UNIVERSITY OF CALIFORNIA
RIVERSIDE

Spider Silk Adaptations: Sex-Specific Gene Expression and Aquatic Specializations

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Evolution, Ecology, and Organismal Biology

by

Sandra Magdony Correa-Garhwal

March 2018

Dissertation Committee:
Dr. Cheryl Hayashi, Co-Chairperson
Dr. Mark Springer, Co-Chairperson
Dr. John Gatesy
Dr. Paul De Ley

The Dissertation of Sandra Magdony Correa-Garhwal is approved:

_____

_____

_____

Committee Co-Chairperson

_____

Committee Co-Chairperson

University of California, Riverside

Acknowledgements

I want to express my deepest gratitude to my advisor, Dr. Cheryl Hayashi, for her admirable guidance, patience, and knowledge. I feel very fortunate to have her as my advisor and could simply not wish for a better advisor. Thank you very much to my dissertation committee members Dr. John Gatesy, Dr. Mark Springer, and Dr. Paul De Ley, for their advice, critiques, and assistance.

I thank my fellow lab mates. I specially thank the graduate student Cindy Dick, Post-Doctoral Associates Crystal Chaw, Thomas Clarke, and Matt Collin, and undergrads from the Hayashi lab who provided assistance with planning and executing experiments helping me move forward with my dissertation.

Thanks to all the funding sources, Army Research Office, UC MEXUS, Dissertation Year Program Fellowship from the University of California, Riverside, Dr. Janet M. Boyce Memorial Endowed Fund for Women Majoring in the Sciences from UCR, and Lewis and Clark Fund for Exploration and Field Research from American Philosophical Society.

I would also like to thank Angela Simpson, Cor Vink, and Bryce McQuillan for aiding in the collection of *Desis marina*. Thanks to Peter Maddison for information, assistance, and hospitality during collecting trips at Kauri Point Reserve, New Zealand. Thanks Marc Janssen and Luc Crevecoeur for helping with the collection of *Argyroneta aquatica* in Neerpelt, Belgium. Thank you to the Kisailus lab for their help with structural analyses.

I would like to thank my family, for their support and encouragement through my career. Specially my husband Rahul Garhwal and my mother Teresa Arroyave who have always given me support and above all believed in me since the beginning. Words are not enough to show my gratitude.

Chapter 1 was reproduced with permission from Correa-Garhwal, S.M., Chaw, R.C., Clarke III, T.H., Ayoub, N.A., Hayashi, C.Y. 2017. Silk gene expression of theridiid spiders: implications for male-specific silk use.

ABSTRACT OF THE DISSERTATION

Spider Silk Adaptations: Sex-Specific Gene Expression and Aquatic Specializations

by

Sandra Magdony Correa-Garhwal

Doctor of Philosophy, Graduate Program in Evolution, Ecology, and Organismal Biology
University of California, Riverside, March 2018
Dr. Cheryl Hayashi, Co-Chairperson
Dr. Mark Springer, Co-Chairperson

Spiders are very diverse and inhabit most terrestrial environments. Spiders are renowned for their silk usage, which they rely on for an array of essential, fitness-related tasks such as reproduction, dispersal and prey-capture. Spider silks are proteinaceous and are largely composed of structural proteins called spidroins. Spidroins are encoded by a gene family that has undergone dramatic proliferation. To date, whether specific molecular modifications of silks are associated with success in particular habitats is still unknown. To better understand the specializations of silk sequence and expression levels, I examined the molecular composition of spidroins from terrestrial and aquatic (marine and freshwater) spiders to determine how aquatic spiders are using their silk to thrive in wet environments. I also compared silk gene expression levels of males and females within and across species.

Sex-biased silk expression was observed in genes associated with male- and female-specific tasks (e.g., egg-case production by females, wandering for mates by males). While all females highly expressed egg-case silk genes, male silk gene expression

differed across species. Comparison of silks from a semi-aquatic spider to a terrestrial counterpart did not provide obvious evidence for unique specializations of silks that function underwater. Rather, the silks of the semi-aquatic spider appear to have been preadapted for aquatic use. Characterization of silk molecules from a terrestrial cribellate spider (constructs webs with cribellar silk), led to the identification of new protein sequence motifs, including a putative gene for cribellar silk. Finally, by comparing the spidroins from multiple species that independently evolved water-association, I found that they share a sequence motif composed of hydrophobic amino acids [$(GV)_n$]. My research also addressed some of the structural aspects of silk fibers used by different spiders in terrestrial and aquatic settings. Characterization of silks from terrestrial, aquatic, and marine semi-aquatic spiders shows that spiders have developed different silk-related strategies to match their habitat and lifestyle. In addition to contributing to an integrated understanding of spider silks, my research expands the potential of spider silk as a novel biomaterial.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

RPKM: reads per kilobase per million reads
AcSp: aciniform spidroin
AmSp: ampullate spidroin
CrSp: cribellar spidroin
PySp: pyriform spidroin
TuSp: Tubuliform spidroin

Introduction

An adaptation can be defined as a trait that evolved via natural selection for a particular use which enhances the fitness of the organism with that trait (Andersen, 1995; Gould and Vrba, 1982). Adaptations help organisms survive in their ecological niches. Adaptive traits are found at many levels of biological organization, including structural and molecular. Structural adaptations are those that change the morphology of the organism and can be studied by direct observation (Futuyma, 2009). Molecular adaptations are perceived as changes in nucleotide and protein sequence contributing to phenotypic variation. These changes could be in gene sequence, gene regulation, and/or the origin of new genes, which usually arrives via gene duplication. Thus, protein-coding genes with functions that have been identified as adaptations are referred to as adaptations themselves (for examples see Kondrashov, 2012).

Spiders produce silk throughout their life and they use silk not only for prey-capture for also for other tasks such as reproduction and dispersal. Silk use in spiders has also allowed them to inhabit nearly all terrestrial environments (Bond and Opell, 1998). Numerous synapomorphies that unite spiders as a monophyletic group are related to silk synthesis including abdominal silk glands, spinnerets (modified appendages), and the modified pedipalpi in males used for sperm transfer (Coddington et al., 2004). Given the fundamental connection of silk to spider ecology, spider silks and the systems that produce it have been used as phylogenetic characters (Blackledge et al., 2009, 2003; Bond and Opell, 1998; Coddington, 1989; Platnick et al., 1991). Despite the importance of silk in the survival of spiders, it is still unclear whether specific molecular

modifications have contributed to their success in diverse habitats. Therefore, in my research I set out to investigate molecular adaptations of spider silk.

Spider silk proteins are produced and stored as a liquid dope in the lumen of silk glands. The liquid silk travels through ducts, and then is secreted as fibers or glues via spigots located on the spider's spinnerets (Coddington, 1989; Gosline et al., 1986; Hayashi, 2002). Glands are grouped according to their morphology, and each gland type produces a different silk type. There is extensive variation in the number and morphology of silk gland types across the phylogeny of spiders. For example, tarantulas and their kin (Mygalomorphae) spin one or two multi-purpose silks from morphologically indistinct glands (Palmer et al., 1982). In contrast, orb-web weaving spiders (Orbiculariae) spin up to seven functionally distinct silks from morphologically differentiated silk glands: aciniform, aggregate, flagelliform, major ampullate, minor ampullate, pyriform, and tubuliform (Guerette et al., 1996).

In addition to different gland types and functions, each silk type has specialized mechanical properties that are partly due to the molecular organization of spider silk proteins. Spidroins (a contraction of "spider fibroins"), are the main component of spider silks and are encoded by a single gene family (Guerette et al., 1996). The majority of spidroins are very large proteins (e.g., 1,700-6,000 amino acids; Ayoub et al., 2013, 2007; Chen et al., 2012; Hayashi and Lewis, 2000) and their primary structure consists of ensemble repeats confined by non-repetitive amino and carboxyl terminal domains (Garb et al., 2010; Gatesy et al., 2001). Tandem repeats can consist of short, iterated amino acid sequence motifs. For example, major ampullate silk (dragline) is dominated by poly A

(strings of alanine), GGX, and GPGXX amino acid motifs (where X is a restricted subset of amino acids; (Ayoub et al., 2007; Xu and Lewis, 1990; Zhang et al., 2013). These amino acid motifs are thought to contribute to the outstanding mechanical properties of dragline silks (Hayashi et al., 1999; Sponner et al., 2007). Each silk has unique combinations of amino acid motifs tailored for specific functions. For example, because primary structure informs secondary and other higher-level structures, poly A confers strength while GPGXX facilitates reversible stretch (Blackledge and Hayashi, 2006; Hayashi and Lewis, 1998).

Silk biology in spiders is vital to the understanding of how spiders adapt to changing environments. Yet, to date, information regarding sex and habitat specific silk use and silk gene expression is virtually unknown. To investigate how important silk is to spiders my work examines the silk spinning structures (spinnerets, spigots) and silk genetics (gene characterization and expression) of spiders that have modified silk use to meet specialized silk requirements. I examined differences and similarities of silks from both sexes within and across species. Additionally, I look into the molecular composition of terrestrial and aquatic (marine and freshwater) spider species and how they are using their silk to thrive in wet environments. This research shows the importance of silks for all spiders, regardless of environment or sex, to achieve essential functions related to their lifestyle.

In Chapter one, to investigate silk gene sex-biased expression in spiders I examine silk synthesis by mature males and females from three cob-web weaving species (Theridiidae). I quantify the relative expression of silk genes and determine differences

and similarities of silk use within and between species. The results show differences in expression levels between sexes for silk gene proteins associated with sex-specific tasks. In addition, silk gene expression of male spiders across species shows a species-specific adaptive divergence of silk use by theridiid males.

In Chapter two, to investigate whether silk sequences of semi-aquatic spiders have molecular modifications linked to aquatic environments I describe the silk genes expressed by the fishing spider *Dolomedes triton* (Pisauridae). Furthermore, I examined the characteristics of the submersible egg sac that allows function under water. Gene characteristics of *D. triton* spidroins show no evidence of adaptive modification unique to functioning in water, suggesting that spider silk in general are adapted to support survival in semi-aquatic environments. Yet, the *D. triton* egg sac shows water-repellent features such as a different fiber arrangement and elemental composition.

In Chapter three, to characterize the structural proteins of silks I studied the spidroins produced by *Tengella perfuga* (Tengellidae), a spider that makes webs padded with cribellar silk. Cribellar silk greatly differs from other silks; it is composed of many micro-fibrils that act as a dry adhesive. A candidate cribellar silk gene is described in this chapter as well as new combinations of amino acid sequence motifs in most *T. perfuga* spidroins. CrSp identified in this chapter allows for reconstruction of evolutionary relationships of cribellate silk proteins to other spider silk proteins.

In Chapter four, to further investigate the role of silk in spiders that adapt to changing environments, I examined habitat and sex-specific silk use across species. I studied the silk molecular composition of spiders within Dictynoidea that live in three

different environments (Forster, 1970). Specifically, silk genes are characterized from the fully aquatic spider *Argyroneta aquatica,* the marine semi-aquatic spider *Desis marina* and the terrestrial *Badumna longinqua*. I also examine sex-biased expression in aquatic and semi-aquatic spiders. Similar to spider species studied in previous chapters, males and females across species show differences in spidroin gene expression related to changes in lifestyle after sexual maturation. Analysis of silks from several spider species associated with aquatic environments reveals the presence of hydrophobic sequence elements possibly associated with underwater silk use.

My dissertation expands the diversity of the spidroin silk family and provides a general depiction of the role of silk genes in the adaptation of spiders to changing environments. From transcriptomics to proteomics, I describe key characteristics underlying the silk-related strategies of spider to different lifestyles, which offers opportunities for unique biomimetic materials.

References

Andersen, N.M., 1995. Cladistic inference and evolutionary scenarios: locomotory structure, function, and performance in water striders. Cladistics 11, 279–295.

Ayoub, N.A., Garb, J.E., Kuelbs, A., Hayashi, C.Y., 2013. Ancient properties of spider silks revealed by the complete gene sequence of the prey-wrapping silk protein (AcSp1). Mol. Biol. Evol. 30, 589–601. https://doi.org/10.1093/molbev/mss254

Ayoub, N.A., Garb, J.E., Tinghitella, R.M., Collin, M.A., Hayashi, C.Y., 2007. Blueprint for a high-performance biomaterial: full-length spider dragline silk genes. PloS One 2, e514. https://doi.org/10.1371/journal.pone.0000514

Blackledge, T.A., Coddington, J.A., Gillespie, R.G., 2003. Are three-dimensional spider webs defensive adaptations? Ecol. Lett. 6, 13–18.

