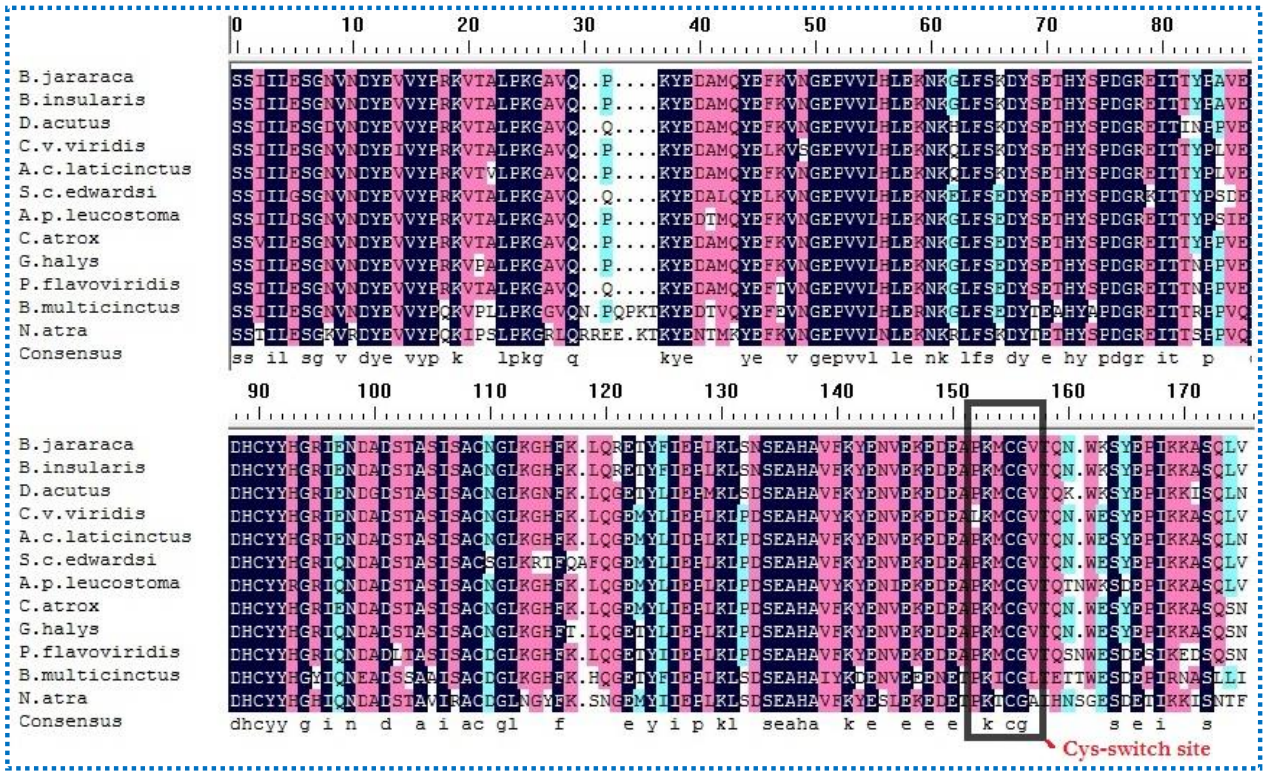

Figure 2. Multiple sequence alignment of propeptide (87.36% identity) of SVMPs sequences: black color indicates positions with fully conserved residue; pink color indicates that one of the following high scoring groups is fully conserved; cyan color indicates that one of the following moderate scoring groups is fully conserved; white color indicates that one of the following 'weaker' scoring groups is fully conserved.

