Sivarathri Siva Rajesh et al /J. Pharm. Sci. & Res. Vol.3(12), 2011,1612-1618

ISSN:0975-1459
Journal of Pharmaceutical Sciences and Research
www.jpsr.pharmainfo.in

TFTX: A Computational Tool for Predicting Subfamilies of Three-Finger Toxins from the Venom of Elapid Snakes

Sivarathri Siva Rajesh and Thirunavukkarasu Sivaraman*

School of Chemical and Biotechnology, SASTRA University,
Thanjavur – 613401,Tamil Nadu, India.
* sivaram@scbt.sastra.edu

Abstract
Venoms from *elapidae* snakes are rich sources of over 100 protein toxins and most of them have been grouped into six superfamilies. Of the six superfamilies, three-finger toxin (3FTx) family contains most toxic principles such as cardiotoxins and α-neurotoxins. The proteins belonging to these two members of 3FTx superfamily are highly similar in their primary, secondary and tertiary structures. Notwithstanding their extreme structural similarities, they are drastically differing from each other in their biological functions. Thus, authentic functional assays are prerequisites for unambiguously annotating these protein toxins and the assays demand highly pure protein samples, sound experimental knowledge and sophisticated instrumentations. In these backgrounds, we have developed a computational tool, TFTX, which identifies cardiotoxins and α-neurotoxins on the basis of their primary structures. The robustness of the tool has been validated using authentic sequences of cardiotoxins and α-neurotoxins reported in the literature. Moreover, the TFTX is powerful to differentiate cardiotoxins and α-neurotoxins of snake venoms from the three-finger proteins of various organisms that are not toxic in nature. The uniqueness of the TFTX has also been dealt through a comparative analysis of the tool with existing computational means such as 'NTXpred server', 'SUPERFAMILY server' and 'phylogenetic methods' for the structural classifications. The program is publicly available through the web server at *http://feat.sastra.edu/TFTX.html*

Keywords:
Annotations, α-neurotoxins, cardiotoxins, server, three-finger toxins and venoms.

1. INTRODUCTION

The venoms of elapid snakes are mixture of over 100 bioactive polypeptides and enzymes [1]. Most of the biomacromolecules have been grouped into any one of six superfamilies [2]: three-finger toxin (3FTx) family, proteinase inhibitors family, C-type lectins family, phospholipase A2 (PLA2) family, serine proteinases family and metalloproteinases family. The proteins belonging to each superfamily have their unique three-dimensional structures and functions. However, each member (subfamily) of a family may drastically differ in their target sites/molecules on which they elicit their functions. For instance, three-finger toxin family have following subfamilies: cardiotoxins (CTXs), α-neurotoxins (α-NTXs), κ-bungarotoxins, muscarinic toxins, fasciculins, calciseptine and dendroaspins [3]. Of these, CTXs and α-NTXs (which are further classified into short α-neurotoxins (SNTXs) and long α-neurotoxins (LNTXs)) are the major components of 3FTx family and the lethality of the cobra venoms are attributed to these protein toxins [4]. Despite the high degree of similarities in their three-dimensional (3D) structures of these members, to each other, CTXs exhibit hemolytic activity and depolarization of muscle cells whereas α-NTXs act on the acetylcholine receptors at the post-synaptic level of the neuromuscular junction [5]. The lacuna on the biological activities of these protein toxins is still so as to understand how the structurally homologous proteins drastically differ in their functions [6,7].

In order to annotate the sequences of CTXs and α-NTXs at structural and functional levels, the 3D structures and biological activities of the polypeptides need to be well characterized, respectively. Unfortunately, many three-finger protein toxins (3FTxs) have been reported without appropriate annotations and few 3FTxs have ambiguous annotations in the literature. Moreover, the molecular moulds of 3FTxs are not restricted to venoms of snakes only because other organisms including human beings have proteins (which are not toxic in nature) depicting the folds similar to that of 3FTxs [8,9]. Thus, 3FTxs from snake venoms should be differentiated from 3FTxs of other organisms for the purpose of having unambiguous annotations. It should be mentioned that servers such as NTXpred [10], SUPERFAMILY [11] and phylogenetic methods [12] are unable to unambiguously differentiate the CTXs from α-NTXs of snake venoms (refer details in the 'results and discussion' part). In the present study, we have developed a software tool, TFTX, which differentiate CTXs, SNTXs and LNTXs from each other on the basis of their primary structures. The program accounts 46 unique parameters derived from the sequences of the