Blackledge, T.A., Hayashi, C.Y., 2006. Unraveling the mechanical properties of composite silk threads spun by cribellate orb-weaving spiders. J. Exp. Biol. 209, 3131–3140. https://doi.org/10.1242/jeb.02327

Blackledge, T.A., Scharff, N., Coddington, J.A., Szüts, T., Wenzel, J.W., Hayashi, C.Y., Agnarsson, I., 2009. Reconstructing web evolution and spider diversification in the molecular era. Proc. Natl. Acad. Sci. U. S. A. 106, 5229–5234. https://doi.org/10.1073/pnas.0901377106

Bond, J.E., Opell, B.D., 1998. Testing adaptive radiation and key innovation hypotheses in spiders. Evolution 52, 403–414. https://doi.org/10.2307/2411077

Chen, G., Liu, X., Zhang, Y., Lin, S., Yang, Z., Johansson, J., Rising, A., Meng, Q., 2012. Full-length minor ampullate spidroin gene sequence. PLoS ONE 7, e52293. https://doi.org/10.1371/journal.pone.0052293

Coddington, J.A., 1989. Spinneret silk spigot morphology: evidence for the monophyly of orbweaving spiders, Cyrtophorinae (Araneidae), and the group Theridiidae plus Nesticidae. J. Arachnol. 17, 71–95.

Coddington, J.A., Giribet, G., Harvey, M.S., Prendini, L., Walter, D.E., 2004. Arachnida, in: Assembling the Tree of Life. Edied by Cracraft J, Donoghue M. Oxford University Press, New York, pp. 296–318.

Forster, R.R., 1970. The spiders of New Zealand. Part III. Desidae, Dictynidae, Hahniidae, Amaurobioididae, Nicodamidae. Otago Museum Bulletin, pp. 1–184.

Futuyma, D., 2009. Evolution, Second Edition, 2nd Edition edition. ed. Sinauer Associates, Inc., Sunderland, Mass.

Garb, J.E., Ayoub, N.A., Hayashi, C.Y., 2010. Untangling spider silk evolution with spidroin terminal domains. BMC Evol. Biol. 10, 243. https://doi.org/10.1186/1471-2148-10-243

Gatesy, J., Hayashi, C., Motriuk, D., Woods, J., Lewis, R., 2001. Extreme diversity, conservation, and convergence of spider silk fibroin sequences. Science 291, 2603–2605. https://doi.org/10.1126/science.1057561

Gosline, J.M., DeMont, M.E., Denny, M.W., 1986. The structure and properties of spider silk. Endeavour 10, 37–43. https://doi.org/10.1016/0160-9327(86)90049-9

Gould, S.J., Vrba, E., 1982. Exaptation-A Missing Term in the Science of Form. Paleobiology 8, 4–15.

Guerette, P.A., Ginzinger, D.G., Weber, B.H.F., Gosline, J.M., 1996. Silk properties determined by gland-specific expression of a spider fibroin gene family. Science 272, 112–115. https://doi.org/10.1126/science.272.5258.112

Hayashi, C.Y., 2002. Evolution of spider silk proteins: insight from phylogenetic analyses. EXS 209–223.

Hayashi, C.Y., Lewis, R.V., 2000. Molecular architecture and evolution of a modular spider silk protein gene. Science 287, 1477–1479.

Hayashi, C.Y., Lewis, R.V., 1998. Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks. J. Mol. Biol. 275, 773–784. https://doi.org/10.1006/jmbi.1997.1478

Hayashi, C.Y., Shipley, N.H., Lewis, R.V., 1999. Hypotheses that correlate the sequence, structure, and mechanical properties of spider silk proteins. Int. J. Biol. Macromol. 24, 271–275.

Kondrashov, F.A., 2012. Gene duplication as a mechanism of genomic adaptation to a changing environment. Proc. R. Soc. B Biol. Sci. 279, 5048–5057.

Palmer, J.M., Coyle, F.A., Harrison, F.W., 1982. Structure and cytochemistry of the silk glands of the mygalomorph spider *Antrodiaetus unicolor* (Araneae, Antrodiaetidae). J. Morphol. 174, 269–274. https://doi.org/10.1002/jmor.1051740303

Platnick, N.I., Griswold, C.E., Coddington, J.A., 1991. On missing entries in cladistic analysis. Cladistics 7, 337–343.

Sponner, A., Vater, W., Monajembashi, S., Unger, E., Grosse, F., Weisshart, K., 2007. Composition and hierarchical organization of a spider silk. PLOS One 2, e998.

Xu, M., Lewis, R.V., 1990. Structure of a protein superfiber: spider dragline silk. Proc. Natl. Acad. Sci. 87, 7120–7124.

Zhang, Y., Zhao, A.-C., Sima, Y.-H., Lu, C., Xiang, Z.-H., Nakagaki, M., 2013. The molecular structures of major ampullate silk proteins of the wasp spider, *Argiope bruennichi*: A second blueprint for synthesizing de novo silk. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 164, 151–158.

Chapter 1

Silk Gene Expression of Theridiid Spiders: Implications for Male-Specific Silk Use

Abstract

Spiders (order Araneae) rely on their silks for essential tasks, such as dispersal, prey capture, and reproduction. Spider silks are largely composed of spidroins, members of a protein family that are synthesized in silk glands. As needed, silk stored in silk glands is extruded through spigots on the spinnerets. Nearly all studies of spider silks have been conducted on females; thus, little is known about male silk biology. To shed light on silk use by males, we compared silk gene expression profiles of mature males to those of females from three cob-web weaving species (Theridiidae). We de novo assembled species-specific male transcriptomes from *Latrodectus hesperus*, *Latrodectus geometricus*, and *Steatoda grossa* followed by differential gene expression analyses. Consistent with their complement of silk spigots, male theridiid spiders express appreciable amounts of aciniform, major ampullate, minor ampullate, and pyriform spidroin genes but not tubuliform spidroin genes. The relative expression levels of particular spidroin genes varied between sexes and species. Because mature males desert their prey-capture webs and become cursorial in their search for mates, we anticipated that major ampullate (dragline) spidroin genes would be the silk genes most highly expressed by males. Indeed, major ampullate spidroin genes had the highest expression in *S. grossa* males. However, minor ampullate spidroin genes were the most highly expressed spidroin genes in *L. geometricus* and *L. hesperus* males. Our expression profiling results suggest species-specific adaptive divergence of silk use by male theridiids.

Introduction

Sexual dimorphism is a phenomenon resulting in significant differences in how males and females interact with their environment. In spiders, males and females show extraordinary sexual size dimorphism, with some males being dwarfs in comparison to females (Vollrath and Parker, 1992; Vollrath, 1998; Hormiga et al., 2000; Schütz and Taborsky, 2003, 2005). Often, this dwarfism in male spiders is associated with a nearly parasitic relationship between males and females, with males feeding on prey captured by females and living on female spider webs (Vollrath, 1998).

Differences between the sexes in spiders have been recognized regarding size, venom, and behavior, but differences in silk use remain largely unknown (Atkinson, 1981; Vollrath and Parker, 1992; de Oliveira et al., 1999; Binford, 2001; Binford et al., 2016). While there have been many studies that characterize spider silk genes, they have been based almost entirely on female spiders (Tian and Lewis, 2005; Zhao et al., 2006; Perry et al., 2010; Correa-Garhwal and Garb, 2014). For many species, including cob-web weavers (Theridiidae), this is because males mature at smaller body sizes than females, and males tend to have shorter lifespans (Figure 1.1A; Kaston, 1970; Andrade, 2003). Furthermore, female spiders have more types of silk glands than males (Kovoor and Peters, 1988; Peters, 1992; Park and Moon, 2002; Moon and An, 2006; Moon, 2008). Thus, virtually nothing is known regarding silk use and silk gene expression in male spiders.

Silk spinning in spiders involves a highly specialized system of genes, proteins, glands, and behaviors. An individual spider can have multiple types of silk glands that

can be categorized according to their morphology (e.g., aciniform, aggregate, flagelliform, major ampullate, minor ampullate, pyriform, and tubuliform silk glands; Vollrath, 1992). Each gland type produces a unique, gland-specific proteinaceous silk spinning dope. As needed, silk dope exits the glands through ducts, and then is extruded from external spigots on the spider's spinnerets (Gosline et al., 1986; Coddington, 1989). Silk spigots vary in size and shape and each spigot is associated with a single silk gland. Spider silk spigots are identified based on morphological characteristics such as relative size, number, and position on the spinnerets. Spigots are named according to the silk gland connected to them. For example, major ampullate spigots, which are located on the anterior lateral spinnerets, are attached to the major ampullate silk glands (Coddington, 1989).

From spider silk gland cDNA studies, it has been established that silk gland types differ in their expression of silk structural protein encoding genes. The predominant silk proteins are spidroins (a contraction of "spider fibroins"; Hinman and Lewis, 1992), which are encoded by a gene family (Guerette et al., 1996). Each silk gland type can express a particular set of spidroins. For example, major ampullate silk glands predominately express major ampullate spidroins, whereas silk protein expression in aciniform silk glands is dominated by aciniform spidroins. Each spidroin type forms task-specific silk fibers. Major ampullate spidroins are the main constituent of draglines, and aciniform spidroins form the silk used in prey wrapping.

Besides spidroins, other structural proteins have been identified as silk components, such as egg case proteins 1 and 2 (ECP-1 and ECP-2, used in egg case silk;

Hu et al., 2005, 2006), spider coating peptides 1 and 2 (SCP-1 and SCP-2. used as a gluey coating on some silk types; Hu et al., 2007), and aggregate silk factors 1 and 2 (AgSF1 and AgSF2, used in web-scaffolding; Vasanthavada et al., 2012). The genes that encode ECP-1, ECP-2, SCP-1, SCP-2, AgSF1, and AgSF2 are also differentially expressed among silk gland types. *ECP-1* and *ECP-2* are co-expressed with *tubuliform spidroin 1* (*TuSp1*) in tubuliform silk glands, and *SCP-1*, *SCP-2*, *AgSF1,* and *AgSF2* are most highly expressed in aggregate silk glands (Hu et al., 2005, 2006, 2007; Casem et al., 2010; Vasanthavada et al., 2012).

To investigate the silk biology of male spiders, we quantified the relative silk gene expression levels of mature male spiders versus those of conspecific females from three theridiid species, *Latrodectus hesperus* (Western black widow), *Latrodectus geometricus* (brown widow), and *Steatoda grossa* (false black widow). The presence of silk spigots corresponding to major ampullate, minor ampullate, pyriform, and aciniform glands have been documented in mature male and female theridiids, and thus we expected both sexes to express genes encoding spidroins associated with those glands (Figure 1.1B) (Moon and An, 2006). By contrast, females, but not males, have functional spigots for tubuliform, aggregate, and flagelliform glands. Accordingly, we expected only females to express genes associated with tubuliform, aggregate, and flagelliform silk glands. Finally, we predicted that in males, major ampullate spidroin genes will be highly expressed relative to other spidroin genes because males may need more major ampullate silk than other silks for functions related to wandering behaviors such as draglines and bridging lines.

Materials and methods

Construction and Sequencing of RNA-Seq Libraries

RNA-Seq library construction and sequencing were described by Clarke et al. (2015). Briefly, total RNA from two males was separately extracted for each focal species (*L. hesperus, L. geometricus,* and *S. grossa*) as biological replicates. Total RNA was also extracted from multiple tissue types from females for each focal species. The specific tissues from females included were aciniform+flagelliform, anterior aggregate, posterior aggregate, major ampullate, minor ampullate, pyriform, and tubuliform silk glands, as well as total silk (combination of all silk gland types). The tissue-specific assemblies also included the cephalothorax, venom glands, and ovaries. Indexed RNA-Seq libraries were made from each RNA extraction with either the mRNA Sequencing Sample Preparation Kit (Illumina, Inc., San Diego, CA, USA) as described by Clarke et al. (2014), or the TruSeq RNA v2 kit (Illumina) by the Johns Hopkins University Deep Sequencing and Microarray Core Facility. Libraries were paired-end sequenced for 75 or 100 cycles in each direction on a Genome Analyzer I or II (as in Clarke et al., 2014), or an Illumina HiSeq 2000 at the University of California, Riverside IIGB Genomics Core Facility. NCBI Sequence Read Archive (SRA) accession numbers are SRR1539570, SRR1539569, and SRR1539523 for *L. hesperus*, *L. geometricus*, and *S. grossa,* respectively.

De novo Transcriptome Assemblies

Assembly of de novo transcriptomes was described by Clarke et al. (2015). Briefly, raw sequence reads were processed before assembly; adaptors were clipped, arthropod ribosomal RNA and low quality reads were removed following Clarke et al. (2014). Reads were then de novo assembled using Trinity with default parameters (Grabherr et al., 2011) resulting in male transcriptomes and tissue-specific female transcriptomes for each species (see Section 2.1; Clarke et al., 2015). No k-mer normalization was performed, following Clarke et al. (2014).

