

De Novo Transcriptome Assembly for Venom Gland in Two Species of Spiders (*Sinopoda Pengi* and *Trichonephila Clavata*)

Li jun Ding

Dali university

Xiu mei Wu

Dali university

Cheng gui Zhang

Dali university

Peng fei Gao

Dali university

Yan Zhang

Yunnan Provincial Academy of Science and Technology

Yu Zhao

Dali university

Zi zhong Yang (✉ yangzizhong@Dali.edu.cn)

Dali University <https://orcid.org/0000-0002-2143-8226>

Research Article

Keywords: spider venom gland, RNA-seq, encode venom, *Sinopoda pengi*, *Trichonephila clavata*

Posted Date: December 10th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1135110/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Natural molecules from spider venom are considered potential drugs for diseases including cancer and pain, as well as the development of new biological insecticides for agricultural use. During coevolution in the long-term predator-prey game, spiders have formed a huge molecular diversity of toxins. As of March 1 of 2021, a total of 49,243 spider species had been described, but studies of venom have been performed in only a few hundred of these species due to the difficulty of collecting venom. Two technologies have helped partially dealing with this limitation in the recent past: the screening of cDNA libraries constructed from venom gland mRNAs and the heterologous expression of the coded peptides for functional characterization. In this study, transcriptomic analysis was performed to describe the predicted toxins of *Sinopoda pengi* (hereafter *S. pengi*) and *Trichonephila clavata* (hereafter *T. clavata*). The Trinity assembly result in 163,418 transcripts, 114,127 unigene of *S. pengi* and 125,099 transcripts, 87,084 unigene of *T. clavata*. A total of 22 and 24 unigenes were identified which were predicted to inhibitor cysteine knot (ICK) toxins from *S. pengi* and *T. clavata*, respectively. In summary, molecular templates with potential application value in medical and biological fields were obtained by classifying and characterizing presumed venom components, which lays a foundation for the further study of venom.

1. Introduction

More than 100,000 species distributed across all major phyla of the animal kingdom have developed sophisticated machinery to produce poison or venom to defend themselves from predators, to fight against competitors, or to immobilize and digest prey [1–2]. Spiders are the seventh most diversified order of animals and the most diversified group of venomous predators in terms of the number of species, occupying most ecological niches [3]. Spider venoms are complex cocktails produced from the holocrine glands in the chelicerae. During coevolution in the long-term predator-prey game, venomous animals have formed a huge molecular diversity of animal toxins [2]. These toxin molecules usually act on the key physiological protein elements of the target organism, such as cell membrane receptors and ion channels (effectively distinguishing between different membrane receptors and ion channel subtypes) [4]. Toxins have become an important source of molecular tools for dissecting physiological processes [5] and are taken as lead molecules for the development of drugs targeting a range of conditions, such as chronic pain, cancer, stroke, and autoimmune disease [6].

Sinopoda pengi (Sparassidae; Fig. 1A) is a wandering spider with a widespread distribution in Yunnan, China, that lives on walls near houses or in rock cracks along roadsides. These spiders are ambush predators that do not spin webs but instead rely on their strong chelicerae and powerful venom to paralyze or kill their prey during hunting. They hunt at night, remaining quietly on walls waiting for prey to appear and then killing the prey when it passes by. They prey on nocturnal insects such as small beetles and crickets. In contrast, *Trichonephila clavata* (Araneidae; Fig. 1B) usually produces large, strong nets around hedges or bushes to facilitate prey capture, in addition to deploying its venom arsenal. These spiders are active during the day, and their web often contains the remains of butterflies, bees and other

insects. These two species are relatively widespread spiders with different feeding habits, and their venom composition have not yet been studied with transcriptomic.

Spider venom received little attention for quite some time due to technological limitations and its low impact on human health until its enormous medical potential was discovered. As of March 1 of 2021, a total of 49,243 spider species had been described, but studies of venom have been performed in only a few hundred of these species [7]. Considering that individual venoms may contain hundreds of unique compounds and that the venoms of even comparatively well-studied lineages remain largely unexplored [2], animal venoms constitute an enormous unexplored natural library of bioactive compounds [8]. One of the reasons for the gap in venom knowledge resides in the difficulties imposed by the need to collect a large number of specimens from natural environments to “milk” their venom in large-enough quantities to allow the isolation of the less represented components [9]. The emergence and development of sequencing techniques has provided a solution to this awkward situation, and the toxin information available in the ArachnoServer and UniProt databases has increased rapidly in the last decade. Through transcriptome sequencing, we can determine the expression of the components associated with venom from just a few samples. Based on this approach, spider venom has been speculated to contain neurotoxins as well as other components, such as enzymes with serine proteinase, metalloprotease activity [10–12]. Here, only a few spiders are required to perform a detailed venom composition analysis via transcriptomic techniques.