protein toxins and employs various scoring tactics to identify the toxins. The functions and robustness of the program have been elaborately dealt in 'methods' and 'results and discussion' sections of the article. The TFTX is implemented using CGI-PERL language [13], which is platform-independent, and HTML [14] is used as the front-end of the program. We believe that the TFTX will be useful for toxicologists and structural biologists to unambiguously annotate the 3FTxs for which sequences are known. The program is publicly available at *http://feat.sastra.edu/TFTX.html*

2. METHODS

The amino acid sequences of CTX, SNTX, LNTX and other 3FTxs were collected from proteins primary databases of NCBI (*www.ncbi.nlm.nih*) and SWISSPROT (*www.swissprot.com*). As on July 2011, 269 amino acid sequences of CTXs and α-NTXs from elapid snakes reported in the literature were collected and they were subjected to SignalP [15] and ProP [16] to eliminate signal peptides and propeptides that may present in the sequences, respectively. All the pre-processed sequences (SQs) were then classified on the basis of their literature annotations and multible sequence alignments performed using MultAlign [17] as follows: CTX (80 SQs), CTX-homologous (11 SQs), SNTX (65 SQs), SNTX-homologous (27 SQs), LNTX (50 SQs), LNTX-homologous (32 SQs) and NTX-like proteins (4 SQs). Using authentic sequences of CTXs and α-NTXs, 46 parameters were defined for each sub-family and the parameters were divided into four categories like general parameters (such as sequence length, molecular weight, aliphatic index [18,19], hydropathy index [20], acid-base ratio and net charge at neutral pH), residue-specific parameters (frequency of occurrence of each standard amino acid in the sequences), position-specific parameters and signature peptides. Presences of certain amino acids at particular positions were found to be unique for each subfamily of 3FTX-superfamily and we could successfully derive 12 parameters under the position-specific category. For instance, CTXs have Ser residue invariantly at position 46/47; SNTXs have Ser and Cys residues at positions 8 and 17, respectively; LNTXs have two conserved Cys residues at 45/46/47 and 62/63/64 positions. Similarly, we derived 8 signature peptides [21] that are uniquely representing either the CTXs or the SNTXs/LNTXs and they were KRGC, GKNLC, KTCP, LKC, ATCP, ERGC, HRG and CNN. Of these parameters, KRGC, GKNLC, KTCP & LKC represent CTXs; ERGC, HRG & CNN represent SNTXs and ATCP represents LNTXs. All the parameters that were taken into considerations for the purpose of classifications of the three subfamilies of 3FTx-superfamily have been listed in the web server *http://www.feat.sastra.edu/TFTX.html* and the calculated score values of each parameter (discussed below) have also been provided in the web server.

The TFTX program is implemented in cgi-perl language [13] and the program requires only the primary sequence of the protein toxin as an input. The program accepts both FASTA and simple plain text format (without any white space) of protein sequences representing single-letter codes of amino acids. The program also accepts multiple sequences prepared in the FASTA format. The program then calculates two types of scores for each given sequence: individual score percentage (ISP) and relative score percentage (RSP), which are calculated using following mathematical expressions.

ISP = (Total score of the given sequence to be CTX/SNTX/LNTX / Maximum score of a typical sequence to be CTX/SNTX/LNTX) * 100

……..(1)

RSP = (ISP of the given sequence to be CTX/SNTX/LNTX / Sum of ISP of the given sequence to be CTX, SNTX and LNTX) * 100

…… (2)

Wherein, score is the probability of a parameter considered in the total number of sequences of CTX/SNTX/LNTX and its mathematical expression is shown, herein.

Score = Numbers of sequences fulfilling a parameter / total number of sequences …… (3)

It can be inferred from the calculations that the ISP and the RSP reveal about the overall percentage similarity of the input sequence within a subfamily and among the three subfamilies, respectively. The TFTX predicts that the input sequence belongs to a subfamily when the ISP and RSP of the query sequence are >=50% to the subfamily and the RSP values of the input sequence belonging to the subfamily must be at least 20% greater than that of other sub-families considered in the study. This is based on the criteria that the given sequence must have at least 50% characteristic properties of a particular subfamily and further to ascertain that the sequence may show maximum of 30% similarity only to the features of other subfamilies. The program predicts the given sequence as unclassified when the above mentioned conditions are not satisfactorily accounted.