Clarke et al. (2015) combined the individual male and tissue-specific female transcriptomes for each species using CAP3 (Huang and Madan, 1999). Since CAP3 combines multiple transcripts into a single transcript, maintaining isoform information for one or more transcripts became untenable so we only used the longest isoform for each transcript from each tissue-specific female and male library. This resulted in three species-specific, global transcriptomes that included assembled male individual and female tissue-specific sequencing reads (TSA accession numbers for *L. hesperus, L. geometricus, and S. grossa*: GBJN00000000, GBJM00000000, and GBJQ00000000, respectively). In addition to these global transcriptomes, the present study uses the transcriptomes assembled from male sequencing reads of each species (TSA accession numbers GFDB00000000, GFCZ00000000, GFDC00000000 for *L. hesperus*, *L. geometricus*, and *S. grossa,* respectively). Assembly quality was approximated using N50 (see Table S1.1). The completeness of the transcriptomes was assessed with CEGMA

(Parra et al., 2007) and BUSCO (Simão et al., 2015). Identified putative artifacts such as chimeras, bacterial sequences, and sequencing errors were removed.

Annotation and Expression Profiling

A database of silk proteins was created using spidroins, spider coating peptides SCP-1 and SCP-2, egg case silk proteins ECP-1 and ECP-2, and aggregate silk factors AgSF1 and AgSF2 that were downloaded from the NCBI nr protein and UniProtKB/Swiss-Prot databases in July 2015 (Table S1.2). BLASTX searches (e-value < 1e-5) implemented in Geneious version 8.1.8 (Kearse et al., 2012) against the silk protein database were used to infer protein translations and identify silk gene homologs in each assembly. Identified silk gene homologs were then manually inspected, and regions coding for the (N)- and (C)-terminal domains of published silk genes were used for expression analyses (Table S1.3).

Paired-end reads from male libraries were mapped to the appropriate male transcriptome and paired-end reads from female silk glands were mapped to the corresponding global transcriptome using the short read aligner, Bowtie (version 1.1.1; Langmead, 2010). Bowtie parameters allowed for permissive read mapping and were set to obtain the best unique match accounting for paired-end data (Langmead, 2010). SAMtools (Li et al., 2009) were used to count the mapped reads. Reads mapping to the terminal domain regions of silk gene transcripts were counted and used to calculate the reads per kilobase of transcript per million mapped reads (RPKM). Counts for mapping the reads from male biological replicates (two replicates per species) were averaged.

Biological replicates were found to have strong positive correlation (Figure S1.1). Identified silk transcripts from each library with at least ten mapped reads and RPKM > 1 were kept for analysis. Comparison of expression levels of male silk genes was done using the Kruskal–Wallis *H*-test implemented in the statistical R package (R Core Team, 2013).

Results and discussion

Silk Gene Expression in Females and Males

Assembly statistics for the three male transcriptomes, including total number of sequencing reads, assembled contigs (contiguous sequences), and N50 lengths, were similar (Table S1.1). We were successful in identifying silk gene transcripts in our transcriptome assemblies (Figure 1.2). Previously characterized spider silk gene transcripts were mostly partial length (Starrett et al., 2012; Clarke et al., 2014, 2015; Sanggaard et al., 2014) because obtaining complete silk transcripts is challenging due to the large size of spidroin-coding sequences (~5–19 kbp) and the internally repetitive sequence of spidroin and other silk genes (Hu et al., 2005, 2006; Ayoub et al., 2007, 2013; Chen et al., 2012; Chaw et al., 2014). Thus, as expected, our identified silk gene transcripts were mostly partial length, usually containing coding sequences for amino (N)- or carboxyl (C)-terminal domains adjacent to varying amounts of repetitive region sequence.

While the repetitive nature of spidroin genes can lead to inaccurate estimation of gene expression in RNA-Seq analyses (Chaw et al., 2016), terminal domains can

differentiate silk genes (Garb et al., 2010) and thus, the terminal portions of the transcripts were used to assess the expression of silk genes. To investigate silk gene expression in males, we mapped sequencing reads from males to male transcriptomes that included whole individual males, and we mapped sequencing reads from female silk glands (combination of all silk gland types) to global transcriptome assemblies that included sequencing reads from males and females. Transcription of silk genes associated with aciniform, major ampullate, minor ampullate, and pyriform glands was found in both sexes (Figure 1.2). Silk genes associated with the flagelliform silk gland were expressed exclusively in female silk glands. Males lack functional aggregate spigots and tubuliform spigots (Figure 1.1B) and thus were not expected to express genes associated with these glands. Surprisingly, however, the aggregate silk gland-associated genes *AgSp1* (*aggregate spidroin 1*), *AgSF2*, and *SCP-1*, as well as the tubuliform silk gland-associated gene *TuSp1* were expressed in males, although at a lower level than in conspecific females (Figure 1.2; Table S1.4). Overall, there was a significant difference between the expression levels of the expected silk genes *AcSp1, MaSp*s (*MaSp1, MaSp2, MaSp3*), *MiSp,* and *PySp1* and those of the silk genes that were not expected in males (*AgSp1*, *AgSF2, SCP-1,* and *TuSp1*) (Kruskal–Wallis *H*-test: P-value < 0.05; Figure 1.3).

Male theridiids lack aggregate and tubuliform silk glands (Moon and An, 2006). Aggregate glands secrete sticky glue that is used to subdue prey, and tubuliform glands produce silk fibers that are a primary component of egg case wrapping (Gosline et al., 1986; Tillinghast et al., 1992; Garb and Hayashi, 2005; Tian and Lewis, 2005; Hu et al., 2005, 2006, 2007; Choresh et al., 2009; Vasanthavada et al., 2012). Lack of aggregate

silk glands is consistent with a roving lifestyle where males do not capture prey, and males lack tubuliform glands because males do not spin egg sacs. Expression of silk genes that are thought to be restricted to aggregate and tubuliform glands could indicate a lack of complete transcriptional shutdown of silk genes after males reach sexual maturity. Alternatively, co-expression of expected with unexpected silk genes is also possible. Co-expression of silk genes can be facilitated by selection in regulatory regions (Ayoub et al., 2007). It has previously been shown that some silk genes are expressed in multiple types of silk glands. For example, *TuSp1* transcript expression was detected in the major ampullate glands of *L. hesperus* females (Lane et al., 2013). Thus, there is precedent for *TuSp1* silk genes to be expressed at low levels in the major ampullate glands of male theridiids.

Silk Gene Expression in Male Spiders

Current knowledge of spider behavior suggests that, upon sexual maturity, males use silk differently from females. Soon after hatching, cob-web weaver spiderlings disperse by aerial ballooning with silken lines that catch the wind (Suter, 1991; Foelix, 2011). Thereafter, juvenile theridiids of both sexes use silk to construct retreats and prey-capture webs, and both sexes trail a dragline as they walk or dangle in mid-air. Draglines are considered to be primarily composed of major ampullate silk (Gosline et al., 1984). After reaching sexual maturity, males enter a mate-searching phase, when they desert their juvenile prey-capture webs and search for a female that is receptive to mating (Moore, 1977; Foelix, 2011).

Expression of *AcSp1, MaSp*s*, MiSp,* and *PySp1* differed among the males of each species (Figure 1.4). Expression of *MaSp*s in *S. grossa* was consistent with our expectations. *S. grossa* males had higher relative expression of major ampullate spidroin genes (*MaSp*s) than of other silk genes or than conspecific females (Figure 1.4A). However, in *L. hesperus* and *L. geometricus*, we found minor ampullate spidroin genes (*MiSp*) to have a higher relative expression in male spiders than other spidroin genes or than *MiSp* in conspecific females (Figure 1.4B and C). This finding suggests extensive use of minor ampullate silk by *Latrodectus* males. For several species in the family Araneidae, minor ampullate silk is known to be used in orb-web construction, dragline silk, and as a "bridging line" between distant destination points (Work, 1981; Peters, 1990). For female *Latrodectus* spiders, minor ampullate silk has also been implicated in web construction and prey capture (Benjamin and Zschokke, 2002; LaMattina et al., 2008).

The draglines of female *L. hesperus* consist of two fibers spun from major ampullate glands that are often supplemented by two fibers spun from the minor ampullate glands (Hsia et al., 2011). A single female *L. hesperus* major ampullate silk fiber is ~3.3 μm in diameter (Lawrence et al., 2004) and a minor ampullate silk fiber is ~1.12 μm in diameter (Vienneau-Hathaway et al., 2017). Differences in fiber diameter of major and minor ampullate silk have been shown for other spider species (Blackledge and Hayashi, 2006).

We observed two different dragline compositions in *Latrodectus* males. Draglines laid down by males wandering over surfaces were found to consist of two fibers, each ~580 nm in diameter (Figure S1.2). In contrast, draglines produced by falling males had

the narrow (~580 nm) diameter fiber sometimes supplemented with wider (~1000 nm) diameter fibers. The narrow and wide diameter fibers we observed in male *L. hesperus* draglines are likely minor ampullate and major ampullate silk fibers, respectively. Unlike in female *L. hesperus,* it appears that male *L. hesperus* use minor ampullate silk in draglines when walking, and only supplement the minor ampullate silk with major ampullate silk occasionally when falling. Greater use of minor ampullate silk in male *Latrodectus* draglines would be consistent with the high expression level of *MiSp* genes in *Latrodectus* males (Figure 1.4 B and C).

Minor ampullate silk could also be used by male spiders in the construction of sperm webs. Upon maturation, males make a small web in which they deposit sperm for transfer to their copulatory organs, the pedipalps (Foelix, 2011). Sperm-web silk is of unknown glandular origin. In some species, it is thought that the sperm web is made of epiandrous silk from epiandrous glands, which are glands associated with ventral abdominal spigots that are not on the spinnerets (Knoflach, 1998). Other silk types may also be involved in sperm web construction. It has been shown that minor ampullate silk does not undergo supercontraction (breakage of hydrogen bonds that help protein alignment in silk fibers) when exposed to water (Parkhe et al., 1997). Sperm deposits contain sperm and seminal fluid (acidic microsubstances and proteinaceous substances in a basic aqueous matrix) that could alter the behavior of silk fibers with regard to supercontraction (Michalik and Uhl, 2005). Because minor ampullate silk retains its structure in contact with moisture, minor ampullate silk is a candidate for use in the sperm web. If minor ampullate silk fibers were part of the sperm web, we would expect

MiSp genes to have higher expression levels than other silk genes when males prepare to construct sperm webs. This could be consistent with the high expression levels of *MiSp* in our male cob-web weavers (Figure 1.4).

Minor ampullate silk has also been implicated in prey immobilization by theridiids (LaMattina et al., 2008). Thus, expression of minor ampullate silk genes could indicate that mature male *Latrodectus* are catching prey. The few reports of mature males constructing prey-capture webs have involved orb-web building spiders (e.g. *Uloborus* sp., *Cyclosa* sp.; Eberhard, 1977). During orb-web construction, these spiders use minor ampullate silk in the temporary spiral, which serves as a guide for the sticky prey-capture spiral. Theridiid spiders do not make an orb web but instead construct a three-dimensional cob-web. In cob-webs, prey are arrested by the gum-footed lines, which extend from the supporting structure of the cob web to the ground. Gum-footed lines consist of two types of threads: one thicker than the other (Benjamin and Zschokke, 2002; Blackledge et al., 2005). The glandular origin of these two fiber types is not certain; some researchers have suggested that gum-footed threads are made exclusively from major ampullate silk fibers (Blackledge et al., 2005), while other researchers have suggested that the thin threads may be composed of minor ampullate silk (Benjamin and Zschokke, 2002). Given that current knowledge is based solely on the silks spun by female spiders, it is possible that *Latrodectus* males use minor ampullate silk for tasks thought to be dominated by major ampullate silk. It is also possible that, like araneid spiders, theridiid males use minor ampullate silk as bridging lines. If so, this could explain the higher expression of *MiSp* genes in *Latrodectus* males compared to other silk

genes. Further studies of key developmental stages of male and female spiders are needed to examine the temporal expression pattern of silk genes.

