

Type of the Paper (Article)

SVDB: a Comprehensive Domain Specific Database of Snake Venom Toxins Generated Through NCBI

Md. Monir Hossain ¹, Ataul Haque ², Sayeda Zannatul Sakina Mazid ³, Abira Khan ³, Tomalika Rahmat Ullah ³, Sarker Tanveer Ahmed Rume ² and Jesmin ^{3,*}

¹ Department of Computer Science & Engineering, Dhaka City College, Dhaka-1205, Bangladesh; saif.monir@yahoo.com

² Department of Computer Science & Engineering, University of Dhaka, Dhaka-1000, Bangladesh; ataul.haque@gmail.com; rume@cse.du.ac.bd

³ Department of Genetic Engineering & Biotechnology, University of Dhaka, Dhaka-1000, Bangladesh; zannatulsakina@gmail.com; abira.khan@du.ac.bd; tomalika91@du.ac.bd; jesmin@univdhaka.edu

* Correspondence: jesmin@univdhaka.edu; Tel.: +88-02-9661920-7819

Abstract: Venoms that drip from the fangs of snakes are incredibly complex chemical cocktails of compounds, with different proteins and enzymes, including a large variety of toxins like myotoxins, cardiotoxins, hemotoxins, and neurotoxins and their countless combinations. In addition to their use in the treatment of snake bites in humans, they have numerous therapeutic and medicinal applications. Potential use of snake venom includes excessive bleeding, stroke, neurological disorders, cancer, diabetes and aging. Therefore, a proper understanding of snake venom toxin and facilitating their use is of utmost importance. In this paper, we describe a novel database, called SVDB, for storage, dissemination and analysis of snake venom and toxins related information. SVDB has autonomous links to NCBI databases to pull relevant information both on-demand and asynchronous ways to facilitate data integration. SVDB includes authentic, non-redundant, up-to-date scientific information on literature, sequences, structures, small molecules, taxonomy and many more. SVDB portal also provides external links to tools like BLAST, CLUSTAL, Swiss-model, phylogeny and other toxin related resources. The architecture of SVDB information fetching, linking and structuring is unique and can be implemented to any domain specific generic data collection pipeline through the NCBI. The database is publicly available at <https://www.snakevenomdb.org>.

Keywords: venom; toxins; database; NCBI

Key Contribution: A generic pipeline has been designed/modeled through NCBI for fetching, authentic, non-redundant domain specific data of user's choice from millions of tons of information of NCBI. In this case, all information regarding snake venom & toxins from sequence level to taxonomy. SVDB is a comprehensive repository of snake venom toxins, a useful data resource for alternative & complementary medicine and future drug designing.

1. Introduction

Snakes have evolved with a stunning range of venomous compounds worldwide. Venom that drips from the fangs of snakes is nature's most efficient killer, works very fast and is highly specific (1, 2). Snake venom contains dozens to hundreds of different proteins, peptides and other molecules. For snakes, these toxins are principally for prey acquisition, digestion and defence against predators. Ironically, the properties that make venom deadly are also what make it so valuable in medicine (1-3). The therapeutic promise and medicinal use of snake venoms are enormous. Many of the proteins and enzymes in venoms have interesting biological properties and can play important roles in various biological functions such as blood coagulation, blood pressure regulation, and transmission

of the nervous or muscular signals. Some of these proteins and enzymes have been successfully used in many pharmacological purposes and are useful drugs (4, 5).