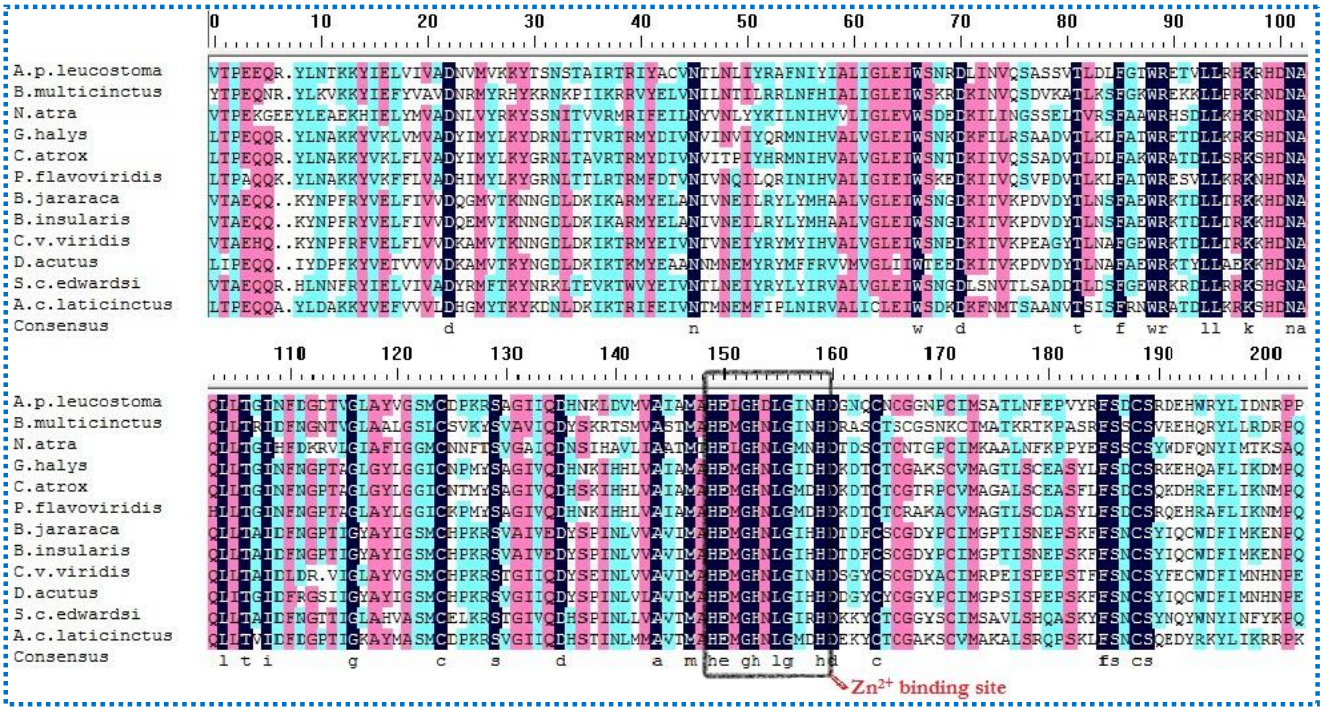

Figure 3. Multiple sequence alignment of metalloproteinase domain (66.63% identity) of SVMPs sequences: black color indicates positions with fully conserved residue; pink color indicates that one of the following high scoring groups is fully conserved; cyan color indicates that one of the following moderate scoring groups is fully conserved; white color indicates that one of the following 'weaker' scoring groups is fully conserved.

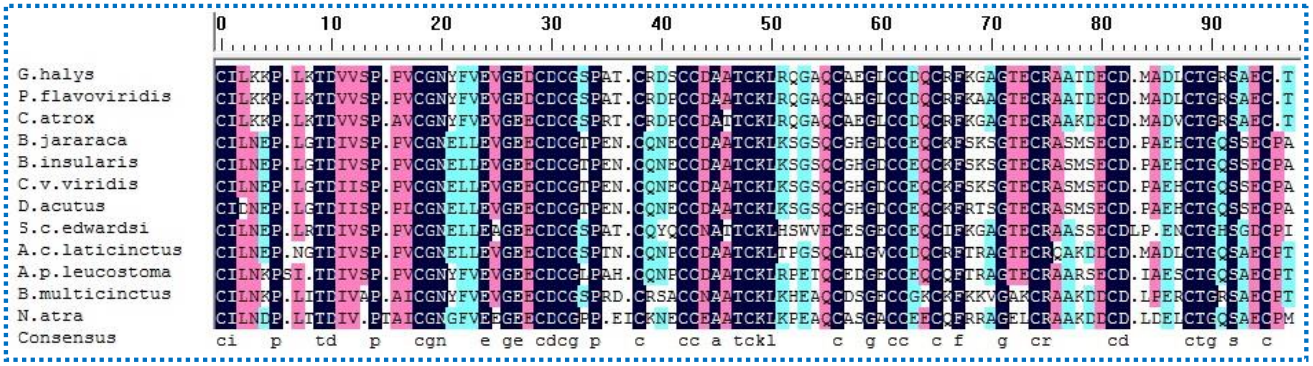

Figure 4. Multiple sequence alignment of desintegrin domain (78.63% identity) of SVMPs sequences: black color indicates positions with fully conserved residue; pink color indicates that one of the following high scoring groups is fully conserved; cyan color indicates that one of the following moderate scoring groups is fully conserved; white color indicates that one of the following 'weaker' scoring groups is fully conserved.

region were more changeable due to synonymous and non-synonymous mutation (Figure 3) and have very low identity; whereas signal peptide domain region was very less changeable and has the highest percentage similarity among different SVMPs sequences. The remaining other three domains: propeptide (Figure 2), desintegrin (Figure 4) and cyestine-rich domains (Figure 5) were moderately changeable and showed moderate percentage identity.

Phylogenetic tree and homology indicate that metalloproteinase domain has very high distance relationship (Figure 8A and B) among twelve SVMPs; on

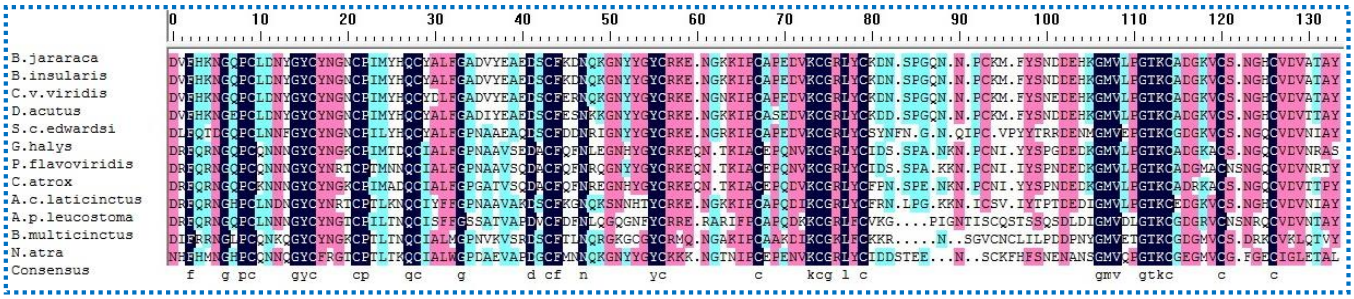

Figure 5. Multiple sequence alignment of cyestine-rich domain (75.70% identity) of SVMPs sequences: black color indicates positions with fully conserved residue; orange color indicates that one of the following high scoring groups is fully conserved; blue color indicates that one of the following moderate scoring groups is fully conserved; white color indicates that one of the following 'weaker' scoring groups is fully conserved.