2. Materials And Methods

2.1. Biological sample and RNA extraction

Specimens were collected in Dali, China (25°40'41"N, 100°8'59"E; and 25°40'52"N, 100°10'20"E) and were identified as *S. pengi* and *T. clavata* by Professor Zi-Zhong Yang of Dali University. The spiders were kept at the Yunnan Provincial Key Laboratory of Entomological Biopharmaceutical R&D at room temperature under a natural light-dark cycle, fed mealworms on a weekly basis and provided with petri dishes as drinking vessels.

Venom glands were removed from the cephalothorax of each spider (Fig. 2) and total RNA was extracted from the venom glands of 5 *S. pengi* and 5 *T. clavata* using TRIzol, following manufacturer instructions, respectively (Invitrogen Life Technologies, USA). The concentration and purity of the RNA were checked with a Nanodrop 2000 (Thermo Fisher Scientific, USA). For RNA sequencing, RNA integrity was assessed by standard denaturing agarose gel electrophoresis.

2.2. cDNA library preparation and Illumina sequencing

Total RNA was used for the construction of RNA-seq libraries, and we performed paired-end sequencing on an Illumina NovaSeq 6000 system (Origin-gene Biopharm, China) following the vendor's recommended protocol.

After adaptor clipping, the quality of the raw reads was assessed with FastQC v0.11.4 (www.bioinformatics.babraham.ac.uk/bugzilla/). The quantification of the raw data was performed using FastqStat.jar V1.0, followed by filtering to obtain clean data, which were analyzed using Cutadapt v1.16. (<http://cutadapt.readthedocs.io/>). All reads were submitted to NCBI with the login numbers SRR14235845 and SRR15356212 for *S. pengi* and *T. clavata*, respectively.

2.3. Assembly and annotation

The clean data were de novo assembled with Trinity v.2.6.6 using the standard protocol [13]. After the transcripts were assembled by Trinity, the functional annotation of these transcripts was performed. Before annotation, the open reading frame (ORF) prediction method of Trinity was used to predict the amino acid sequences of all assembled transcripts.

Annotation of de novo assembled transcripts was performed using Blastx (<http://blast.ncbi.nlm.nih.gov/Blast>) based on the following databases: Nr (NCBI nonredundant protein sequences, <https://www.ncbi.nlm.nih.gov/>), Pfam (Protein family, <http://pfam.sanger.ac.uk/>), GO (Gene Ontology, <http://www.geneontology.org>), and KEGG (Kyoto Encyclopedia of Genes and Genomes, <http://www.genome.jp/kegg/>). Finally, gene expression levels were estimated in transcripts per kilobase million (TPM) values by Salmon v0.11.3.

2.4. Bioinformatic analysis

The unigenes used BLASTX searched against a toxin-related subdatabase of UniProt (<https://www.uniprot.org/program/Toxins>), best hits were defined using E-value cutoff of 1e-5 and selecting the best score. These unigenes were translated using the ORFfinder (<https://www.ncbi.nlm.nih.gov/orffinder>).

To remove potentially nontoxic orthologs, all candidate toxin unigenes were used as queries for the local BLASTX search (e-value = 1e-5) against two databases: (1) animal-reviewed proteins in the UniprotKB and (2) toxin protein database, as explained previously. Unigenes were filtered out from the candidate toxin gene dataset if the BLAST score of the best hit from the animal protein database was higher than the animal toxin database results [14].

The sequences which may encode for toxins were used for multiple sequence alignment and performed the alignments with Clustal Omega (version: 1.2.4), without a conserved region, were discarded. The signal peptide was predicted using the SignalP-5.0 program (<http://www.cbs.dtu.dk/services/SignalP/>); pro-peptides were determined by using ProP software (<http://www.cbs.dtu.dk/services/ProP/>).

3. Results And Discussion

3.1. RNA extraction and RNA-Seq

The total amounts of RNA obtained from fifteen adult female *S. pengi* and *T. clavata*, respectively. mRNA was isolated, enriched, fragmented and reverse transcribed into cDNA. Then, the cDNA libraries were sequenced using an Illumina NovaSeq 6000 system (LC Bio, China). The sequencing of the PE libraries produced 47,271,254 and 39,592,392 reads, respectively (Table S1).

3.2. De novo transcriptome assembly and functional annotation

There is no available reference genome for *S. pengi* or *T. clavata*, so Trinity needs to be used for de novo assembly. After assembly, we obtained 163,418 transcripts, 114,127 unigene of *S. pengi* and 125,099 transcripts, 87,084 unigene of *T. clavata* by using Trinity. After statistical analysis, transcripts were annotated in the NR, GO, KEGG, and Pfam databases (Supplementary Figs S1–3, Table S2–3).

3.3. Possible toxins identified at the transcriptional level

Detailed toxin information is provided in the supplementary materials (Supplementary Tables S4–5). Each transcript annotated as a toxin is discussed based on the existing literature, followed by a categorical description of the toxin.