The National Center for Biotechnology Information (<https://www.ncbi.nlm.nih.gov>) was founded more than two decades ago with a mission to provide researchers and scientific community access to authentic, up-to-date biomedical data. NCBI hosts a substantial number of publicly available databases relevant to biotechnology and biomedicine and is an important resource for Bioinformatics tools and services that allow developers to access and manipulate NCBI data in their applications (6). At present, exponential growth of biological data poses huge challenges in data deposition, integration, retrieval, curation, translation and demands new data access techniques that will allow users to automatically perform the homogenization and provide easy access to the most relevant data. Here, we present SVDB, a new resource for all about snake venom and toxins related information from literature, taxonomy to molecular sequence data of genes, mRNA, protein, structure, and genome, together with external web links for data analysis like BLAST, CLUSTALW, Phylogeny, SWISS-model among many others. Users of all categories from layman to molecular biologist, biotechnologist have the flexibility to perform simple searches to detailed structural-functional investigation valuable for future drug designing. This database is the most complete resource for up-to-date data for molecular biologists, informaticians, toxicologists and anyone interested in snake venom & their components. SVDB is an open-source and publicly accessible at <https://www.snakevenomdb.org>.

2. Materials and Methods

System design and implementation: The snake venom database (SVDB) has been implemented as a website, which has both server and database component. The front end of the web interface runs on a Linux-based Apache Web server (<https://httpd.apache.org/>) and PHP (<http://www.php.net/>). PHP supports all major Web servers and databases and is used for server-side scripting. It also provides multiple layers of security to prevent threats and malicious attacks. For data storage and management, MySQL (<http://www.mysql.com>) was used, as in combination with PHP, it provides both faster service and user friendly features to configure the web services. The interface of SVDB was built using CSS (<https://www.w3.org/Style/CSS/Overview.en.html>) for HTML elements to be presented on the Web. In SVDB, the time-based job scheduler CRON (8, 9) was used for scheduling tasks to run on the server. Automated searches and fetching of venom and toxin related data through NCBI (<https://www.ncbi.nlm.nih.gov/>) were scheduled to be executed periodically using crontabs. All resources used to develop SVDB are open source.

Automated Data fetching through NCBI: The core component of SVDB is its 'data repository' that allows organizing and storage of the categorized data together with an efficient retrieval system. To build the repository, surveys were conducted on the NCBI ERD flowchart, to identify classes of categories and then tags were assigned to organize its content (Supplementary Figure S1). In SVDB, all relevant information on venom, toxins, snakes, family, species, genes, sequences and small compounds among many others were then collected and stored. All data were pipelined through NCBI which is a reliable source and we used API provided by NCBI (Entrez Programming Utilities) to search and fetch records (<https://www.ncbi.nlm.nih.gov/books/NBK25498/>) (10).

Data acquisition from NCBI is very tricky as it involves heterogeneous data formats. To avoid complexity, in SVDB structuring, data from all different databases were integrated into one database where all NCBI records are stored using RDBMS model (11, 12), Common fields like unique ID (UID), title, accession number, etc., that are conserved across most of the databases of NCBI, were fetched and stored in a single common summary table (summery_tbl). Records of different databases were identified, pooled and structured under dbname column (Supplementary Figure S1). For example, taxonomy, data stored in a separate table and its unique ID was stored in the main common table. In addition, uncommon relevant information was organized and stored in other separate tables based on EAV (Entity Attribute Value) model (13, 14).

NCBI automatically updates data in every 24 hrs, so SVDB has been programmed for automated scheduled data fetching (crawling) at fixed time intervals using Corn (Figure 1). The crawled data was then normalized and automatically stored locally in SVDB portal. SVDB portal thus provides the most up to date data and much faster data access to users for the same query than that can be performed through NCBI. In addition, SVDB provides the user the flexibility to perform search using NCBI from within the system. When displaying results, if any, record that is new and not found in SVDB, it will be instantly fetched from NCBI. All crawled records will then be stored in the SVDB portal, so the next time it can be accessed through without any API call to NCBI. The detailed strategies for SVDB data, crawling implementation are depicted in Algorithm 1, 2 and 3.

