International Research Journal of Biological Sciences _____ ISSN 2278-3202
Vol. **7(3),** 9-23, March **(2018)** *Int. Res. J. Biological Sci.*

Transcriptomic analysis of *Agelenopsis naevia* (Aranae: Agelenidae) Venom Gland

Ahmed J.[*], Shehu D.M. and Ndams I.S.
Department of Zoology, Ahmadu Bello University Zaria, Nigeria
jamilamaaruf@yahoo.com

Available online at: www.isca.in, www.isca.me
Received 28th November 2017, revised 19th February 2018, accepted 5th March 2018

Abstract

The venom of spider is made up of toxins with varying biological activities whose selectivity and affinity for various receptors and ion channels are yet to be sufficiently explored and exploited. Recently, transcriptomics have been employed as a tool to reveal the molecular diversity and structure of animal venoms across species. Thus, this study was carried out to determine the transcripts coding for toxins in the venom gland of Agelenopsis naevia, collected in open gardens of Ahmadu Bello University Zaria, Nigeria. Venom glands were isolated using micro dissection followed by mRNA extraction. A cDNA library was constructed and pair-end sequencing was carried out. A total of 11,167,123 reads were generated which were assembled into 33,182 sequences. Fourty eight (48) transcript coded for proteins/peptides amongst which are sphingomyelinase-D, hyaluronidase, astacin-like metalloproteases, techylectin and cystine knot toxins. The results provide insight towards the discovery of novel potential bioinsecticides and/or drug leads from Agelenopsis Naevia venom for agro-allied/pharmaceutical applications.

Keywords: *Agelenopsis Naevia,* transcriptomics, venom, mRNA, cDNA.

Introduction

Spiders are venomous arthropods with an estimate of 120,000 species[1]. Majority of spiders including *Agelenopsis naevia* employ a fatal mixture of toxins which they use to overcome their prey. Most spiders feed on other arthropods while few species feed on other animals such as reptile, fish and birds[2]. Consequently, spider venom is made up of toxins which targets vertebrate and invertebrate species[3] and have different biological activities. Some spider toxins have been used pharmacologically to probe structure and function of receptors and ion channels/ potential drug leads[4-6] while others are being developed as environmental friendly bioinsecticides[7,8]. Unfortunately, during the recent past, only a handful of spider venoms have been explored extensively, as such only about 0.01% of venom toxins have been identified[3,9-11].

Agelenopsis naevia belongs to the family Agelenidae which is a large family of Araneomorph spiders commonly referred to as grass spiders[12]. This spider shares the same genus with *Agelenopsis aperta* whose venom has been extensively characterized[13-15]. Members of this family build webs with funnels in various ground habitats[16]. Many Agelenids have light brown bodies with paired darker longitudinal stripes on the cephalothorax (and sometimes on the abdomen as well). They have long and conspicuous posterior spinnerets that extend well beyond the tip of the abdomen. *Agelenopsis* have spinnerets nearly twice the size of the basal segment which is a distinguishing feature of members of this genus[17]. Agelenids trap their prey on their webs using a barrier of silk above the web; they then come out of their web almost immediately, envenomate the prey and return to safety, where they feed.

Spider venom toxins are produced in combination. Based on estimate of 100,000 species and 200 peptide toxins per venom, there are over 10 million toxins in the spider venoms which are biologically active[18]. Transcriptomics have been extensively used as a tool to uncover the diverse molecular and evolutionary relationships of venom toxins[19-23]. Despite the high throughput techniques that are available for exploring spider venom diversity, only 1544 toxins from 100 spiders have been deposited in Arachno Server 2.0 database[24]. The importance of spider venom has recently been explored by scientist globally. However, there is dearth information on spiders' especially their venom in different geographical regions like the tropics. As such this study was design to provide baseline information on *A. naevia* venom transcriptomes which could be engineered towards pharmaceutical and/or agro-allied applications.

Methodology

Spider collection and identification: The samples of *A. naevia* were collected in open gardens of Ahmadu Bello University, Zaria (11°9'N and 7°39'E), Nigeria. Zaria is located at an altitude of 199m above sea level within the Northern Guinea Savannah ecological zone of Nigeria[25]. Funnel webs of *A. naevia* were irrigated with water to force spiders out. *Agelenopsis naevia* were caught and handpicked into cleaned sample bottle for identification. They were identified using the keys of Dippenaar-Schoeman and Jocqué[26] Whitman-Zai *et al.*[27] and World Spider Catalogue[28].

Venom extraction: Fifteen *A. naevia* were electro-milked for venom to stimulate the production of mRNA. After three days they were anesthetized by freezing at −20°C for 5mins. The dissection area and equipment was cleaned with RNase away to avoid contamination by RNases. A petri dish was placed on the microscope stage and filled with 1X saline sodium citrate buffer.

The anesthetized spider was placed in the petri dish and the cephalothorax was immediately separated from the abdomen. The cephalothorax was held with forceps and the cuticle linking lateral aspects of the chelicerae was cut. This was followed by grasping and gently tugging of the chelicerae back and forth to pull out the venom glands. The venom glands were then separated from the chelicerae[29] and placed in RNA later which was stored at 4^{O}C until use.

Figure-1: *Agelenopsis naevia* venom apparatus.

Messenger RNA (mRNA) Isolation, Complementary DNA (cDNA) Synthesis and Sequencing: Isolation of total RNA (TRNA) from venom gland was done using Gene JET^{TM} RNA purification kit while Mag JET mRNA enrichment kit was used to isolate messenger RNA (mRNAs) from total RNAs according to users' guide. The construction of cDNA library was accomplished according to the method of Luna-Ramírez *et al*.[30] with slight modification. Isolated mRNAs were subjected to polyA enrichment followed by fragmentation with fragment buffer treatment. The first strand of cDNA was synthesized using random hexamer with mRNA as templates. Buffer, dNTPs, RNAase and DNA polymerase 1 were used to synthesize the second strand. This was followed by adenylation of the 3' ends, adapter ligation and DNA fragment enrichment. The double strand cDNAs were further subjected to real time PCR (qPCR) quantification using KAPA SYBR fast qPCR kit (following user guide) and pair-end sequencing was carried out using illumine NextSeq 500.

Sequence analyses: Clean read were generated by removing adaptor fragments. They were subjected to quality control analysis to measure some parameters using FastQC software[31]. Transcriptome assembly of paired end reads was achieved by using Trinity RNA-seq software[32]. This was followed by mapping raw reads to assembled transcriptomes using Tophat[33] on galaxy online software package.

The resulting contigs were searched against public database containing toxins from spiders using the BLASTx algorithm with the expect value (e-value) set to$<10^{-5}$. This was done to identify sequences homologous to query sequences. Sequences producing hits in NCBI were further searched against arachno server[24] to screen for transcript coding for venom toxins. Gene Onthology (Go terms) for sequences presenting hits to toxins in venom were obtained using Blast2go software[34]. ExPasy-Translate tool program[35] was used to translate nucleotide sequence to unveil precursor peptides. The signal peptide, propeptide and mature toxin were determined using SpiderP 1.0 program[24]. Multiple alignment was carried out using CLUSTALX v2.0. The phylogenetic tree of spider toxin families was constructed using MEGA 6.0 (maximum likelihood) running 1000 bootstrap replicates[36].

Results and discussion

The resulting paired end reads were 11167123 with 35-76nt sequence length. Average percentage GC content was 44.5%. Table-1 shows 90.1% right and 90.8% left reads were mapped to assembled transcriptomes with 35.3% right reads and 35.4% left read having multiple alignment. Resulting reads from this study were deposited in Sequence Read Archive of the Genebank with accession number SRP073132. The reads were assembled into 33,182 sequences of which 11,251 produce hits with sequence in the gene bank. Since our main interest is on spider toxins, sequences producing hits were further filtered by searching arachnoserver[24] of which 48 putative proteins were identified (Table-2).

Some of these toxins were found to be lacking signal and/or propeptide sequence (Table-2). Only sequences presenting hits in arachnoserver where annotated. The GO database comprised three ontology domains; Molecular function (MF), Cellular component (CC) and Biological process (BP). The most abundant Go-term in this study is the cellular component (Figure-2). The pie chart (Figure-3) shows categories of toxins identified in transcriptomes of *A. naevia.* Twelve categories were identified. These include enzymes (15%), neurotoxin (7%), cystine knot (17%), lycotoxins (21%), venom allergens (8%), ankyrin repeat toxins (10%), calcium dependent toxins (4%), scytotoxin (6%), BPTI-kunitz (6%), ctenitoxin (2%), IGFBP (2%) and TCTP1-1 (2%).

Table-1: Summary of assembled statistic of *Agelenopsis naevia* venom transcripts.

Measure	Right Reads	Left Reads
Input(nt)	11167123	11167123
Mapped (nt)	10063923	10144990
% Mapped	90.1	90.8
Multiple Alignment	3551626	3593126
% Multiple Alignment	35.3	35.4

Table-2: Identified protein sequences encoded by transcripts.