Venom Gene Expression in Male Spiders

If males are catching prey, it is reasonable to expect that they may also be producing venom. To explore this possibility, we screened our male assemblies for venom gene transcripts. Spider venom is a complex mixture of proteins, peptides, and low-molecular-weight components produced in paired cheliceral venom glands. Although mixtures vary across species, venoms generally contain molecules classified as neurotoxins, cytotoxins, and antimicrobial factors (Schulz, 1997; Tedford et al., 2004; Escoubas, 2006; Binford et al., 2009). Major components of most spider venoms are short, 50–80 aa long peptides with multiple cysteine residues that form disulfide bonds (Kuzmenkov et al., 2013). These short peptides are known as inhibitory cysteine knot (ICK) toxins because they form small "knots" that typically act to block or modify neuronal ion channels (Kuzmenkov et al., 2013). In addition to ICK peptides, theridiid venom contains neurotoxins known as latrotoxins (Kiyatkin et al., 1990; Garb and Hayashi, 2013), cysteine-rich secretory proteins (CRISPs) (Fry et al., 2009), and low molecular weight proteins (LMWPs) also known as latrodectins (Pescatori et al., 1995; Volkova et al., 1995). We identified contigs for several known theridiid venom components in the male transcriptome of each spider species, including latrotoxins, latrodectins, ICK peptides, and CRISPs (Garb and Hayashi, 2013; Haney et al., 2014,

2016). The identification of venom gene transcripts in male spiders supports the contention that males are capable of capturing prey.

Although males are expressing venom genes, they are expressing only a subset of the venom components expressed by conspecific females (Table S1.5). This simpler venom composition could indicate differences in feeding behavior, venom use, or both. Differences in feeding behavior between males and females have been observed in wolf spiders (*Pardosa milvina*, *Rabidosa punctulata, Rabidosa rabida, Schizocosa ocreata, Schizocosa rovneri*), where females attack and consume more prey than males (Walker and Rypstra, 2001, 2002). Furthermore, sex-linked variability in venom has been proposed for brown recluse spiders (*Loxosceles intermedia*), where females produce larger volumes of venom with higher potency than males (de Oliveira et al., 1999). Similarly, theridiid females are expected to need larger quantities and more potent venom than males because females have to catch more prey to support their larger body sizes (body length: ♀ 6–15 mm and ♂ 2–8 mm), longer lifespans, and egg production (Kaston, 1970).

It should be mentioned that venoms may have other functions besides prey capture, such as defense against predation (Olivera, 1999; Sher and Zlotkin, 2009; Kutsukake et al., 2004). We found that *L. geometricus* and *L. hesperus* males express latroinsectotoxin genes, which are known to selectively affect insect neurons (Magazanik et al., 1992). While latroinsectotoxin could aid in prey capture, it could also be effective for avoidance of insect predators. Further studies such as observational experiments are

needed to broadly examine venom usage of male spiders for prey capture and defense against predators.

Conclusions

We found that for all three focal species, males expressed silk genes associated with aciniform, major ampullate, minor ampullate, and pyriform silk glands, consistent with morphological studies that showed the presence of silk spigots for these gland types in mature males (Figs. 1.1, 1.2 and 1.4). We also detected very low expression of genes associated with tubuliform and aggregate glands in males. In males of all three species, we detected *TuSp1*; in *L. geometricus* and *L. hesperus* males only, we detected expression of *AgSp1* and *AgSF2* (*aggregate spidroin 1* and *aggregate silk factor 2*); and in *L. hesperus* males only, we detected *SCP-1* (*spider coating peptide 1*; Figs. 1.2 and 1.3; Table S1.4). Expression of tubuliform and aggregate gland-associated silk genes was unexpected because males lack functional spigots for these gland types. However, these unexpected genes were expressed at much lower levels than the expected genes (Figure 1.3). In contrast, female conspecifics expressed very high levels of the silk genes associated with egg case construction (*TuSp1, ECP-1, ECP-2*) and aggregate glue secretions (*AgSp1, AgSF1, AgSF2, SCP-1, SCP-2*; Table S1.4). Because mature theridiid males wander in search of females, we anticipated that genes for major ampullate spidroins would be the highest expressed silk genes. This prediction was upheld in *S. grossa,* but *Latrodectus* males had a higher expression of minor ampullate spidroin genes (Figure 1.4). High minor ampullate silk gene expression suggests that minor ampullate

silk may be widely used by *Latrodectus* males for purposes such as sperm webs, draglines, bridging lines, and prey capture. Expression of venom genes by males further supports our hypothesis that males are producing proteins to support a variety of tasks that could include prey capture and predator evasion (Table S1.6). In summary, our analyses of transcriptomes have provided new insights into the biology of male spiders. Future studies are needed to examine the temporal and spatial localization of silk gene expression as males and females become sexually mature.

Figures

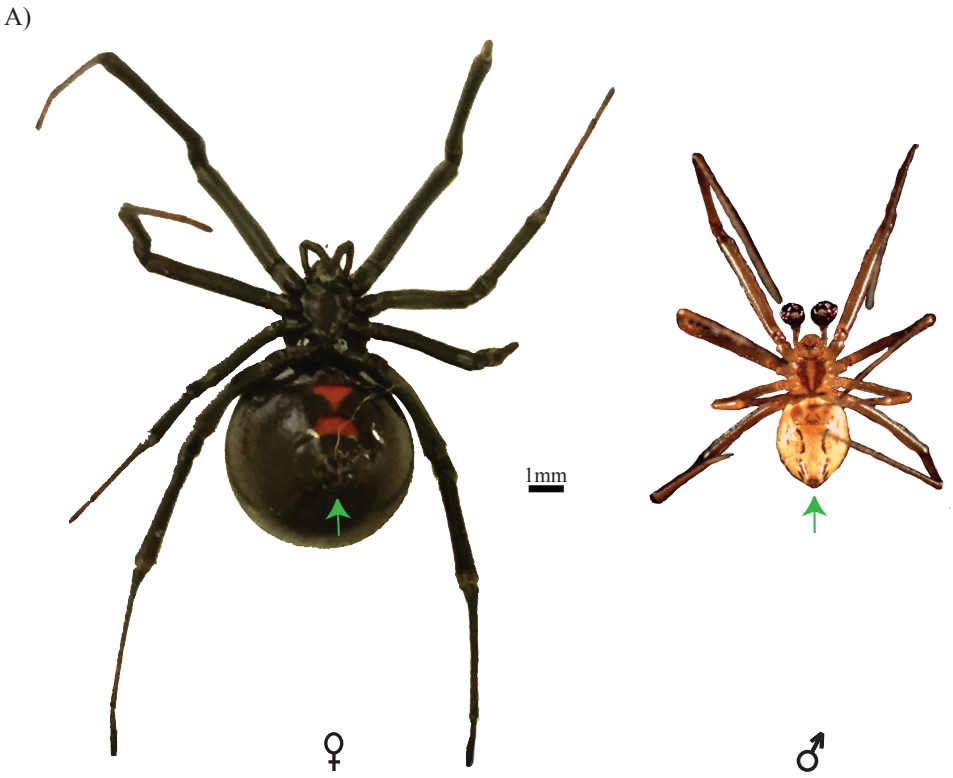


Figure 1.1. Spinnerets and spigots of female and male *Latrodectus hesperus*. (A) *L. hesperus* female (left) and male (right). Ventral view, anterior to the top. Spinnerets indicated by green arrows. Scale bar = 1 mm. (B) Spigots associated with the spinnerets in female (left) and male (right) spiders after Moon and An (2006). Diagrams not to scale. Vertical line indicates sagittal midline, only one side of the spinneret region is shown. Anterior lateral spinnerets (ALS) in beige, posterior median spinnerets (PMS) in light blue, and posterior lateral spinnerets (PLS) in pale pink. Spigots abbreviated as follows: Ac, aciniform gland spigots (orange); Ag, aggregate gland spigots (grey); Fl, flagelliform gland spigots (grey); Py – pyriform gland spigots (blue); Td, non-functional remnants of aggregate and flagelliform gland spigots (light brown). "Am" and "Cy" of Moon and An (2006) are shown here as: Am = MAmp, major ampullate gland spigot (pink) and mAmp, minor ampullate gland spigot (red); Cy = Tub, tubuliform gland spigots (purple).

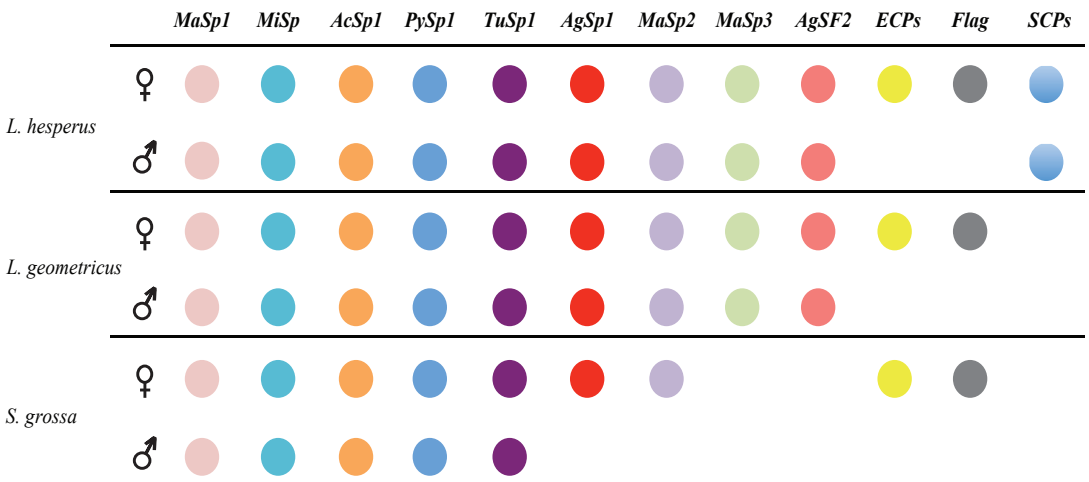


Figure 1.2. Silk gene expression in females and males of three focal species. Detection of gene expression with RNA-Seq indicated by filled circles under each gene name. Silk genes abbreviated as *MaSp1* (*major ampullate spidroin 1*), *MiSp* (*minor ampullate spidroin*), *AcSp1* (*aciniform spidroin 1*), *PySp1* (*pyriform spidroin 1*), *TuSp1* (*tubuliform spidroin 1*), *MaSp2* (*major ampullate spidroin 2*), *AgSp1* (*aggregate spidroin 1*), *MaSp3* (*major ampullate spidroin 3*; shown in Chaw et al., 2015 as *MaSp'*), *AgSF2* (*aggregate silk factor 2*), *ECPs* (*egg case protein 1* and *2*), *Flag* (*flagelliform spidroin*), and *SCPs* (*spider coating peptide 1* and *2*).

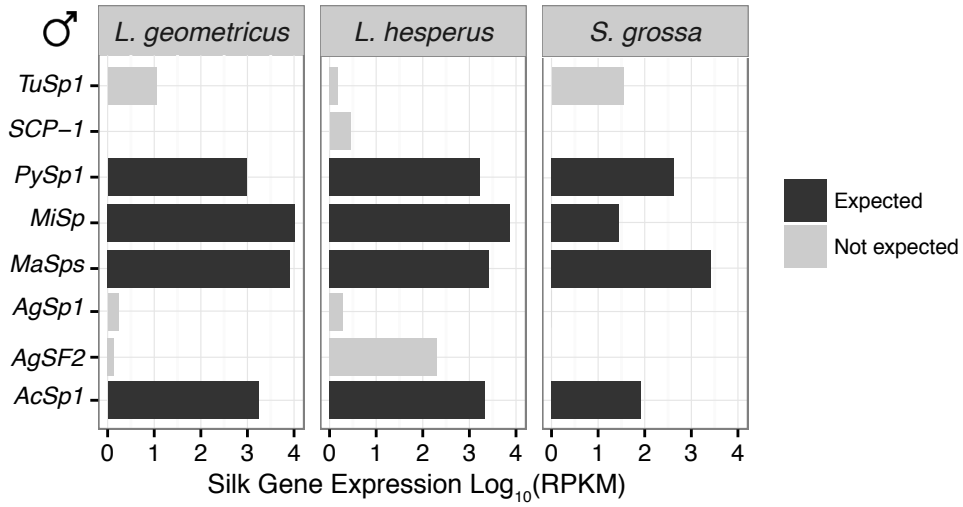


Figure 1.3. Silk gene expression in male spiders from three focal species. Silk genes expected to be expressed (black): *MaSps* (sum of *MaSp1*, *MaSp2*, *MaSp3*), *MiSp*, *AcSp1*, and *PySp1*. Silk genes not expected (grey): *AgSp1*, *AgSF2*, *SCP-1*, and *TuSp1*. Silk genes abbreviated as in Figure 1.2. Average expression shown as $\log_{10}$ of reads per kilobase of transcript per million mapped reads (RPKM) of male libraries mapped to male transcriptomes.

