Toxicon: X 1 (2019) 100005

Contents lists available at ScienceDirect

Toxicon: X

journal homepage: www.journals.elsevier.com/toxicon-x

Can we resolve the taxonomic bias in spider venom research?

Volker Herzig[a,*], Glenn F. King[a], Eivind A.B. Undheim[b,**]

[a] Institute for Molecular Bioscience, The University of Queensland, St. Lucia, QLD 4072, Australia
[b] Centre for Advanced Imaging, The University of Queensland, St. Lucia, QLD 4072, Australia

ARTICLE INFO

Keywords:
Spider
Venom
Peptide
Diversity
Taxonomic
Transcriptomics
Proteomics

ABSTRACT

The rate of discovery of new spider species greatly exceeds the rate of spider venom characterisation, leading to an increasing number of species with unstudied venoms. However, recent advances in proteomics and genomics that enable the study of venoms from smaller species has expanded the accessible taxonomic range. Thus, although the number of unstudied spider venoms is likely to further increase, future research should focus on the characterisation of venoms and toxins from previously unstudied spider families.

Spiders are among the most diverse and speciose venomous animals (King and Hardy, 2013). There are currently 47,807 recognised extant species divided into 118 families (World Spider Catalog, 2018), with the number of described species increasing at an average rate of ~800/year (based on the last 10 years recorded in the World Spider Catalog). Spiders are classified into two suborders (King, 2004): Mesothelae, comprised of a single family, and Opisthothelae, which contains the remaining 117 families. The Opisthothelae are further divided into the infraorders Mygalomorphae (20 families) and Araneomorphae (97 families), with the latter infraorder containing 93.5% of all extant species (Garrison et al., 2016).

Unfortunately, the incredible taxonomic diversity of spiders is poorly reflected in research on spider venoms. As of October 29, 2018, the databases **ArachnoServer**, a publicly available database on spider-venom proteins and peptides (Pineda et al., 2018), and **VenomZone** (https://venomzone.expasy.org/), containing all the manually curated UniProtKB/Swiss-Prot entries, together contained data on 1946 toxins from only 28 of the 118 extant spider families. Nothing is known about toxins from the remaining 90 taxonomic families (Fig. 1A). At the species level, the statistics are even more dire, with those databases containing toxins from 61 mygalomorph and 66 araneomorph species, representing only 2.0% and 0.1% of all extant species from these infraorders, respectively (Fig. 1B). Overall, only 0.3% of all extant spider species are represented in those databases. Thus, we have barely scratched the tip of the iceberg of spider-venom peptide diversity.

There are several reasons behind this taxonomic bias. The most important is the size of the spiders, which is a major factor in determining venom yield (Herzig, 2010). Venom research is traditionally pharmacology-driven, with toxin discovery often based on bioactivity screening of crude or fractionated venoms. This is a sample-costly approach that typically requires hundreds of micrograms to several milligrams of venom. Unfortunately, most spiders are too small for the commonly used method of venom extraction via mild electrical stimulation of the chelicerae, which contracts the muscles surrounding the venom gland and causes secretion of the venom (Herzig and Hodgson, 2008). An alternative method is the dissection of the venom glands with subsequent extraction of venom (Rash et al., 2000) (which in case of small spiders requires many specimens to obtain sufficient venom yields).

The generally larger size of mygalomorph spiders thus explains why a larger percentage of described toxins are from mygalomorph species. Spiders from the family Theraphosidae (commonly referred to as tarantulas) are known to provide large venom yields and therefore, not surprisingly, they represent 34.6% of all spider species with toxins listed in the ArachnoServer and VenomZone databases. Based on a dataset of 1067 theraphosids from > 300 species that were milked by one of the authors (V.H.), their median venom yield increases with increasing size of the spider (Fig. 1C), reaching 5.3 mg in the largest specimens with > 30 mm prosoma length (used as an indication of the size class, for details see (Herzig, 2010)). The top three venom yielding spiders (all females) were: *Acanthoscurria geniculata* (28.4 mg), *Poecilotheria metallica* (27.2 mg), and *Lasiodora cf klugi* (22.9 mg). Needless to say, comparable venom yields are not obtainable from the vast majority of spider species.