S/N	Sequence	Blast description
1	MSTTACHSQLTLTSENDKFRNPPWWNRRTKLERTLCFVTVTSLLMLSVMAAALAVF GYLYQKNIDSRQTIYVPPNTERLVSSNASAAHLCLTPGCVKAAAEVLSNLDESVDPC DDFYRFSCGGWLERHAISEDKSSVSVFSEVQDDLNLKLRVLLDKKLDGSEPQFVQMI KDMYDTCMDVKSIEKAGAEPLKKVLKSLGGWPVVEGERWDGSQFDWVDSLISFRQ LGYSHDILMDLSVTADYRNNTIHIIDLDQTSLGMPDRTYLVRGLNDTSTAAYFHLMV KSAKLLGADPKTVENELLKALHFEIALANFSLPREERRNISKLYNKYTVGKLRDLVP QIPWLKYFNGLLNDEIGEDEPLIVAVPDFIVKFADLITLTEKRVVANYMLWRVVAQS LSILSKDWRDIAQEYSSVITGKSQEEPRWEQCLSSLTGSLGIALSSYYVRHYFKEESKE AALELVRYIHSEFLNILQDIDWMDEVTKREAKAKANTIQSYIGYPKELLNDTAVMEL YENLTITRRGYFDNIVGLRKWSTDYAFGQLRKPNIRGEWKKHARAAVVNAFYNSLE NSIEFPAGILQNAFFNKDRPHYMNFGAIGYVIGHEITHGFDDRGRQFDKDGNNVNW WDETTDELFKQKAQCIIDQYGNYTAENGMQVNGVNTQGENIADNGGLKQAYWAY HSWVKDNGPEPLLPGLKYSQSQMFWISAASVWCGKYRPEVLKLRIISGSHSPPQFRV VGPMSNLAEFAKEFDCPDHSPMNKKHKCRVW	Neprilysin-1
2	MWLYKRDTCIRFVPRTNEANYIRIFPGQGCYSHVGLTLVGAQPVSLGQGCGYMGTVI HELAHALGFYHEQNRSDRDDWLTIYWDNIKEGMEPQFMKLKPDQNQLLTPFDYDSI MLYGSYTFSKEYGKLKTMEGKGGMFLKDVIRKYFMSKSDIQRIKMLYKCTN	Astacin-like metalloprotease
3	MPAGCIQFSSNRFPWLDDMPGLSEDCLYLNVWAPKGFRGKKSVMFWIHSGGFRISSI RVEYFDGLILAALGDVVVVTVNYRLGPQGFLYSATEDAPGNMGLLDLVEALKWVR NNIESFGGDKDSITVFGHSSGGIAAGMFCVSPLTRGLFQRAIMQSGSATNVDGEDND RNFRVSQQLAEEVGCADDANSLAEEPEKVVECLSNVDPFILSKALSKVTANSPRSYY PRFGDEVLPKNPRQAFIDGDFAAVDLLIGTNNGDGATLITKSMKDVFGFFGEKDPQIN MTFGESIIRKKFKDYPDPETVVQQYLGNVSEDDYHTIRQQVFNASGDYARTCPSVYM AESFAKKDALVYYYYFTRHNSAMPWAPWMLAAHFDETIFVFGRPLAHPDDYTAGE VALSKRLLRAWSDFAKTGIPSGGEWPLYSKENPRYQILDVDKSGTGQGIHGTNCDFF RPHFGF	Acetylcholinesterase-1
4	MISDPLFEETPGNIPAALKLSSKRQVFKGGWTHENKMQILQLHNIYRGNVTPPAGNM AYMEWDDELEELAQKWADGCVFEHGMVANNYPDTIGQNLYIGTDPTGYLGLYLW YEEYHHYDLRKENCEPYKQCGHYVQMAWAESTKLGCAMTKCNMRFYVVCHYS	U21-aranetoxin-Av1b
5	MAVEQLREVCKSRGANGIIGLSRCFRLMDDNRDRKLDLEEMKLGLQEYGANMDAG QVAELFKEIDKDGSGTVSLEEFLRAVRPPLSQDRLDIIERAFAKMDKTGDGSVTIDDL KGVYSAKEHPDVIDGKKTEDEVLEEFLQCFDTPNKGDGKVSKQEFIDYYTGVSSNID LDEYFVTMMEKAWQL	Ef-hand protein 1
6	*MGHPSVVSLLLFWGAAVDTIDA*EGRTVLSIAAAQGDPDTVRQLLDRGLDEMHRDNGG WTPLHYAAFEGHTEACGLLLEAGAKITEVDNDGRIPLILAAQEGHTALVGKLLDHTP SMIDCRAHDGKTALRIATLEGHKETVELLINHGADLNYKDADGRSTLYLLALENRTD MAEFLLSKGADVGARDLEGRTALHVSAWQGHTDMVELLLNHGADVNAVDNDYRT ALQSAAWQGHVSVVRLLLERNAVVDHVCNQGATALCIAAQEGHLEVVKALLEYGA DPSHADQFGRNPIRVASKGGHNNVIRLLEEYMANQHDVGNNFSGSVSTTSLTSASTA	Alpha-Latrotoxin-Lt1a

S/N	Sequence	Blast description
	ETKPCSAILCQPGVSAQSPIESPESTVDKRRSFISNQSSSKSSSNLTNSTNKSSSKTHSQP PQQISQDSCSNLHAALAIQNEKAVGSQMTFTQQLQQCTRNRNRISRLLSPLSEPHSPV PSPPQSPLSDIQGPVQSTSPSGEGAYNVPHMPHSTHVPIVIPPIIHESSRQSHLPRRISET MNPIDSSHPIPRPTEPRIRRNGIVTNPNYKGNNIMSAVPVKNSPVKQSTAVNNHSGQIK SKHISQYEEPLKTTAMPFKKETHL	
7	MLACTNDNVEIIEFLLSHGADLSIKNKDGWTPFHIAARGGHINVIHCLLNADPNVWN TTSKNGRTPLHTAALHGHLETVEKILSYDNEYKNVQDSCGTTPLMDAASGNHKHIV DYL	Alpha-Latrotoxin-Lha1a
8	MTNKEGVVDAGSGNQEQTNIHVESDSNASFLRAARAGNVEKVLEYLKGSIDINTSN ANGLNALHLASKEGHVAVVVELVKRGANVNAATKKGNTPLHIASLAGQEEVVKILI RNGANANVQSQNGFTPLYMAAQENHDGVVKFLLANGANQSLATEDGFTPLAVALQ QGHDKVVAVLLENDAKGRVRLPALHIAAKKDDCKAAALLLQNDHDPDVTSKSGFT PLHIAAHYGNENIASLLLDKNADVNYSAKHQITPLHVAAKWGKSNMVTLLIDKGAK IDASTRDGLTPLHCAARSGHDQVVDLLLEKGAPITAKTKNGLAPLHMASQGDHVDS ARILLFHKAPVDDVTVDYLTALHVAAHCGHVRVAKLLLDRKADPNARALNGFTPLH IACKKNRIKVVELLLKHGASIEATTESGLTPLHVASFMGCMNIVIYLIQHGANPDVPT VRGETPLHLAARANQSDIIRILLRNGAHVDAS	Alpha-Latroinsectotoxin-Lt1a
9	MEDIFHWCREGNAFQVRVWLDDTENDLNLGDDHGFSPLHWSAKEGRANIVEMLIL RGARINATNMGDDTALHLASAHGHRDIVHMLLRSKADINAVNEHGNTPLHYACFW GYQAIAEDLINNGALVAICNKFGETPLEKCKGQMAQKLHELAVQLGQDTKKVPFKD ENWLGTKTRSRDATLSRHAGINMKEISFEVKLANTPSGETWKGLWQKNYIVAKILKL RETPPRVTRDFNDEYPRLRIFSHPNVLPVIGCCVSPPSHIIINQYMPFGSLYTVLHKDTG IVVDTAQALKFAIDICRGMAFLHTLDPLVPRLYLSSKHVMIDDELTARVNMADAKFS FQEKGKMYSPAWMAPEALQKKQDDINVKAADMWSYAILLWELATREVPFSDLSPM ETGMKIALEGLRITIPPGISQHMSRLIRICMNEDAGKRPTFDMILPILEKMK	Delta latroinsectotoxin-Lt1a
10	MSELEICRLAYEGKFNILKDKIDKDASYLIARDGSGRIPLHWAASGSQTNIVNYLLQQ GSPVDDADDTNWTPLMIASSVGRADIVIALINRGADVNAINQTGQSSLHYAASKNRE DVAKILLDNHAHINAADNMGSTPLHRAASKGNMKILKLFIEEYSNQLDLNQKDSVG NTPLHLACEEERMEAVKLLIEAGANTDTLNKEKKNPIEMARPEFRPVIRRMIESAIST	Alpha-Latrocrustotoxin-Lt1a
11	*MKVYILFAVVAVAIA*RKQCGENEHYNGCGTDCPLTCDNYDNPPKFCNLMCRIGCECDQ GYVRSSDGRCVRPEECPARVQRQALEQTCTEKPDPGMCLAYIPSWYYDPESGECKEF VYGGCQGNGNRYKTEEECLSKCQAPKRAPQACLEEKVTGPCKAYFPR	U32-aranetoxin-Av1a
12	*MLSRVLIFLTFWGFLSCRASA*DDSQCTSLLLKNITDETQSSIVVILLNNAKSLLENAKKY ATMKTRCTATETLCDKPMDCADVKEKNPFAQSGVYKIWPRNRVIKGPIEVYCDMEE DGGAWTVLQRRGDFDSDKEYFYEDWQAYKKGFGDLSKDFWLGNDNIFALTNQRSN IVKFTLTDWEKNQTYATYDEFWIDNEAHKYAMHVSGYRGTAGEHSCT	Techylectin-1
13	MVCVRKKKECKNPEVNDTPPGTCFTGSRKHGQRACIKEGEKAGDCQCCGKWSYCS CPLFGALGCSCICARDKPKRSVLG	U9-agatoxin-Ao1a
14	*MMIAMKHLLLATIVAFLFVQALA*AVPYPPPLDRNLSEDYNENVDM YMKADKRACIRRGGGCDGKPNDCCPNSSCRCNLWGTNCRCERAGLFQQWGK	U8-agatoxin-Ao1a
15	*MLVSKRIMKTIPALILICLLHTSIA*SILVDTEETSDEFTVETAEAPEQARKTCLSTGSECDG SKDDCQCCGAYVECKCPFGINWPSVLGP CKCSMDHLGTCLAKQKCPNKHEWGGGDCKTPRKPRYG	Omega-agatoxin-Aa1a
16	*MKLTIFTGLVLLVVVSLIEA*EAESERSCITLEKVCTEKKGNCCSG LKCDCYQRYEDGVLKGKRCWCLENNVRYARDKK	U6-lycotoxin-Ls1g
17	*MSSKVQAVLFV*VGLLTFLVVQAKKEISDNTELGRDRCIEVYKRCEWNGTPCCENRACI CSWIGTNCKCKRKWIGK	U16-lycotoxin-Ls1a
18	*MIRFLFISALLIIAVHS*FAIEDNEDLFLKDAEDEIIPEEARSSLPPGVCDGNKDDCQCFGE NYKCGCPFLWPMRSGKCHCKKGWRT	U2-lycotoxin-Ls1d
19	MRSGKCHCKKGWRWTRIKKRSCRNRYEWLR	U2-lycotoxin-Ls1c
20	*MNPRVFSILLILGIATCVLA*GGFCPKSRHPTCNLGYKINDCCAQSDCRMGSVCCVEGCG NVCRQESDTAIGEKFVDGSECQLGHYPKRWYEFW	U20-lycotoxin-Ls1b

