

Evolutionary Distance and Conserved Domain Analysis of Divergent Phylogenetic Lineages from Genus *Naja*

Sherkhane AS and Gomase VS*

The Global Open University, India

Abstract

Naja Naja is one of the poisonous snakes in the genus *Naja* of Elapids family; commonly called Indian cobras and are mostly found in Asia and Africa. They are highly venomous species having a cocktail of chemicals that have different effects may be speed the heart rate, blood pressure and interfere with the normal function of the nervous system and potentially block $\alpha 7$ homo-oligomeric neuronal AChRs causing paralysis. The aim of the present study is to identify the origin of neurotoxin, prediction of structure and function of neurotoxin by multiple sequences analysis and observing the conserved pattern of amino acid residues and to construct the phylogenetic tree for organizing evolutionary history of *N. Naja* of genus *Naja* from Elapidae family.

Keywords: Neurotoxins; *Naja* genus; Multiple sequence alignment; Conserve domain; Phylogenetic analysis

Introduction

Naja Naja is one of the poisonous snakes in the genus *Naja* of Elapids family and commonly called Indian cobras and are mostly found in Asia and Africa [1]. Elapidae family approximately consists of 300 venomous snakes in 62 genera [2]. The genus *Naja* consists of currently 26 species of cobra of which 11 inhabit Asia and 15 occur in Africa [3,4]. Proteins from *Naja Naja* are potent postsynaptic neurotoxins [5]. Neurotoxins that acts by binding to the nicotinic acetylcholine receptors in the postsynaptic membrane of skeletal muscles [6] causing severe local pain, swelling immediately after bite; dark discoloration, necrosis, paralysis and even death [7-10]. In this research work, we study the origin and evolution of neurotoxin from *N. Naja* by multiple sequence alignments that provide the functional information of conserved sequence regions of neurotoxin from *Naja Naja*, phylogenetic analysis shows taxonomical classification, identifying and naming new members of protein families that derived from a common ancestor [11-15].

Materials and Methods

Sources and sequence information of genus *Naja* neurotoxins

Thirteen species has taken from genus *Naja* of Elapidae family, in which targeted neurotoxins protein data were used to observe molecular resemble of related protein by phylogenic analysis (Table 1) [16,17].

Accession	Description	Identity %	E Value	Total Score
P60814.1	<i>Naja naja</i>	100%	5e-55	176
Q9YGI4.1	<i>Naja atra</i>	98%	1e-52	170
O42255.1	<i>Naja sputatrix</i>	97%	5e-52	168
P82935.2	<i>Naja kaouthia</i>	88%	2e-40	139
Q9YGI2.1	<i>Naja naja</i>	72%	4e-35	125
P25679.2	<i>Naja kaouthia</i>	85%	7e-31	114
O93422.1	<i>Naja atra</i>	72%	1e-27	106
P01399.1	<i>Naja annulifera</i>	67%	2e-24	97.8
P85520.1	<i>Naja oxiana</i>	69%	6e-24	96.3
P25680.1	<i>Naja nivea</i>	68%	5e-23	94.0
P01401.1	<i>Naja haje haje</i>	62%	3e-21	89.4
P01400.1	<i>Naja melanoleuca</i>	62%	2e-20	87.0
Q9W717.1	<i>Naja naja</i>	54%	8e-20	86.3

Table 1: Sequences producing significant alignments.

Multiple sequence alignment of genus *Naja* neurotoxins

Multiple sequence alignment [MSA] is conducted by COBALT, which aligns thirteen neurotoxin protein sequences of similar *Naja* genus using a combination of distance matrix and approximate parsimony methods. Numerical setting method is used to study the relative entropy threshold, in bits, that must be met for an alignment column to be displayed in red. A larger number indicates higher degree of conservation. The relative entropy is computed as: $\sum_i f_i \log_2 (f_i / p_i)$, where i is residue type, f_i is residue frequency observed in the multiple alignment column, and p_i is the background residue frequency. Identity setting used for only columns with one residue type will be colored in red [18].