After successful completion of a run, TFTX generates outputs in txt and xls formats for each query sequence submitted and the outputs can be

directly downloaded to system drives for purposes of recording the data. In the case of input files containing multiple sequences, txt file of each sequence will be packed and compressed into a single zip file as 'summary.txt'. The summary file displays the following data for each sequence submitted one after another in an order: Accession number, Amino acid sequence in FASTA format, Sequence length, Molecular weight, Amino acids distributions, Gravy value, Aliphatic index, Acid-base ratio, Net charge of the protein at neutral pH, Score values (ISP & RSP), Sub-family predictions and Rationalization.

Phylogenetic tree was constructed for 269 sequences of 3FTx reported from elapid snakes using MEGA 5.05 (*http://en.bio-soft.net/tree/MEGA.html*) computational tool [22]. Maximum parsimony method was employed to generate the tree, which was further refined using 'bootstraping' method with default parameters except number of iterations, which was set to 50 to obtain a reliable phylogenetic tree for the 3FTxs. Servers such as NTXPred (*www.imtech.res.in/raghava/ntxpred)* and SUPERFAMILY (*http://supfam.org*) were also employed to predict structural classifications of the 3FTxs considered in the present work and the data outputs of the servers have been compared with that of the TFTX program described in the research article.

3. RESULTS AND DISCUSSIONS

Cardiotoxins and α-neurotoxins from the venoms of elapid snakes are homologous proteins belonging to 3FTx-superfamily but they are drastically differing in their biological functions from each other [23,24]. In other words, the 3FTxs have similar molecular folds with multiple functions [1,2]. Obviously, structural and functional characterizations are prerequisites in order to classify them into right subfamilies. Herein, we have developed a computational tool, TFTX, to differentiate CTXs, SNTXs and LNTXs from each other using an array of parameters derived from their amino acids sequences. The flowchart in Fig. 1 outlines the overall functions of the program on calculating scores and classifying the sub-families of given protein sequences.