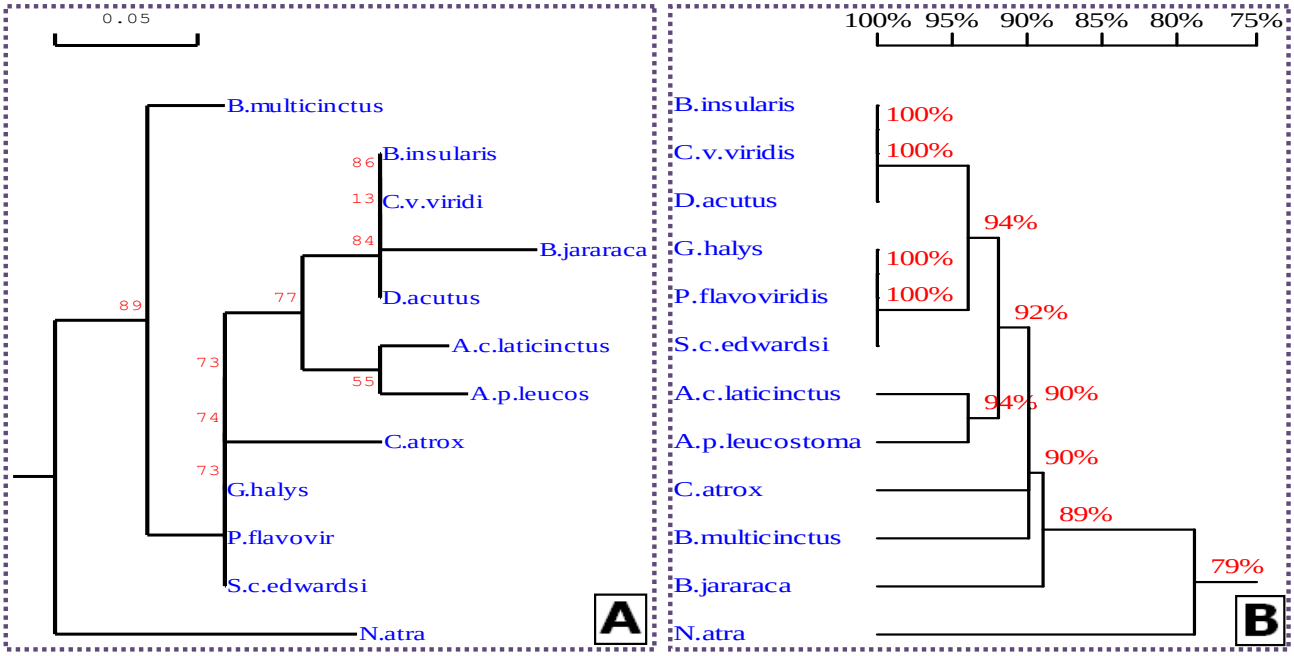

Figure 6. (A) Phylogenetic tree construction using by DNAMAN software of signal peptide of 12 SVMPS sequences from different venomous snake species using the bootstrap approach by taking 1,000 replicates and the bootstrap analysis indicates strong support **(B)** Homology construction of signal peptide using by DNAMAN software of 12 SVMPS sequences from different venomous snake species.

the other hand, signal peptide has very close distance relationship (Figure 6A and B) while the other remaining three domains (Propetide, Desintegrin domain and Cyestine rich domain) show moderate distance relationship (Figures 7A, 7B, 9A, 9B, 10A and 10B respectively)

In the analysis of amino acid composition, the percentage of cysteine residues in majority of the SVMPs sequences lies in the range of 6.1-6.7%; SVMPs sequences of *B. multicinctus*, *C. v. viridis* and *A. p. leucostoma* show a significant increase with values of 6.7, 6.6 and 6.5 percent, respectively (Table 2). The highest quantity of cysteine residues in *B. multicinctus* and *C. v. viridis* SVMPs sequences might be correlated with presence of cysteine switch motif and role of these SVMPs s in pathological conditions. These gelatinases have been early associated with several disorders such as carcinomas, cardio-vascular and so on. Highly significant presence of cysteine suggests its role as a critical residue for SVMPs activity and thus these SVMPs may be investigated for possible role in diseased

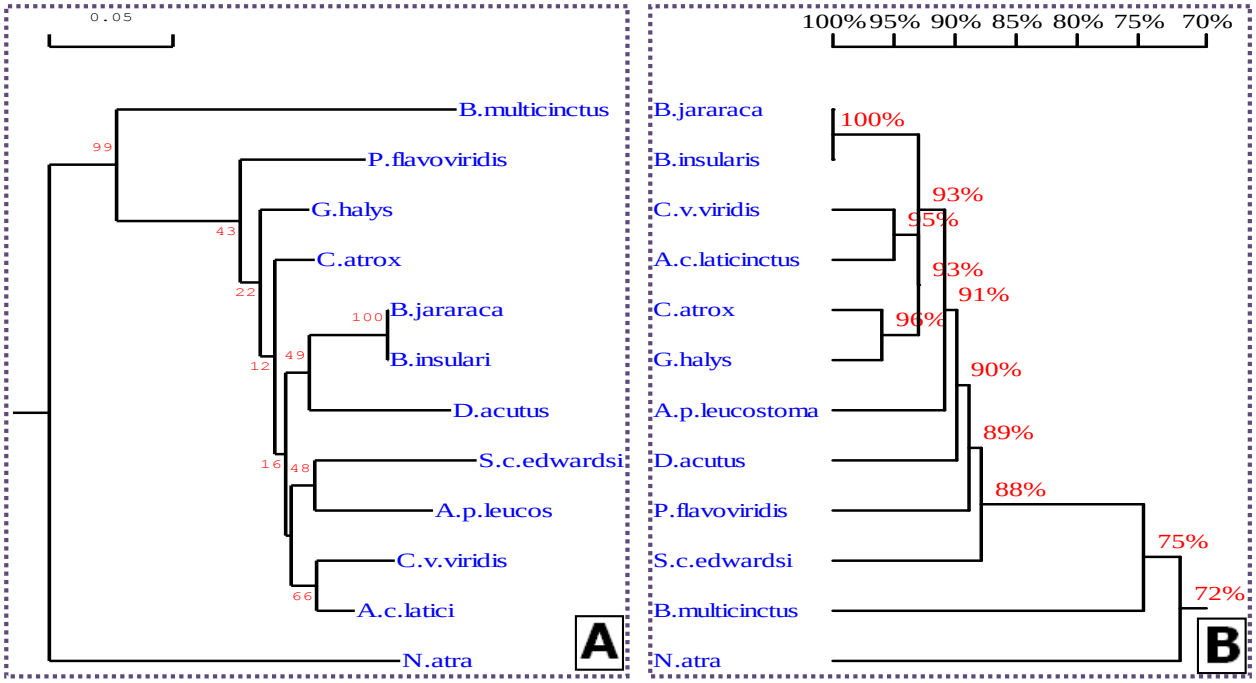

Figure 7. (A) Phylogenetic tree construction using by DNAMAN software of propeptide of 12 SVMPS sequences from different venomous snake species using the bootstrap approach by taking 1,000 replicates and the bootstrap analysis indicates strong support **(B)** Homology construction of signal peptide using by DNAMAN software of 12 SVMPS sequences from different venomous snake species.