Construction of a phylogenetic tree for neurotoxin from genus *Naja* of Elapidae family

Phylogenetic analyses were performed by fast minimum evolution algorithm and Neighbor Joining algorithms to allow the reconstruction phylogenetic tree of the molecular evolutionary history of various aligned sequences that are useful to align highly evolved gene families clearing evolutionary relationships such as multiple actin proteins [19,20]. Trees were obtained by the methods fast minimum evolution algorithm and Neighbor Joining algorithms. Evolutionary distance is studied by Grishin (protein) model [21,22] and distance between two sequences modeled as expected fraction of amino acid substitutions per site given the fraction of mismatched amino acids in the aligned region and can be computed for fraction of mismatched amino acids larger than 0.75 [23-26].

Results and Interpretation

Evolutionary distance

This study, thirteen neurotoxin protein from genus *Naja* is

*Corresponding author: Gomase VS, The Global Open University, Nagaland, India, Tel: 91-998-777-0696; E-mail: gomase.viren@gmail.com

Received May 21, 2014; Accepted August 20, 2014; Published August 25, 2014

Citation: Sherkhane AS, Gomase VS (2014) Evolutionary Distance and Conserved Domain Analysis of Divergent Phylogenetic Lineages from Genus *Naja*. J Data Mining Genomics Proteomics 5: 156. doi:10.4172/2153-0602.1000156

Copyright: © 2014 Sherkhane AS, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

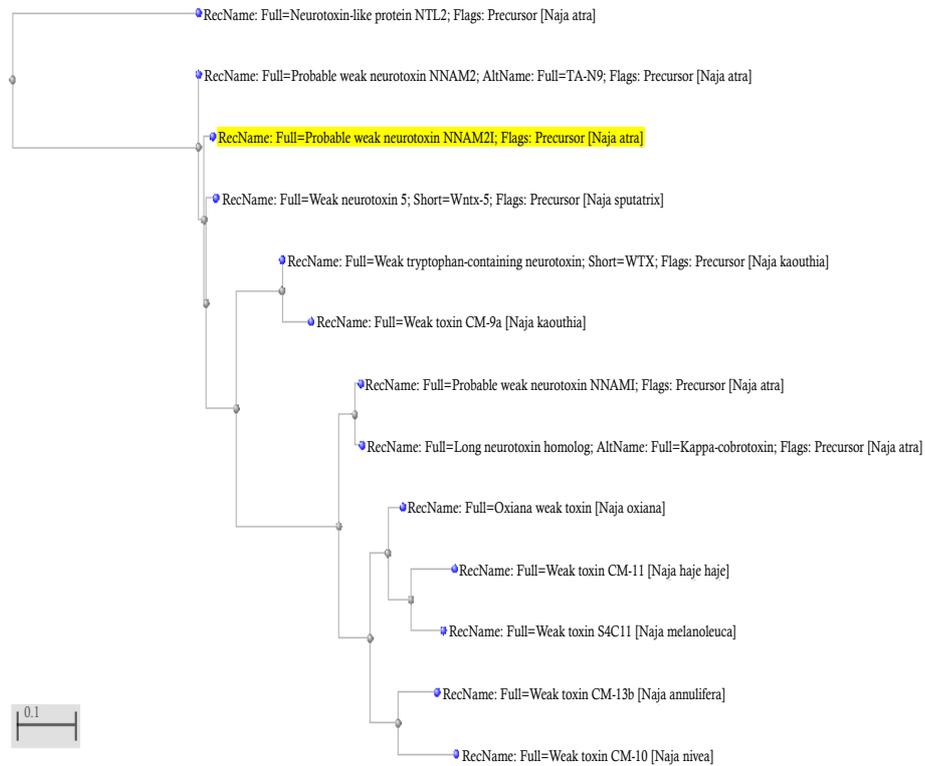


Figure 1: Rectangle tree - Fast minimum evolution algorithm - Phylogenetic study of genus *Naja* with the help of rendering tree showing the evolutionary difference with *N. Naja* in the rectangular shaped rooted tree, root is places in the longest edge.