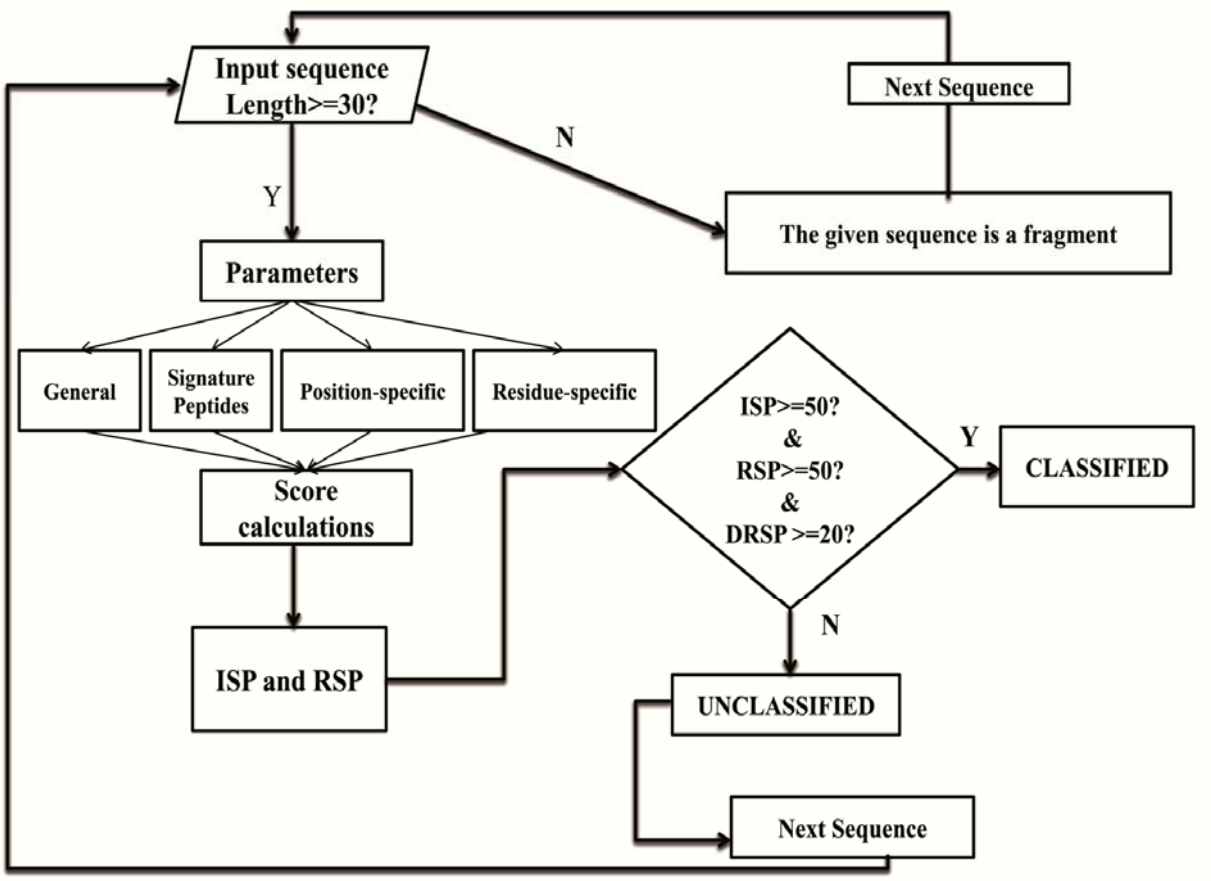

Figure 1: The flowchart outlines the overall functions of the TFTX program.

The TFTX declines the input sequences having less than 30 residues for further analyses because those sequences may probably small peptides or may be originated as fragments from large protein molecules and the sequences will moreover not be compatible in length to fold into 3FTx-superfamily. In general, 3FTxs are single polypeptide chains comprising of 59-74 aminoacids with minimum of 4 disulfide bridges [25]. Of the 269 primary sequences of cardiotoxins and α-neurotoxins reported in the literature, 195 sequences showed unambiguous annotations and remaining 74 sequences are either with no or ambiguous annotations. In order to validate the robustness of the TFTX, we first predicted the subfamilies of all 195 authentic sequences using the program and the results are depicted in the Table 1.

A quick inspection to the table strikingly shows that the program predicts the exact subfamilies for all authentic sequences suggesting the 46 parameters considered for identifying the CTXs and the α-NTXs are unique and good enough to achieve the task with high degree of confidence levels.

We have then used the TFTX to predict the subfamilies of sequences of CTXs, SNTXs, LNTXs and NTXs having ambiguous annotations. The data analyzes are shown in the Table 2. From the Table 2, it can be inferred that the TFTX predicts 4 sequences out of the 11 CTX-homologous as CTXs; 3 sequences out of the 27 SNTX-homologous as SNTX; 17 sequences out of 32 LNTX-homologous as LNTX. While the program identifies a few numbers of sequences as CTXs, SNTXs and LNTXs from their respective homologous, it also alerts that homologous sequences to the authentic sequences of a subfamily need not necessarily belong to the subfamily. Furthermore, 4 three-finger toxins from elapid snakes were reported as neurotoxins and their subfamilies were left unaddressed. The TFTX predicts that all of them are belonging to neither CTXs nor α-NTXs and show them as 'unclassified'. It implies that the 4 NTXs may be either k-toxins or muscarinic toxins. Also, there is no reason to neglect a new subfamily that may well represent these toxins. This particular aspect, as we have also discussed in the homologous sequences, is the interesting features of the program, as these computational findings may facilitate to detect novel subfamilies in the superfamily of snake venoms.

Comparative analyzes of data outputs of the TFTX with that of servers such as NTXpred [10] and SUPERFAMILY [11], reveals that TFTX is unique in its functions on identifying the subfamilies of proteins belonging to 3FTx-superfamily of elapid snake venoms.

Table 1: Comparing the literature annotations of CTXs, SNTXs and LNTXs of elapid snake venoms with the annotations of the toxins as predicted using the TFTX program (swissprot accession numbers of the sequences are given in the parenthesis).