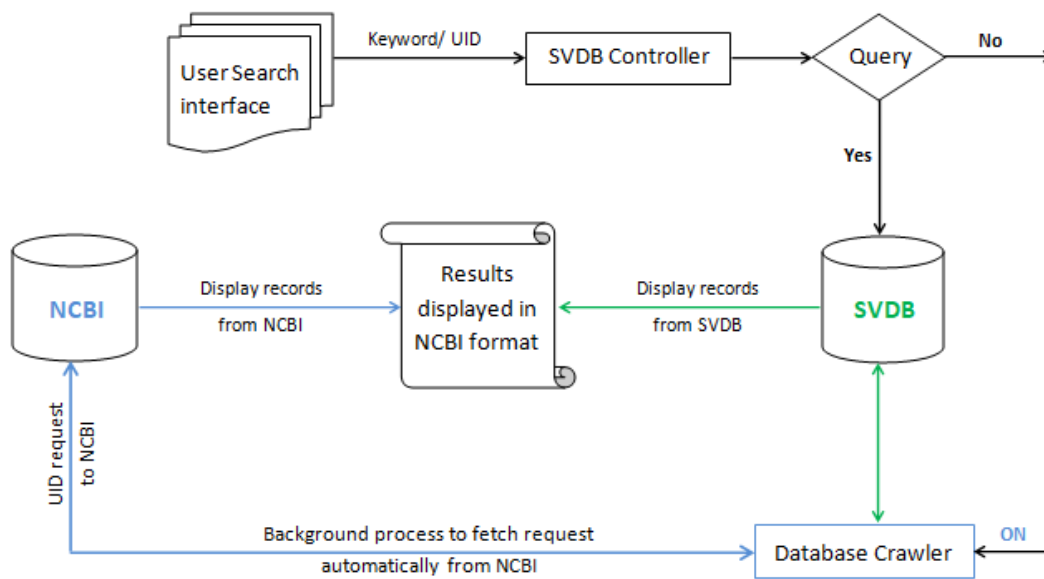


Figure 1. The architecture of SVDB.

Algorithm 1: Automatically Fetch UID Details

```

1: procedure AUTOFETCH_RESULT(uid_list)
2:   [uid list = UIDs whose details are not yet fetched]
3:   Start Cron job
4:   for Each id in uid_list do
5:     id_detail ← Request NCBI for details of id
6:     Store id_detail in DB
7:     flag ← Existence of Related Links of id
8:     [flag = 1 means related link found, otherwise 0]
9:
10:    if flag = 1 then
11:      Save Related Links to SVDB
12:      Update status of id
13:    end if
14:  end for
15: end procedure
  
```

Algorithm 2: Fetch UIDs Related to Query Keywords

```

1: procedure UID_FETCH(query)
2:
  
```

```

3:   for Each keyword in query do
4:     status ← Find keyword in SVDB repository
5:           [If found, status=1; otherwise 0]
6:
7:     if status = 0 then
8:       uid ← Fetch UID of keyword from NCBI
9:
10:      Save uid to SVDB repository
11:
12:     end if
13:   end for
14: end procedure

```

Algorithm 3: Search and Display Results

```

1: procedure Search
2:   Search Local SVDB Database for UIDs of keywords
3:   Check for query keywords in the DB
4:           [DB = SVDB Database]
5:   for Each new(not found in DB) keyword do
6:     uid ← UID found in NCBI
7:     uid_list ← uid_list + uid
8:
9:   end for
10:
11:   Invoke Procedure AUTOFETCH_REULT (uid_list)
12:
13:   for Each uid in uid_list do
13:     Display the summary
14:   end for
15: end procedure

```

Data sources: In SVDB, different types of data from literature to taxonomy, sequence to structure, from small molecules to LD₅₀ venom toxicity were accessed and extracted from various databases like PubMed, Nucleotide, Gene, Protein, Structure, Genome, Taxonomy, Conserved Domains, OMIM, PubChem compounds, substance, bioassay, Popset, GEO datasets and many more are accessed through NCBI (<https://www.ncbi.nlm.nih.gov/>) to construct a central repository of all snake venom and toxin related information (Supplementary Figure S2, Table 1).