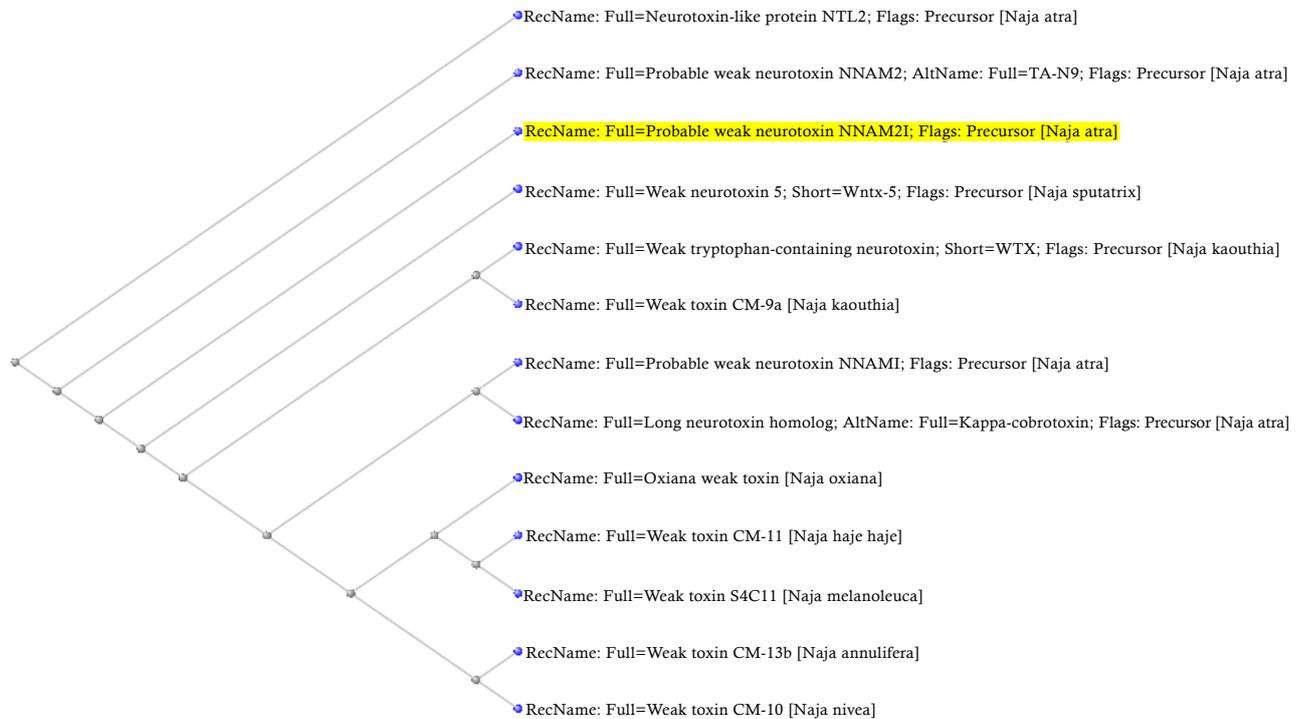


Figure 2: Slanted tree - Grishin (protein) model- Phylogenetic study of genus *Naja* with the help of rendering tree. Slanted tree is Similar to rectangle, but with triangular tree shape.

summarized to study the evolutionary distance. The identification of the origin of neurotoxin protein from genus *Naja*, multiple sequences analysis, observing the conserved amino acid residues and reconstruct the phylogenetic tree specify the evolutionary history, relationship of *N. Naja* a with different species (Table 1). Rectangle tree shows rectangular shaped rooted tree, where root is places in the longest edge. Fast minimum evolution algorithm produce un-rooted tree such as ones shown as radial or force in the tabs below. The rooted trees are created by placing a root in the middle of the longest edge (Figures 1-4). Slanted tree shows similar to rectangle, but with triangular tree shape. Neighbour Joining algorithms produce un-rooted tree such as ones shown as radial or force in the tabs below. The rooted trees are created by placing a root in the middle of the longest edge.