	Protein toxins	Literature annotations	TFTX predictions
A	**Cardiotoxins (CTXs)** (**80 sequences:** Q9W6W6, P60306, P07525, Q9W6W9, Q98965, Q91124, P60304, P01442, P60301, P01443, P62375, P80245, Q98957, Q98956, P79810, Q98958, Q98959, Q98960, Q98962, P60307, P60308, P60309, Q98961, Q91126, Q91996,P49122, P49123, Q9W716, P01453, P62394, P62390, P01455, P01462, P01459, P01461, P01464, P01457, P01465, P01466, P01460, P01454, P60305, Q9DGH9, P60303, P01445, P01446, P24779, P01448, P01467, P01469, P01470, P01452, P25517, P01447, P01440, P24780, P01456, P01463, P01458, Q9PS34, P01451, P01441, Q9PS33, P01468, P83345, O93471, Q9PST4, Q9PST3, O93472, P60302, O93473, O73856, O73857, P60310, O73858, O73859, Q02454, P01471, P24776 & P24777.)	Cardiotoxins	Cardiotoxins
B	**Short-neurotoxins (SNTXs)(65 sequences:** P68417, P68418, P01422, , P25675, P01420, P01421, P60770, P80958, P01424, P01431, P01432, P68419, P01423, P01427, P01426, P60773, P60774, Q9PSN6, P60772, Q9YGJ6, Q9YGJ5, O57326, O57327, P14613, P82849, P60771, P59275, P59276, P01425, P01433, P34075, P34076, P01434, A8S6A4, A6MFK6, A8HDJ9, Q7T2I1, Q9YGC2, P10457, Q9YGW9, Q9YGC4, Q9YGX0, Q9YGW8, Q9YGC7, P25495, P60775, P25496, P10459, Q90VW1, P10455, P10458, P10460, Q7T2I5, P10456, Q9YGX1, P80548, P86095, P0CAR1, A8HDK0, Q45Z11, P0CB06, P25497, A8HDJ4, A8HDJ5 & A8HDJ6)	Short-neurotoxins	Short-neurotoxins
C	**Long-neurotoxins (LNTXs)** (**50 sequences:** P01389, P25674, P01383, P01388, P25668, P25669, P25671, P25672, P25673, P01390, P01382, P01391, O42257, Q53B54, P01387, Q2VBP5, Q2VBP4, P01386, Q53B53, Q53B59, Q2VBP3, P07526, P80156, Q53B58, Q53B57, Q53B56, P80965, P82662, Q2VBP8, Q2VBP6, C5ILC5, P25670, P15815, P34074, P01385, P34073, P01385, A8S6A8, A8S6B0, A6MFK4, A6MFK5, P0C8R6, P01379, Q7T2I3, P01384, A8HDK7, A8HDK9, A8HDK8, P13495 & A8HDK4)	Long-neurotoxins	Long-neurotoxins

Table 2: Predicting the subfamilies of 3FTx homologous of snake venoms using the TFTX (swissprot accession numbers of the sequences are given in the parenthesis).

	Protein toxins	Literature annotations	TFTX Prediction
A	**CTX homologous** (**11 sequences:** Q2VBN7, Q53B46, Q2VBN5, Q2VBN4, Q69CK0, Q2VBN8, Q91136, Q91137, Q91135, Q91137, Q9PW19 & P60311)	Ambiguous annotations	Cardiotoxins (**4 sequences :** Q91136,Q91137,Q91135 & P60311
B	**SNTX homologous** (**27 sequences:** Q2VBP1, Q2VBP0, Q53B52, Q2VBN9, Q53B50, Q53B48, Q53B47, Q2VBP2, P86421, P86422, P86097, P86094, A8HDK2, Q9W7K2, Q9W7K1, Q9W7K0, Q9W7J9, Q9W7J7, Q9W7J6, Q53B49, P86420, Q9PUB7, Q9PRI1, P58370, P43445, B2BRS3 & B2BRS2)	Ambiguous annotations	Short-neurotoxins (**3sequences:** P86420,B2BRS3 & B2BRS2)
C	**LNTX homologous** (**32 sequences:** P0C8R7, P0C8R8, P86096, P86423, P86098, P86424, P86099, Q53B55, Q9YGH9, O12963, P15818, O93422, A8N285, P14612, Q9W7J5, B2BRQ7, B2BRR1, B2BRR7, B2BRS0, B2BRR4, B2BRR6, B2BRQ9, B2BRQ8, B2BRR8, B2BRR9, B2BRR5, B2BRQ5, B2BRQ6, B2BRR2, B2BRR3, B2BRR0 & B2BRS1)	Ambiguous annotations	Long-neurotoxins (**17 sequences:** P0C8R7,A8N285,Q9W7J5, B2BRR1, B2BRR7, B2BRR4, B2BRR6,B2BRQ9,B2BRQ8, B2BRR8, B2BRR5,B2BRQ5,B2BRQ6, B2BRR2, B2BRR3, B2BRR0 & B2BRS1)
D	**NTX homologous** (**4 sequences:** Q9W717, Q70WS8, Q7ZT13 & Q800Y3)	Ambiguous annotations	Unclassified