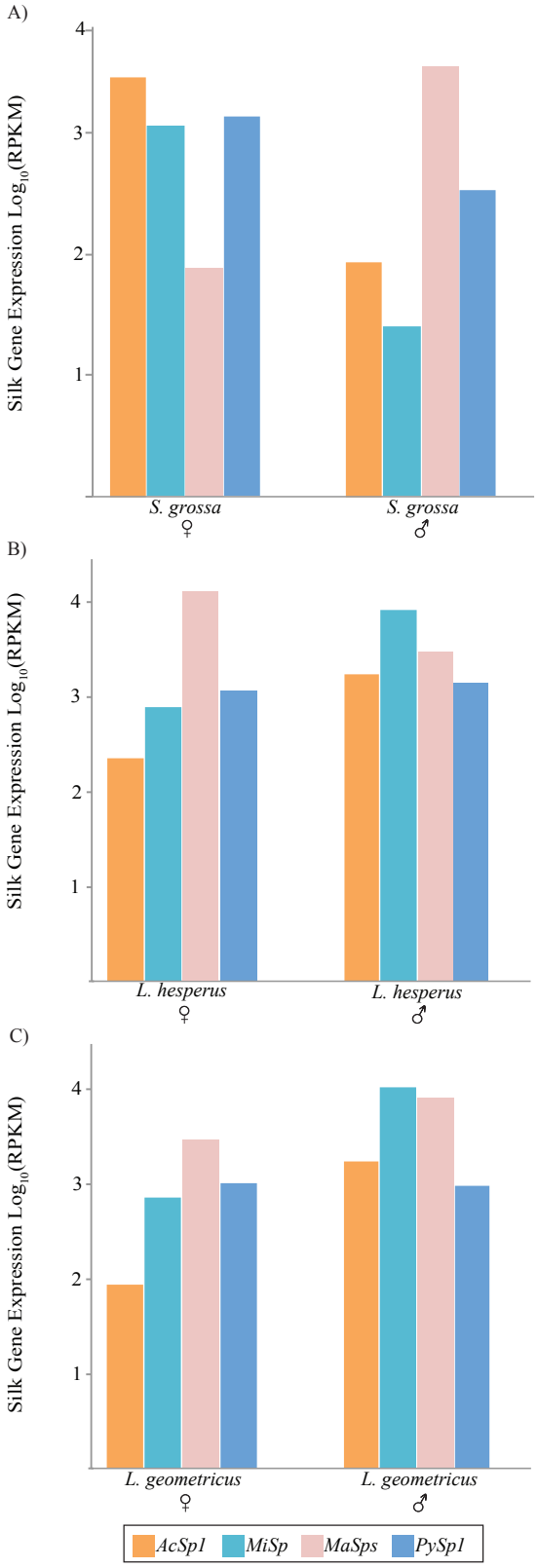

A)
Silk Gene Expression Log10(RPKM)
4
3
2
1
S. grossa
♀
S. grossa
♂
B)
Silk Gene Expression Log10(RPKM)
4
3
2
1
L. hesperus
♀
L. hesperus
♂
C)
Silk Gene Expression Log10(RPKM)
4
3
2
1
L. geometricus
♀
L. geometricus
♂
AcSp1
MiSp
MaSps
PySp1

Figure 1.4. Silk gene expression in female (left) and male (right) silk glands of the focal species (A) *Steatoda grossa*, (B) *Latrodectus hesperus*, and (C) *Latrodectus geometricus*. *AcSp1*, *MiSp*, *MaSps* (sum of *MaSp1*, *MaSp2*, and *MaSp3*), and *PySp1* are shown. Average expression of male reads mapped to species-specific male transcriptomes and total silk gland female reads mapped to species-specific global transcriptomes shown in $\log_{10}$ reads per kilobase of transcript per million mapped reads (RPKM). Silk genes abbreviated as in Figure 1.2. Total RPKM of silk genes shown in Figure 1.2 for *L. hesperus*: ♀ 36,251, ♂ 21,432; *L. geometricus*: ♀ 14,250, ♂ 14,146; and *S. grossa*: ♀ 11,832, ♂ 3,259.

References

Andrade, M.C.B., 2003. Risky mate search and male self-sacrifice in redback spiders. Behav. Ecol. 14, 531–538.

Atkinson, R., 1981. Comparisons of the neurotoxic activity of the venom of several species of funnel web spiders (*atrax*). Aust. J. Exp. Biol. Med. 59, 307–316.

Ayoub, N.A., Hayashi, C.Y., 2008. Multiple recombining loci encode MaSp1, the primary constituent of dragline silk, in widow spiders (*Latrodectus*: Theridiidae). Mol. Biol. Evol. 25, 277–286.

Ayoub, N.A., Garb, J.E., Tinghitella, R.M., Collin, M.A., Hayashi, C.Y., 2007. Blueprint for a high-performance biomaterial: full-length spider dragline silk genes. PloS One 2, e514.

Ayoub, N.A., Garb, J.E., Kuelbs, A., Hayashi, C.Y., 2013. Ancient properties of spider silks revealed by the complete gene sequence of the prey-wrapping silk protein (AcSp1). Mol. Biol. Evol. 30, 589–601.

Benjamin, S.P., Zschokke, S., 2002. Untangling the tangle-web: web construction behavior of the comb-footed spider *Steatoda triangulosa* and comments on phylogenetic implications (Araneae: Theridiidae). J. Insect Behav. 15, 791–809.

Binford, G.J., 2001. Differences in venom composition between orb-weaving and wandering Hawaiian *Tetragnatha* (Araneae). Biol. J. Linn. Soc. 74, 581–595.

Binford, G.J., Bodner, M.R., Cordes, M.H., Baldwin, K.L., Rynerson, M.R., Burns, S.N., Zobel-Thropp, P.A., 2009. Molecular evolution, functional variation, and proposed nomenclature of the gene family that includes sphingomyelinase D in sicariid spider venoms. Mol. Biol. Evol. 26, 547–566.

Binford, G.J., Gillespie, R.G., Maddison, W.P., 2016. Sexual dimorphism in venom chemistry in *Tetragnatha* spiders is not easily explained by adult niche differences. Toxicon 114, 45–52.

Blackledge, T.A., Hayashi, C.Y., 2006. Silken toolkits: biomechanics of silk fibers spun by the orb web spider *Argiope argentata* (Fabricius 1775). J. Exp. Biol. 209, 2452–2461.

Blackledge, T.A., Summers, A.P., Hayashi, C.Y., 2005. Gumfooted lines in black widow cobwebs and the mechanical properties of spider capture silk. Zoology 108, 41–46.

Casem, M.L., Collin, M.A., Ayoub, N.A., Hayashi, C.Y., 2010. Silk gene transcripts in the developing tubuliform glands of the Western black widow, *Latrodectus hesperus*. J. Arachnol. 38, 99–103.

Chaw, R.C., Zhao, Y., Wei, J., Ayoub, N.A., Allen, R., Atrushi, K., Hayashi, C.Y., 2014. Intragenic homogenization and multiple copies of prey-wrapping silk genes in *Argiope* garden spiders. BMC Evol. Biol. 14, 31.

Chaw, R.C., Correa-Garhwal, S.M., Clarke, T.H., Ayoub, N.A., Hayashi, C.Y., 2015. Proteomic evidence for components of spider silk synthesis from black widow silk glands and fibers. J. Proteome Res. 14, 4223–4231.

Chaw, R.C., Arensburger, P., Clarke, T.H., Ayoub, N.A., Hayashi, C.Y., 2016. Candidate egg case silk genes for the spider *Argiope argentata* from differential gene expression analyses. Insect Mol. Biol. 25, 757–768.

Chen, G., Liu, X., Zhang, Y., Lin, S., Yang, Z., Johansson, J., Rising, A., Meng, Q., 2012. Full-length minor ampullate spidroin gene sequence. PLoS ONE 7, e52293.

Choresh, O., Bayarmagnai, B., Lewis, R.V., 2009. Spider web glue: two proteins expressed from opposite strands of the same DNA sequence. Biomacromolecules 10, 2852–2856.

Clarke, T.H., Garb, J.E., Hayashi, C.Y., Haney, R.A., Lancaster, A.K., Corbett, S., Ayoub, N.A., 2014. Multi-tissue transcriptomics of the black widow spider reveals expansions, co-options, and functional processes of the silk gland gene toolkit. BMC Genom. 15, 365.

Clarke, T.H., Garb, J.E., Hayashi, C.Y., Arensburger, P., Ayoub, N.A., 2015. Spider transcriptomes identify ancient large-scale gene duplication event potentially important in silk gland evolution. Genome Biol. Evol. 7, 1856–1870.

Coddington, J.A., 1989. Spinneret silk spigot morphology: evidence for the monophyly of orbweaving spiders, Cyrtophorinae (Araneidae), and the group Theridiidae plus Nesticidae. J. Arachnol. 17, 71–95.

Correa-Garhwal, S.M., Garb, J.E., 2014. Diverse formulas for spider dragline fibers demonstrated by molecular and mechanical characterization of spitting spider silk. Biomacromolecules 15, 4598–4605.

de Oliveira, K.C., Gonçalves de Andrade, R.M., Giusti, A.L., da Silva, W.D., Tambourgi, D.V., 1999. Sex-linked variation of *Loxosceles intermedia* spider venoms. Toxicon 37, 217–221.

Eberhard, W.G., 1977. The webs of newly emerged *Uloborus diversus* and of a male *Uloborus* sp. (Araneae: Uloboridae). J. Arachnol. 4, 201–206.

Escoubas, P., 2006. Molecular diversification in spider venoms: a web of combinatorial peptide libraries. Mol. Divers. 10, 545–554.

Foelix, R., 2011. Biology of Spiders. Oxford University Press, Oxford.

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., de la Vega, R.C.R., 2009. The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. Annu. Rev. Genomics Hum. Genet. 10, 483–511.

Garb, J.E., Hayashi, C.Y., 2005. Modular evolution of egg case silk genes across orb-weaving spider superfamilies. Proc. Natl. Acad. Sci. U. S. A. 102, 11379–11384.

Garb, J.E., Hayashi, C.Y., 2013. Molecular evolution of α-latrotoxin, the exceptionally potent vertebrate neurotoxin in black widow spider venom. Mol. Biol. Evol. 30, 999–1014.

Garb, J.E., Ayoub, N.A., Hayashi, C.Y., 2010. Untangling spider silk evolution with spidroin terminal domains. BMC Evol. Biol. 10, 243.

Gosline, J.M., Denny, M.W., DeMont, M.E., 1984. Spider silk as rubber. Nature 309, 551–552.

Gosline, J.M., DeMont, M.E., Denny, M.W., 1986. The structure and properties of spider silk. Endeavour 10, 37–43.

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., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652.

Guerette, P.A., Ginzinger, D.G., Weber, B.H.F., Gosline, J.M., 1996. Silk properties determined by gland-specific expression of a spider fibroin gene family. Science 272, 112–115.

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

Haney, R.A., Clarke, T.H., Gadgil, R., Fitzpatrick, R., Hayashi, C.Y., Ayoub, N.A., Garb, J.E., 2016. Effects of gene duplication, positive selection, and shifts in gene expression on the evolution of the venom gland transcriptome in widow spiders. Genome Biol. Evol. 8, 228–242.

Hinman, M.B., Lewis, R.V., 1992. Isolation of a clone encoding a second dragline silk fibroin. *Nephila clavipes* dragline silk is a two-protein fiber. J. Biol. Chem. 267, 19320–19324.

Hormiga, G., Scharff, N., Coddington, J.A., 2000. The phylogenetic basis of sexual size dimorphism in orb-weaving spiders (Araneae, Orbiculariae). Syst. Biol. 49, 435–462.

Hsia, Y., Gnesa, E., Jeffery, F., Tang, S., Vierra, C. 2011. Spider silk composites and applications. In: Cuppoletti J. (Ed.), Metal, Ceramic and Polymeric Composites for Various Uses, vol. 2. InTech, Rijeka, pp. 303–324.

Hu, X., Kohler, K., Falick, A.M., Moore, A.M., Jones, P.R., Sparkman, O.D., Vierra, C., 2005. Egg case protein-1. A new class of silk proteins with fibroin-like properties from the spider *Latrodectus hesperus*. J. Biol. Chem. 280, 21220–21230.

Hu, X., Kohler, K., Falick, A.M., Moore, A.M., Jones, P.R., Vierra, C., 2006. Spider egg case core fibers: trimeric complexes assembled from TuSp1, ECP-1, and ECP-2. Biochemistry (Mosc.) 45, 3506–3516.

Hu, X., Yuan, J., Wang, X., Vasanthavada, K., Falick, A.M., Jones, P.R., La Mattina, C., Vierra, C.A., 2007. Analysis of aqueous glue coating proteins on the silk fibers of the cob weaver, *Latrodectus hesperus*. Biochemistry (Mosc.) 46, 3294–3303.