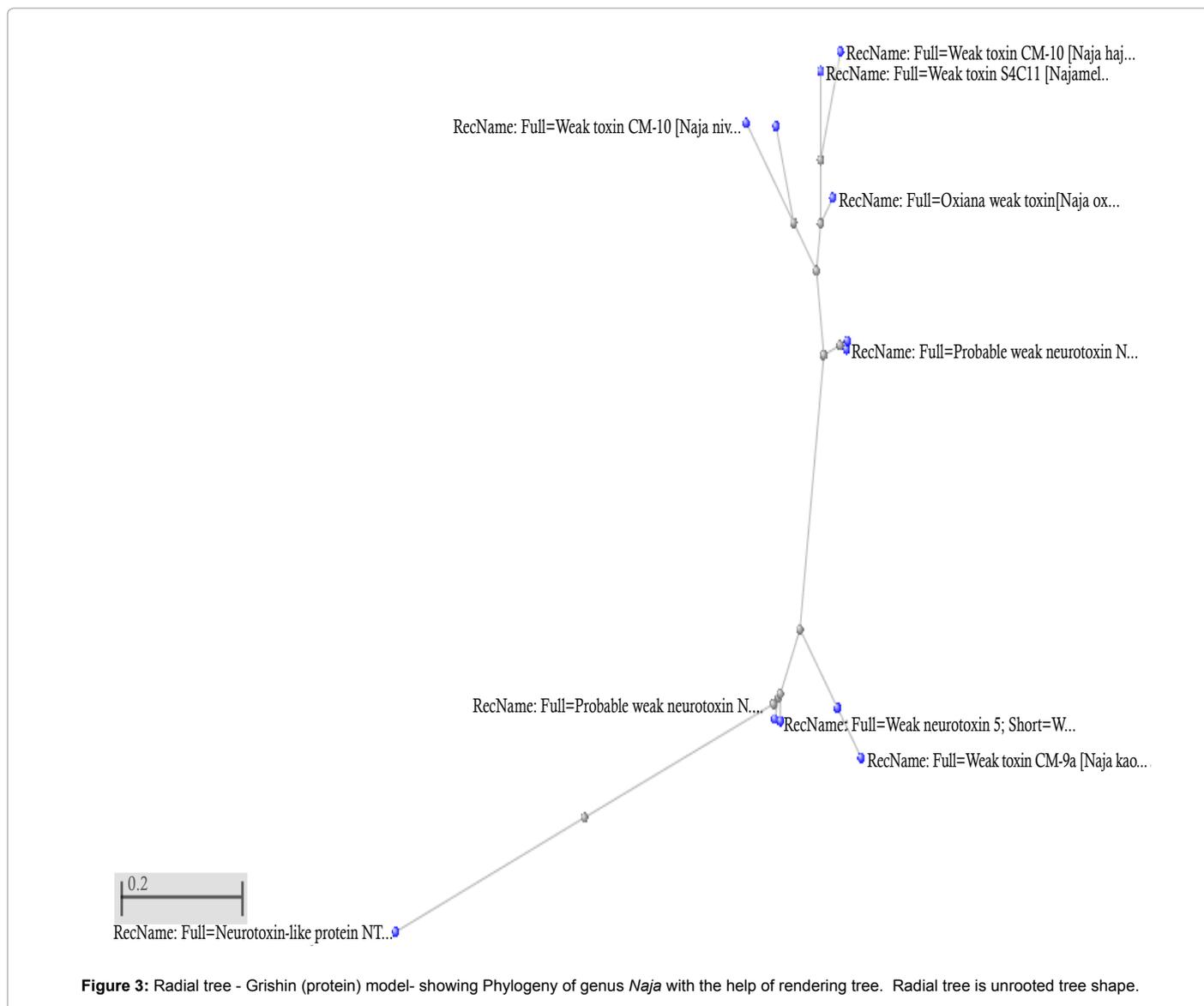
MSA

Multiple sequence alignment analysis shows columns with no gaps are colored in blue or red. The red color indicates highly conserved regions and blue indicates less conserved ones. The Conservation analysis can be used to select a threshold for determining which

columns are colored in red (Figure 5). Multiple sequence alignment identify conserved motifs and to predict functional role in the variable sites as well as conserved sites show the sequence divergence profile of these neurotoxin proteins, which demonstrate the sequence enrichment strategy of these sequences for adaptation to different physiological systems. Here we observed that from all sequences of neurotoxin proteins that Cys(c), Thr (T), Asn (N) (Hydrophilic amino acid) Phe(F), Gly(G), Ala(A), Pro(P) (hydrophobic amino acid), Lys (K), Arg(R), Positive charged, Asp(D), Negative charged which is conserved in all peptides having a common ancestor. That all of these peptides share eight highly conserved cysteines which were involved in the formation of β -strands are almost conserved. Cysteine (C) is conserved in all sequences at 8 sites. Multiple sequence alignment is carried out by COBALT of *Naja* genus.