S/N	Sequence	Blast description
21	*MNPRVYSVLLILGIATCVLA*GGFCPKSRHPTCDLGYKINDCCAQSDCRMGSVCCVEGC GNVCRQESDTAIGEKFVDGSECQLGHVYPKRWYEFW	U20-lycotoxin-Ls1c
22	*MKLSIFFVLLLVAIAYCEH*ENLETEDDFETGDFEAEETLPAAPEEREKNCIRHKDSCMH NDKGCCFPWSCSCWDQKVRTGSNKTVRKCQCRFTFWT	Omega-lycotoxin-Gsp(267)1a
23	*MKFLVLFGALFLTLISYTST*EPTDDFDSYLDDGTFEADDIPSLMAEKARDKACTPRYHE CTDDRHSCCRSELFKDVCTCFYPEGGEQGRAGKKRALYMPTTQTSQVYRKGR	U3-lycotoxin-Ls1y_2
24	*MKYQVLFGVVFLTLLSYCSS*EIDGDFESYADEKSVQADDLTSLFMNEQNRKDKENCIPK HHECTSDRHGCCRGHMFKYKCHCKFINDKKEENDRCACITPGIHKAAELVVQ	U5-lycotoxin-Ls1a
25	*MNSKIFAILLLLALSVCVLS*DEYCPKPADTYCHWKHMRNDCCD EDCSVGSLCCKDACGNFCKRPVADPAGERPDPAEACELGYY	U15-lycotoxin-Ls1d_1
26	MRIFKDIITGDEMFTDSSKYKLVDDCLYEVECRHVQRRLGDVQLEGANPSQEEADEG TEEIVESGLDLVLNQRLVETGFSKNDFKSYLKMYTKSLQDKWKEMGKSDSEMAEAK SKFTEAVKKVLPKIGDLQFFMGESSNPDGLIALLEYRENASGDETPIMIFFKHGLEEEK V	GTx-TCTP1
27	*MEGNMLMFLLVLSSIGLALS*NPLVNPNSGNTPQCLLPPTSGLCLAYIPSFYYNPSTEKCD TFIYGGCQGNANRFATRDECLAKCGGFKQERDESFKLIHDEEQSSSDSEQQVEDICSL SSKMGPCRAMMPRFYFNGDKCENFIYGGCQGNGNNFKTLEECQQKCGSTSSSAEEA EADNSEETEKASDEAGSQEASVKSDEESDNKTEN	U19-barytoxin-Tl1a
28	*MMLKALLLLSVAVVSYA*IVCPHDYCSKVDCEELTDCSESNGMRVREKGSFCNCCDICV KVLGEGERCQPEGEFLGVIITSECGKDLYCDYETRRCIRT	U16-pisautoxin-Dm1a
29	*MKLAIFLLFSLVVLVVA*SESVEENKREDRPEQQRACADLNEKCTK GDNCSCCGDRGKCDCNWPNKPGCYCMRGGPIDAIIKKFEC	U5-pisautoxin-Dm1a_3
30	*MKVTLALILLASLLCLAQS*QWCGDGECGEGQCCSGGFYHRYCRRYSDEGEPCERPNKF NEYKTACPCKEGLFCSVIARCQRN	U17-pisautoxin-Dm1a
31	*MKCIIALTILATLVVAIQG*KYCSRSSECGEGMCCTGGSFNRHCQALADNGRPCQRPND YDSYSTGCPCKEGLICSVINYCQEA	U11-pisautoxin-Dm1a
32	*MRTSTLIGLCTVVVLLLTIADISGA*DEVDSQNIPEERGYCAEKNI RCDDIHCCTGLKCKCDSGGYNCVCRKG	U8-pisautoxin-Dm1a
33	*MIKYILISSLLVVAVYS*FAVENEEEIAQDAEDEVMPEEARSLPPG SECDGNKDDCQCYGKWHKCGCPLFGGKCTCQKGYKYTCITKLSCPNKGEWGLDW RSEESERSPC	U7-pisautoxin-Dm1c
34	*MAIKMFALLLLTSFLIGHEVQS*SVYTGEEDTSDGYCDIQFFERIP VGETGFNDEECLEASCYSNSYEVRGCGSVHNEDERCTLVEGKGHYPDCCPTLKCNTI DGE	U13-SYTX-Sth1a
35	*MLPPFPLSLAWTLLAAARTFA*DTGLRSLLIHERNQELLDHLGTGY EIVYPHQIRKDWSRGLSTALAGTVLHVPETTLLIETYENKLQLDLELNTKLFVPTLVQ LLYPKQASPVPIHKLPENCYYHAKVRNHPEAKAALQTCHGIRGVIYLDGKTFFIHPLH GNHSGKHPHLLYHYFPEDDLRCGSTGAPDGDPSESGARHEEKDDWHRNKFIEMAIV LDQSLMEKFRTPLAEAVPAALEIVNHVDLLYRPLNTSVSLVYAEMWMENQMSLSSD DLPQDVLQRFQEYAARRIARISVDTTHLLTGLELKDGIRGISFPDSICTQNAVGISQVR HLFEPHVTAFVLSHMVGHNLGMDHDHPGCECPDRAGCIMEDKVPILASTILSECSVQ AYQRTLSRGYGACLYNMPAMRVSICGNNILEEPEQCDCGTVEECLRSNPCCDPITCR LIKHAQCSSGPCCRKCKLLSPDHLCRPSKGECDIPEFCDGTHEQCPEDVFMRNGAVCS GGQGYCFRGECPVLQQQCRDIWGTGSESADSVCYERLNVQGTPNGNCGTDGRGGIV KCPTQHAGCGFLQCRSGQKTPVSSEDERVHFVVHKIAAQGTVHECKARTLPTSELSH GLVKDGTKCGYQKLCFNQTCTSLKSLVSSKCPLEDVQNTCSGHGMCTNANTCFCED GWKGEDCGTVDEDWRPDYEYDPAHNNASSDHHPYGDSEMRLNPNLAVTVITASTV VGVGLVLVLLLLFFYRRP	U1-filistatoxin-Kh1b
36	*MKFVILLLAVAVVAVYA*ERPKSDCEKHRESAEKTGTIMKLIPKCKENGDYEELQCYRDS KFCVCYDKKGHAASPISTKLKECGCYLKRKQKLDSGHDNAYIPQCDDDGSWVKMQ CWDYNGSCWCVDKDGERVGDIKSGPERLNCE	U24-ctenitoxin-Pn1a
37	MASASAEEGDEAALQAEETPRKGCLEIGEVCDGDKDDCQCCRSNGFCNCSWIFNHC TCQVGDASKSYGVCLWKQKNCPNKPGKCTKPCTNRRCNRRRNGK	Omega-ctenitoxin-Pn3a

S/N	Sequence	Blast description
38	MNSFLLAVAIPLFFMLITVQSYSEPDCSEPCDRSYCPTIPPCKCGTYLDYCKCCSFCLG CPGEKCNLVGHQKCANGTTCKAREGAEGMELFNGPGTCR	U12-SYTX-Sth1b
39	*MKKAFLIILLIVGLLCPIDG*RLNRFKRLIRLDLNNPLTRDVCDYPL DPGRCFSAHRRFYFNGDTCEQFTYKGCLGNSNNFRTIEECMLKCAHKLDGGSGEIED NEHFSFSLFG	Kappa-theraphotoxin-Hs1a_3
40	*MIAPNVPLFSFLLVTWVWTAEA*KLTKSKYHGRSIPQRELDTRFSNTKKKIVLIHNFYRA KVEPPASNMLEMTWHKGAQRAAQGWAEACELLTHDDPLGRWVEDYGSCGQNIFV SSTPVDWMDVAKAWFMEHKNFTYGSGKNDLARIGHY	CRISP-1-*Grammostola rosea*
41	*MRVFLLFFLVNIIVIYAVA*QTSTECPELFLRYSSTHSYCKPANPKC KILRSFVNDEEKNTIVKIHNSYRSKLAIGEEQNLPTAADMLQMVWDDELAAVAQKH AEQCLFEHDCANCRRVQNFGVGQNLAIEEVFNSPSIPNSNWPSAIKSWYDEIAYFSSK AIDPFIAPPPEPTYGHFSQLFWASSWRIGCGFVLHQDRNMYKKLYVCNYGPAGNAVS QSMYKMGSPCSACPVNSCCDKTCGTKAEYSGLCKITDNKAPNYKPIGPYIFYCDFNN QSDCKSFVTGINKWLIFNTLSGSYLSITLHGGEESSMTFQNKIKPTTSDFCFVVMYRK GPNEAGQEEANTAIETFSGNGFEFTQNLPTFDGDVSQQFSQFSMTLSWNMETEFKITF AVPAGKPDQFLEIQSIYATEGLCKQ	CRISP-2-*Grammostola rosea*
42	*MKCAFLFLALFAVLILDVEC*TSPVRCAEECDKSMCKSVECPCGKYLDQCGCCEYCYK CPGEECLTMANELCTEGYKCEHAKGTSGPDYWNKPGRCIKME	U11-SYTX-Sth1a
43	*MAPSGILILFMMFTSCVC*DSRRPFYIIAHMVNSIEQIRRYLDLGAN VIESDIQFFPNGSVREAYHGFPCDCFRTCTRSANLREFLTYVRNITDPYSPGSYSNRMV MQFFDLKLSGVSDKNAAGRDIARHVIDYLWDKKRKQEIRALIYIDSVEDRDALEGFV EEFSDRGLESRLKD	Sphingomyelinase-D
44	MRLSTSCGINAVVSQMVGGWSDVEDTDSSDMQKLVHFAAEKISSSMGRPNKLQPIE VLSAKQQVVAGMKYKFIMRIGDPGCKIPKNIEECNLESCPEIHKCAATVWERTWLND IKLLKYSCEKEENFCTS	Cystatin-1
45	*MMWISILLLALAVPAVRC*ERKCDECNAEKCQPPSEECLAGLVKDLCDCCYVCGRREGE LCDGDFLPIPYRNRGHGPCGEHLECRPRTDLAPGDPPEAMCVCLKTETLCGSDGKTY ENECQLTEARYKQRDGLRAMHRGPCRSAPKITSPPEEASNYTGGNIAMSCEATGWPI PVFEWRVDIGDGNTIPLPSDDPKVSVQTRGGPSKYEVTSWLQLLSIQPKDDATYWCI ARNDEGESSAAARVVVLDFRGSKTADSGRNNDL	IGFBP_rP1-1-*Cupiennius salei*
46	MTAYCPDPYKRYSAEHTFCLGRSPCRPYRRGVSPWLKKELLREHNQLRSKVAGGNT YGVDHLPRATNMLEMVWDDELADIAQKWVDRCERRSDCDQCRRVGRFGVGQMIM EFSGKFEGEDLGLEFFSYHQGLMSFRKEQVTRFSGNPKTSQLLWAKTWRMGCGFGD FQYSGDGRRYSLIVCNYGPKGNVEGEEVYKAGSVCSACPANSCCGDACKRHKLQSS YHGLCKVIDENLPPEGNTPHPKSGDELFYCGFSGEEDCGHSVDGVDRWIRNVSTGGT WLSTTLGNYGYTILRFHQPIVSKSGKLCLSIRTRSGPLEAGQDYRYKLTGSMDEEGRY STTMVFPDAEKSVEHQFHTDYFQPGDFPKNREVKLSLSFSVDAMGPKQYFEIERIVA REKDCPKKTITEGSGFAVTR	CRISP-1-*Lycosa singoriensis*
47	MNSLVSFAKRAAMESRIFFINLLVIASGFEIYWNVPSGQCVHNYKMSFVQLLRTYGIR VNEGDRFQGNHVTIFYEAQLGLYPKILKSGKMENGGPQQGNLDQHLEKASKDIAKVI PWPDFDGLGVIDWEAWRPTWEFNWEPLRIYQTKSIELVKQMYPTAKDPFIKEMAEQ QWENSTKSYMLETLRLAKKLRPRASWCYYLFPDCYNYVGKKPKDFQCSASIRKGND KLQWLWKDSTALCPSIYVYENQLDRYSLDQRTWRDNEKLREALRVASRTAKIYPYV NYFDKELIPEDEMWRMLAQAAAVGGNGAVIWGSSASVRSEQHCKD LKQYIISTLGPAAEKVAWRSNLCSKQICNNQGRCTFPDDDYAKAWKLFIDDTVPFYA GNITCRCSENYTGRFCEKKNQLHFSNEILNFNDDENTERNQKIDFIQVSKV	Hyaluronidase-1
48	MVDYGLTEEQVAEFKEAFLLFDKDSDGMITAAELGVVMRSLGQRPSEMELRNMVH MVDKDGNGTIEFDEFLSMMSKKLQDSDSETELREAFRVFDKNGDGFISPEELRHVMT NLGEKLTDEDVEDMIREADLDGDGLVNYSEFVFILTSAK	Calmodulin A-like