3.3.1. ICK-like spider venom peptides

Cysteine-rich peptides are the best investigated venom components and are believed to exist in most spider venoms. They modulate a broad range of channels and receptors on the membranes of excitable cells (e.g., nerves and muscles) [15]. Ion channels on the membranes of excitable cells are responsible for proper signal transduction. Calcium channels are involved in neurotransmitter release from presynaptic cells, voltage-gated sodium channels enable action potential transmission along excitable cells, and voltage-gated potassium channels are crucial for restoring a resting state in depolarized cells [15]. Toxins interact with these targets to disrupt normal channel function, which can affect breathing or heart function, leading to symptoms such as convulsions, paralysis, and eventually death. Abundant ion channel toxins were identified by annotating the transcripts of *S. pengi* and *T. clavata*. The families of the predicted ICK toxins are described in detail below.

Group I (C-C-CC-C-C)

Group I contained three unigenes in *T. clavata*, two of which (DN1625_c0_g1, DN16263_c0_g1) were similar to Mu-Sparatoxin Hv2 (UniProt A0A088BP94) and one (DN5056_c0_g1) was similar to Kappa-Sparatoxin Hv1a (UniProt P58425) (Fig. 3A). These reference toxins belong to the neurotoxin 10 (HwtX-1) family. Mu-sparatoxin-Hv2 from *Heteropoda venatoria*, is insecticidal toxin potently and irreversibly blocks Na_v channels in cockroach dorsal unpaired median (DUM) neurons (IC₅₀=833.7 nM) [16]. Na_v channels are crucial in the generation and transmission of action potentials in the central nervous system, peripheral nervous system, heart, smooth muscle and skeletal muscle, so the function of Na_v channels is crucial, and they are targets of many toxins. There are nine α subtypes, known as Na_v1.1-

Na_v1.9 in mammals, each with different tissue distributions and functions [17]. The reference toxin Kappa-Sparatoxin Hv1a is a potassium and calcium channel blocker. K_v channels play an important role in human physiology, including the regulation of neurotransmitter release, heart rate, insulin secretion, nerve cell secretion, skeletal muscle contraction, etc. Numerous channelopathies arising from mutations in these channels or from autoimmune attack on the channels have been characterized [18–19]. Potassium channels are the target of a variety of toxic animal toxins, and the potassium channel inhibitors found in spider venom thus far mainly act on K_v1, K_v2 and K_v4 channels, although a few act on other channels, such as K_v3, K_v7 and K_v11 channels [20]. In *S. pengi*, Group I contained 9 unigenes (Fig. 3B), two of them showed homology with kappa-sparatoxin-Hv1a and seven were aligned to Mu-sparatoxin-Hv2.

Group II (C-C-CC-CXC-CXC)

In *T. clavata*, Group II included 5 sequences (Fig. 4A), correspond to U3-lycotoxin-Ls1b (UniProt B6DCQ7), LSTX-D6 (UniProt B6DCU5), U1-lycotoxin-LS1B (UniProt B6DCK2), U19-Lycotoxin-LS1B (UniProt B6DD58) from *Lycosa singoriensis* and U8-agatoxin Ao1a (UniProt Q5Y4U4) from *Agelena orientalis*, respectively. The ion channel activity of these reference peptides is not known.

In *S. pengi*, Group II is composed of 5 sequences (Fig. 4B). One of them resembled Omega-Segestritoxin-Sf1a (UniProt P58605) from *Segestria florentina*, another resembled U10-Ctenitoxin-Pr1a (UniProt P84000) and the other three resemble U6-Lycotoxin-LS1g (UniProt B6DCV8) from *Lycosa singoriensis*. Omega-Segestritoxin-Sf1a is potent and selective blocker of N-type voltage-gated calcium channels (Cav2.2/CACNA1B), also blocks vertebrate Cav2.1/CACNA1A (P/Q-type) and Cav1.2/CACNA1C (L-type) channels at very high concentration (2 micromolar) [21].

Groups III (C-CC-CC-CC-C-C-C), IV (C-C-CC-C-C-CXC-C-C), V (C-C-CC-C-CC-C-C-C) and VI (C-C-CC-CXC-CXC-CC)

In *T. clavata*, Group III contained only one sequence (DN18027_c0_g1) (Fig. 5), which had 66.3% identity with Omega-agatoxin-1A from *Agelenopsis aperta*. Omega-agatoxins are antagonists of voltage-gated calcium channels, subtypes of omega-agatoxins are distinguished as type I (omega-Aga-IA, omega-Aga-IB, and omega-Aga-IC), type II (omega-Aga-IIA and omega-Aga-IIB), and type III (omega-Aga-IIIA and omega-Aga-IIIB). All except type III toxins block calcium channels in insect motor nerve terminals and in neuronal cell bodies at nanomolar concentrations [22]. No sequence from *S. pengi* in this family.