Huang, X., Madan, A., 1999. CAP3: a DNA sequence assembly program. Genome Res. 9, 868–877.

Kaston, B.J., 1970. Comparative biology of American black widow spiders. Trans. San Diego Soc. Nat. Hist. 16, 33–82.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649.

Kiyatkin, N.I., Dulubova, I.E., Chekhovskaya, I.A., Grishin, E.V., 1990. Cloning and structure of cDNA encoding α-latrotoxin from black widow spider venom. FEBS Lett. 270, 127–131.

Knoflach, B., 1998. Mating in *Theridion varians* Hahn and related species (Araneae: Theridiidae). J. Nat. Hist. 32, 545–604.

Kovoor, J., Peters, H.M., 1988. The spinning apparatus of *Polenecia producta* (Araneae, Uloboridae): structure and histochemistry. Zoomorphology 108, 47–59.

Kutsukake, M., Shibao, H., Nikoh, N., Morioka, M., Tamura, T., Hoshino, T., Ohgiya, S., Fukatsu, T., 2004. Venomous protease of aphid soldier for colony defense. Proc. Natl. Acad. Sci. U. S. A. 101, 11338–11343.

Kuzmenkov, A.I., Fedorova, I.M., Vassilevski, A.A., Grishin, E.V., 2013. Cysteine-rich toxins from *Lachesana tarabaevi* spider venom with amphiphilic C-terminal segments. Biochim. Biophys. Acta 1828, 724–731.

LaMattina, C., Reza, R., Hu, X., Falick, A.M., Vasanthavada, K., McNary, S., Yee, R., Vierra, C.A., 2008. Spider minor ampullate silk proteins are constituents of prey wrapping silk in the cob weaver *Latrodectus hesperus*. Biochemistry (Mosc.) 47, 4692–4700.

Lane, A.K., Hayashi, C.Y., Whitworth, G.B., Ayoub, N.A., 2013. Complex gene expression in the dragline silk producing glands of the Western black widow (*Latrodectus hesperus*). BMC Genomics 14, 846.

Langmead, B., 2010. Aligning short sequencing reads with Bowtie. Curr. Protoc. Bioinform. 32, 11.7.1–11.7.14.

Lawrence, B.A., Vierra, C.A., Moore, A.M., 2004. Molecular and mechanical properties of major ampullate silk of the black widow spider, *Latrodectus hesperus*. Biomacromolecules 5, 689–695.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079.

Magazanik, L.G., Fedorova, I.M., Kovalevskaya, G.I., Pashkov, V.N., Bulgakov, O.V., Grishin, E.V., 1992. Selective presynaptic insectotoxin (α-latroinsectotoxin) isolated from black widow spider venom. Neuroscience 46, 181–188.

Michalik, P., Uhl, G., 2005. The male genital system of the cellar spider *Pholcus phalangioides* (Fuesslin, 1775) (Pholcidae, Araneae): development of spermatozoa and seminal secretion. Front. Zool. 2, 12.

Moon, M.-J., 2008. Microstructure of the silk apparatus in the coelotine spider *Paracoelotes spinivulva* (Araneae: Amaurobiidae). Entomol. Res. 38, 149–156.

Moon, M.-J., An, J.-S., 2006. Microstructure of the silk apparatus of the comb-footed spider, *Achaearanea tepidariorum* (Araneae: Theridiidae). Entomol. Res. 36, 56–63.

Moore, C.W., 1977. The life cycle, habitat and variation in selected web parameters in the spider, *Nephila clavipes* Koch (Araneidae). Am. Midl. Nat. 98, 95–108.

Olivera, B.M., 1999. *Conus* venom peptides: correlating chemistry and behavior. J. Comp. Physiol. A 185, 353–359.

Park, J.-G., Moon, M.-J., 2002. Fine structural analysis of the silk apparatus in the funnel-web spider, *Agelena limbata* (Araneae: Agelenidae). Entomol. Res. 32, 223–232.

Parkhe, A.D., Seeley, S.K., Gardner, K., Thompson, L., Lewis, R.V., 1997. Structural studies of spider silk proteins in the fiber. J. Mol. Recognit. 10, 1–6.

Parra, G., Bradnam, K., Korf, I., 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics 23, 1061–1067.

Perry, D.J., Bittencourt, D., Siltberg-Liberles, J., Rech, E.L., Lewis, R.V., 2010. Piriform spider silk sequences reveal unique repetitive elements. Biomacromolecules 11, 3000–3006.

Pescatori, M., Bradbury, A., Bouet, F., Gargano, N., Mastrogiacomo, A., Grasso, A., 1995. The cloning of a cDNA encoding a protein (latrodectin) which co-purifies with the α-latrotoxin from the black widow spider *Latrodectus tredecimguttatus* (Theridiidae). Eur. J. Biochem. 230, 322–328.

Peters, H.M., 1990. On the structure and glandular origin of bridging lines used by spiders for moving to distant places. Act. Zool. Fenn. 190, 309–314.

Peters, H.M., 1992. On the spinning apparatus and the structure of the capture threads of *Deinopis subrufus* (Araneae, Deinopidae). Zoomorphology 112, 27–37. doi:10.1007/BF01632992

R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Sanggaard, K.W., Bechsgaard, J.S., Fang, X., Duan, J., Dyrlund, T.F., Gupta, V., Jiang, X., Cheng, L., Fan, D., Feng, Y., Han, L., Huang, Z., Wu, Z., Liao, L., Settepani, V., Thøgersen, I.B., Vanthournout, B., Wang, T., Zhu, Y., Funch, P., Enghild, J.J., Schauser, L., Andersen, S.U., Villesen, P., Schierup, M.H., Bilde, T., Wang, J., 2014. Spider genomes provide insight into composition and evolution of venom and silk. Nat. Commun. 5, 3765.

Schulz, S., 1997. The chemistry of spider toxins and spider silk. Angew. Chem. Int. Ed. Engl. 36, 314–326.

Schütz, D., Taborsky, M., 2003. Adaptations to an aquatic life may be responsible for the reversed sexual size dimorphism in the water spider, *Argyroneta aquatica*. Evol. Ecol. Res. 5, 105–117.

Schütz, D., Taborsky, M., 2005. Mate choice and sexual conflict in the size dimorphic water spider *Argyroneta aquatica* (Araneae, Argyronetidae). J. Arachnol. 33, 767–775.

Sher, D., Zlotkin, E., 2009. A hydra with many heads: protein and polypeptide toxins from hydra and their biological roles. Toxicon 54, 1148–1161.

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212.

Starrett, J., Garb, J.E., Kuelbs, A., Azubuike, U.O., Hayashi, C.Y., 2012. Early events in the evolution of spider silk genes. PloS One 7, e38084.

Suter, R.B., 1991. Ballooning in spiders: results of wind tunnel experiments. Ethol. Ecol. Evol. 3, 13–25.

Tedford, H.W., Sollod, B.L., Maggio, F., King, G.F., 2004. Australian funnel-web spiders: master insecticide chemists. Toxicon 43, 601–618.

Tian, M., Lewis, R.V., 2005. Molecular characterization and evolutionary study of spider tubuliform (eggcase) silk protein. Biochemistry (Mosc.) 44, 8006–8012.

Tillinghast, E.K., Townley, M.A., Wight, T.N., Uhlenbruck, G., Janssen, E., 1992. The adhesive glycoprotein of the orb web of *Argiope aurantia* (Araneae, Araneidae). In: Viney, C., Case, S.T., Waite, J.H. (Eds.), Materials Research Society Symposium Proceedings, vol. 292. MRS, Pittsburgh, pp. 9–23.

Vasanthavada, K., Hu, X., Tuton-Blasingame, T., Hsia, Y., Sampath, S., Pacheco, R., Freeark, J., Falick, A.M., Tang, S., Fong, J., Kohler, K., La Mattina-Hawkins, C., Vierra, V., 2012. Spider glue proteins have distinct architectures compared with traditional spidroin family members. J. Biol. Chem. 287, 35986–35999.

Vienneau-Hathaway, J.M., Brassfield, E.R., Lane, A.K., Collin, M. A., Correa-Garhwal, S.M., Clarke, T.H., Schwager , E.E., Garb, J.E., Hayashi, C.Y., Ayoub, N.A. Duplication and concerted evolution of MiSp-encoding genes underlie the material properties of minor ampullate silks of cobweb weaving spiders. BMC Evol. Biol., 17, 78.

Volkova, T.M., Pluzhnikov, K.A., Woll, P.G., Grishin, E.V., 1995. Low molecular weight components from black widow spider venom. Toxicon 33, 483–489.

Vollrath, F., 1992. Spider webs and silks. Sci. Am. 266, 70–76.

Vollrath, F., 1998. Dwarf males. Trends Ecol. Evol. 13, 159–163.

Vollrath, F., Parker, G.A., 1992. Sexual dimorphism and distorted sex ratios in spiders. Nature 360, 156–159.

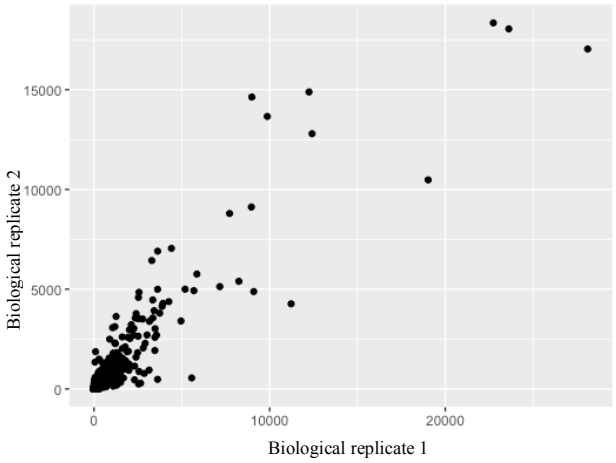
Walker, S.E., Rypstra, A.L., 2001. Sexual dimorphism in functional response and trophic morphology in *Rabidosa rabida* (Araneae: Lycosidae). Am. Midl. Nat. 146, 161–170.

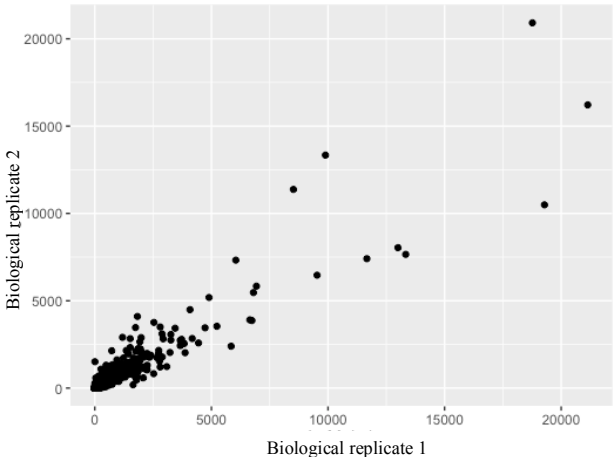
Walker, S.E., Rypstra, A.L., 2002. Sexual dimorphism in trophic morphology and feeding behavior of wolf spiders (Araneae: Lycosidae) as a result of differences in reproductive roles. Can. J. Zool. 80, 679–688.

Work, R.W., 1981. Web components associated with the major ampullate silk fibers of orb-web-building spiders. Trans. Am. Microsc. Soc. 100, 1–20.

Zhao, A.-C., Zhao, T.-F., Nakagaki, K., Zhang, Y.-S., Sima, Y.-H., Miao, Y.-G., Shiomi, K., Kajiura, Z., Nagata, Y., Takadera, M., Nakagaki, M., 2006. Novel molecular and mechanical properties of egg case silk from wasp spider, *Argiope bruennichi*. Biochemistry (Mosc.) 45, 3348–3356.

Supplementary Figures





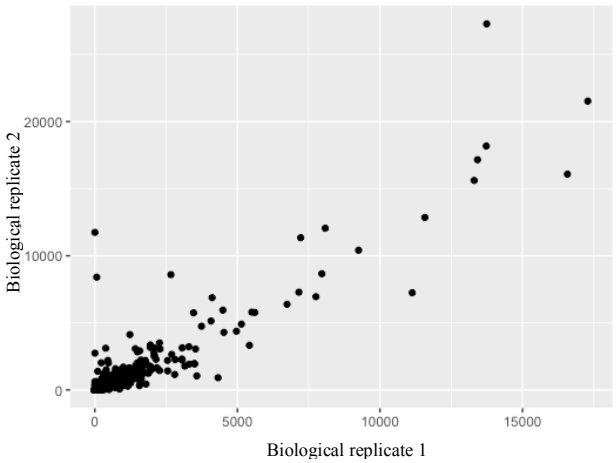


Figure S1.1. Correlation plots of expression between male biological replicates for (A) *Latrodectus hesperus*, (B) *Latrodectus geometricus* (C) *Steatoda grossa*. In the scatter plots, each point corresponds to reads per kilobase of transcript per million mapped reads (RPKM). Pearson correlation is shown.