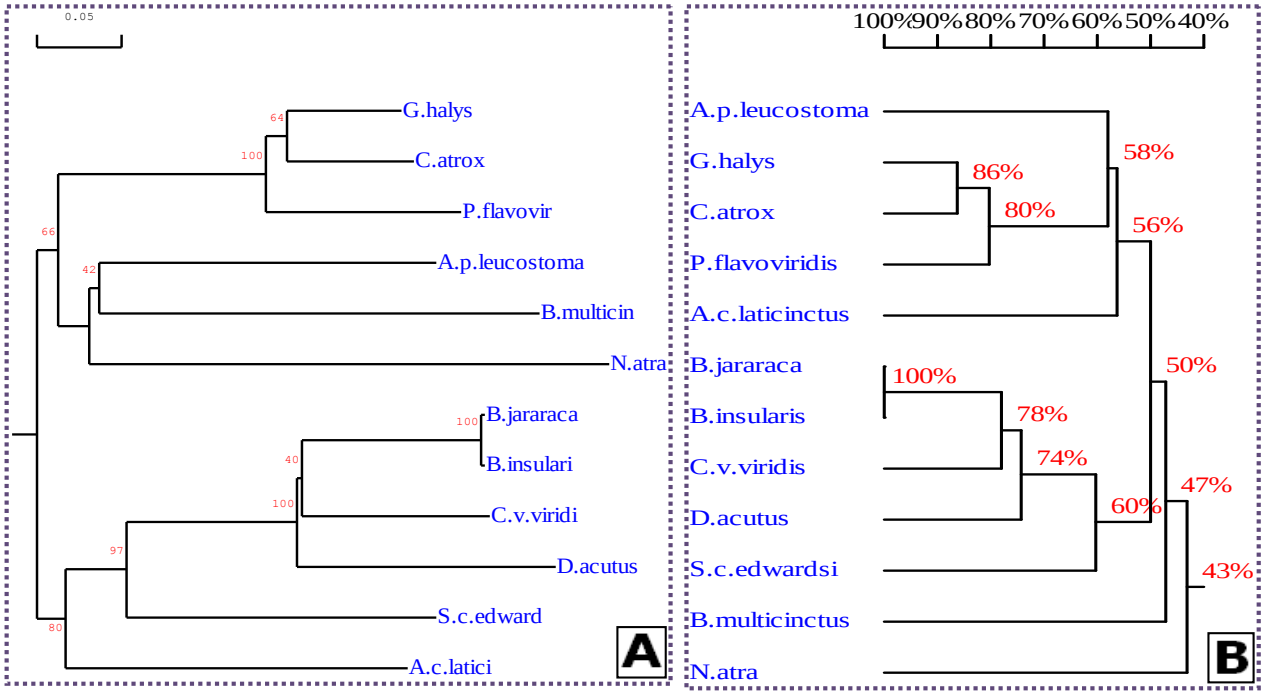

Figure 8. (A) Phylogenetic tree construction using by DNAMAN software of metalloproteinase domain of 12 SVMPS sequences from different venomous snake species using the bootstrap approach by taking 1,000 replicates and the bootstrap analysis indicates strong support **(B)** Homology construction of signal peptide using by DNAMAN software of 12 SVMPS sequences from different venomous snake species.