Tools & additional resources: SVDB provides Web links to other important databases and resources on snake venom & toxins like LD₅₀ toxicity (<http://snakedatabase.org/pages/ld50.php>), T3DB: the toxin & toxin target database (15), WHO database of venomous snakes (<http://apps.who.int/bloodproducts/snakeantivenoms/database/>), VenomKB (<http://www.venomkb.org/>), UniProt entries for venom toxins (<https://www.uniprot.org/help/uniprotkb>), reptile databases (<http://www.reptile-database.org/>), NABTSCT (<http://www.nabtsct.net/questionnaire.htm>) and many more. Also, users have access to Web based analytical tools like BLAST (16), CLUSTAL (17), PSIPRED (18), SWISS model (19), Phylogeny (20), Swiss PDBviewer (21), RasMol (<http://rasmol.org/>). Cn3D

Table 1. Summary of major data sources of SVDB (as of May 2018).

Information type	Source	Number of Records	Total
Literature	PubMed	25,295	53,680
	PMC	28,307	
	PubMed Health	47	
	OMIM	31	
Sequence	Nucleotide (GenBank)	429,103	649809
	Protein (GenBank, UniProt)	137,525	
	Transcript (EST, GenBank)	83,178	
	Gene (GenBank, RefSeq)	18,314	
	Genome (GenBank)	3	
3D Structure	PDB	523	1,784
	PubChem Substance	706	
	PubChem Bioassay	555	
Integrated systems	BioSystems	609	660
	BioProject	51	
Expression	GEO Dataset	157	157
Evolutionary relatedness	PopSet	1312	1312

(<https://www.ncbi.nlm.nih.gov/Structure/CN3D/cn3d.shtml>), among many others and resources like KEGG (<https://www.genome.jp/kegg/>), REACTOME (<https://reactome.org/>), STRING (https://string-db.org/cgi/input.pl?_ga=1.95771156.929347311.1423123536), and DIP (<http://dip.doe-mbi.ucla.edu/dip/Main.cgi>) for understanding high-level functions of the biological systems through this portal that would be especially helpful for researchers keen to study the mechanism of action and structure-functional details of snake venom components and toxins. The Taxonomical data on classification and nomenclature of five major venomous snake families: Colubridae, Elapidae, Viperidae, Atractaspididae and Hydrophiidae could also be accessed through SVDB. Furthermore, in SVDB, snake bite treatment section provides an ample resource links to the users for easy accessing general information on snakebite, their sign & symptoms, prevention, treatment and first-aid. In addition, a special section has been dedicated to the venomous snakes of Bangladesh their classification, types, images and habitats.

3. Results & Discussion

3.1 Web Interface

Snake venom database (SVDB), is a customized, domain specific, non-redundant data repository built to communicate with and extract all snake venom and toxin related information from millions of data that can be accessed through NCBI. The Web interface of SVDB allows users to browse, search, analyse and download all snake venom and toxin related information from a number of primary data resources of NCBI (Figure 2 & 3). There is an option of choosing from 20 different databases available on the pull-down menus: in the simple 'Search' menu (located at the top right hand corner) of the web interface. SVDB users can simply perform searches by entering appropriate keywords and phrases into the search box, like snake name, venom, toxin name, venom components, accession no/UID or 3D structural information.

A. Home page with searchable database options

B. PubMed Search results

C. Gene

D. EST

E. Structure

F. Protein

G. Phylogenetic studies

H. Small compounds

I. Transcriptome


J. Conserved domains

K. Genome


Figure 2. Web portal and examples of some key elements of SVDB's user interface.

The left hand sidebar of the Home page contains nine menus: like literature, sequence & structure, pathways, venom components, venom toxins and others. Each menu contains sub-menus. Users can

browse relevant databases in SVDB by clicking on the relevant sub-menus (Figure 3). Alternatively, users can click on the left side of the navigation bar to view the corresponding page or entry. Relevant data are also available for bulk download in several formats from the SVDB FTP site. Through the default filter setting of SVDB, data from NCBI is automatically filtered down and only the most relevant subject-oriented data will be displayed. So, sophisticated queries can also be performed by restricting searches to specific fields and combining terms with Boolean operators (AND, OR, NOT) using the following syntax: term [field] OPERATOR term [field]. In SVDB, 'Advanced Search' and 'Limits' page options are also available on the home page to assist users to quickly construct multi-part, user defined field queries.