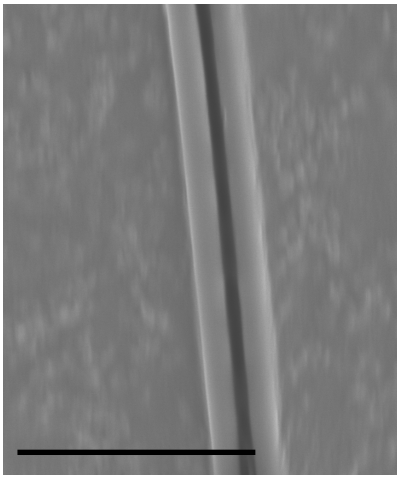


Figure S1.2. Scanning electron micrograph of *Latrodectus hesperus* male dragline. Two fibers are shown; draglines are typically two fibers, one emerging from each spigot. Fiber diameter is 0.58 μm. Scale bar = 5 μm.

Table S1.1. Summary of male and global de novo transcriptome assemblies.

	No. of raw reads	No. of cleaned reads	No. of Trinity contigs	Total transcriptome length (bp)	N50 (bp)
***Latrodectus hesperus* global**	728,100,105	436,656,078	191,314	131,935,700	1,324
***Latrodectus hesperus* male**	69,398,954	41,529,868	32,316	27,390,428	1,430
***Latrodectus geometricus* global**	644,421,794	436,552,222	152,807	100,426,900	1,080
***Latrodectus geometricus* male**	53,505,807	31,325,972	31,913	28,602,801	1,357
***Steatoda grossa* global**	495,063,992	282,229,204	161,843	117,928,200	1,482
***Steatoda grossa* male**	61,171,521	44,846,134	40,382	30,427,603	1,281

Table S1.2. Sequences used in Blastx searches to identify spider silk sequences.

Protein name	NCBI Accession ID	Species
Aciniform Spidroin-Putative	AZAQ01111412.1	*Stegodyphus mimosarum*
Aciniform Spidroin 1	AHK09813.1	*Argiope argentata*
Aciniform Spidroin 1	ADM35668.1	*Araneus ventricosus*
Aciniform Spidroin 1	ADM35669.1	*Argiope amoena*
Aciniform Spidroin 1	AAR83925.1	*Argiope trifasciata*
Aciniform Spidroin 1	AFX83566.1	*Latrodectus geometricus*
Aciniform Spidroin 1	AFX83567.1	*Latrodectus geometricus*
Aciniform Spidroin 1	AFX83568.1	*Latrodectus geometricus*
Aciniform Spidroin 1	AFX83557.1	*Latrodectus hesperus*
Aciniform Spidroin 1	AFX83558.1	*Latrodectus hesperus*
Aciniform Spidroin 1	AFX83559.1	*Latrodectus hesperus*
Aciniform Spidroin 1	AFX83560.1	*Latrodectus hesperus*
Aciniform Spidroin 1	AFX83561.1	*Latrodectus hesperus*
Aciniform Spidroin 1	AFX83562.1	*Latrodectus hesperus*
Aciniform Spidroin 1	AFX83563.1	*Latrodectus hesperus*
Aciniform Spidroin 1	AFX83565.1	*Latrodectus hesperus*
Aciniform Spidroin 1	ADG57593.1	*Parawixia bistriata*
Aciniform Spidroin 1	ABD61598.1	*Uloborus diversus*
Aciniform Spidroin 1-Like	ABW24499.1	*Latrodectus hesperus*
Aciniform Spidroin 1 C-Terminal Region Variant 1, Partial	AHK09802.1	*Argiope trifasciata*
Aciniform Spidroin 1 C-Terminal Region Variant 1, Partial	AHK09800.1	*Argiope aurantia*
Aciniform Spidroin 1 C-Terminal Region Variant 1, Partial	AHK09793.1	*Argiope argentata*
Aciniform Spidroin 1 C-Terminal Region Variant 10, Partial	AHK09811.1	*Argiope trifasciata*
Aciniform Spidroin 1 C-Terminal Region Variant 11, Partial	AHK09812.1	*Argiope trifasciata*
Aciniform Spidroin 1 C-Terminal Region Variant 2, Partial	AHK09803.1	*Argiope trifasciata*
Aciniform Spidroin 1 C-Terminal Region Variant 2, Partial	AHK09801.1	*Argiope aurantia*
Aciniform Spidroin 1 C-Terminal Region Variant 3, Partial	AHK09804.1	*Argiope trifasciata*
Aciniform Spidroin 1 C-Terminal Region Variant 3, Partial	AHK09795.1	*Argiope argentata*
Aciniform Spidroin 1 C-Terminal Region Variant 4, Partial	AHK09805.1	*Argiope trifasciata*
Aciniform Spidroin 1 C-Terminal Region Variant 4, Partial	AHK09796.1	*Argiope argentata*
Aciniform Spidroin 1 C-Terminal	AHK09806.1	*Argiope trifasciata*

Region Variant 5, Partial		
Aciniform Spidroin 1 C-Terminal Region Variant 5, Partial	AHK09797.1	*Argiope argentata*
Aciniform Spidroin 1 C-Terminal Region Variant 6, Partial	AHK09807.1	*Argiope trifasciata*
Aciniform Spidroin 1 C-Terminal Region Variant 6, Partial	AHK09798.1	*Argiope argentata*
Aciniform Spidroin 1 C-Terminal Region Variant 7, Partial	AHK09808.1	*Argiope trifasciata*
Aciniform Spidroin 1 C-Terminal Region Variant 7, Partial	AHK09799.1	*Argiope argentata*
Aciniform Spidroin 1 C-Terminal Region Variant 8, Partial	AHK09809.1	*Argiope trifasciata*
Aciniform Spidroin 1 C-Terminal Region Variant 9, Partial	AHK09810.1	*Argiope trifasciata*
Aciniform Spidroin 1 N-Terminal Region Variant 1, Partial	AHK09776.1	*Argiope trifasciata*
Aciniform Spidroin 1 N-Terminal Region Variant 1, Partial	AHK09770.1	*Argiope aurantia*
Aciniform Spidroin 1 N-Terminal Region Variant 1, Partial	AHK09763.1	*Argiope argentata*
Aciniform Spidroin 1 N-Terminal Region Variant 10, Partial	AHK09785.1	*Argiope trifasciata*
Aciniform Spidroin 1 N-Terminal Region Variant 11, Partial	AHK09786.1	*Argiope trifasciata*
Aciniform Spidroin 1 N-Terminal Region Variant 12, Partial	AHK09787.1	*Argiope trifasciata*
Aciniform Spidroin 1 N-Terminal Region Variant 14, Partial	AHK09789.1	*Argiope trifasciata*
Aciniform Spidroin 1 N-Terminal Region Variant 2, Partial	AHK09777.1	*Argiope trifasciata*
Aciniform Spidroin 1 N-Terminal Region Variant 2, Partial	AHK09771.1	*Argiope aurantia*
Aciniform Spidroin 1 N-Terminal Region Variant 2, Partial	AHK09764.1	*Argiope argentata*
Aciniform Spidroin 1 N-Terminal Region Variant 3, Partial	AHK09778.1	*Argiope trifasciata*
Aciniform Spidroin 1 N-Terminal Region Variant 3, Partial	AHK09772.1	*Argiope aurantia*
Aciniform Spidroin 1 N-Terminal Region Variant 3, Partial	AHK09765.1	*Argiope argentata*
Aciniform Spidroin 1 N-Terminal Region Variant 31, Partial	AHK09788.1	*Argiope trifasciata*
Aciniform Spidroin 1 N-Terminal Region Variant 4, Partial	AHK09779.1	*Argiope trifasciata*
Aciniform Spidroin 1 N-Terminal Region Variant 4, Partial	AHK09773.1	*Argiope aurantia*
Aciniform Spidroin 1 N-Terminal Region Variant 4, Partial	AHK09766.1	*Argiope argentata*
Aciniform Spidroin 1 N-Terminal Region Variant 5, Partial	AHK09780.1	*Argiope trifasciata*

Aciniform Spidroin 1 N-Terminal Region Variant 5, Partial	AHK09774.1	*Argiope aurantia*
Aciniform Spidroin 1 N-Terminal Region Variant 5, Partial	AHK09767.1	*Argiope argentata*
Aciniform Spidroin 1 N-Terminal Region Variant 6, Partial	AHK09781.1	*Argiope trifasciata*
Aciniform Spidroin 1 N-Terminal Region Variant 6, Partial	AHK09775.1	*Argiope aurantia*
Aciniform Spidroin 1 N-Terminal Region Variant 6, Partial	AHK09768.1	*Argiope argentata*
Aciniform Spidroin 1 N-Terminal Region Variant 7, Partial	AHK09782.1	*Argiope trifasciata*
Aciniform Spidroin 1 N-Terminal Region Variant 7, Partial	AHK09769.1	*Argiope argentata*
Aciniform Spidroin 1 N-Terminal Region Variant 8, Partial	AHK09783.1	*Argiope trifasciata*
Aciniform Spidroin 1 N-Terminal Region Variant 9, Partial	AHK09784.1	*Argiope trifasciata*
Aciniform Spidroin 1, Partial	AAR83925.1	*Argiope trifasciata*
Aciniform Spidroin 1, Partial	AHK09794.1	*Argiope argentata*
Aciniform Spidroin 1, Partial	AHK09792.1	*Argiope trifasciata*
Aciniform Spidroin 1, Partial	AHK09791.1	*Argiope aurantia*
Aciniform Spidroin 1, Partial	AHK09790.1	*Argiope argentata*
Aggregate Gland Silk Factor 1	AFP57565.1	*Latrodectus hesperus*
Aggregate Gland Silk Factor 2	AFP57562.1	*Latrodectus hesperus*
Aggregate Gland Silk Factor 2-Like	AFP57559.1	*Latrodectus hesperus*
Aggregate Spidroin-1	AMK48674.1	*Araneus diadematus*
Aggregate Spidroin-1	AMK48675.1	*Argiope argentata*
Aggregate Spidroin-1	AMK48676.1	*Latrodectus hesperus*
Aggregate Spidroin-1	AMK48677.1	*Nephila clavipes*
Aggregate Spidroin-1	AMK48678.1	*Parasteatoda tepidariorum*
Aggregate Spidroin-1	AMK48679.1	*Steatoda grossa*
Aqueous Glue Droplet Peptide-1	ABO09798.1	*Latrodectus hesperus*
Aqueous Glue Droplet Peptide-2	ABO09799.1	*Latrodectus hesperus*
Dragline Silk Fibroin 1	P46802.1	*Araneus bicentenarius*
Dragline Silk Fibroin 1	AAV48932.1	*Cyrtophora moluccensis*
Dragline Silk Fibroin 1	AAV48934.1	*Cyrtophora moluccensis*
Dragline Silk Fibroin 1	AAV48935.1	*Cyrtophora moluccensis*
Dragline Silk Fibroin 1	AAV48951.1	*Cyrtophora moluccensis*
Dragline Silk Fibroin 1	AAV48940.1	*Macrothele holsti*
Dragline Silk Fibroin 1	P19837.3	*Nephila clavipes*
Dragline Silk Fibroin 1	AAV48920.1	*Nephila pilipes*
Dragline Silk Fibroin 1	AAV48921.1	*Nephila pilipes*