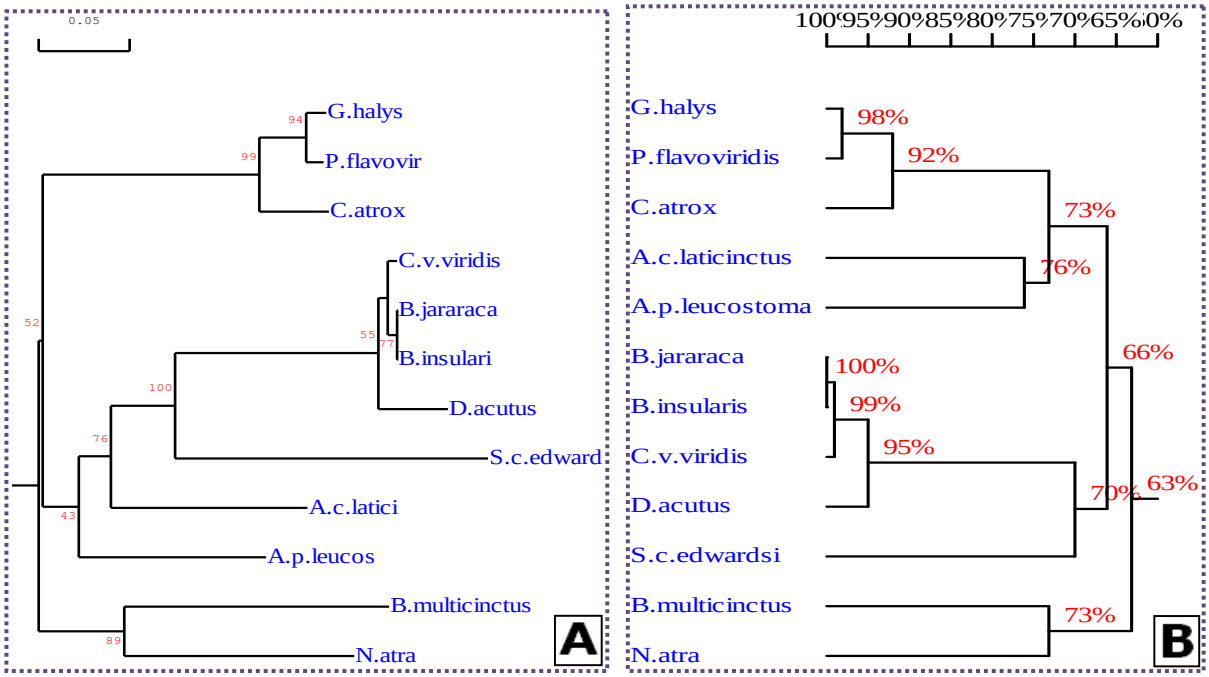

Figure 9. (A) Phylogenetic tree construction using by DNAMAN software of desingtegrin domain of 12 SVMPS sequences from different venomous snake species using the bootstrap approach by taking 1,000 replicates and the bootstrap analysis indicates strong support **(B)** Homology construction of signal peptide using by DNAMAN software of 12 SVMPS sequences from different venomous snake species.

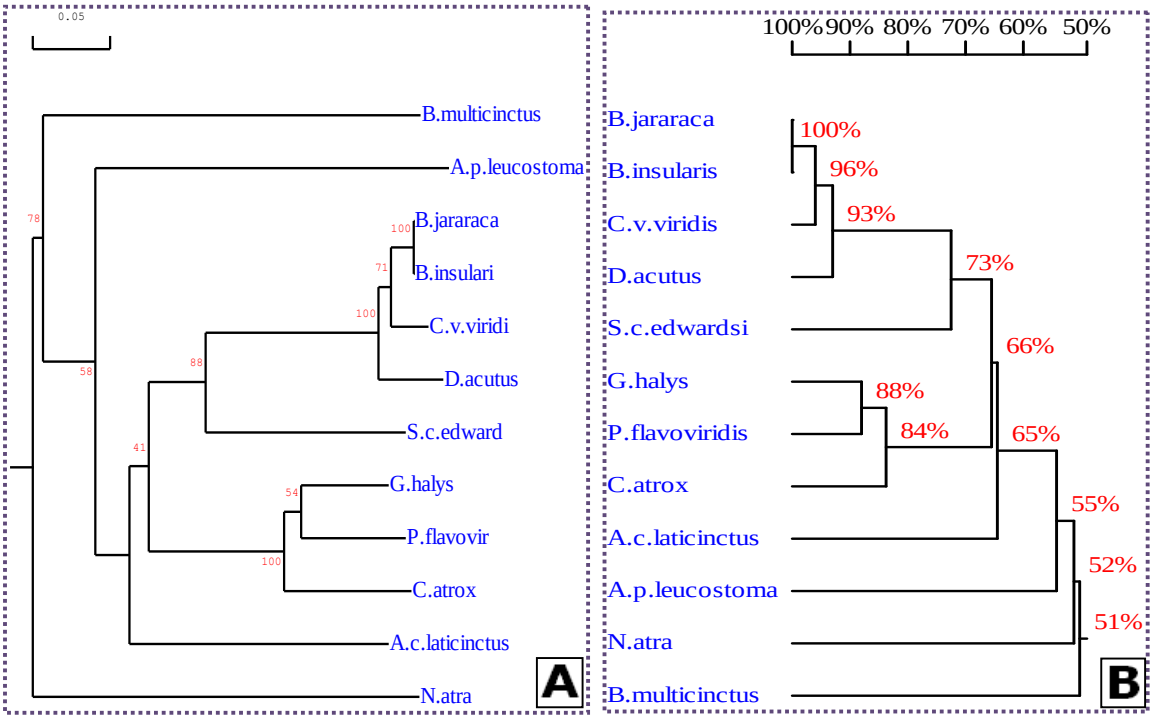

Figure 10. (A) Phylogenetic tree construction using by DNAMAN software of cyestine-rich domain of 12 SVMPS sequences from different venomous snake species using the bootstrap approach by taking 1,000 replicates and the bootstrap analysis indicates strong support **(B)** Homology construction of signal peptide using by DNAMAN software of 12 SVMPS sequences from different venomous snake species.

Table 2. Amino acid compositions of twelve SVMPs sequences of different venomous snake species (in %).

Species	Ala	Arg	Asn	Asp	Cys	Gln	Glu	His	Gly	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
A. c. laticinctus	5.8	3.5	6.3	6.8	6.3	3.9	5.8	6.0	2.4	5.2	6.8	7.9	2.9	3.1	5.2	5.6	5.6	0.5	4.5	6.0
A. p. leucostoma	6.0	4.7	6.5	6.5	6.5	4.4	6.0	6.8	2.1	6.0	7.0	4.6	1.6	3.1	4.9	6.2	5.7	0.7	4.4	6.2
D. acutus	5.1	2.0	5.7	6.6	6.4	3.1	7.9	7.5	3.1	6.1	5.2	7.4	3.1	3.0	5.2	5.9	3.9	0.7	6.4	5.7
G. halys	7.5	3.6	6.2	6.4	6.4	4.3	5.9	7.0	2.6	4.8	7.0	6.6	2.5	2.6	4.8	5.2	4.9	0.5	4.9	6.2
S. c. edwardsi	5.9	3.9	6.5	5.2	6.2	4.2	7.2	7.0	2.6	5.1	7.5	5.1	1.6	2.8	4.2	6.2	4.9	1.0	6.5	6.2
N. atra	5.6	3.1	6.9	5.0	6.1	3.1	8.5	7.4	3.1	6.1	6.9	6.4	2.4	3.5	4.5	5.6	5.6	0.6	4.3	5.0
B. jararaca	6.1	2.1	6.4	6.2	6.4	3.3	7.5	7.0	3.1	4.9	5.9	7.2	2.5	3.1	5.2	5.9	4.1	0.7	5.9	6.4
B. insularis	6.1	2.1	6.4	6.2	6.4	3.4	7.5	6.9	3.1	4.9	5.9	7.2	2.5	3.1	5.2	5.9	4.1	0.7	5.9	6.4
C. atrox	7.0	3.9	5.7	6.2	6.2	4.1	6.1	7.0	2.8	4.4	6.7	6.6	2.6	3.0	5.1	5.2	5.2	0.5	5.1	6.4
C. v. viridis	5.6	2.5	6.2	5.9	6.6	3.1	8.3	7.1	3.3	5.1	6.6	6.6	2.6	2.6	4.9	5.9	4.1	0.7	6.1	6.2
P. flavoviridis	7.5	3.9	6.5	6.4	6.4	4.9	5.6	6.4	2.8	5.2	6.9	6.4	2.5	2.9	4.6	4.9	5.4	0.5	4.4	6.0
B. multicinctus	6.0	6.0	5.9	5.2	6.7	3.4	6.0	6.7	2.3	5.5	7.3	8.3	1.6	2.4	4.7	5.7	5.0	0.5	4.6	6.0