NB: Signal peptide- italics, Propeptide- Underline, matured toxin- Bold

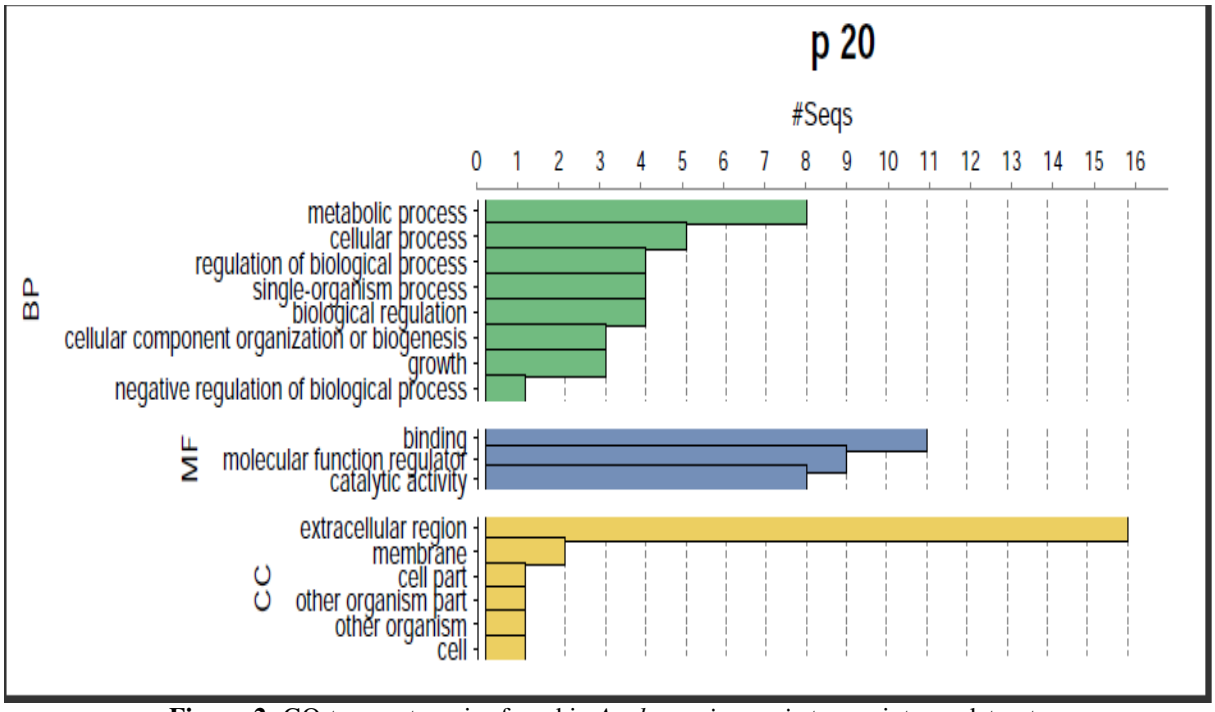

Figure-2: GO-term categories found in *Agelenopsis naevia* transcriptome dataset.

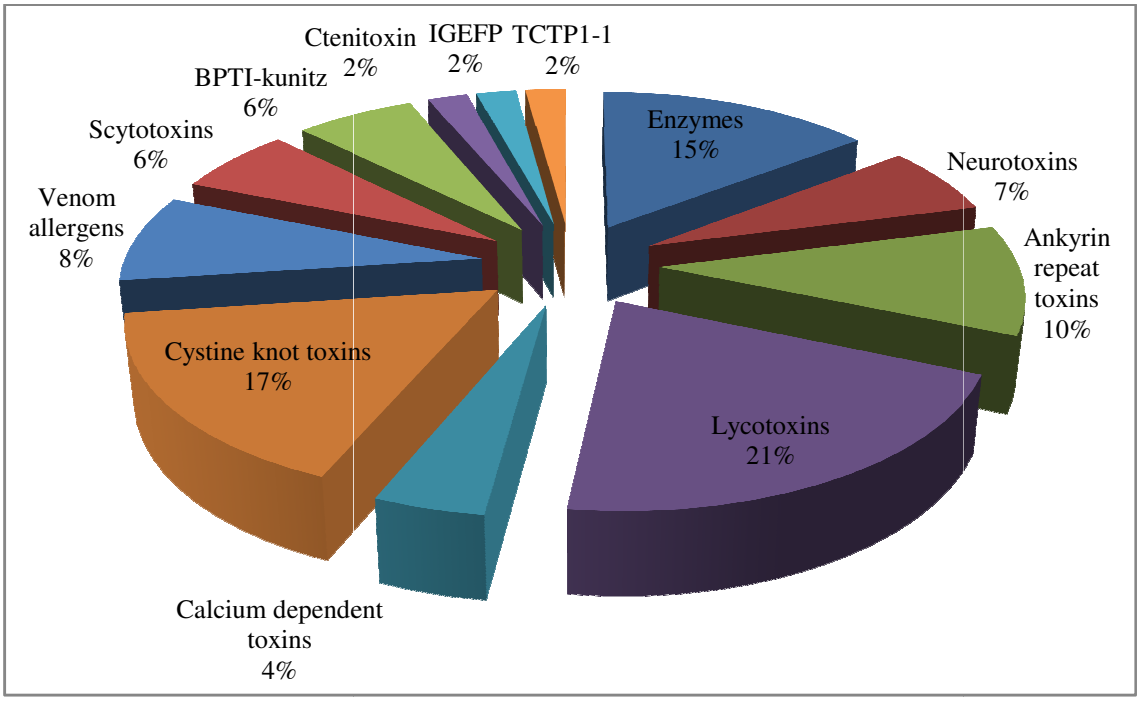

Figure-3: Percentage spider toxins and enzymes found in transcriptomes of *A. naevia* venom gland

Multiple alignments of some sequences with similar mined sequences shows conserved cysteine region and other regions (*). Conserved cysteine region is more abundant in putative U8-agatoxin alignment which is a cytine knot toxin (Figure-4). Figure-5 shows phylogenetic relationship among some sequences identified in this study including mined sequences. The tree was rooted to *Strigamia maritime* as an out-group. Seven distinct clades were formed with sequences clustering with closely related sequences. However, alpha-latroinsectotoxin-Lt1a did not cluster with the remaining ankyrin sequences but clustered with metalloproteases. *Agelenopsis naevia* venom gland transcript showed homology with those from other spiders, scorpions, tick, termite and centipede.