Conserved domain analysis

Molecular study of *N. Naja* shows conserved domains and having one of snake toxin superfamily with user query added Superfamily (Figures 6 and 7). Snake toxin domains are present in short and long



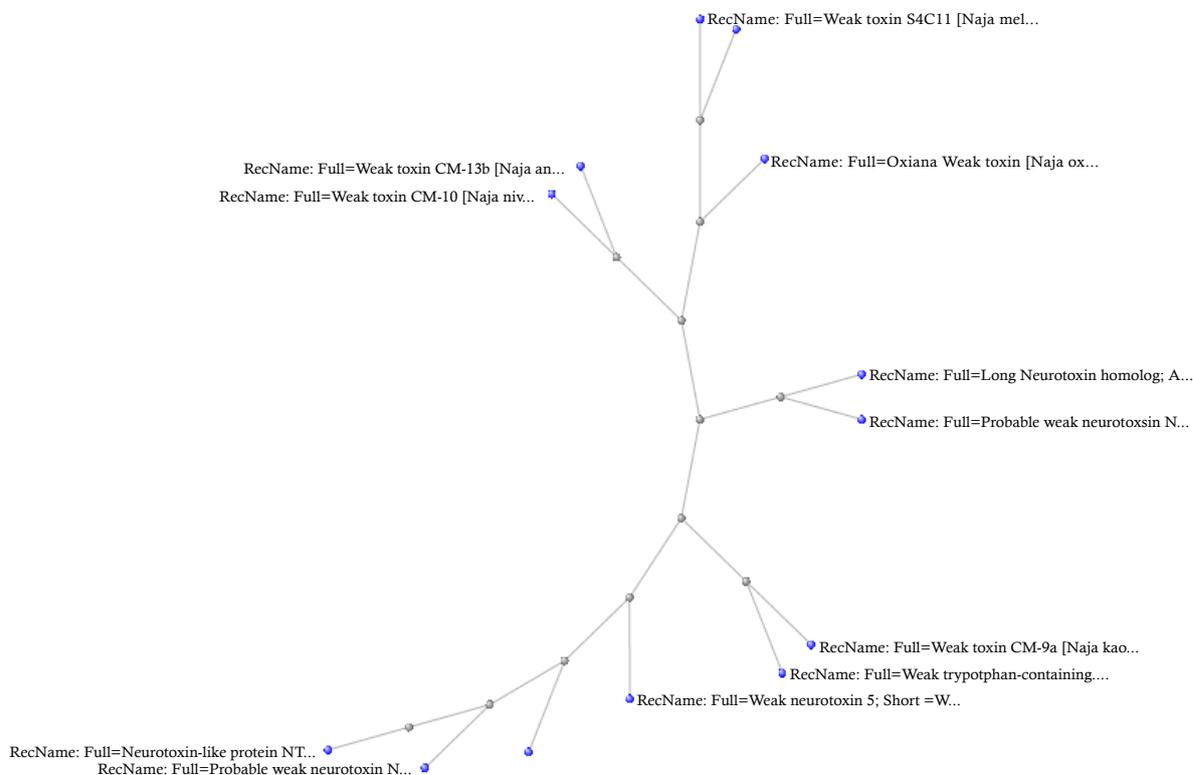
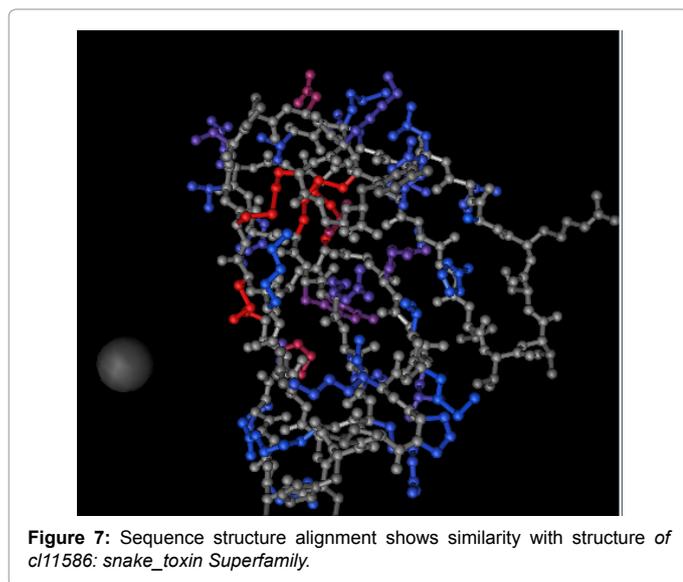
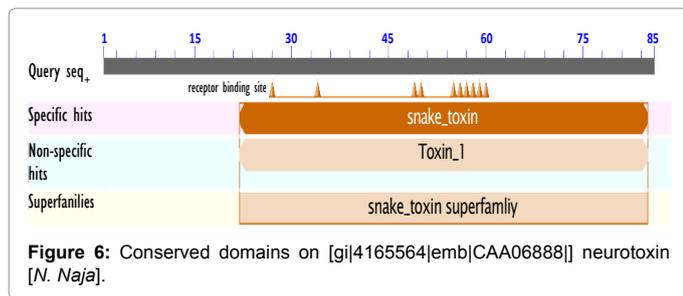