In *T. clavata*, group IV contained 9 unigenes (Fig. 6A), 5 of which displayed similarity to toxin CSTX-20 from *Cupiennius salei* (UniProt B3EWT5). One of which presented identity to U19-ctenitoxin-Pn1a (UniProt P83997) from *Phoneutria nigriventer*. Three other sequences were similar to U3-aranetoxin-Ce1a3 (UniProt Q8MTX1) from *Caerostris extrusa*. In *S. pengi*, one sequence is similar to U33-theraphotoxin-Cg1b (UniProt B1P1J0) from *Chilobrachys guangxiensis* and one to U8-theraphotoxin-Hhn1d (UniProt D2Y2E4) from *Cyriopagopus hainanus*, the other 5 unigenes correspond to CSTX-20 (Fig. 6B).

In *T. clavata*, group V consisted of 4 unigenes (Fig. 7), two of them showed homology with U20-lycotoxin-Ls1c (UniProt B6DD62) from *Lycosa singoriensis*. The other two sequences were similar to U15-lycotoxin-Ls1a (UniProt B6DD43) from *Lycosa singoriensis* and U7-agatoxin-Ao1a (UniProt Q5Y4V9) from *Agelena orientalis*, respectively. No sequence from *S. pengi* in this family.

In *S. pengi*, group Ⅱ contained only one sequence (Fig. 8), the sequence had homology to U16-lycotoxin-Ls1b (UniProt B6DD53C) from *Lycosa singoriensis*. No sequence from *T. clavata* in this family.

Group Ⅲ (C-C-C-CC-CXC-CXC-C-C-C)

In *T. clavata*, Group Ⅳ was represented by only two sequences (DN21003_c0_g1, DN51867_c0_g1), which had identity with U9-agatoxin-Ao1a (UniProt Q5Y4U3) from *Agelena orientalis*. No sequence from *S. pengi* in this family (Fig. 9).

3.3.2. Enzymes

Enzymes are very common components of the venom of poisonous animals such as poisonous snakes, scorpions and centipedes. Proteases are enzymes that hydrolyze the amide bonds of the peptide units of polypeptides and proteins. The overall purpose of such an enzyme arsenal co-injected with toxins into a prey's tissue seems clear: by destroying the barriers imposed by the extracellular matrix and cell membranes, the toxins can quickly reach their targets [23]. Additionally, the proteolytic activity of some of these enzymes facilitates subsequent preoral digestion. As proteins play important roles in the maintenance of homeostasis, proteases are vital regulators of physiological processes. Many pathological conditions of humans and animals have been linked to the malfunctioning of this category of enzymes; as a result, they are seen as attractive targets for drug discovery [24].

Unigenes related to metalloproteinases were found in *S. pengi* and *T. clavata*. The role of metalloproteinases in venoms is not fully understood, but it may be a diffusion factor promoting the diffusion of other venoms. It may be involved in proteolysis of other venomous toxins, or it may aid extra-oral digestion of prey [25].

In *S. pengi*, we found unigenes may encoding acetylcholinesterase. Acetylcholine-mediated neurotransmission is fundamental for nervous system function. Acetylcholinesterase (AChE) hydrolyses and inactivates acetylcholine, thereby regulating the concentration of the transmitter at the synapse [26]. In venom, the toxic role of this enzyme is unclear; it could reduce musculatory control by rapidly hydrolyzing acetylcholine, or it works synergistically with alkaline phosphatase to paralyze prey through hypotension.

4. Conclusion

Through RNA-seq, we performed a detailed analysis of all predicted toxins of *S. pengi* and *T. clavata*. The results can serve as a reference for other spider toxicological studies and provide valuable molecular templates for research and therapeutic applications.

Declarations

Acknowledgments

This work was funded by The Special Program of Science and Technology of Yunnan Province (202002AA100007), Yunnan Education Department (2021Y401).

Conflicts of Interest The authors have declared that no competing interests exist.

Ethical approval All applicable international, national, and/or institutional guidelines for the care and use of animals were followed.