* Corresponding author.
** Corresponding author.
E-mail addresses: v.herzig@uq.edu.au (V. Herzig), e.undheim@uq.edu.au (E.A.B. Undheim).

https://doi.org/10.1016/j.toxcx.2018.100005

Available online 02 January 2019

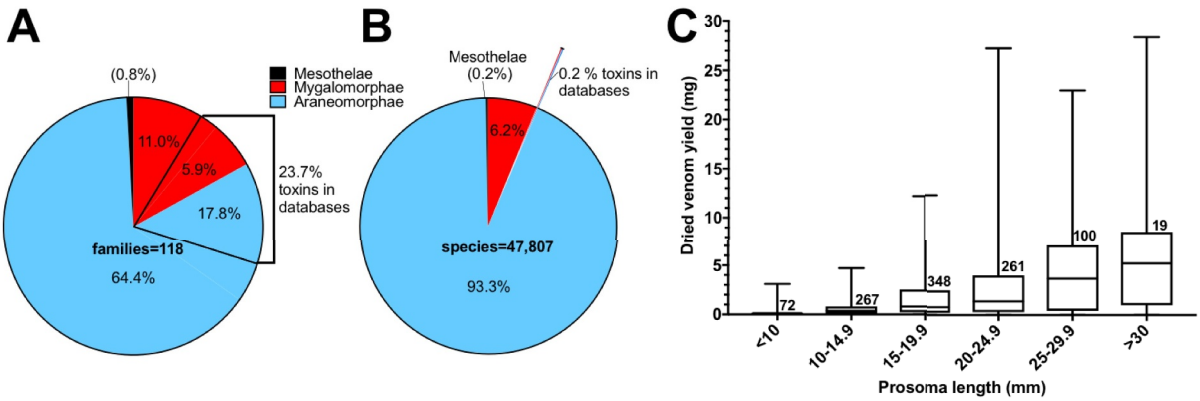

Fig. 1. Taxonomic representation of described spider-venom peptides across all spider families (**A**) and species (**B**). The total percentage of known toxins based on venom peptides in the ArachnoServer and VenomZone databases from the respective taxonomic level is also indicated. For better visibility, in panel **B** the percentage of species covered by these databases is represented by the tiny section that has been taken out of the circle of all the 99.8% of species with no known peptide toxins. These figures also include the family Uloboridae with 283 species (0.6%), which do not have venom. (**C**) Venom yields from 1067 theraphosid spiders divided into six size classes based on prosoma length (cumulative data obtained by V.H. over an 8-year period). The box-and-whisker plots indicate the median venom yield (line inside box), the 25th and the 75th percentile (bottom and top of box) as well as the 5th and 95th percentile (the whiskers). The number of analysed milkings per size class is indicated on top of each box.

A sequence-driven approach to peptide-toxin discovery (Mobli et al., 2017) obviates the problem of venom availability. This approach makes use of bioinformatics to mine the increasing amount of publicly available sequence data that has resulted from the exponential reduction in cost of DNA sequencing over the past decade (Fig. 2). Using a combination of bioinformatics tools and existing functional annotations of known toxins, peptides with potentially interesting or novel pharmacologies can be identified and subsequently produced for functional testing using recombinant technology (Klint et al., 2013). The reduction in sequencing costs, combined with reduced sample requirements, have also increased the scope and number of phylogenomic studies, and this has resulted in a greatly improved taxonomic diversity of publicly available spider whole-body transcriptomic datasets (Fernandez et al., 2018). The Sequence Read Archive (NCBI) currently contains 491 spider transcriptome biosamples (440 with > 1000 MBases) from 278 species (251 species > 1000 MBases). Although these data represent a substantial resource for taxonomically unbiased exploration of spider toxins, accurate annotation of venom components is a non-trivial task, even in venom-gland specific transcriptome data (Smith and Undheim, 2018). Fortunately, continuing increases in the sensitivity and resolution of mass spectrometers (Robinson et al., 2017) has made it increasingly affordable and experimentally feasible to study venom from species that were traditionally considered inaccessible.

Other reasons for taxonomic bias in spider-venom research include the availability and proper identification of spiders. Spider diversity and size increase with increasing ambient temperature, and therefore toxinologists working in colder climates may be disadvantaged in terms of finding local species large enough for venom extraction. Alternatively, sourcing spiders from other (warmer) countries usually involves the application of permits for collection, export and import, which can be time-consuming. This might explain why most spider venom research is performed in regions with a warmer climate (usually the tropics or subtropics). In regions with a more temperate climate, some researchers use spiders sourced from foreign countries (Kuhn-Nentwig et al., 2004). In Europe and the USA, there is a large and still growing sector of private enthusiasts that keep arachnids as pets (Hauke and Herzig, 2017), and most species available via the pet trade are sourced from tropical or subtropical regions.