Table 3: Functional annotations of proteins adopting three-finger folds from *homo sapiens* using 'SUPERFAMILY' server and 'TFTX' computational tool (swissprot accession numbers of the sequences are given in the parenthesis).

	Proteins	Literature annotations	SUPERFAMILY server predictions	TFTX predictions
A	**3FTxs from Homo sapiens** (**7 Sequences**: Q9BZG9, P13987, P00749 - 3 chains, P36897 & P37173)	Lynx1 (Q9BZG9), CD59 (P13987), Urokinase (P00749), TGF-beta1 (P36897) and TGF-BETA TYPE II (P37173)	Snake -toxin like superfamily	Unclassified

The NTXpred predicts the neurotoxic properties of given sequences, in general and the program does not classify CTXs from the α-NTXs as well not differentiating various types of NTXs from each other. SUPERFAMILY is a web-accessible database and predicts domains and superfamilies of given protein sequences. All the 269 sequences of 3FTxs considered in the present study were subjected to the SUPERFAMILY server and the program predicts all of them as 'snake-toxin like superfamily' only. It is important to mention that proteins adopting three-finger folds are also present in many organisms other than snakes [26]. For instance, we collected 7 protein sequences of three-finger folds from *homo sapiens* and despite their structural architectures, they are not toxic proteins as we do observe in the case of CTXs and α-NTXs of snake venoms [27]. The 7 proteins were subjected to the SUPERFAMILY and the TFTX programs for further structural classifications. The SUPERFAMILY predicts them as 'snake-toxin like superfamily', whereas, the TFTX predicts them as 'unclassified' representing that none of them are snake venom proteins (Table 3). It is now obvious that TFTX has the merits to differentiate 3FTxs of snake venoms from 3FTxs of various organisms, which may be non-toxic in nature. We also attempted grouping the 269 sequences of the 3FTxs using phylogenetic methods and the data is shown in Fig. 2. The phylogenetic tree of the sequences was constructed using MEGA 5.05, which employed maximum parsimony method followed by bootstrapping refinements. As one can observe, the tree fails to show well-defined clades representing CTXs, SNTXs and LNTXs. Moreover, the sequences of CTXs, SNTXs and LNTXs are randomly shuffled in the tree suggesting that grouping of those toxins on the basis of evolutionary relationship itself was not a straightforward approach. Thus, in addition to evolutionary parameters, the identification of subfamilies of 3FTXs of snake venoms requires more stringent parameters that could be derived from many facets of structural architectures of the proteins at primary, secondary and tertiary levels. In these backgrounds, we demonstrated that the TFTX, which is right now in its first version, is capable of identifying the subfamilies of the three-finger toxins of snake venoms in a robust manner.