References

1. Calvete, JJ (2017) Venomics: integrative venom proteomics and beyond. *Biochem. J.* 474, 611-634. <https://doi.org/10.1042/BCJ20160577>
2. Schendel, V, Rash, LD, Jenner, RA, Undheim, EAB (2019) The Diversity of Venom: The Importance of Behavior and Venom System Morphology in Understanding Its Ecology and Evolution. *Toxins (Basel)* 11, 666-687. <https://doi.org/10.3390/toxins11110666>
3. Coddington, JA, Levi, HW (1991) Systematics and evolution of spiders (Araneae). *Annu. Rev. Ecol. Syst.* 22, 565-592. <https://doi.org/10.1146/annurev.es.22.110191.003025>
4. Sunagar, K, Casewell, NR, Varma, S, Kolla, R, Antunes, A, Moran, Y (2016) *Deadly Innovations: Unraveling the Molecular Evolution of Animal Venoms*. Springer Netherlands. https://doi.org/10.1007/978-94-007-6416-3_27
5. Akef, HM (2018) Anticancer, antimicrobial, and analgesic activities of spider venoms. *Toxicology Research* 7, 381-395. <https://doi.org/10.1039/c8tx00022k>
6. Peigneur, S, Tytgat, J (2018) *Toxins in Drug Discovery and Pharmacology*. *Toxins (Basel)* 10, 126-129. <https://doi.org/10.3390/toxins10030126>
7. Pineda, SS, Chaumeil, P-A, Kunert, A, Kaas, Q, Thang, MWC, Li, L, Nuhn, M, Herzig, V, Saez, NJ, Cristofori-Armstrong, B, Anangi, R, Senff, S, Gorse, D, King, GF (2017) ArachnoServer 3.0: an online resource for automated discovery, analysis and annotation of spider toxins. *Bioinformatics* 34, 1074–1076. <https://doi.org/10.1093/bioinformatics/btx661>
8. Herzig, V, King, GF, Undheim, EAB (2019) Can we resolve the taxonomic bias in spider venom research? *Toxicon* 1, 100005. <https://doi.org/10.1016/j.toxcx.2018.100005>
9. Romero-Gutierrez, T, Peguero-Sanchez, E, Batista, CVF, Ortiz, E, Cevallos, MA, Possani, LD (2017) A Deeper Examination of *Thorellius atrox* Scorpion Venom Components with Omic Technologies. *Toxins (Basel)* 9, 399-421. <https://doi.org/10.3390/toxins9120399>
10. Tang, X, Zhang, Y-Q, Hu, W-J, Xu, D-H, Tao, H, Yang, X-X, Li, Y, Jiang, L-P, Liang, S-P (2010) Molecular Diversification of Peptide Toxins from the Tarantula *Haplopelma hainanum* (*Ornithoctonus hainana*)

- Venom Based on Transcriptomic, Peptidomic, and Genomic Analyses. *J Proteome Res* 9, 2550–2564. <https://doi.org/10.1021/pr1000016>
11. Diego-García, E, Cologna, CT, Cassoli, JS, Corzo, G (2015) Spider Transcriptomes from Venom Glands: Molecular Diversity of Ion Channel Toxins and Antimicrobial Peptide Transcripts, Spider Venoms. https://doi.org/10.1007/978-94-007-6646-4_17-1
 12. Paiva, ALB, Mudadu, MA, Pereira, EHT, Marri, CA, Guerra-Duarte, C, Diniz, MRV (2019) Transcriptome analysis of the spider *Phoneutria pertyi* venom glands reveals novel venom components for the genus *Phoneutria*. *Toxicon* 163, 59-69. <https://doi.org/10.1016/j.toxicon.2019.03.014>
 13. Grabherr, MG, Haas, BJ, Yassour, M, Levin, JZ, Thompson, DA, Amit, I, Adiconis, X, Fan, L, Raychowdhury, R, Zeng, Q, Chen, Z, Mauceli, E, Hacohen, N, Gnirke, A, Rhind, N, Palma, Fd, Birren, BW, Nusbaum, C, Lindblad-Toh, K, Friedman, N, Regev, A (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29, 644-652. <https://doi.org/10.1038/nbt.1883>
 14. Mitpuangchon, N, Nualcharoen, K, Boonrotpong, S, Engsontia, P (2021) Identification of Novel Toxin Genes from the Stinging Nettle Caterpillar *Parasa lepida* (Cramer, 1799): Insights into the Evolution of Lepidoptera Toxins. *Insects* 12. <https://doi.org/10.3390/insects12050396>
 15. Langenegger, N, Nentwig, W, Kuhn-Nentwig, L (2019) Spider Venom: Components, Modes of Action, and Novel Strategies in Transcriptomic and Proteomic Analyses. *Toxins* 11, 611-654. <https://doi.org/10.3390/toxins11100611>
 16. Xiao, Z, Zhang, Y-X, Zeng, J, Liang, S-P, Tang, C, Liu, Z-H (2018) Purification and Characterization of a Novel Insecticidal Toxin, μ -sparatoxin-Hv2, from the Venom of the Spider *Heteropoda venatoria*. *Toxins (Basel)* 10, 233-244. <https://doi.org/10.3390/toxins10060233>
 17. Israel, MR, Tay, B, Deuis, JR, Vetter, I (2017) Sodium Channels and Venom Peptide Pharmacology. *Adv Pharmacol* 79, 67-116. <https://doi.org/10.1016/bs.apha.2017.01.004>
 18. Norton, RS, Chandy, KG (2017) Venom-derived peptide inhibitors of voltage-gated potassium channels. *Neuropharmacology* 127, 124-138. <https://doi.org/10.1016/j.neuropharm.2017.07.002>
 19. Tajti, G, Wai, DCC, Panyi, G, Norton, RS (2020) The voltage-gated potassium channel $K_v1.3$ as a therapeutic target for venom-derived peptides. *Biochem. Pharmacol.* 181, 114146. <https://doi.org/10.1016/j.bcp.2020.114146>
 20. Tabakmakher, VM, Krylov, NA, Kuzmenkov, AI, Efremov, RG, Vassilevski, AA (2019) Kalium 2.0, a comprehensive database of polypeptide ligands of potassium channels. *Sci Data* 6, 73-80. <https://doi.org/10.1038/s41597-019-0074-x>
 21. Newcomb, R, Palma, A, Fox, J, Gaur, S, L, K. (1995) SNX-325, a novel calcium antagonist from the spider *Segestria florentina*. *Biochem* 34, 8341-8347. <https://doi.org/10.1021/bi00026a015>
 22. Bindokas, VP, Venema, VJ, Adams, ME (1991) Differential antagonism of transmitter release by subtypes of omega-agatoxins. *J Neurophysiol* 66, 590. [https://doi.org/10.1016/0165-5728\(91\)90061-B](https://doi.org/10.1016/0165-5728(91)90061-B)