Regardless of where spiders are sourced, proper identification of the specimen remains a major challenge for most toxinologists. Toxinologists are usually not experts in spider taxonomy, so the help of an expert arachnologist is often essential in order to properly identify spiders prior to publication. Unfortunately, spider taxonomists are rare and the time-scale for taxonomic projects does not align well with the more rapid progression of toxinology projects. Insufficient or incorrect identification of spiders used for research makes it difficult for other researchers to confirm the results. Correct identification is especially important for studies that use toxin sequences in a phylogenetic context to understand the taxonomic relationship between spiders. On the other hand, studies focused on the pharmacological effects of venom toxins might be less affected by incorrect specimen identification, as toxins can be produced recombinantly or by chemical synthesis, making the outcomes independent of the source specimen. Nevertheless, it might be useful (wherever possible) to deposit a voucher specimen of spiders used for toxinological research in museum collections to enable later access by research colleagues in cases where the spider identification has led to questionable outcomes.

Fortunately, molecular approaches provide a potential solution to also this issue, for example through identifying morphologically similar or cryptic species by DNA barcoding (Astrin et al., 2016). DNA barcoding also provides an easily accessible taxonomic reference for subsequent studies, which is particularly useful following taxonomic revisions that lead to the splitting of species. Additionally, recent improvements in the sensitivity of analytical techniques means that there is a reduced requirement for pooling specimens, which has the potential for inadvertent inclusion of multiple cryptic species.

Given the advanced proteomic and transcriptomic techniques that we already have available, we suggest that a stronger focus of future efforts in the field should be directed towards studying venoms from spider families (or subfamilies), from which no toxins have been described. This will at least help to reduce the current taxonomic bias in spider venom research at a family level. Examples of entire spider families for which the first transcriptome-derived toxin sequences were recently added to ArachnoServer include Barychelidae (*Trittame loki*; Undheim et al., 2013)), Miturgidae (*Cheiracanthium punctorium*; Vassilevski et al., 2010) Pisauridae (*Dolomedes mizhoanus*; Jiang et al., 2013), Scytodidae (*Scytodes thoracica*; Zobel-Thropp et al., 2014) and Zodariidae (*Lachesana tarabaevi*; Kozlov et al., 2006). In contrast, taxonomic bias at the species level is unlikely to be resolved in the foreseeable future due to the rate at which new species are discovered. With an annual increase in newly described spider species of ~800/year (World Spider Catalog, 2018), the taxonomic bias at the species level will likely get even worse. Given that the number of extant species may be as high as 170,000 (Coddington and Levi, 1991), it is unlikely