conditions. Further analysis of the amino acid composition can help to place amino acid presence at remarkable level and be correlated with precise pathological conditions (Shckorbatov et al., 2008).

Other physico-chemical parameters also signify the behavior of SVMPs in different conditions (Table 3). pH values for majority of the SVMPs (*A. c. laticinctus, D. acutus, B. jararaca, B. insularis, C. v. viridis*, *A. p. leucostoma, G. halys, S. c. edwardsi, N. n. atra, P. flavoviridis & C. atrox*) lie in the acidic range (pH<7); while for the only one SVMPs sequence (*B. multicinctus*), it increases in the alkaline range (pH>7). In addition, the instability index SVMPs sequences of different snake species of *A. c. laticinctus, D. acutus, B. jararaca, B. insularis, C. v. viridis*, *G. halys, S. c. edwardsi, N. n. atra & C. atrox* were stable (Instability index <40), but the remaining were unstable metalloproteinases (Instability index <40) (Table 3). Secondary structural analysis indicates a pre-dominance of random coils, followed by α-helices, extended strands and β-turns in 12 SVMPs sequences while the extended strands exceed α-helices in *A. c. laticinctus, A. p. leucostoma, D. acutus, N. n. atra, B. jararaca, B. insularis, B. multicinctus* and *C. v. viridis* (Table 4). This is very useful to predict three dimensional structures of proteins and can also help in approximation of some aspects of protein function and their classification into families (Rost, 2001).

Furthermore, the Motif Scan tool predicts the presence of a cysteine switch, a zinc protease and desintegrin motif in SVMPs sequences which have been the subject of discussion in various literatures (Table 5). The cysteine switch regulates activity of SVMPs sequences via complex formation between cysteine residue of prodomain and zinc atom of catalytic domain (Van Wart et al., 1990). Cys-switch site (PKMCGV) motif is present in the propeptide (Figure 2) and blocks the active site of metalloproteinase domain; and when this peptide is removed metalloproteinase is active. The primary sequence motif HExxH is present in the catalytic domain of zinc-dependant SVMPs sequences. The two conserved histidine residues coordinate the zinc atom and the glutamic acid residue is a member of the active site of enzyme (Devault et al., 1988). The zinc binding region signature has been characterized as (uncharged)-(uncharged)-H-E-(uncharged) -(uncharged)-H-(uncharged)-(hydro phobic) (Jongeneel et al., 1989). Zinc protease motif is present within the catalytic domain (metalloproteinase domain) of SVMPs sequences (Table), playing a pivotal role in the collagen binding region of these enzymes.

Conclusion

Intensive characterization and comparative analysis of the SVMPs sequence of proteins with the help of numerous bio-computational tools yielded new insights and perspectives which can

Table 3. Physico-chemical parameters of twelve SVMPs sequences of different venomous snake species.

Species	No. of A.A.	M.W (Da)	pI	"-" charged residues	"+" charged residues	Extinction coefficient	Instability index	Aliphatic index	GRAVY
A. c. laticinctus	620	69512.2	6.14	78	71	60595	34.26	69.66	-0.452
A. p. leucostoma	613	68212.0	5.17	77	57	64605	41.15	74.91	-0.355
D. acutus	610	68542.4	5.03	88	57	82485	37.11	65.84	-0.462
G. halys	610	67651.7	5.71	75	62	63575	39.28	71.64	-0.389
S. c. edwardsi	612	68930.7	5.27	76	55	94975	34.61	72.96	-0.399
N. atra	621	69402.5	5.25	84	59	64605	39.88	70.98	-0.411
B. jararaca	610	68213.0	5.15	84	57	78015	35.29	66.80	-0.451
B. insularis	610	68284.0	5.15	84	57	78015	34.76	66.80	-0.456
C. atrox	610	67960.1	5.90	75	64	65065	36.86	69.07	-0.417
C. v. viridis	609	68364.1	5.01	88	55	79505	40.85	69.15	-0.449
P. flavoviridis	612	68191.4	5.97	73	63	59105	42.25	72.21	-0.396
B. multicinctus	614	68988.1	8.73	69	88	60720	42.57	73.68	-0.487

Table 4. Secondary structural features of twelve SVMPs sequences of different venomous snake species (in %).