Snake Venom Database
SVDB



Department of Genetic Engineering & Biotechnology, University of Dhaka
Dhaka-1000, Bangladesh

Home	Taxonomy	Snake bite treatment	Tools	Other Resources	Help
SVDB Resources		Welcome		Toxicity of venomous snakes	
Literature		The Snake venom database (SVDB) contains information on venomous snakes, their compositions and functions. There are approximately 375 venomous snake species worldwide, sources of biological information which are known to cause envenomation, diabetes, bacterial infections.		All	
Reactions & Pathways		KEGG		PubMed	
Venom Components		DIP		Advanced Search	
Venom Toxins		STRING		Data Summary	
Venom Applications		REACTOME		Updated June 9, 2018	
Venomous Snakes of Bangladesh		RasMol		Sequences (GenBank) 253,310	
		RasMol		Proteins (GenBank & UniProt) 100,767	
				3D Structures (PDB) 523	
				Genes (RefSeq) 17,204	

Taxonomy

Mainly there are five venomous snake families:

- Colubridae
- Elapidae
- Viperidae
- Atractaspididae
- Hydrophiidae

Family: Elapidae

- Subfamily: Acanthophinae
 - Acanthophis
 - Acanthophis antarcticus (death adder)
 - Acanthophis cryptamydros (Kimberley death adder)
 - Acanthophis haseketi
 - Acanthophis cf. laevis
 - Acanthophis levis
 - Acanthophis praelongus (Northern death adder)
 - Acanthophis pyrrhus
 - Acanthophis rugosus
 - Acanthophis cf. rugosus
 - Acanthophis wellii
 - Acanthophis sp. Seram
- Aspidelaps
- Austrelaps
- Notechis
- Oxyuranus
- Pseudochis
- Pseudonaja
- Rhinocrotophalus
- Subfamily: Bungarinae
 - Bungarus
 - Boulengerina

Snakebite treatment

- Bangla guide book (pdf)
- Summary table of important resources

Site	Description
Wikipedia link: Snakebite	General information about snakebite: sign and symptoms, prevention, treatment and first-aid
Centers for disease control and prevention (CDC)	Habitat and types of venomous snakes types, symptoms and first aid associated with snake bite and other resource links.
Healthline	Identifying venomous snakes and snake bites, symptoms, treatment and prevention of snake bite
NHS choices	Information about snake bites and its prevalence, symptom, treatment, prevention and antivenom
World Health Organization (WHO)	About snake antivenom and its use, snake bite causes and effects, about WHO guidelines on production, control and regulation of snake venoms and antivenom
MedlinePlus	Snake bite causes and symptoms (Rattlesnake, Cottonmouth and copperhead, Coral snake), first aid, precaution and prevention
Department of wildlife and Fisheries	Snake bite occurrence, snake venom apparatus and venom discharge process, toothmark patterns for different types of snake bites, bite treatment and prevention
HealthDirect	Identification of venomous snakes, snake bites and their first

Biomedicine

Snake venom components and their applications in biomedicine.

- Summary table

Type of compound	Action on the body	Family/Species
Acetylcholinesterases (AChE)	Might cause tetanic paralysis	Colubridae, Elapidae
Arginine esterase	Predigest prey	Viperidae
Bradykinin-potentiating peptides (BPP)	Pain, hypertension, immobilize prey	Viperidae
Croto lactin	Modulate platelet activity, prevent clotting	Viperidae
Cysteine-rich secretory proteins (CRISP)	Induce hypothermia, immobilize prey	Colubridae, Elapidae, Viperidae
Disintegrins	Inhibit platelet activity, promote hemorrhaging	Viperidae
Hyaluronidases	Increase interstitial fluidity, aiding the dissemination of venom from the bite site	Elapidae, Viperidae
L-amino acid peptidase (LAAC)	Cell damage/apoptosis	Elapidae, Viperidae
Metalloproteinases	Hemorrhage, myonecrosis, believed to	Atractaspididae,