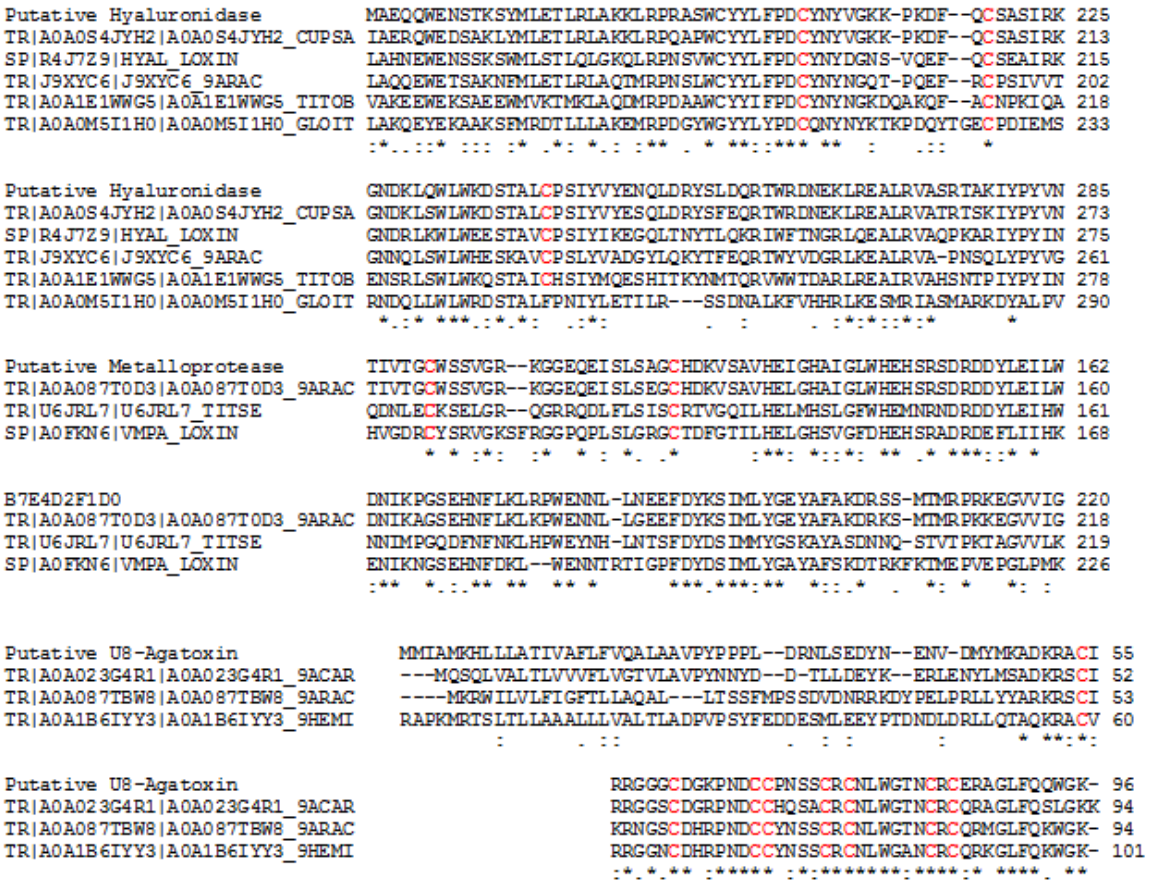

Figure-4: Multiple sequence alignment of some transcripts found in *A. naevia* gland showing conserved regions.

Only approximately 44% of the sequences produce hits with sequence in the gene bank indicating that many spider specific venom transcripts are yet to be discovered. Details of each category of proteins/peptides identified in the transcriptomes of *A. naevia*, is as follows:

Lycotoxins: This is the most abundant (21%) category found in the transcriptomes of *A. naevia* venom gland. Ten sequences homologous to lycotoxins were identified. These are: U6-lycotoxin-Ls1g, U16-lycotoxin-Ls1a, U2-lycotoxin-Ls1d, U2-lycotoxin-Ls1c, U20-lycotoxin-Ls1b, U20-lycotoxin-Ls1c, Omega-lycotoxin-Gsp (267)1a, U3-lycotoxin-Ls1y_2, U5-lycotoxin-Ls1a and U15-lycotoxin-Ls1d_1. Lycotoxins were first discovered in venom of *Latrodectus tredecimguttatus*. They share a characteristic inhibitor cysteine kont (ICK) like motif and may have neurotoxic activity on ion channel[37]. The high abundance of lycotoxins may indicate their functional significance. Signal peptides were predicted in all the sequences except in sequence homologous to U2-lycotoxin-Ls1c. Propeptide were also predicted in putative U6-lycotoxin-Ls1g, U2-lycotoxin-Ls1d, Omega-lycotoxin-Gsp(267)1a and U3-lycotoxin-Ls1y_2. Similar findings were reported by He *et al.*[38] from venom gland transcript of *Latrodectus tredecimguttatus.*

Cystine knot toxins: This group of toxin consists of putative agatoxin and pisautoxin that were identified in this study. These are sequences homologous to U8-agatoxin-Ao1a, U9-agatoxin-Ao1a, omega-agatoxin-Aa1a, U5-pisautoxin-Dm1a_3, U17-pisautoxin-Dm1a, U11-pisautoxin-Dm1a, U8-pisautoxin-Dm1a and U7-pisautoxin-Dm1c. All these sequences have predicted signal peptide except for putative U9-agatoxin-Ao1a. They also contain 8-12 cysteine residues. Cysteine is used to form several disulphide bridges (1-7) in venom peptides known as cystine knot[7]. This knot creates highly stable proteins which are able to resist extreme pH, organic solvents, proteases and high temperatures[39,40]. Omega-agatoxin-Aa1a is an insecticidal toxin from *Agelenopsis aperta* venom which inhibits neuromuscular transmission in both vertebrates and insects by blocking voltage-gated calcium (Cav) channels on presynaptic terminals, reduces the content of quantum at neuromuscular junctions of fly and also affects insect neurons by blocking calcium action potentials[14,15]. U8-agatoxin-Ao1a and U9-agatoxin-Ao1a are toxins identified from the venom of *Agelena orientalis* with a yet to be discovered activity. Homologous pisautoxins reported in this study are toxins reported from venom of *Dolomedes mizhoanus* also with a yet to be discovered activity[41].

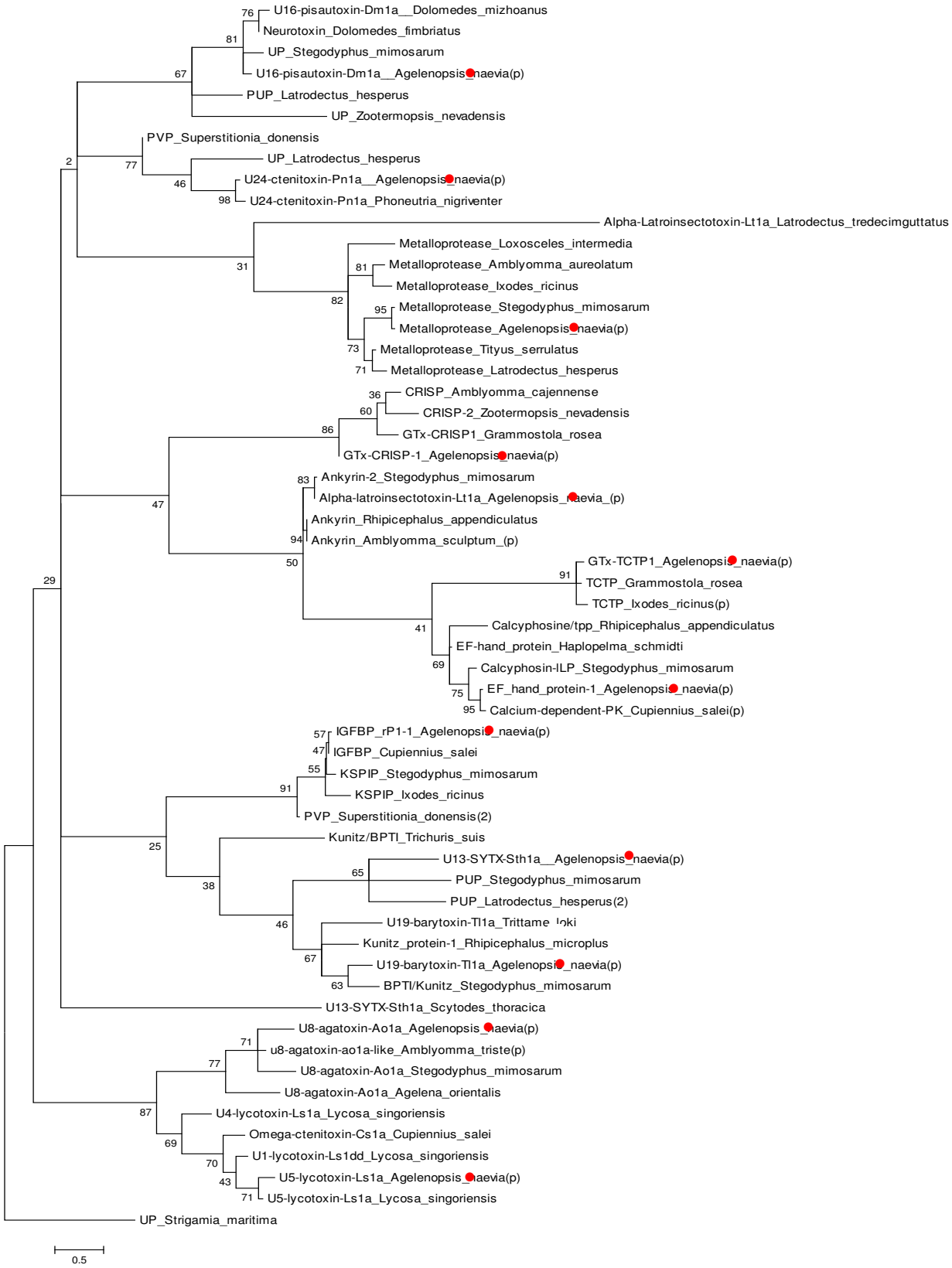

Figure-5: Phylogenetic tree of *Agelenopsis naevia* venom gland transcripts including mined sequences.

Enzymes: Enzymes have been reported from the venom gland transcript of many spiders. Here we report six enzymes making 15% of the total toxins reported. These include: Cystatin-1, sphingomyelinase-D, hyaluronidase, astacin-like metalloproteases, techylectin, neprilysin and acetylcholinesterase. Enzymes have been reported from the venom gland transcripts of scorpions and spiders[30,42]. Astacin belong tometzincin clan of metallopeptidase zymogens that is proteolytic following activation[43]. They were first reported from *Loxosceles* venom by da Silveira *et al*.[42]. In *Loxosceles* venom, astacin degrade matrix protein suggesting the role it plays in spreading venom components[44]. Acetylcholinesterase and neprilysin have been reported from the venom gland transcripts of *Trittameloki*[45]. Acetylcholinesteraseact at synapses by reducing the amount of acetylcholine while neprilysin play a role in degradation of extracellular matrix. These enzymes aid in envenomation process with neprilysin promoting the activity of acetylcholinesterase and other neurotoxins by aiding their access to synapses[8]. Techylectin-5A was first reported from horse-shoe crab (*Tachypleustridentatus*). It was found to have haemagglutination activity that is Ca^{2+} dependent against human A, B, and O types of erythrocytes and also had activity against gram negative and gram positive bacteria[46]. Cystatin are characterized as inhibitors of cysteine proteinases having activity against papain and legumain. It acts by controlling intracellular degradation of protein[47]. Cystatin have been reported from venom gland transcripts of tarantula, *Chilobrachys jingzhao*[19], *Loxosceles laeta*[48]and *Araneus ventricosus*[47]. Venom hyaluronidase is a spreading factor which increases the dermonecrotic effect of dermonecrotic toxin. It is able to degrade purified hyaluronic acid and chondroitin sulphate and has no activity on dermatansulphate and heparin sulphate[42,49]. Sphingomyelinase D catalyzes the hydrolysis of sphingomyelin to form ceramide 1-phosphate and choline. It also aid in the release of the pleiotropic lipid mediator lysophosphatidic acid by cleaving lysophophatidylcholine. It is responsible for demonecrotic lesions and serious systemic effect following bites in human by spiders. It could also be neurotoxic to prey through indirect inhibition of ion channels[50,51].