Figure 4: Force tree - Grishin (protein) model- showing Phylogeny of genus *Naja* with the help of rendering tree. Force tree is similar to radial, where nodes are pushed away from one another for better presentation.

<input checked="" type="checkbox"/>	P60814	1	MKTLPLTLVWVTIVCLDLGYTLTCLNCPFMFCGKFKQICRN	GEKICFKK	KLHQRPFSLRYIRGCAATCPG	TKPRDMVECCS	80
<input checked="" type="checkbox"/>	Q9YGI4	1	MKTLPLTLVWVTIVCLDLGYTLTCLNCPFMFCGKFKQICRN	GEKICFKK	KLHQRPFSLRYIRGCAATCPET	KPRDMVECCS	80
<input checked="" type="checkbox"/>	O42255	1	MKTLPLTLVWVTIVCLDLGYTLTCLNCPFMFCGKFKQICRN	GEKICFKK	LQRRPFSRLRYIRGCAATCPG	TKPRDMVECCS	80
<input checked="" type="checkbox"/>	P82935	1	MKTLPLTLVWVTIVCLDLGYTLTCLNCPFMFCGKFKQICRN	GEKICFKK	KLHQRPFSLRWYIRGCAATCPV	GKPYEMIECCS	80
<input checked="" type="checkbox"/>	Q9YGI2	1	MKTLPLTLVWVTIVCLDLGYTLTCLICPEKYCNKVHTCLN	GEKICFKK	YDQRKLLGKRYIRGCAATCPV	RKPREIVECCS	80
<input checked="" type="checkbox"/>	P25679	1	-----LTCLNCPFMFCGKFKQICRN	GEKICFKK	KLHQRPFSLRWYIRGCAATCPV	GKPYEMIECCS	58
<input checked="" type="checkbox"/>	Q93422	1	MKTLPLTLVWVTIVCLDLGYTLTCLICPEKYCNKVHTCLN	GEKICFKK	YDQRKLLGKRYIRGCAATCPV	RKPREIVECCS	80
<input checked="" type="checkbox"/>	P01399	1	-----LTCFNCPVYCNRFHTCRN	GEKICFKR	FNERKLLGKRYIRGCAATCPV	AKPREIVECCS	59
<input checked="" type="checkbox"/>	P85520	1	-----LTCLICPEKYCNKVHTCRN	GENICFKR	FYEGNLLGKRYIRGCAATCP	EAKPREIVECCS	59
<input checked="" type="checkbox"/>	P25680	1	-----LRCLNCPVYCNRFHTCRN	GENICFKR	FYEGNLLGKRYIRGCAATCP	EAKPREIVECCS	59
<input checked="" type="checkbox"/>	P01401	1	-----LTCLICPEKYCNKVHTCRN	GENICFKR	FYEGNLLGKRYIRGCAATCP	EAKPREIVECCT	59
<input checked="" type="checkbox"/>	P01400	1	-----LTCLICPEKYCNKVHTCRN	GENICFKR	FYEGNLLGKRYIRGCAATCP	EAKPREIVECCS	59
<input checked="" type="checkbox"/>	Q9W717	1	MKTLPLTLVWVTIVCLDLGYTRLCLSDYSISFSETIETIC	PDGHNFC	FKKFKGIRLFPWYIRGCAATCP	KAERVYVDCCA	80
<input checked="" type="checkbox"/>	P60814	81	TDRCNR	86			
<input checked="" type="checkbox"/>	Q9YGI4	81	TDRCNR	86			
<input checked="" type="checkbox"/>	O42255	81	TDRCNK	86			
<input checked="" type="checkbox"/>	P02935	01	TDKCNR	06			
<input checked="" type="checkbox"/>	Q9YGI2	81	TDKCNH	86			
<input checked="" type="checkbox"/>	P25679	59	TDKCNR	64			
<input checked="" type="checkbox"/>	Q93422	81	TDKCNH	86			
<input checked="" type="checkbox"/>	P01399	60	TDRCNH	65			
<input checked="" type="checkbox"/>	P85520	60	TDKCNH	65			
<input checked="" type="checkbox"/>	P25680	60	TDKCNR	65			
<input checked="" type="checkbox"/>	P01401	60	TDRCNK	65			
<input checked="" type="checkbox"/>	P01400	60	TDKCNH	65			
<input checked="" type="checkbox"/>	Q9W717	81	RDKCNR	86			