Other Resources

Currently available important databases & venom resources on snake venom & toxins:

Name	Description
VenomDB	Therapeutic uses for animal venom & venom toxins
Toxin Peptides	Toxin Peptides
Venom Toxins	UniProt entries for snake venom toxins
The reptile database	Taxonomy, distribution, external links
WHO database of venomous snakes	Taxonomy, distribution, anti-venoms
ENTRC Snake Database	Common information, distribution
Clinical toxicology resources	Snake taxonomy, snake Venoms venom/venomous terminology, snakebite clinical effects, snakebite diagnosis, snakebite treatment, anti-venom products
International Society on Toxicology - Global Snakebite Initiative	Global snakebite statistics
Snake planet	Common information
Swainland's reptiles database	Information about different species
Australian Reptile Online Database	Taxonomy, common information of Australian snake species
Online Guide to Florida Snakes	Common information on Florida snakes, bibliography
Reptile Knowledge	Taxonomy, common information

Figure 3. Analysis tools and other online resources accessed through SVDB portal.

In addition, SVDB provides links to a collection of tools like BLAST, Phylogeny, Swiss-model on the pull-down menus: under 'Tools' menu on the top bar of the web interface for users for users to carry out further structural investigations on venom toxins. The 'Other resources' menu on the top bar of SVDB provides users access other important venom & toxin related databases including T3DB, VenomKB and LD₅₀ toxicity levels of venomous snakes. The 'Taxonomy' menu provides taxonomical classification and nomenclature data of major venomous snake families. Furthermore, the 'Snake bite

treatment' menu provides links to general information on snakebite, their symptoms, treatment and first-aid (Figure 2 & 3). Together, the SVDB website provides a complete information reservoir for snake venom components and toxins for all users from beginners to the most advanced researcher.

3.2 Performance Evaluation

In SVDB portal, all data are collected from NCBI which is an authentic and up to date resource. NCBI provided APIs (Entrez Programming Utilities) were used to fetch snake venom & toxin related data for SVDB. The amount of data in NCBI is huge & the search results of every query always pulled cross-referred data. The architecture of SVDB allows retrieving and display search results which is both fast and highly relevant to the user query (Figure 1). The Web performance of SVDB and NCBI were evaluated and compared using WEBPAGETEST (<http://www.webpagetest.org>), and GTmetrix (<https://gtmetrix.com>), two benchmark tools to evaluation performance of web portals. The basis of comparison is the load time of the search result page, in the web browser. For example, the keyword "snake venom" was searched in the PubMed database of both NCBI and SVDB and the outcome of the performances were compared (Figure 4, Table 2). SVDB uses its domain specific repositories to fetch records. As the database is pre-processed and especially structured for storing snake venom related data only, it takes less time to perform searches and display results than performing the same search in NCBI (Table 2). Additionally, SVDB page is very light weight, hence consumes less bandwidth resulting in faster load time (Table 2, Figure 4). Performance tests were also evaluated using different keywords & selecting different databases of SVDB & NCBI. In each case, similar performance patterns were observed. The observed differences in performance clearly indicated that the data retrieval from SVDB is faster than NCBI.

Table 2. Performance comparison of SVDB and NCBI.

Parameters	PubMed [NCBI]	SVDB
Webpagetest results:		
Load time	2.976s	1.4s
Page size	424 KB	70 KB
Display results	Items:1 to 20 of 20144	Items:1 to 20 of 20144
GTmetrix results:		
Load time	2.6s	1.1s
Page size	430 KB	70.4 KB
PageSpeed score	B (81%)	A (93%)
YSlow score	C (73%)	A (94%)
Display results	Items:1 to 20 of 20144	Items:1 to 20 of 20144



Figure 4. Performance comparison of SVDB and NCBI.