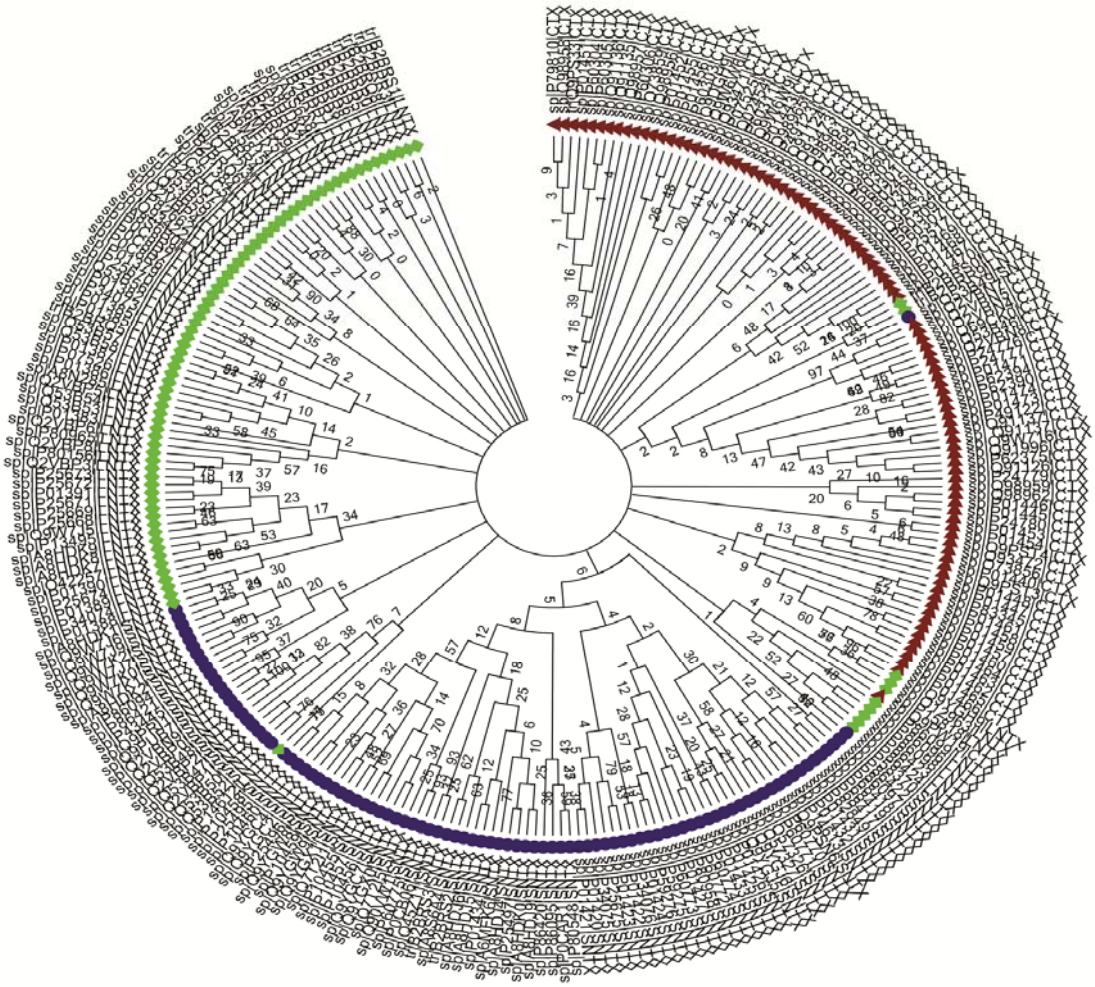

Figure 2: Phylogenetic tree of 269 sequences representing CTXs, SNTXs and LNTXs from the elapid snake venoms, is depicted. The tree was constructed using MEGA 5.05, which employed 'maximum parsimony method' followed by 'bootstrapping' refinements. The colour codes, brown, blue and green denote CTXs, SNTXs and LNTXs that are distributed in the phylogenetic tree.

4. CONCLUDING REMARKS

In the present study, we describe a novel computation tool, TFTX, on predicting the subfamilies of protein toxins belonging to 3FTx-superfamily of elapid snake venoms. To date, there were no programs reported in the literature for identifying subfamilies of snake venom proteins and hence, TFTX is the first program to address the subfamilies classifications, to our best knowledge. The applications of TFTX extend beyond the identification of subfamilies of snake toxins: (i) it provides few hints to explore novel subfamilies; (ii) it differentiates snake toxins from proteins with similar folds present in other organisms; (iii) unique structural features of each subfamilies are uncovered (iv) the overall similarities of each subfamily with others could be well probed. Foreseeing the potential applications of the TFTX in toxicology and structural biology, we do anticipate a great scope to improve the software tool at many different angles. Right now, the program classifies CTXs, SNTXs and LNTXs only, though many subfamilies such as k-toxins, muscarinic toxins, weak toxins, fasciculins, calciseptin and dendroaspin are documented to present in the 3FTx-superfamily of snake venoms. In the near future, the program would be developed with high degree of versatility for identifying all subfamilies present in each superfamily of snake toxins by including more numbers of stringent and novel parameters.