Snake species	A helix	310 Helix	Pi Helix	β Bridge	Extended strand	β Turn	Bend region	Random coil	Ambiguous states	Other states
A. c. laticinctus	22.42%	0.00%	0.00%	0.00%	24.68%	11.13%	0.00%	41.77%	0.00%	0.00%
A. p. leucostoma	16.48%	0.00%	0.00%	0.00%	29.53%	9.95%	0.00%	44.05%	0.00%	0.00%
D. acutus	23.28%	0.00%	0.00%	0.00%	27.21%	10.00%	0.00%	39.51%	0.00%	0.00%
G. halys	26.23%	0.00%	0.00%	0.00%	23.61%	9.67%	0.00%	40.49%	0.00%	0.00%
S. c. edwardsi	29.90%	0.00%	0.00%	0.00%	24.18%	8.82%	0.00%	37.09%	0.00%	0.00%
N. atra	23.67%	0.00%	0.00%	0.00%	24.96%	10.79%	0.00%	40.58%	0.00%	0.00%
B. jararaca	21.97%	0.00%	0.00%	0.00%	25.41%	10.82%	0.00%	41.80%	0.00%	0.00%
B. insularis	22.62%	0.00%	0.00%	0.00%	25.08%	10.49%	0.00%	41.80%	0.00%	0.00%
P. flavoviridis	25.65%	0.00%	0.00%	0.00%	25.33%	10.78%	0.00%	38.24%	0.00%	0.00%
B. multicinctus	21.82%	0.00%	0.00%	0.00%	25.57%	10.59%	0.00%	42.02%	0.00%	0.00%
C. v. viridis	23.81%	0.00%	0.00%	0.00%	25.62%	11.82%	0.00%	38.75%	0.00%	0.00%
C. atrox	26.89%	0.00%	0.00%	0.00%	22.62%	22.62%	0.00%	40.82%	0.00%	0.00%

be used to identify accelerated evolution of SVMPs sequence of proteins of different venomous snake species that play a crucial role in pathological conditions. In this study, multiple sequence alignment, phylogenetic tree, homology, physico-chemical, secondary structural and functional analysis of SVMPs sequence of proteins of different venomous snake species was carried out. The findings through this study may be used by researchers working on metalloproteinase of SVMPs in the context of any experimental system. So, from the identity comparison we can say that metalloproteinase domain is more diverse and under the evolutionary pressure. The amino acid composition shows a considerably high percentage of cysteine residues in *B. multicinctus* and *C. v. viridis* of SVMPs sequences, which might be a key player in pathological conditions. Future studies with the help of experimental research and test need to be carried out to validate this proposal. This study may be taken as a prototype for similar *in silico* investigational studies with regard to other large proteins families, where such comparative analysis might aid in giving a direction and help to rationalize the conduct of experimentation; it will also be very helpful to develop new drugs.

Conflicts of interest

The authors have not declared any conflict of interests.

ABBREVIATIONS

SVMP, Snake venom metalloproteinase sequences;

Table 5. Motifs found in of twelve SVMPs sequences of different venomous snake species.

Snake species	Motif found	Motif ID	Description	Start	End
A. c. laticinctus	ZINC_PROTEASE	PS00142	*Neutral zinc metallopeptidases, zinc-binding region signature.*	**332**	**341**
	DISINTEGRIN_1	PS00427	*Disintegrins signature*	**443**	**462**
A. p. leucostoma	ZINC_PROTEASE	PS00142	*Neutral zinc metallopeptidases, zinc-binding region signature.*	**333**	**342**
	DISINTEGRIN_1	PS00427	*Disintegrins signature*	**444**	**463**
D. acutus	ZINC_PROTEASE	PS00142	*Neutral zinc metallopeptidases, zinc-binding region signature.*	**331**	340
	DISINTEGRIN_1	PS00427	*Disintegrins signature*	**442**	461
G. halys	ZINC_PROTEASE	PS00142	*Neutral zinc metallopeptidases, zinc-binding region signature.*	**332**	341
	DISINTEGRIN_1	PS00427	*Disintegrins signature*	**443**	462
S. c. edwardsi	ZINC_PROTEASE	PS00142	*Neutral zinc metallopeptidases, zinc-binding region signature.*	**333**	342
	DISINTEGRIN_1	PS00427	*Disintegrins signature*	**444**	463
N. atra	ZINC_PROTEASE	PS00142	*Neutral zinc metallopeptidases, zinc-binding region signature.*	**339**	348
	DISINTEGRIN_1	PS00427	*Disintegrins signature*	**450**	469
B. jararaca	ZINC_PROTEASE	PS00142	*Neutral zinc metallopeptidases, zinc-binding region signature.*	**331**	340
	DISINTEGRIN_1	PS00427	*Disintegrins signature*	**442**	461
B. insularis	ZINC_PROTEASE	PS00142	*Neutral zinc metallopeptidases, zinc-binding region signature.*	**331**	340
	DISINTEGRIN_1	PS00427	*Disintegrins signature*	**442**	461
P. flavoviridis	ZINC_PROTEASE	PS00142	*Neutral zinc metallopeptidases, zinc-binding region signature.*	**333**	342
	DISINTEGRIN_1	PS00427	*Disintegrins signature*	**444**	463
B. multicinctus	ZINC_PROTEASE	PS00142	*Neutral zinc metallopeptidases, zinc-binding region signature.*	**338**	347
	DISINTEGRIN_1	PS00427	*Disintegrins signature*	**449**	468
C. v. viridis	ZINC_PROTEASE	PS00142	*Neutral zinc metallopeptidases, zinc-binding region signature.*	**330**	339
	DISINTEGRIN_1	PS00427	*Disintegrins signature*	**441**	460
C. atrox	ZINC_PROTEASE	PS00142	*Neutral zinc metallopeptidases, zinc-binding region signature.*	**332**	341
	DISINTEGRIN_1	PS00427	*Disintegrins signature*	**443**	462

NCBI, National Center for Biotechnology Information; **MEGA**, molecular evolutionary genetic analysis.

ACKNOWLEDGEMENTS

We are thankful to Md. Zohorul Islam (Student of Genetic Engineering and Biotechnology Department, University of Rajshahi, Rajshahi-6205, Bangladesh) for their valuable support and discussions related to this research and we are also grateful to Protein Science Laboratory for instrument facility.