Figure 5: Multiple sequence alignment by COBALT of genus *Naja*. Here columns with no gaps are colored in blue or red. The red color Cys(c),Thr (T),Asn (N) hydrophilic polar, Phe(F),Gly(G),Ala(A),Pro(P) hydrophobic nonpolar,Lys (K),Arg(R), Positive charged, Asp(D),Negative charge indicates highly conserved columns and blue indicates less conserved ones. The Conservation Setting can be used to select a threshold for determining, which columns are colored in red.



Description	Pssmld	Multi-dom	E-value
snake_toxin[cd00206], Snake toxin domain, present in short and long neurotoxins, cytotoxins and short toxins, ...	119411	no	8.44e-20
Toxin_1[pfam00087], Snake toxin; A family of venomous neurotoxins and cytotoxins. Structure is small, disulfide-rich,	249576	no	5.93e-17

Table 2: List of domain hits.

neurotoxins, cytotoxins and short toxins, and in other miscellaneous venom peptides (Table 2). The toxin acts by binding to the nicotinic acetylcholine receptors in the postsynaptic membrane of skeletal muscles and preventing the binding of acetylcholine, thereby blocking the excitation of muscles. This domain contains 60-75 amino acids that are fixed by 4-5 disulfide bridges and is nearly all beta sheet; it exists as either monomers or dimers.

Conclusion

Thirteen neurotoxin proteins of genus *Naja* from Elapidae family are summarized the identical regions. Using multiple sequences analysis and phylogenetic tree we observe the conserved residues to specify the evolutionary history and analysing sequence structure relationship of neurotoxin among *Naja* species. Efficient utilization of Polar, nonpolar, positively and negatively charged amino acids and their distribution in toxin sequence make them a killer element in snake venom. Comparative analyses specify that the neurotoxin demonstrates how proteins are generated within the nature's testing ground for tailor-made biologic needs. Evolutionary studies of neurotoxin sequence of *Naja* genus found the common ancestor of all the *Naja* species. In

future, different neurotoxin may be converted in laboratories through protein engineering to design synthetic peptide vaccine that have a much positive role.