23. Kuhn-Nentwig, L, Stöcklin, R, Nentwig, W (2011) Venom Composition and Strategies in Spiders: Is Everything Possible? *Adv. Insect Physiol.* 40, 1-86. <https://doi.org/10.1016/b978-0-12-387668-3.00001-5>
24. Agbowuro, AA, Huston, WM, Gamble, AB, Tyndall, JDA (2018) Proteases and protease inhibitors in infectious diseases. *Med Res Rev* 38, 1295-1331. <https://doi.org/10.1002/med.21475>
25. Silveira., RB, Wille, AC, Chaim, OM, Appel, MH, Silva, DT, Franco, CR, Toma, L, Mangili, OC, Gremski, W, Dietrich, CP, Nader, HB, Veiga, SS (2007) Identification, cloning, expression and functional characterization of an astacin-like metalloprotease toxin from *Loxosceles intermedia* (brown spider) venom. *Biochem. J.* 406, 355-363. <https://doi.org/10.1042/BJ20070363>
26. Hermona Soreq, Seidman, S (2001) Acetylcholinesterase—new roles for an old actor. *Perspectives (Montclair)* 2, 294-302. <https://doi.org/10.1038/35067589>

Figures



Figure 1

Spiders and their prey. A. *S. pengi* hunting a cockchafer, B. *T. clavata* preying on a butterfly.

Figure 2

Isolation of spider venom glands. A1. Dissecting the venom glands of *T. clavata*, A2 venom glands of *T. clavata*, B1. Dissecting the venom glands of *S. pengi*, B2. venom glands of *S. pengi*.

3A

```

      |         |         |         |         |         |
      10        20        30        40
T.clavata_DN1625_c0_g1  EDDED CGKYFYF CKK - TA E C C KDWV C H L . . . . G M C K W D S G W R K
T.clavata_DN5056_c0_g1  . . . . S C G T I W H Y C G T D Q S E C C E G W K C S R . . . . Q L C K Y V I D W G K
sp|A0A088BP94|Mu-sparatoxin-Hv2  . . D D D C G K L F A D C T S - D S D C C E N W V C S K T G F V K N I C K Y N F G . . .
sp|P58425|Kappa-sparatoxin-Hv1a  . . . . D C G T I W H Y C G T D Q S E C C E G W K C S R . . . . Q L C K Y V I D W . .
T.clavata_DN16263_c0_g1  - D D Q Y C S Y P H S P C D K - Y P C R C P G Y E C Y D . . . . G M C G E I R R R D .

```

3B

```

      |         |         |         |         |         |
      10        20        30        40
TRINITY_DN17662_c0_g1  . . . . S C G T I W H Y C G T D Q S D C C E G W K C S L . . . . Q L C K Y V I D W G K
TRINITY_DN33845_c0_g1  . . . . S C G T I W H Y C G T D Q S E C C E G W K C S R . . . . Q L C K Y V I D W G K
sp|P58425|Kappa-sparatoxin-Hv1a  . . . . D C G T I W H Y C G T D Q S E C C E G W K C S R . . . . Q L C K Y V I D W . .
TRINITY_DN883_c0_g2     - Q D D D C S D L Y G S C R E - - L P C C P S Y V C L G . . . . N I C V I T V R . . .
TRINITY_DN9047_c0_g1   - D D E Y C S Y P E F S C R R - S E N C C P G F M C Y D . . . . N M C T M . . . . .
TRINITY_DN23341_c1_g1  - D D E Y C S Y P E F S C R R - S E N C C P G Y M C Y N . . . . Y M C T L I . . . . .
TRINITY_DN35302_c0_g1  E D D E D C V K Y F Y F C K K - T A E C C K V W V C H L . . . . G M C K W D S G W R K
TRINITY_DN41330_c0_g1  . - T D D C G W L L S S C K S - S K T C C E N W V C S K . . . . G L C K Y R I G . . .
TRINITY_DN52140_c0_g1  E D D E D C G K Y F Y F C K K - T A E C C K D W V C H L . . . . G M C K W . . . . .
TRINITY_DN68838_c1_g1  . . . . V C G T H F S L C D K - E S H C C E G Y T C I L W . . . . C R K N L - W . .
sp|A0A088BP94|Mu-sparatoxin-Hv2  . . D D D C G K L F A D C T S - D S D C C E N W V C S K T G F V K N I C K Y N F G . . .