REFERENCES

Birrell GW, Earl S, Wallis TP, Masci PP, de Jersey J, Gorman JJ, Lavin MF (2007). The diversity of bioactive proteins in Australian snake venoms. Mol. Cell. Proteomics 6:973-986.

Bjarnason JB, Fox JW (1994). Hemorrhagic metalloproteinase from snake venom. Pharmacol. Theor. 62:325-372.

Bode W, Gomis-Rüth FX, Stockler W (1993). Astacins, serralysins, snake venom and matrix metalloproteinases exhibit identical zinc-binding environments (HEXX-HXXGXXH and Met-turn) and topologies and should be grouped into a common family, the 'metzincins'. Febs Lett. 331:134-140.

Brust A, Sunagar K, Undheim EA, Vetter I, Yang DC, Casewell NR, Jackson TNW, Koludarov I, Alewood PF, Hodgson WC, Lewis RJ, King GF, Antunes A, Hendrikx I, Fry BG (2013). Differential evolution

and neofunctionalization of snake venom metalloprotease domains. Mol. Cell Proteomics. 12(3):651-663.
Casewell NR (2012). On the ancestral recruitment of metalloproteinases into the venom of snakes. Toxicon, 60(4):449-454.
Costa Jde O, Fonseca KC, Garrote-Filho MS, Cunha CC, de Freitas MV, Silva HS, Araújo RB, Penha-Silva N, de Oliveira F (2010). Structural and functional comparison of proteolytic enzymes from plant latex and snake venoms. Biochimie. 92(12):1760-1765.
Devault A, Sales V, Nault C, Beaumont A, Roques B, Crine P, Boileau G (1988). Exploration of the catalytic site of endopeptidase 24.11 by site-directed mutagenesis Histidine residues 583 and 587 are essential for catalysis. FEBS Lett. 231(1):54-58.
Fry BG, Wüster W, Kini RM, Brusic V, Khan A, Venkataraman D, Rooney AP (2003). Molecular evolution and phylogeny of elapid snake venom three-finger toxins. J. Mol. Evol. 57(1):110-129.
Geourjon C, Deleage G (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput. Appl. Biosci. 11(6):681-684.
Hite LA, Jia LG, Bjarnason JB, Fox JW (1994). cDNA sequences for four snake venom metalloproteinases: structure, classification, and their relationship to mammalian reproductive proteins. Arch. Biochem Biophys. 308(1):182-191.
Islam MM, Haque ME, Dutta AK, Islam MA, Khalekuzzaman M, Sikdar B (2013). A Computational Study on Metalloproteinase in the Snake Venomous of Different Rattle Snake Species in Comparison with the Asian Snake Species. Int. J. Pharm. Bio. Sci. 4(4):741-748.
Islam Z, Islam MM, Saha S, Jahangir CA, Basak B, Islam MN, Islam MS, Paul S, Khalekuzzaman M (2015). Identification and Computational Analysis of Chicken Alpha-1 Collagen Sequences. Int. J. Sci. Eng. Res. 6(1):217-221.
Jongeneel CV, Bouvier J, Bairoch A (1989). A unique signature identifies a family of zinc-dependent metallopeptidases. FEBS let. 242(2):211-214.
Libra M, Scalisi A, Vella N, Clementi S, Sorio R, Stivala S, Demetrios A, Spandidos, Mazzarino C (2009). Uterine cervical carcinoma: Role of matrix metalloproteinases. Int. J. Oncol. 34:897-903.
Obermüller N, Morente N, Kränzlin N, Gretz N, Witzgall R (2001). A possible role for metalloproteinases in renal cyst development. Renal. Physiol. 280(3):540-550.
Pepper MS (2001). Role of the matrix metalloproteinase and plasminogen activator–plasmin systems in angiogenesis. Arterioscler Thromb. Vasc. Biol. 21(7):1104-1117.
Reza MA, Swarup S, Kini RM (2006). Molecular Evolution Caught In Action: Gene Duplication and Evolution of Molecular Isoforms of Prothrombin Activators in *Pseudonaja textilis* (Brown Snake). J. Thromb. Haemost. 4:1346-1353.
Roly ZY, Hasan SN, Ferdaus KMKB, Reza MA (2014). Predicted structure model of Bungarotoxin from Bungarus fasciatus snake. Bioinformation 10(10):617.
Roly ZY, Islam MM, Reza MA.(2014). A comparative in silico characterization of functional and physicochemical properties of 3FTx (three finger toxin) proteins from four venomous snakes. Bioinformation 10(5):281-287.
Rost B (2001). Review: protein secondary structure prediction continues to rise. J. Struct. Biol. 134(2):204-218.
Ryan CA, Pearce G (2003). Systemins: a functionally defined family of peptide signals that regulate defensive genes in Solanaceae species. Proc Natl Acad Sci. 100(2):14577-14580.
Shckorbatov YG, Zhuravleva LA, Navrotskaya VV, Miroshnichenko EV, Montvid PY, Shakhbazov VG, Sutushev TA (2008). Chromatin structure and the state of human organism. Cell Biol. Internat. 2005(29):77-81.
Tamura K, Dudley J, Nei M, Kumar S (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24(8):1596-1599.
Vanlaere I, Libert C (2009). Matrix metalloproteinases as drug targets in infections caused by gram-negative bacteria and in septic shock. Clin. Microbiol. Rev. 22(2):224-239.
Van Wart HE, Birkedal-Hansen H (1990). The cysteine switch: a principle of regulation of metalloproteinase activity with potential applicability to the entire matrix metalloproteinase gene family. Proc. Natl. Acad. Sci. 87(14):5578-5582.
Wallace AM, Sandford AJ (2002). Genetic polymorphisms of matrix metalloproteinases: functional importance in the development of chronic obstructive pulmonary disease? Am. J. Pharmacogenomics 2(3):167-175.