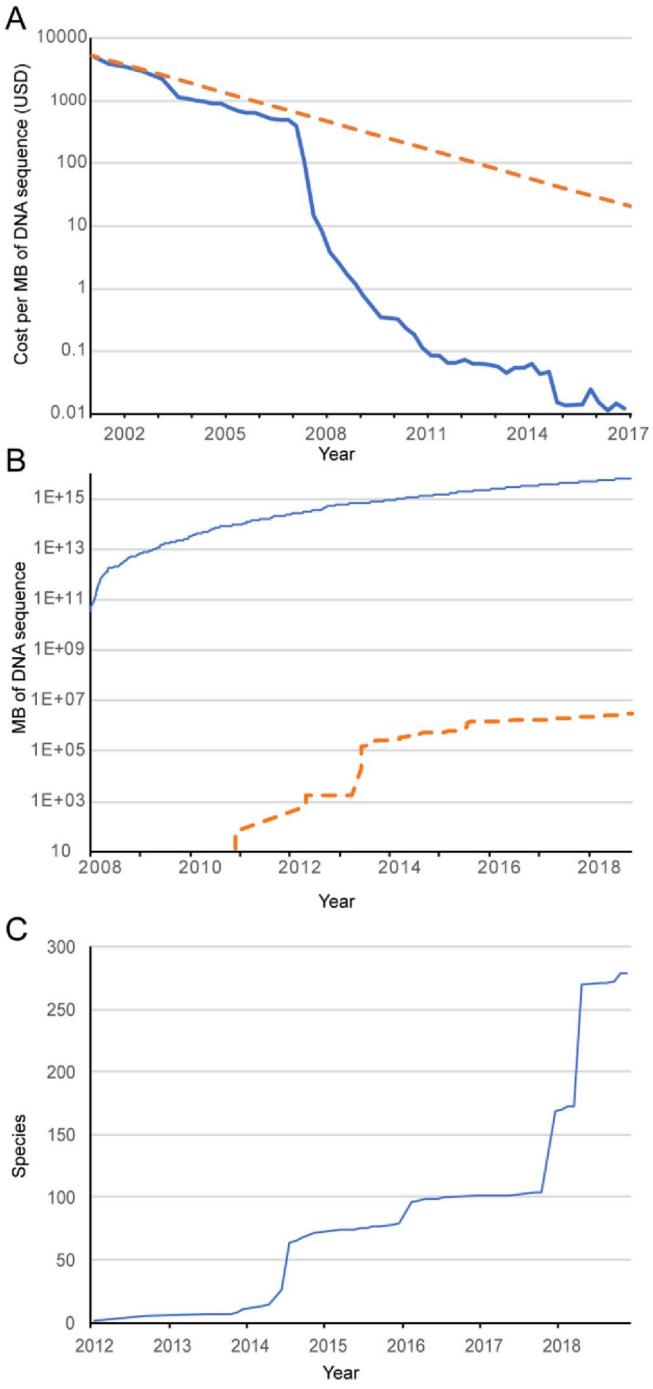

Fig. 2. Decrease in sequencing costs and increase in sequence data in the NCBI Sequence Read Archive (SRA). (**A**) Actual sequencing costs (blue solid line) are those estimated by the National Human Genome Research Institute (Wetterstrand, 2018). Costs predicted by Moore's Law (i.e., a two-fold reduction in cost every second year) is shown by the dashed orange line. The blue line in panel (**B**) shows the publicly available megabases of DNA sequence, while the orange line indicates the publicly available megabases of spider (Araneae) DNA in the SRA obtained by extracting all available datasets. (**C**) Number of spider species with transcriptomes in the SRA. SRA data are from the SRA website (https://www.ncbi.nlm.nih.gov/sra/docs/sragrowth/; accessed 13/11/2018).

that the toxinology field will be able to keep pace with the growth in recognised spider diversity. Therefore, closing the taxonomic gap at the family level over the next decade should be the main focus of the field in order to provide us with a better understanding of the evolution of spider venoms as well as an abundance of new spider-venom peptides with novel structures (Undheim et al., 2015) and potential applications as drugs, insecticides, or molecular tools (King, 2011; King and Hardy, 2013; Saez and Herzig, 2019).

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We acknowledge financial support from the Australian Research Council (ARC Linkage Grant LP140100832 to G.F.K. and DECRA Fellowship DE160101142 to E.A.B.U.) and Australian National Health & Medical Research Council (Principal Research Fellowship APP1044414 to G.F.K.)

Transparency document

Transparency document related to this article can be found online at https://doi.org/10.1016/j.toxcx.2018.100005.