Ankyrin repeat toxins (ANK Superfamily): These are toxin belonging to ANK super family having several ANK domain repeat. The ANK domain is made up of 33 amino acids approximately with helix-loop-helix in its structure. This super family contains five families (ANK family, δ-LIT-Lt1a, α-LTX-Lt1a family 1, α-LTX-Lt1a family 2 and α-LIT-Lt1a) of which five toxins were identified in this study. These are: putative Alpha-Latrotoxin-Lt1a, Alpha-Latrotoxin-Lha1a, Alpha-Latroinsectotoxin-Lt1a and Delta latroinsectotoxin-Lt1a. These families are characterized mainly by the distribution and number of ANK domains which is linked to their function. For instance, it was reported that consecutive repeats of ANK domains aid in protein-protein interactions and may also be involved in directing protein binding to receptors[52]. Orlova *et al*.[53] reported that when ANK domain repeats are present in the central and C-terminal regions, they play a role in pore formation through the assembly of tetramer complexes to induce the release of exhaustive neurotransmitter in vertebrates. α-LTX-Lt1a and α-LIT-Lt1a, were reported to be the main neurotoxins responsible for envenomation in human by binding neuronal cell receptors on the presynaptic plasma membrane. Similar mechanism was observed with δ-LIT-Lt1a, which had activity in insects, but not in vertebrates[54]. Also δ-LTX-Lt1a lacks three key residue (C34, C91, C413) which are conserved in α-LTX-Lt1a family 1 and α-LTX-Lt1a family 2 which might explain the differences in their functions[38]. Alpha-latrotoxin is the main component of *Loxosceles* venom toxic to vertebrate which is responsible for human envenomation[55]. This toxin binds to neurexin 1α, latrophilin1 and receptor-like protein tyrosine phosphatase to induce neurotransmitter exocytosis. It is said to act through two calcium-dependent mechanism and a yet to be identified calcium-independent mechanism[56]. δ-latroinsectotoxin and α-latrocrustotoxin are toxins from *Loxoceles* venom that induces neurotransmitter release in insect neuromuscular junction[56] and neurotransmitter exocytosis in crustacean[57] respectively. Both toxins are presumed to induce neurotransmitter release similar to α-latrotoxin[56].

Venom allergens (SCP family): Four sequences homologous to venom allergens: CRISP-1-*Lycosa singoriensis*, CRISP-1-*Grammostola rosea* and CRISP-2-*Grammostola rosea* and U21-aranetoxin-Av1b, were identified in this study. Venom allergens belong to SCP family which contains a cysteine rich motif found in secretory proteins. They are important in the construction of extracellular matrix, branching morphogenesis and ion channel regulation in fertility[58]. Silva *et al*.[59] reported the ability of venom allergen to inhibit ryanodine receptors. Venom allergens have been reported from the venom of ant, wasp, spider, scorpion and centipede[30, 38, 60, 61, 62]. The presence of these toxins in various arthropods that are venomous might suggest that they are toxins that were likely inherited from a common ancestral arthropod.

Neurotoxins: Putative neurotoxin sequences homologous to Omega-ctenitoxin-Pn3a, U1-filistatoxin-Kh1b and U16-pisautoxin-Dm1a were also found in venom gland transcript of *A. naevia.* Signal peptide were predicted in putative U16-pisautoxin-Dm1a while propeptide were predicted in putative Omega-ctenitoxin-Pn3a sequence. Omega-ctenitoxin-Pn3a is an irreversible antagonist of Cav2.1/CACNA1A and Cav2.2/CACNA1B calcium channels, displays a partially reversible block of Cav2.3/CACNA1E calcium channels with no effect on Cav3/CACNA1 calcium channels[63]. U16-pisautoxin-Dm1a is a putative toxin whose molecular target is yet to be discovered[41]. U1-filistatoxin-Kh1b is toxin from venom of *Kukulcania hibernalis* suggested to be neurotoxic based on the sequence homologies to U_1-filistatoxin-Kh1a and U_1-filistatoxin-Kh2a from the same species[24].

BPTI-kunitz toxins: Three putative BPTI (Bovine pancreatic trypsin inhibitor)-kunitz toxins were identified in this study which made up 4% of the total toxins reported. These are

Kappa-theraphotoxin-Hs1a_3, U32-aranetoxin-Av1a and U19-barytoxin-Tl1a. Kunitz type toxin in venomous animals was first identified in snake in 1974 (Swissprot No:P00976) after which many have been reported from snakes, sea anemones, cone snails, spiders and scorpion[64-68]. Proteins inhibiting serine proteases have kunitz domain as the active site. The kunitz type motif usually has 50-60 amino acid residues. BPTI-kunitz toxin also known as aprotinin inhibits a spectrum of serine proteases including trypsin, chymotrypsin, plasmin and kallikrein. Kappa-theraphotoxin-Hs1a inhibits both serine proteases and voltage-gated potassium channel (Kv). The toxin is more active on Kv1.1/KCNA1 with 78% inhibition than on Kv1.2/KCNA2 and Kv1.3/KCNA3 with 10% and 28% inhibition respectively[4,69]. U19-barytoxin-Tl1a was discovered in 2014 which has two BPTI-kunitz domains whose biological activity is yet to be discovered[45].

Scytotoxins: Sequences homologous to Scytotoxins were identified from venom gland transcripts of *A. naevia*. These are: U11-SYTX-Sth1a, U15-SYTX-Sth1a and U12-SYTX-Sth1b. These toxins were identified from scytodes spiders[70] and their molecular function is not yet known.

Calcium dependent toxins: These toxins made up 4% of the total toxins identified. These are sequence homologous to EF-hand protein-1 and Calmodulin A-like. They are characterized by having Efh domain which contains approximately 40 residues and is involve in binding intracellular calcium. Calmodulin functions in regulating the process of protein transport and secretion and have been reported from transcripts of *L. singoriensis*[23].

Ctenitoxins: A single sequence homologous to U24-ctenitoxin-Pn1a from venom gland transcripts of *Phoneutria nigriventer* was identified in this study. Ctenitoxins were first identified in venom gland transcripts of *Latrodectus trecimgattatus*. These toxins are known to contain TY (thyroglobulin) domains. The ctenitoxin sequence identified in this study has two TY domains with a predicted signal peptides (Table-2). Ctenitoxins with predicted signal peptides have been reported by He *et al*.[38] from venom gland transcripts of *Latrodectus trecimgattatus.* Proteins containing TY domains have been proposed to be cysteine proteinase inhibitors[21].

GTx-TCTP1: Putative toxin belonging to translationally-controlled tumor protein (TCTP) super family was also identified. GTx-TCTP1 was discovered from venom of *Grammostola rosea*. The toxin causes edema, enhances vascular permeability and is likely related to the inflammatory activity of the venom[71].

Insulin-like growth factor-binding protein-related protein 1 (IGFBP): Sequence homologous to IGFBP_rP1-1-*Cupiennius salei* was also identified in transcripts of *A. naevia*. IGFBP_rP1-1-*Cupiennius salei* was isolated from the hemocytes of the spider *Cupiennius salei*[72]. It is assumed to have serine-type endopeptidase inhibitor activity based on GO annotation (GO:0004867).

Phylogenetic Relationship of *Agelenopsis naevia* venom toxins: The phylogenetic tree constructed includes a single member of each category identified in this study including mined sequences. Based on basic local alignment search tool, homologs of *A. naevia* transcripts were found in snakes, scorpions, wasps, termites, bug, *Daphnia* etc. indicating that these toxins have high variation in origin and complex evolutionary context[38,73]. *Agelenopsis naevia* venom gland sequence clustered mainly with that of *Stegodyphus mimosarum* indicating close relationship. *Stegodyphus mimosarum* is a social spider belonging to Eresidae family whose venom is not very toxic to humans which is similar with the venom of *A. naevia*. Both Eresidae and Agelenidae belong to entelegynes group of spiders[72]. Putative astacin-like metalloproteases sequence clustered with that of *Stegodyphus mimosarum* with 84.8% identity and 95% node reliability indicating that they are orthologs. Putative α-latroinsectotoxin-Lt1a clustered with ankyrin containing sequences from ticks; *Rhipicephalus appendiculatus*, *Amblyommasculptum*, and *Stegodyphus mimosarum* with 89.1%, 91.2% and 96.5% identity respectively. This sequence was not only homologous to sequences from spider and ticks but was also homologous to sequence from *Drosophila* spp, fly, bug, parasitic wasps, birds, mammals etc.[38]. However, alpha-latroinsectotoxin-Lt1a did not cluster with other ankyrin sequences on the tree but with metalloproteases. This may indicate close evolutionary relationship with metalloprotease than with other ankyrin sequences in this study. Lycotoxins and agatoxins clustered together likely because these toxins are rich in cysteine which is used to form cysteine knot. Homologs of lycotoxins where only found in four spider species (*Lycosasingoriensis*, *Cheiracanthium punctorium*, *Cupienniussalei* and *Lachesanatarabaevi*). Putative GTx-CRISP from *A. naevia* clustered with that from *Grammostola rosea*, *Amblyomma cajennense* and *Zootermopsis nevadensis*. This is similar to the findings of Haney *et al*.[73]. Kunitz type toxin (KTT) from *A. naevia* clustered with that from other spiders and tick. Kunitz type toxins have been reported to form a clade between snakes KTT's and primitive KTT's[67]. These authors reported a new active site in HWTX-XI (KTT) toxin which transforms from S1-protease inhibitors to channel blockers of which the evolution may probably involve three stages: old functional molecular, bi-functional toxin and new function toxin. Why venomous animals have protease inhibitors remain unclear, however, it is safe to suggest that selective pressure as a result of prey proteases might have necessitate the need for protease inhibitors in venom since their role is to protect venom proteins from prey proteases[65,67].