```

Figure 3

Alignments of cysteine-rich peptide toxins mature sequences from group 1 (C-C-CC-C-C) in *T. clavata*. All alignment was performed with Clustal OMEGA. Conserved cysteines are marked in blue.

4A

```

      |         |         |         |         |         |         |         |         |         |
      10        20        30        40        50        60        70        80
T.clavata_DN2914_c0_g1  S E T C V E E N H E C T R H R H S C C R S K - - M F K F K C T C F Y K E G N D T S P . . . . T D E D I C T C Q Q V W Y H A L A E D L I D R G T Q . . . L L K N W I G F G . .
sp|B6DCQ7|U3-lycotoxin-Ls1b  A K E C T P R F Y D C S R D R H S C C R S E - - L F K D V C T C F Y P E G G D . . . . . N - E V C T C Q Q P K H L K Y M E K A A D K A K K F G G K I K K W F G . . .
T.clavata_DN5074_c0_g1  S E D C I P K N H E C T K N R H G C C R S S - - L F K Y K C K C F Y K E G K S G K P . . . . T N K E M C T C Q L P G N M R F V E E I V . . . T A . . . V K K S . . . . .
sp|B6DCU5|LSTX-D6        A K A C T P R Y Y D C S H D R H S C C R S S - - M F K D V C T C F Y P E G G D . . . . . N K E V C T C Q Q P K H L K Y M E K A T D . . . . . K I K N L F G . . .
T.clavata_DN17357_c0_g1   . . E C I P E N K D C K S N R N G C C P T N K W L Q E S Y C H C - Y L A A K D K K R Y N T V S K A E E K C F D . . . . . T Q V S I F G - G A K
sp|B6DCK2|U1-lycotoxin-Ls1b  - K E C I P K H H E C T S N K H G C C R G N - - F F K Y K C C C T T V V T Q D G E - . . . . . Q T E R C F C G T P P H H K A A E L V V . . . . . G F G K K I F G . . .
T.clavata_DN48781_c0_g2   - E T C A E A Y K S C D - R V - K C C K A R . . . . . P C C . . . . . N I L G Y L C T C K K P L . . . . . S E I L T D W F G . . .
sp|B6DD58|U19-lycotoxin-Ls1b  . . G C S G A Y K R C S - S S Q R C C E G R . . . . . P C C . . . . . S A I N S N C K C R K T Y . . . . . T E L F K E Y F G . . . K
T.clavata_DN1126_c0_g1   . . S C I R R G G S C D H R P S D C C F S S . . . . . S C R C . . . . . N L W G T N C R C Q R . . . . . M G L F Q K W - G . . . K

```

4B

```

      |         |         |         |         |         |         |         |         |         |
      10        20        30        40        50        60
TRINITY_DN1_c0_g1      - E C I P E N K D C K S N . . . . R N G C C P T N K W L Q E S Y C H C Y L A A K D K K R Y N T V S K A E E K C F C D T Q V S I F G G A K . .
sp|P58605|Omega-segestritoxin-Sf1a  G S C I E S G K S C T H S R S M K N G L C C P K S . . . . . R C N C R Q I Q H R H - D Y - - L G K R K Y S C R C S . . . . .
TRINITY_DN919_c0_g2    E S C S P Q E N R C F K E . . . . G P K C G G F . . . . . D C R C Y T . . . . . P I V N G V K E E V V C W C N E P N V L Y E Y A Y K V
TRINITY_DN14378_c0_g3  - Q C I K L E T E C T K N . . . . K G N C C G G H . . . . . G C R C Y R . . . . . K Y V N G V K A E V R C W C I E N D V T Y E P R I E A
TRINITY_DN26733_c0_g1  - Q C P H I E Q E C T I K . . . . K G I C C N E E . . . . . E C R C Y R . . . . . K Y I N D V E G G V R C W C T N V D E T Y L P V . . .
sp|B6DCV8|U6-lycotoxin-Ls1g        - A C I P L E K E C T K T . . . . P G N C C S G L . . . . . K D C Y R . . . . . R F E Q G V A K G V Q C W C I E K D V T Y K G V . . .
TRINITY_DN896_c0_g1    E T C A E A Y K S C D R V . . . . K C C K A R . . . . . P C C N I L G . . . . . Y L C T C K K P L S E I L T D W F G
sp|P84000|U10-ctenitoxin-Pr1a      - S C A D A Y K S C D S L . . . . K C C N N R . . . . . T C M C S M I G . . . . . T N C T C R K K . . . . .