References

- Habermehl GG, Habermehl G (1981). Venomous animals and their toxins New York. Springer Verlag : 81-83.
- Golay P, Smith HM, Broadley DG, Dixon JR, McCarthy C, et al. (1993) Endoglyphs and Other Major Venomous Snakes of the World A checklist. Azemiops herpetological data center : 478.
- Kelly C, Barker N, Villet M, Broadley D (2009). Phylogeny, biogeography and classification of the snake superfamily Elapoidea: a rapid radiation in the late Eocene. Cladistics 25: 38–63.
- Wallach V, Wuster W, Broadley DG (2009). In praise of subgenera: taxonomic status of cobras of the genus *Naja* Laurenti (Serpentes: Elapidae). Zootaxa. 2236: 26–36.
- Endo T, Tamiya N (1987). Current view on structure function relationship of postsynaptic neurotoxins from snake venoms. Pharmacol Ther 34: 403–451.
- Yang CC (1974). Chemistry and evolution of toxins in snake venoms. Toxicon 12: 1-43.
- Rees B, Bilwes A (1993). Three-dimensional structures of neurotoxins and cardiotoxins. Chem Res Toxicol 6: 385-406.
- TsetlinV I, Hucho F (2004). Snake and snail toxins acting on nicotinic acetylcholine receptors: fundamental aspects and medical applications. FEBS Lett 557: 9-13.
- Rossetto O, Rigoni M, Montecucco C (2004). Different mechanism of blockade of neuroexocytosis by presynaptic neurotoxins. Toxicol Lett 149: 91-101.
- Dufton MJ, Hider RC (1983). Conformational properties of the neurotoxins and cytotoxins isolated from Elapid snake venoms. CRC Crit Rev Biochem 14: 113-171.
- Durbin R, Eddy S, Krogh A and Mitchison G (2002). Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids. Cambridge University Press, UK.
- Fry BG (2005). From genome to „venome“: molecular origin and evolution of the snake venom proteome inferred from phylogenetic analysis of toxin sequences and related body proteins. Genome Res 15: 403-420.
- Chang L, Lin S, Wang J, Hu WP, Wu B, et al. (2000) Structure-function studies on Taiwan cobra long neurotoxin homolog. Biochim Biophys Acta 1480: 293-301.
- Gomase VS, Waghmare SB, Jadhav BV, Kale KV (2009). Functional analyses of the binding ability of neuro peptide f (NPF) form *Moniezia expansa*. Gene Therapy and Molecular Biology 13: 64-70.
- Gomase VS, Kale KV, Tagore S, Hatture SR (2008). Proteomics: Technologies for Protein Analysis. Current Drug Metabolism. 9: 213-220.
- http://www.ncbi.nlm.nih.gov
- (2013) Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 41: D8-D20.
- Papadopoulos JS, Agarwala R (2007). COBALT: constraint-based alignment tool for multiple protein sequences. Bioinformatics 23: 1073-1079.
- Desper R, Gascuel O (2004). Fast and accurate phylogeny reconstruction algorithms based on the minimum-evolution principle. Mol Biol Evol 21: 587-598.
- Saitou N, Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4: 406-425.
- Grishin NV (1995). Approximates the same model as Kimura, but can be computed for fraction of mismatched amino acids larger than 0.75. J Mol Evol 41: 675-679.
- Grishin NV (1995). More general evolutionary model: substitution rates vary for both amino acids and sites. J Mol Evol 41: 675-679.
- Gomase VS and Tagore S (2009). Phylogenomics: evolution and genomics intersection. Int J Bioinformatics Research and Applications 5: 548–563.

24. Waghmare S, Waghmare D, Bhatnagar PS (2013). Species Diversity of Short Horned Grasshopper (Orthoptera: Acrididae) in Selected Grasslands of Solapur District, Maharashtra, India. J Biodivers Endanger Species 1: 110.
25. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, et al. (2011) „CDD: a Conserved Domain Database for the functional annotation of proteins.“ Nucleic Acids Res 39: 225-229.
26. Marchler-Bauer A, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese Scott C, Fong JH, et al. (2009) „CDD: specific functional annotation with the Conserved Domain Database.“ Nucleic Acids Res 37: 205-210.