5. ACKNOWLEDGEMENTS

This work is supported by the research grant (BT/PR13378/GBD/27/262/2009) from the Department of Biotechnology, New Delhi, India. We also express our sincere thanks to Nikunji and Aukash for their contributions on the earlier stage of this research project.

REFERENCES

1. Kini, R.M., *Clin. Exp. Pharmacol. Physiol.* 2002, *29,* 815-822.
2. Kini, R.M., Doley, R., *Toxicon.* 2010, *56(6)*, 855-867.
3. Nandhakishore, R., Yuh, F.P., Yi, Z.Z., Peter, T.H.W., Prakash K., P., Kini, R.M., *FASEB. J*. 2007, *21*, 3686.
4. Kumar, T.K.S., Pandian, S.K., Srisailam, S., Yu, C., *J. Toxicol.Toxin Rev.* 1998, *17 (2)*, 183-211.
5. Kumar, T.K.S., Jayaraman, G., Lee, C.S., Arunkumar, A.I., Sivaraman, T., Samuel, D., Yu, C., *J. Biomol. Struct. Dyn.* 1997, *15(3)*, 431-463**.**
6. Sivaraman, T., Kumar, T. K. S., Tu, Y. T., Peng, H. J., Yu, C., *Arch. Biochem. Biophys.* 1999, *363*(1), 107-15.
7. Sivaraman, T., Kumar, K., Hung, K.W., Yu, C., *Biochemistry*, 2000, *39(30)*, 8705-10.
8. Ploug, M., Ellis, V., *FEBS Lett*. 1994, *349,* 163–8.
9. Galat, A., *Cell. mol. life. sci*. 2008, *65(21)*, 3481-3493.
10. Saha, S., Raghava, G.P.S., *In silico biology*. 2007, *7(4-5)*, 369-387.
11. Wilson, D., Madera, M., Vogel, C., Chothia, C., Gough, J., *Nucleic Acids Res.*, 2007, *35,* D308–D313.
12. Fry, B.G., Wuster, W., Kini, R.M., Brusic, V., Khan, A., Venkataraman, D., Rooney, P., *J. Mol. Evol.* 2003, *57(1)*, 110-29.
13. Moorhouse, M., Paul, B., *Bioinformatics Biocomputing and Perl*. 2004.
14. Wendy, W., *HTML A Beginner's Guide.* 2010.
15. Bendtsen, J. D., Nielsen, H., von Heijne G., Brunak, S., *Journal Mol. Biol.* 2004, 340, 783-795.
16. Duckert, P., Brunak, S., Blom, N., *PEDS*. 107-2004, *17(1)*, 112.
17. Corpet, F., *Nucl. Acids Res.* 1988, *16 (22)*, 10881-10890.
18. Atsushi, I., *J. Biochem.*, 1980, *88,* 1895-1898.
19. Gasteiger, E., Walker, J.M., *Humana Press. (The Proteomics Protocols Handbook).* 2005, 571-607.
20. Kyte, J., Doolittle., *J. Mol. Biol.*, 1982, *157,* 105-132.
21. Khader, S., Lalima, L., Madan., Shivamurthy, V., Balasubramanian, G., Sowdhamini, R., *BMC Bioinformatics*. 2010, *11*, 473.
22. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., *Mol. biol.evol.* 2011, *28(10)*, 2731-2739.
23. Kumar, T.K.S., Sivaraman, T., Yu, C., *ACS symposium Series (Natural and selected synthetic toxins)*, 2000, *745*, 222-248.
24. Sivaraman, T., Kumar T.K.S., Yu.C., *Biochemistry* 1999, *38*, 9899-9905.
25. Tsetlin, V., *Eur. J. Biochem.* 1999, *264,* 281–286.
26. Gumley, T.P., McKenzie, I.F., Sandrin, M.S., *Immunol. Cell Biol.* 1995, *73*, 277–96.
27. Ploug, M., Ellis, V., *FEBS Lett*. 1994, *349*, 163–168.