4. Conclusions & Future Scopes

Snake venom toxins have great diversity in structure, function and evolution; are therefore, invaluable in both basic and applied research. To our knowledge, SVDB is the most comprehensive repository of snake venom and toxins. It provides information from the literature to taxonomy, sequence to structure, LD₅₀ toxicity and many others. Together, SVDB provide links to important tools essential for structural- functional analysis of venom components & toxins, enabling researchers to screen for drug leads. Venom proteins are of considerable interest to the drug discovery and biotechnology communities (22). We believe that more researchers will be focusing on elucidating the biological role of snake venom and toxins in the near future and can take advantage of this categorized domain - specific data portal for further structural investigations. We believe that SVDB will be a valuable resource for snake venom & toxin community.

The architecture of SVDB ensures automatic growth & update of data as new records will continue to be added and updated in every 24 hours at NCBI. So, the user will always be able to access the most up to date, authentic and non redundant data from SVDB portal. This architecture of data extraction, storage & retrieval of SVDB through NCBI (Figure 1) is not limited for the snake venom and toxins only, but is actually very flexible and has the expandability to generate a generic model for any domain specific data that can be accessed through the NCBI. Therefore, we will continue to expand SVDB database with the new publicly available datasets and keep improving. We host SVDB on its own domain (snakevenomdb.org) instead of on an institutional domain: as they tend to change. Dedicated domain name & space improve the site's sustainability. We also believe that feedback from researchers could further enrich this resource so a feedback site been incorporated at the SVDB portal and we make the database online accessible to anyone in the world.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Figure S1: ERD of SVDB, Figure S2: Database contents and workflow of SVDB.

Author Contributions: Conceptualization, Jesmin, Ataul Haque and Md. Hossain; Methodology, Md. Hossain and Ataul Haque; Software, Md. Hossain and Ataul Haque; Data curation, Md. Hossain, Ataul Haque, Sayeda Mazid, Abira Khan, Tomalika Ullah and Jesmin; Formal analysis, Md. Hossain, Ataul Haque, Sayeda Mazid, Abira Khan, Tomalika Ullah and Jesmin; Funding acquisition, Jesmin; Investigation, Md. Hossain, Ataul Haque, Sayeda Mazid, Abira Khan, Tomalika Ullah and Sarker Rume; Project administration, Ataul Haque; Resources, Md. Hossain; Software, Supervision, Sarker Rume and Jesmin; Validation, Ataul Haque, Sayeda Mazid, Tomalika Ullah, Sarker Rume and Jesmin; Visualization, Ataul Haque, Abira Khan, Tomalika Ullah and Jesmin; Writing – original draft, Jesmin, Md. Hossain and Sarker Rume; Writing – review & editing, Md. Hossain, Ataul Haque, Sarker Rume and Jesmin. All authors have read and approved the final manuscript.

Funding: This research was funded by the Research & Development allocation Grant from the Ministry of Science & Technology (MOST), Government of the People's Republic of Bangladesh [39.009.002.01.00.057.2015-2016/1425, 39.00.0000.09.06.79.2017/BC-160/164]; and the Special allocation Grant from the Ministry of Education, Government of the People's Republic of Bangladesh [37.20.0000.004.033.020.2016.7724, 37.20.0000.004.033.020.2017.2147].

Acknowledgments: The authors thank Dr. Hasan Jamil (Department of Computer Science, University of Idaho) for valuable inputs and suggestions that helped improve the manuscript.

Conflicts of Interest: "The authors declare no conflict of interest."