Conclusion

This study reports transcripts from venom gland of *A. naevia* spider of which 48 putative toxins belonging to 12 categories were identified which could be potential candidate(s) for

pharmaceutical and/or agro-allied application. Less than 50% of the transcripts were significantly homologous to sequences in the gene bank, as such further extensive proteomics and genomics research on *A. naevia* and other spiders cannot be over emphasized, so as to improve our understanding of molecular and evolutionary complexities of spider venom.

Acknowledgement

The authors acknowledge TETFund for sponsoring the research, Applied Biological Materials (abm) Canada for doing some quantification/quality analysis free of charge, Prof. M.N Shuaibu and Dr. T.T Gbem for their technical contribution and review.

References

1. Agnarsson I., Coddington J.A. and Kuntner M. (2013). Systematics - progress in the study of spider diversity and evolution. D. Penney (Ed.), Spider research in the 21st century: trends and Perspectives, Siri Scientific Press, Manchester, UK, 58-11. 978-0-9574530-1-2

2. Rash L.D. and Hodgson W.C. (2002). Pharmacology and biochemistry of spider venoms. *Toxicon*, 40(3), 225-254.

3. Escoubas P., Sollod B. and King G.F. (2006). Venom landscapes: Mining the complexity of spider venoms via a combined cDNA and mass spectrometric approach. *Toxicon*, 47(6), 650-663.

4. Liang S. (2004). An overview of peptide toxins from the venom of the Chinese bird spider Selenocosmia huwena Wang [= Ornithoctonus huwena (Wang)]. *Toxicon*, 43(5), 575-585.

5. Siemens J., Zhou S., Piskorowski R., Nikai T., Lumpkin E.A., Basbaum A.I., King D. and Julius D. (2006). Spider toxins activate the capsaicin receptor to produce inflammatory pain. *Nature,* 444, 208-212.

6. Bohlen C.J., Priel A., Zhou S., King D., Siemens J. and Julius D. (2010). A bivalent tarantula toxin activates the capsaicin receptor, TRPV1, by targeting the outer pore domain. *Cell*, 141(5), 834-845.

7. Windley M.J., Herzig V., Dziemborowicz S.A., Hardy M.C., King G.F. and Nicholson G.M. (2012). Spider-Venom Peptides as Bioinsecticides. *Toxins,* 4(3), 191-227.

8. King G.F. and Hardy M.C. (2013). Spider venom peptides: Structure, pharmacology, and potential for control of insect pests. *Annu. Rev. Entomol.*, 58, 475-496.

9. Escoubas P. and Rash L. (2004). Tarantulas: Eight-legged pharmacists and combinatorial chemists. *Toxicon*, 43(5), 555-574.

10. Tedford H.W, Sollod B.L., Maggio F. and King G.F. (2004). Australian funnel-web spiders: Master insecticide chemists. *Toxicon*, 43(5), 601-618.

11. Sollod B.L., Wilson D., Zhaxybayeva O., Gogarten J.P., Drinkwater R. and King G.F. (2005). Were arachnids the first to use combinatorial peptide libraries?. *Peptides*, 26(1), 131-139.

12. Platnick N.I. (2012). The world spider catalog. version 12.0. American Museum of Natural History, 2012. Available online: http://research.amnh.org/iz/spiders/ catalog (accessed on 20 December 2014).

13. Skinner W.S., Adams M.E., Quistad G.B., Kataoka H., Cesarin B.J., Enderlin F.E. and Schooley D.A. (1989). Purification and characterization of two classes of neurotoxins from the funnel web spider, *Agelenopsis aperta*. *J. Biol Chem.*, 264(4), 2150-2155.

14. Adams M.E., Bindokas V.P., Hasegawa L. and Venema V.J. (2004). Omega-agatoxins: novel calcium channel antagonists of two subtypes from funnel web spider (*Agelenopsisaperta*) venom. *J. Biol. Chem.*, 265, 861-867.

15. Adams M.E. (2004). Agatoxins: ion channel specific toxins from the American funnel web spider, *Agelenopsisaperta. Toxicon,* 43(5), 509-525.

16. Guarisco H. (2014). The funnelweb spider genus *Agelenopsis* (Araneae: Agelenidae) in Kansas. *Transactions of the Kansas Academy of Science,* 117(1-2), 79-87.

17. Bennett R.G. and Ubick D. (2005). Agelenidae. In Spiders of North America: *an identification manual*, 2nd ed.; Ubick, D., Paquin, P., Cushing, P.E, Roth, V., Eds.; American Arachnological Society, 56-59. ISBN-13:9780998014609

18. King G.F., Gentz M.C., Escoubas P. and Nicholson G.M. (2008). A rational nomenclature for naming peptide toxins from spiders and other venomous animals. *Toxicon*, 52(2), 264-276.

19. Chen J., Deng M., He Q., Meng E., Jiang L., Liao Z., Rong M. and Liang S. (2008). Molecular diversity and evolution of cystine knot toxins of the tarantula *Chilobrachys jingzhao*. *Cell. Mol. Life Sci.,* 65(15), 2431-2444.

20. Diego-García E., Peigneur S., Waelkens E., Debaveye S. and Tytgat J. (2010). Venom components from *Citharischiuscrawshayi* spider (Family Theraphosidae): Exploring transcriptome, venomics, and function. *Cell Mol. Life Sci.*, 67(16), 2799-2813.

21. Jiang L., Zhang D., Zhang Y., Peng L., Chen J. and Liang S. (2010). Venomics of the spider *Ornithoctonushuwena* based on transcriptomic versus proteomic analysis. *Comp. Biochm. Physiol. Part D Genomics and Proteomics,* 5(2), 81-88.

22. Tang X., Zhang Y., Hu W., Xu D., Tao H., Yang X., Li Y., Jiang L. and Liang S. (2010). Molecular diversification of peptide toxins from the tarantula *Haplopelmahainanum* (*Ornithoctonushainana*) venom based transcriptomic,

peptidomic, and genomic analyses. *J. Proteome Res.,* 9(5), 2550-2564.

23. Zhang Y., Chen J., Tang X., Wang F., Jiang L., Xiong X., Wang M., Rong M., Liu Z. and Liang S. (2010). Transcriptome analysis of the venom glands of the Chinese wolf spider *Lycosasingoriensis*. *Zoology*, 113, 10-18.

24. Herzig V., Wood D.L., Newell F., Chaumeil P.A., Kaas Q., Binford G.J., Nicholson G.M., Gorse D. and King G.F. (2011). Arachno Server 2.0, an updated online resource for spider toxin sequences and structures. *Nucleic Acids Res.,* 39, 653-657. http://www.arachnoserver.org(accessed on 19/04/2016)

25. Google Earth. Available online: http://www.google.com/earth/index.html (accessed on 9/03/2015).

26. Dippenaar-Schoeman A.S. and Jocque R. (1997). African spiders, an identification manual. Biosystematics Division, ARC-Plant Protection Research Institute, Pretoria. Handbook 9. Ultra Litho: Heriotdale Johannesburg, South Africa, 392pp; ISBN 0 621 17544 7

27. Whitman-Zai J., Francis M., Geick M. and Cushing P.E. (2015). Revision and morphological phylogenetic analysis of the funnel web spider genus *Agelenopsis* (Araneae: Agelenidae). *The Journal of Arachnology,* 43, 1-25.

28. Spider (2015). World Spider Catalog. *Natural History Museum Bern*, online at http://wsc.nmbe.ch, *version* 18.5, (accessed on 10/12/2015).

29. Garb J.E. (2014). Extraction of venom and venom gland microdissection from spiders for proteomic and transcriptomic analyses. *J. Vis. Exp.,* 93, 51618.

30. Luna-Ramírez K., Quintero-Hernández V., Juárez-González V.R. and Possani L.D. (2015). Whole transcriptome of the venom gland from *Urodacu syaschenkoi* scorpion. *PLoS ONE*, 10(5), 1-33.

31. Andrews S. (2010). FastQC: A quality control tool for high throughput sequence data. Available online: http://www.bioinformatics.babraham.ac.uk？/projects/fastqc/ (accessed on 9/02/2016).

32. Grabherr M.G., Haas B.J., Yassour M., Levin J.Z., Thompson D.A., Amit I., Adiconis X., Fan L., Raychowdhury R., Zeng Q., Chen Z., Mauceli E., Hacohen N., Gnirke A., Rhind N., di Palma F., Birren B.W., Nusbaum C., Lindblad-Toh K., Friedman N. and Regev A. (2011). Full-length transcriptomeassembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.*, 29(7), 644-652.

33. Trapnell C., Roberts A., Goff L., Pertea G., Kim D., Kelley D.R., Pimentel H., Salzberg S.L., Rinn J.L. and Pachter L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with Top Hat and Cufflinks. *Nat. Protoc.,* 7, 562-578. http://tophat.cbcb.umd.edu/(accessed on 18/02/2016)

34. Conesa A., Götz S., Garcia-Gomez J.M., Terol J., Talon M. and Robles M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, 21(18), 3674-3676.https://www.blast2go.com (accessed on 16/05/2017).

35. Artimo P., Jonnalagedda M., Arnold K., Baratin D., Csardi G., de Castro E., Duvaud S., Flegel V., Fortier A., Gasteiger E., Grosdidier A., Hernandez C., Ioannidis V., Kuznetsov D., Liechti R., Moretti S., Mostaguir K., Redaschi N., Rossier G., Xenarios I. and Stockinger H. (2012). ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res*, 40(W1), W597-W603. http://web.expasy.org/translate(accessed on 18/02/2016).

36. Kumar S., Tamura K. and Nei M. (2004). MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief. Bioinform.,* 5(2), 150-163.

37. Kubista H., Mafra R.A., Chong Y., Nicholson G.M., Beirão P.S., Cruz J.S., Boehm S., Nentwig W. and Kuhn-Nentwig L. (2007). CSTX-1, a toxin from the venom of the hunting spider *Cupienniussalei,* is a selective blocker of L-type calcium channels in mammalian neurons. *Neuropharmacology*, 52(8), 1650-1662.

38. He Q., Duan Z., Yu Y., Liu Z., Liu Z. and Liang S. (2013). The venom gland transcriptome of *Latrodectustredecimguttatus*revealed by deep sequencingand cDNA library analysis. *PLoS ONE*, 8(11), e81357.