```

Figure 4

Alignments of cysteine-rich peptide toxins mature sequences from group (C-C-CC-CXC-CXC) in *S. pengi*.

```

T.clavata_DN18027_c0_g1      . . S L P E G A P C D G D K E D C C C Y G K W H K C R C P W P W E . . . D G P C R C A K G M K H T C I T K L S C P N K D E W G L D W R N E E . E R S P C
sp|P15969|Omega-agatoxin-1A AKALPPGGSVC DGNESDCCKCYGKWHKCRCPWKWHFTGEGPCTCEKGMKHTCITKLHC PNKAEWGLDWRSEESERSPC

```

Figure 5

Alignments of cysteine-rich peptide toxins mature sequences from groups I in *T. clavata*.

Figure 6

Alignments of cysteine-rich peptide toxins mature sequences from group III (C-CC-CC-CC-C-C-C) in *T. clavata* (6A) and *S. pengi* (6B).

```

T.clavata_DN51959_c0_g1      E D P G Y C P D F T . D K V C P F N Y K I N D C C A Q S D C P S Y A V C C V Q P C G N V C R Q M A Y . E P I G T P F V D G T K C E I G H V F P K A W Y E K L W D K I T S G .
T.clavata_DN58023_c0_g1      S H E A Y C P K R Y . T W V C . . V R S I N E C C S D D D C D R G Q F C C Q E N C G N T C Q F T T S F P T D G S K V V F D K R C G V E I . . . . .
sp|B6DD43|U15-Iycotoxin-Ls1a . . D Q Y C P K S S . I T A C K K M N I R N D C C K D D D C T G G S W C A T P C G N F C K Y P T D . R P G G K R A A G G K S C K T S Y V Y . . . . .
T.clavata_DN2521_c0_g1      . . G K Y C P E P K R K E P C P L Y R I N D C C K Q S D C R Q G S T C C K I Q C G N A C Q R E S P V A T S G V P V K D D E P C Q L G R V D K . . . . .
sp|B6DD61|U20-Iycotoxin-Ls1b . . G G F C P K S R . H P Q C N L S Y K I N D C C A Q S D C R V G S V C C V E G C G N V C R A E S D . T P L G E K F V D G S E C K H G H V F P K K W Y Q F W W R V . . . . .
T.clavata_DN11433_c0_g1      S D F D F C P R P I . D D E C P F Q Y K I N D C C A Q H D C R V G D I C C S E P C G S V C R R K S D . T P M G I I F K D G T E C Q L G H V Y P K K W Y Q K A W D N I K G . .
sp|Q5Y4V9|U7-agatoxin-Ao1a  K E F N F C P R A I . D E V C P V K E K R N E C C S S K E C R F G E M C C S E P C G N V C R V K S D . T P L G F P A K E D S N C K V G E I . K K K W Y Q K V W S K I T K W G

```

Figure 7

Alignments of cysteine-rich peptide toxins mature sequences from group IV (C-C-CC-C-C-CXC-C-C) in *T. clavata*.

```

TRINITY_DN854_c0_g1         Q Y S C R S Y G Q A C D G R R P G M S C C T G L G C R C D W I G N S D E C L R H L C C G I Y I . . . . .
sp|B6DD53|U16-Iycotoxin-Ls1b . . S C A K E Y Q W C D . . W N T R P C C D N I S C I C S W I G T N D E C K K G I I D T I K D W Y K G K

```

Figure 8

Alignments of cysteine-rich peptide toxins mature sequences from group V (C-C-CC-C-CC-C-C-C) in *S. pengi*.

```

T.clavata_DN21003_c0_g1      S S C S L P G A R C T D D C D C C I N N V E C D C E S G . . . H C V C I R S S D T S V C E R K K R M C S N P A V S D F P D G P C V V K S R R K R . .
T.clavata_DN51867_c0_g1      . I C K M P G G E C K N D C N C C Y R H T Y C N C P L W G A F G C S C I . H G D S M V C K R K K E Q C S D P M V K D F P R G S C F S G . . . . .
sp|Q5Y4U3|U9-agatoxin-Ao1a  . A C I K E G E K C A G D C C C G K W S Y C S C P L F G A L G C S C I . I G D A M V C V R K K K E C R T S D V M N T P P G G C F S S S K R R H R

```

Figure 9

Alignments of cysteine-rich peptide toxins mature sequences from group (C-C-C-CC-CXC-CXC-C-C-C) in *T. clavata*.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1Assemblystatistics.xlsx](#)
- [TableS2unigeneannotationofS.pengi.csv](#)
- [TableS3unigeneannotationofT.clavata.csv](#)
- [TableS4ToxinsinformationofT.clavata.xlsx](#)
- [TableS5toxinsinformationofS.pengi.xlsx](#)