References

1. King, G.F. The wonderful world of spiders. *Toxicon*. 2004, 43:471-475.
2. Tedford, H.W., Sollod, B.L., Maggio, F. and King, G. F. Australian funnel-web spiders: master insecticide chemists. *Toxicon*. 2004, 43:601-618.
3. Chan, Y.S., Cheung, R.C., Xia, L., Wong, J.H., Ng, T.B., Chan, W.Y. Snake venom toxins: toxicity and medicinal applications. *Appl Microbiol Biotechnol*. 2016, 100(14):6165-81, doi: 10.1007/s00253-016-7610-9.
4. Liu, C.C., Yang, H., Zhang, L.L., Zhang, Q., Chen, B., Wang, Y. Biotoxins for cancer therapy. *Asian Pac J Cancer*. 2014, 15(12):4753-4758
5. Casewell, N.R., Gavin, A., Huttley, G.A, Wüster, W. Dynamic evolution of venom proteins in squamate reptiles. *Nat Commun*. 2012, 3:1066. doi: 10.1038/ncomms2065.
6. NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res*. 2012, 41 (Database issue): D8–D20
7. [BIG Data Center Members](#). The BIG Data Center: from deposition to integration to translation. *Nucleic Acids Res*. 2017, 4:45(D1):D18-D24. doi: 10.1093/nar/gkw1060.
8. Newbie Introduction to cron. Unixgeeks.org. Retrieved 2013-11-06.
9. Crontab, The Open Group Base Specifications Issue 7 - IEEE Std 1003.1, 2013 Edition, The Open Group, 2013, retrieved May 18, 2015
10. Sayers, E. Sample Applications of the E-utilities. 2009, Bethesda (MD): National Center for Biotechnology Information (US)
11. Codd, E., F. Derivability, Redundancy, and Consistency of Relations Stored in Large Data Banks, 1969, Research Report, IBM.

12. Codd, E., F. A Relational Model of Data for Large Shared Data Banks. *Communications of the ACM. Classics*. 1970, 13 (6): 377–87. doi:10.1145/362384.362685.
13. Nadkarni, P. M., Marengo, L., Chen, R., Skoufos, E., Shepherd, G., Miller, P. Organization of Heterogeneous Scientific Data Using the EAV/CR Representation. *J. of the American Medical Informatics Assoc.*, 1999, 6 (6): 478–493, doi:10.1136/jamia.1999.0060478, PMC61391, PMID10579606
14. Marengo, L., Tosches, N., Crasto, C., Shepherd, G., Miller, P., Nadkarni, P. M. Achieving evolvable Web-database bioscience applications using the EAV/CR Framework: Recent advances. *J. of the American Med. Informatics Assoc.* 2003, 10 (5): 444–53, doi:10.1197/jamia.M1303, PMC212781 , PMID12807806
15. Wishart, D., Arndt, D., Pon, A., Sajed, T., Guo, A.C., Djoumbou, Y., Knox, C., Wilson, M., Liang, Y, Grant, J., Liu, Y., Goldansaz, S.A., Rappaport, S.M. T3DB: the toxic exposome database. *Nucleic Acids Res.* 2015, 43 (Database issue): D928-34. 25378312
16. Altschul, S., Gish, W., Miller, W., Myers, E., Lipman D. Basic local alignment search tool. *J. of Mol. Biol.* 1990, 215 (3): 403–410. doi:10.1016/S0022-2836(05)80360-2. PMID 2231712
17. Henna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D. G., Thompson, J. D. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* 2003, 31 (13): 3497–500. doi:10.1093/nar/gkg500. PMC 168907
18. Cuff, J. A., Barton, G. A. Application of multiple sequence alignment profiles to improve protein secondary structure prediction. *Proteins.* 2000, 40 (3): 502–11, doi:10.1002/1097-0134 (20000815) 40:3<502::aid-prot170>3.0.co;2-q. PMID 10861942
19. Benkert, P., Biasini, M., Schwede, T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* 2011, 27, 343-350.
20. Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J.F., Guindon, S., Lefort, V., Lescot, M., Claverie, J.M., Gascuel, O. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* 2008, 1;36 (Web Server issue):W465-9.
21. Guex, N., Peitsch, M.C., Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis* 2009, 30, S162-S173.
22. Harvey, A.L. Toxins and drug discovery. *Toxicon*, 2014, 92:193-200.