39. Colgrave M.L. and Craik D.J. (2004). Thermal, chemical, and enzymatic stability of the cyclotidekalata B1: the importance of the cyclic cystine knot. *Biochemistry,* 43(20), 5965-5975.

40. Fry B.G., Roelants K., Champagne D.E., Scheib H., Tyndall J.D.A., King G.F., Nevalainen T.J., Norman J.A., Lewis R.J., Norton R.S., Renjifo C. and de la Vega R.C. (2009). The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. *Annu. Rev. Genome Hum. Genet.*, 10, 483-511.

41. Jiang L., Liu C., Duan Z., Deng M., Tang X. and Liang S. (2013). Transcriptome analysis of venom glands from a single fishing spider *Dolomedes mizhoanus*. *Toxicon,* 73, 23-32.

42. da Silveira R.B., Chaim O.M., Mangili O.C., Gremski W., Dietrich C.P., Nader H.B. and Veiga S.S. (2007). Hyaluronidases in *Loxoscelesintermedia* (Brown spider) venom are endo-beta-N-acetyl-d-hexosaminidases hydrolases. *Toxicon*, 49(6), 758-768.

43. Cerdà-Costa N. and Gomis-Rüth F.X. (2014). Architecture and function of metallopeptidase catalytic domains. *Protein Science*, 23(2), 123-144.

44. Trevisan-Silva D., Gremski L.H., Chaim O.M., da Silveira R.B., Meissner G.O., Mangili O.C., Barbaro K.C., Gremski W., Veiga S.S. and Senff-Ribeiro A. (2010). Astacin-like metalloproteases are a gene family of toxins present in the venom of different species of the brown spider (genus *Loxosceles*). *Biochimie*, 92(1), 21-32.

45. Undheim E.A., Sunagar K., Herzig V., Kely L., Low D.H., Jackson T.N., Jones A., Kurniawan N., King G.F., Ali S.A., Antunes A., Ruder T. and Fry B.G. (2013). A proteomics and transcriptomics investigation of the venom from the barychelid spider *Trittameloki*(brush-foot trapdoor). *Toxins,* 5(12), 2488-2503.

46. Kairies N., Beisel H.G., Fuentes-Prior P., Tsuda R., Muta T., Iwanaga S., Bode W., Huber R. and Kawabata S. (2001). The 2.0-A crystal structure of tachylectin 5A provides evidence for the common origin of the innate immunity and the blood coagulation systems. *PNAS*, 98(24), 13519-13524.

47. Wan H., Kang T., Kim B.Y., Lee S.K. and Li J. (2014). AvCystatin, a novel cysteine protease inhibitor from spider (*Araneusventricosus*). *Journal of Asian Pacific Entomology*, 18(1), 13-18.

48. deFernandes-padrosa M., Junqueira-de-Azevedo I.L.M., Gonçalves-de-Andrade R., Kobashi L.S, Almeida D.D., Ho P.L. and Tambourgi D.V. (2008). Transcriptome analysis of *Loxosceleslaeta*(Araneae, Sicariidae) spider venomous gland using expressed sequence tags. *BMC Genomics*, 9, 279.

49. Ferrer V.P., de Mari T.L., Gremski L.H., Trevisan Silva D., da Silveira R.B., Gremski W., Chaim O.M., Senff-Ribeiro A., Nader H.B. and Veiga S.S. (2013). A novel hyaluronidase from brown spider (*Loxoscelesintermedia*) venom (Dietrich's Hyaluronidase): from cloning to functional characterization. *PLoSNegl. Trop. Dis.,* 7, E2206-E2206.

50. de Fernandes-Pedrosa M.F., Junqueira de Azevedo I.L.M., Goncalves-de-Andrade R.M., van den Berg C.W., Ramos C.R.R., Ho P.L. and Tambourgi D.V. (2002). Molecular cloning and expression of a functional dermonecrotic and haemolytic factor from *Loxosceleslaeta* venom. *Biochem. Biophys. Res. Commun.*, 298(5), 638-645.

51. vanMeeteren L.A., Frederiks F., Giepmans B.N.G., FernandesPedrosa M.F., Billington S.J., Jost B.H., Tambourgi D.V. and Moolenaar W.H. (2004). Spider and bacterial sphingomyelinases D target cellular lysophosphatidic acid receptors by hydrolyzing lysophosphatidylcholine. *J. Biol. Chem.,* 279, 10833-10836.

52. Luch A. (2010). Mechanistic insights on spider neurotoxins. *EXS*, 100, 293-315, PubMed: 20358687.

53. Orlova E.V., Rahman M.A., Gowen B., Volynski K.E., Ashton A.C., Manser C., van Heel M. and Ushkaryov Y.A. (2000). Structure of alpha-latrotoxin oligomers reveals that divalentcation-dependent tetramers form membrane pores. *Nature Structural Biology,* 7, 48-53.

54. Ashton A.C., Rahman M.A., Volynski K.E., Manser C., Orlova E.V., Matsushita H., Davletov B., van Heel M., Grishin E.V. and Ushkaryov Y. (2000). Tetramerisation of alpha-latrotoxin by divalent cationsisresponsible for toxin-induced non-vesicular release and contributes totheCa(2+)-dependent vesicular exocytosis from synaptosomes. *Biochimie*, 82(5), 453-468.

55. Ushkaryov Y. (2002). α-Latrotoxin: from structure to some functions. *Toxicon,* 40, 1-5.

56. Rohou A., Nield J. and Ushkaryov Y.A. (2007). Insecticidal toxins from black widow spider venom. *Toxicon*, 49(4), 531-549.

57. Volynskiĭ K.E., Volkova T.M., Galkina T.G., Krasnoperov V.G., Pluzhnikov K.A., Khvoshchev M.V. and Grishin E.V. (1999). Molecular cloning and primary structure of cDNA fragment for alpha-latrocrustatoxin from black widow spider venom. *BioorgKhim,* 25, 25-30.

58. Gibbs G.M., Roelants K. and O'Bryan M.K. (2008). The CAP superfamily: cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1proteins--roles in reproduction, cancer, and immune defense. *EndocrRev.*, 29, 865-897.

59. Silva E.C., Camargos T.S., Maranhão A.Q., Silva-Pereira I., Silva L.P., Possani L.D. and Schwartz E.F. (2009). Cloning and characterization of cDNA sequences encoding fornew venom peptides of the Brazilian scorpion *Opisthacanthus cayaporum*. *Toxicon*, 54(3), 252-261.

60. Liu Z.C., Zhang R., Zhao F., Chen Z.M., Liu H.W., Wang Y.J., Jiang P., Zhang Y., Wu Y., Ding J., Lee W. and Zhang Y. (2012). Venomic and transcriptomic analysis ofcentipede scolopendra subspinipes dehaani. *J. Proteome Res.,* 11(12), 6197-6212.

61. Baek J.H., Oh J.H., Kim Y.H. and Lee S.H. (2013). Comparative transcriptome analysis of the venom sac and gland ofsocial wasp Vespa tropica and solitary wasp *Rhynchium brunneum*. *Journal of Asia-Pacific Entomology*, 16(4), 497-502.

62. Bouzid W., Klopp C., Verdenaud M., Ducancel F. and Vétillard A. (2013). Profiling the venom gland transcriptomeofTetramoriumbicarinatum (Hymenoptera: Formicidae): The first transcriptome analysis of an ant species. *Toxicon*, 70, 70-81.

63. Cassola A.C., Jaffe H., Fales H.M., Castro-Afeche S., Magnoli F. and Cipolla-Neto J. (1998). Omega-phonetoxin-IIA: a calcium channel blocker from the spider *Phoneutrianigriventer*. *Pflugers Arch*, 436(4), 545-552.

64. Schweitz H., Bruhn T., Guillemare E., Moinier D., Lancelin J.M., Berees L. and Lazdunski M. (1994). Kalicludines and kaliseptine. Two different classes of sea anemone toxins

forvoltage sensitive K+ channels. *J. Biol. Chem.*, 270, 25121-25126.

65. Zupunski V., Kordis D. and Gubensek F. (2003). Adaptive evolution in the snake venom Kunitz/BPTI protein family. *FEBS Letter*, 547(1-3), 131-136.

66. Dy C.Y., Buczek P., Imperial J.S., Bulaj G. and Horvath M.P. (2006). Structure ofconkunitzin-S1, a neurotoxin and Kunitz-fold disulfide variant from cone snail. *Acta Crystallogr. D Biol. Crystallogr.,* 62, 980-990.

67. Yuan C.H., He Q.Y., Peng K., Diao J.B., Jiang L.P., Tang X. and Liang S.P. (2008). Discovery of a distinct superfamily of Kunitz-type toxin (KTT) from tarantulas. *PLoS ONE*, 3, e3414.

68. Zhao R., Dai H., Qiu S., Li T., He Y., Ma Y., Chen Z., Wu Y., Li W. and Cao Z. (2011). SdPI, the first functionally characterized Kunitz-type trypsin inhibitor from scorpion venom. *PLoS ONE*, 6(11), e27548.

69. Peng K., Lin Y. and Liang S.P. (2006). Nuclear magnetic resonance studies on huwentoxin-XI from the Chinese bird spider *Ornithoctonushuwena*: 15N labeling and sequence-specific 1H, 15N nuclear magnetic resonance assignments. *ActaBiochim. Biophys. Sin*., 38(7), 457-466.

70. Zobel-Thropp P.A., Correa S.M., Garb J.E. and Binford G.J. (2014). Spit and venom from scytodes spiders: a diverse and distinct cocktail. *J. Proteome Res.,* 13(2), 817-835.

71. Kimura T., Ono S. and Kubo T. (2012). Molecular cloning and sequence analysis ofthe cDNAs encoding toxin-like peptides from the venom glands of tarantula *Grammostola rosea*. *International Journal of Peptides,* 1-10.

72. Kuhn-Nentwig L., Largiader C.R., Streitberger K., Chandru S., Baumann T., Kampfer U., Schaller J., Schurch S. and Nentwig W. (2011). Purification, cDNA structure and biological significance of a single insulin-like growth factor-binding domain protein (SIBD-1) identified in the hemocytes of the spider *Cupiennius salei*. *Insect Biochem. Mol. Biol*., 41(11), 891-901.

73. Haney R.A., Ayoub N.A., Clarke T.H., Hayashi C.Y. and Garb J.E. (2014). Dramatic expansion of the black widow toxin arsenal uncovered by multi-tissue transcriptomics and venom proteomics. *BMC Genomics*, 15, 366.