References

Astrin, J.J., Hofer, H., Spelda, J., Holstein, J., Bayer, S., Hendrich, L., Huber, B.A., Kielhorn, K.H., Krammer, H.J., Lemke, M., Monje, J.C., Moriniere, J., Rulik, B., Petersen, M., Janssen, H., Muster, C., 2016. Towards a DNA barcode reference database for spiders and harvestmen of Germany. PLoS One 11, e0162624.
Coddington, J.A., Levi, H.W., 1991. Systematics and evolution of spiders (Araneae). Annu. Rev. Ecol. Syst. 22, 565–592.
Fernandez, R., Kallal, R.J., Dimitrov, D., Ballesteros, J.A., Arnedo, M.A., Giribet, G., Hormiga, G., 2018. Phylogenomics, diversification dynamics, and comparative transcriptomics across the spider tree of life. Curr. Biol. 28, 2190–2193.
Garrison, N.L., Rodriguez, J., Agnarsson, I., Coddington, J.A., Griswold, C.E., Hamilton, C.A., Hedin, M., Kocot, K.M., Ledford, J.M., Bond, J.E., 2016. Spider phylogenomics: untangling the spider tree of life. PeerJ 4, e1719.
Hauke, T.J., Herzig, V., 2017. Dangerous arachnids-fake news or reality? Toxicon 138, 173–183.
Herzig, V., Hodgson, W., 2008. Neurotoxic and insecticidal properties of venom from the Australian theraphosid spider *Selenotholus foelschei*. Neurotoxicology 29, 471–475.
Herzig, V., 2010. Ontogenesis, gender, and molting influence the venom yield in the spider *Coremiocnemis tropix* (Araneae, Theraphosidae). J. Venom Res. 1, 76–83.
Jiang, L., Liu, C., Duan, Z., Deng, M., Tang, X., Liang, S., 2013. Transcriptome analysis of venom glands from a single fishing spider *Dolomedes mizhoanus*. Toxicon 73, 23–32.
King, G.F., 2004. The wonderful world of spiders: preface to the special Toxicon issue on spider venoms. Toxicon 43, 471–475.
King, G.F., 2011. Venoms as a platform for human drugs: translating toxins into therapeutics. Expert Opin. Biol. Ther. 11, 1469–1484.
King, G.F., Hardy, M.C., 2013. Spider-venom peptides: structure, pharmacology, and potential for control of insect pests. Annu. Rev. Entomol. 58, 475–496.
Klint, J.K., Senff, S., Saez, N.J., Seshadri, R., Lau, H.Y., Bende, N.S., Undheim, E.A.B., Rash, L.D., Mobli, M., King, G.F., 2013. Production of recombinant disulfide-rich venom peptides in the periplasm of *E. coli* for structural and functional analysis. PLoS One 8, e63865.
Kozlov, S.A., Vassilevski, A.A., Feofanov, A.V., Surovoy, A.Y., Karpunin, D.V., Grishin, E.V., 2006. Latarcins, antimicrobial and cytolytic peptides from the venom of the spider *Lachesana tarabaevi* (Zodariidae) that exemplify biomolecular diversity. J. Biol. Chem. 281, 20983–20992.
Kuhn-Nentwig, L., Schaller, J., Nentwig, W., 2004. Biochemistry, toxicology and ecology of the venom of the spider *Cupiennius salei* (Ctenidae). Toxicon 43, 543–553.
Mobli, M., Undheim, E.A.B., Rash, L.D., 2017. Modulation of ion channels by cysteine-rich peptides: from sequence to structure. Adv. Pharmacol. 79, 199–223.
Pineda, S.S., Chaumeil, P.A., Kunert, A., Kaas, Q., Thang, M.W.C., Le, L., Nuhn, M., Herzig, V., Saez, N.J., Cristofori-Armstrong, B., Anangi, R., Senff, S., Gorse, D., King, G.F., 2018. ArachnoServer 3.0: an online resource for automated discovery, analysis and annotation of spider toxins. Bioinformatics 34, 1074–1076.
Rash, L.D., King, R.G., Hodgson, W.C., 2000. Sex differences in the pharmacological activity of venom from the white-tailed spider (*Lampona cylindrata*). Toxicon 38, 1111–1127.
Robinson, S.D., Undheim, E.A.B., Ueberheide, B., King, G.F., 2017. Venom peptides as therapeutics: advances, challenges and the future of venom-peptide discovery. Expert Rev. Proteomics 14, 931–939.
Saez, N., Herzig, V., 2019. Versatile spider venom peptides and their medical and agricultural applications. Toxicon 158, 109–126.
Smith, J.J., Undheim, E.A.B., 2018. True lies: using proteomics to assess the accuracy of transcriptome-based venomics in centipedes uncovers false positives and reveals startling intraspecific variation in *Scolopendra subspinipes*. Toxins 10, 96.

Undheim, E.A., Grimm, L.L., Low, C.F., Morgenstern, D., Herzig, V., Zobel-Thropp, P., Pineda, S.S., Habib, R., Dziemborowicz, S., Fry, B.G., Nicholson, G.M., Binford, G.J., Mobli, M., King, G.F., 2015. Weaponization of a hormone: convergent recruitment of hyperglycemic hormone into the venom of arthropod predators. Structure 23, 1283–1292.

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

Vassilevski, A.A., Fedorova, I.M., Maleeva, E.E., Korolkova, Y.V., Efimova, S.S., Samsonova, O.V., Schagina, L.V., Feofanov, A.V., Magazanik, L.G., Grishin, E.V., 2010. Novel class of spider toxin: active principle from the yellow sac spider *Cheiracanthium punctorium* venom is a unique two-domain polypeptide. J. Biol. Chem. 285, 32293–32302.

Wetterstrand, K.A., 2018. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP). http://www.genome.gov/sequencingcostsdata, Accessed date: 24 July 2018.

World Spider Catalog, 2018. Natural History Museum Bern. http://wsc.nmbe.ch, Accessed date: 29 October 2018.

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