Dragline Silk Fibroin 1	AAV48922.1	*Nephila pilipes*
Dragline Silk Fibroin 1	AAV48925.1	*Nephila pilipes*
Dragline Silk Fibroin 1	AAV48926.1	*Nephila pilipes*
Dragline Silk Fibroin 1	AAV48927.1	*Nephila pilipes*
Dragline Silk Fibroin 1	AAV48928.1	*Nephila pilipes*
Dragline Silk Fibroin 1	AAV48941.1	*Nephila pilipes*
Dragline Silk Fibroin 1	AAV48942.1	*Nephila pilipes*
Dragline Silk Fibroin 1	AAV48943.1	*Nephila pilipes*
Dragline Silk Fibroin 1	AAV48944.1	*Nephila pilipes*
Dragline Silk Fibroin 1	AAV48945.1	*Nephila pilipes*
Dragline Silk Fibroin 1	AAV48947.1	*Nephila pilipes*
Dragline Silk Fibroin 1	AAV48948.1	*Nephila pilipes*
Dragline Silk Fibroin 1	AAV48929.1	*Octonoba varians*
Dragline Silk Fibroin 1	AAV48938.1	*Psechrus sinensis*
Dragline Silk Fibroin 1	AAV48939.1	*Psechrus sinensis*
Dragline Silk Fibroin 2	P46804.1	*Nephila clavipes*
Dragline Silk Protein	AAL32375.1	*Nephila clavipes*
Dragline Silk Protein Spidroin 1	AAC04504.1	*Nephila clavipes*
Dragline Silk Protein Spidroin 2	AAL32472.1	*Nephila clavata*
Egg Case Fibroin	AAZ15706.1	*Latrodectus hesperus*
Egg Case Fibroin	ADV40181.1	*Latrodectus hesperus*
Egg Case Fibroin-Like Protein 1	AFM97608.1	*Liphistius malayanus*
Egg Case Fibroin-Like Protein 2	AFM97609.1	*Liphistius malayanus*
Egg Case Fibroin-Like Protein 3	AFM97610.1	*Liphistius malayanus*
Egg Case Fibroin-Like Protein 4	AFM97611.1	*Liphistius malayanus*
Egg Case Fibroin-Like Protein 5	AFM97612.1	*Liphistius malayanus*
Egg Case Fibroin-Like Protein 6	AFM97613.1	*Liphistius malayanus*
Egg Case Silk Protein 1	BAE86855.1	*Argiope bruennichi*
Egg Case Silk Protein 1	AAX92677.1	*Latrodectus hesperus*
Egg Case Silk Protein 2	BAE86856.1	*Argiope bruennichi*
Egg Case Silk Protein 2	ABC68105.1	*Latrodectus hesperus*
Fibroin 1	ABW80562.1	*Aliatypus gulosus*
Fibroin 1	AFM97617.1	*Aphonopelma seemanni*
Fibroin 1	ABW80563.1	*Aptostichus sp. AS217*
Fibroin 1	AFM97622.1	*Atypoides riversi*
Fibroin 1	ABW80565.1	*Bothriocyrtum californicum*
Fibroin 1	ADM14313.1	*Bothriocyrtum californicum*
Fibroin 1	AAK30598.1	*Dolomedes tenebrosus*
Fibroin 1	AAK30600.1	*Euagrus chisoseus*

Fibroin 1	ABW80568.1	*Euagrus chisoseus*
Fibroin 1	AFM97625.1	*Hexura picea*
Fibroin 1	AFM97615.1	*Hypochilus thorelli*
Fibroin 1	AFM97627.1	*Megahexura fulva*
Fibroin 1	AAK30610.1	*Plectreurys tristis*
Fibroin 1	AFM97620.1	*Poecilotheria regalis*
Fibroin 1	AIU80193.1	*Scytodes thoracica*
Fibroin 1a	ABD61591.1	*Deinopis spinosa*
Fibroin 1b	ABD61592.1	*Deinopis spinosa*
Fibroin 2	AFM97618.1	*Aphonopelma seemanni*
Fibroin 2	ABW80564.1	*Aptostichus sp. AS220*
Fibroin 2	AFM97623.1	*Atypoides riversi*
Fibroin 2	ABD61588.1	*Deinopis spinosa*
Fibroin 2	AAK30599.1	*Dolomedes tenebrosus*
Fibroin 2	AFM97616.1	*Hypochilus thorelli*
Fibroin 2	AAK30611.1	*Plectreurys tristis*
Fibroin 2	AFM97621.1	*Poecilotheria regalis*
Fibroin 2	AIU80194.1	*Scytodes thoracica*
Fibroin 3	AFM97619.1	*Aphonopelma seemanni*
Fibroin 4	AAK30613.1	*Plectreurys tristis*
Fibroin-1	AAC47008.1	*Araneus diadematus*
Fibroin-2	AAC47009.1	*Araneus diadematus*
Fibroin-3	AAC47010.1	*Araneus diadematus*
Fibroin-4	AAC47011.1	*Araneus diadematus*
Flagelliform Silk Protein	AAK30594.1	*Argiope trifasciata*
Flagelliform Silk Protein	ABD61590.1	*Deinopis spinosa*
Flagelliform Silk Protein	AAC38847.1	*Nephila clavipes*
Flagelliform Silk Protein	AAF36092.1	*Nephila inaurata madagascariensis*
Flagelliform Silk Protein-1	AAT36347.1	*Araneus ventricosus*
Flagelliform Silk Protein-Like	ABR37273.1	*Nephilengys cruentata*
Major Ampullate Gland Dragline Silk Protein 2	AAN85281.1	*Araneus ventricosus*
Major Ampullate Silk Protein-Putative-a	AZAQ01054438.1	*Stegodyphus mimosarum*
Major Ampullate Silk Protein-Putative-b	AZAQ01171519.1	*Stegodyphus mimosarum*
Major Ampullate Silk Protein-Putative-c	AZAQ01121857.1	*Stegodyphus mimosarum*
Major Ampullate Silk Protein-Putative-d	AZAQ01026555.1	*Stegodyphus mimosarum*
Major Ampullate Silk Protein-	AZAQ01086913.1	*Stegodyphus mimosarum*

Putative-e		
Major Ampullate Silk Protein-Putative-f	AZAQ01026550.1	*Stegodyphus mimosarum*
Major Ampullate Silk Protein-Putative-g	AZAQ01005949.1	*Stegodyphus mimosarum*
Major Ampullate Silk Protein-Putative-h	AZAQ01026555.1	*Stegodyphus mimosarum*
Major Ampullate Silk Protein-Putative-i	AZAQ01107911.1	*Stegodyphus mimosarum*
Major Ampullate Silk Protein-Putative-j	AZAQ01107913.1	*Stegodyphus mimosarum*
Major Ampullate Silk Protein-Putative-k	AZAQ01108596.1	*Stegodyphus mimosarum*
Major Ampullate Spidroin	AAT08436.1	*Agelenopsis aperta*
Major Ampullate Spidroin	ADM14324.1	*Agelenopsis aperta*
Major Ampullate Spidroin	ADM14325.1	*Agelenopsis aperta*
Major Ampullate Spidroin	ADM14315.1	*Diguetia canities*
Major Ampullate Spidroin	ADM14316.1	*Diguetia canities*
Major Ampullate Spidroin 1	AEV46833.2	*Araneus ventricosus*
Major Ampullate Spidroin 1	AAP88232.1	*Argiope amoena*
Major Ampullate Spidroin 1	AAK30591.1	*Argiope aurantia*
Major Ampullate Spidroin 1	AFN54362.1	*Argiope bruennichi*
Major Ampullate Spidroin 1	CAJ00428.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32251.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32252.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32253.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32254.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32255.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32256.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32257.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32258.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32259.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32260.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32261.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32262.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32263.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32264.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32265.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32267.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32268.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32269.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	CAM32270.1	*Euprosthenops australis*

Major Ampullate Spidroin 1	CAM32271.1	*Euprosthenops australis*
Major Ampullate Spidroin 1	AAT08433.1	*Kukulcania hibernalis*
Major Ampullate Spidroin 1	ADM14314.1	*Kukulcania hibernalis*
Major Ampullate Spidroin 1	ABR68856.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1	ABR68857.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1	AAY28935.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1	ABD66602.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1	ABR68856.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1	ABR68857.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1	ADO78764.1	*Latrodectus mactans*
Major Ampullate Spidroin 1	ABC72644.1	*Nephila antipodiana*
Major Ampullate Spidroin 1	AAT75308.1	*Nephila clavipes*
Major Ampullate Spidroin 1	AAT75309.1	*Nephila clavipes*
Major Ampullate Spidroin 1	AAT75310.1	*Nephila clavipes*
Major Ampullate Spidroin 1	AAT75311.1	*Nephila clavipes*
Major Ampullate Spidroin 1	AAT75312.1	*Nephila clavipes*
Major Ampullate Spidroin 1	ACC77633.1	*Nephila clavipes*
Major Ampullate Spidroin 1	AAK30606.1	*Nephila inaurata madagascariensis*
Major Ampullate Spidroin 1	AAK30608.1	*Nephila senegalensis*
Major Ampullate Spidroin 1	AAK30614.1	*Tetragnatha kauaiensis*
Major Ampullate Spidroin 1	AAK30615.1	*Tetragnatha versicolor*
Major Ampullate Spidroin 1	ABD61596.1	*Uloborus diversus*
Major Ampullate Spidroin 1-Like	AAK30595.1	*Argiope trifasciata*
Major Ampullate Spidroin 1-Like	AAZ15320.1	*Latrodectus geometricus*
Major Ampullate Spidroin 1-Like	AAZ15321.1	*Latrodectus geometricus*
Major Ampullate Spidroin 1-Like	AAK30602.1	*Latrodectus geometricus*
Major Ampullate Spidroin 1-Like	ABD24294.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1-Like	ADG57596.1	*Parawixia bistriata*
Major Ampullate Spidroin 1-Like	ADE74592.1	*Peucetia viridans*
Major Ampullate Spidroin 1, Locus 1	ABY67402.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 1	ABY67403.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 1	ABY67412.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 1	ABY67413.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 1	ABY67414.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 1	ABY67415.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 1	ABY67418.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 1	ABY67421.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 1	ABY67423.1	*Latrodectus hesperus*

Major Ampullate Spidroin 1, Locus 2	ABY67406.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 2	ABY67407.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 2	ABY67410.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 2	ABY67411.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 2	ABY67422.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 2	ABY67424.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 2	ABY67425.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 3	ABY67420.1	*Latrodectus geometricus*
Major Ampullate Spidroin 1, Locus 3	ABY67400.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 3	ABY67401.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 3	ABY67404.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 3	ABY67405.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Locus 3	ABY67419.1	*Latrodectus hesperus*
Major Ampullate Spidroin 1, Precursor	CAJ90517.1	*Euprosthenops australis*
Major Ampullate Spidroin 1, Variant 1, Locus 1	ABY67426.1	*Latrodectus geometricus*
Major Ampullate Spidroin 1, Variant 1, Locus 2	ABY67428.1	*Latrodectus geometricus*
Major Ampullate Spidroin 1, Variant 2, Locus 1	ABY67427.1	*Latrodectus geometricus*
Major Ampullate Spidroin 1, Variant 2, Locus 2	ABY67429.1	*Latrodectus geometricus*
Major Ampullate Spidroin 1A, Precursor	ACF19411.1	*Nephila clavipes*
Major Ampullate Spidroin 1B, Precursor	ACF19412.1	*Nephila clavipes*
Major Ampullate Spidroin 2	AAR13808.1	*Argiope amoena*
Major Ampullate Spidroin 2	AAR13809.1	*Argiope amoena*
Major Ampullate Spidroin 2	AAR13810.1	*Argiope amoena*
Major Ampullate Spidroin 2	AAR13811.1	*Argiope amoena*
Major Ampullate Spidroin 2	AAR13812.1	*Argiope amoena*
Major Ampullate Spidroin 2	AAR13813.1	*Argiope amoena*
Major Ampullate Spidroin 2	AAR13814.1	*Argiope amoena*
Major Ampullate Spidroin 2	AAK30592.1	*Argiope aurantia*
Major Ampullate Spidroin 2	AFN54363.1	*Argiope bruennichi*
Major Ampullate Spidroin 2	AAK30596.1	*Argiope trifasciata*
Major Ampullate Spidroin 2	AAZ15371.1	*Argiope trifasciata*
Major Ampullate Spidroin 2	AAZ15372.1	*Argiope trifasciata*
Major Ampullate Spidroin 2	ADM14319.1	*Deinopis spinosa*
Major Ampullate Spidroin 2	CAM32249.1	*Euprosthenops australis*
Major Ampullate Spidroin 2	CAM32272.1	*Euprosthenops australis*

Major Ampullate Spidroin 2	AAK30601.1	*Gasteracantha cancriformis*
Major Ampullate Spidroin 2	AAK30603.1	*Latrodectus geometricus*
Major Ampullate Spidroin 2	AAK30604.1	*Latrodectus geometricus*
Major Ampullate Spidroin 2	ABY67417.1	*Latrodectus geometricus*
Major Ampullate Spidroin 2	AAY28936.1	*Latrodectus hesperus*
Major Ampullate Spidroin 2	ABD66603.1	*Latrodectus hesperus*
Major Ampullate Spidroin 2	ABR68855.1	*Latrodectus hesperus*
Major Ampullate Spidroin 2	ABR68858.1	*Latrodectus hesperus*
Major Ampullate Spidroin 2	ABY67408.1	*Latrodectus hesperus*
Major Ampullate Spidroin 2	ABY67409.1	*Latrodectus hesperus*
Major Ampullate Spidroin 2	ABY67416.1	*Latrodectus hesperus*
Major Ampullate Spidroin 2	ABD24295.1	*Latrodectus hesperus*
Major Ampullate Spidroin 2	AAT75313.1	*Nephila clavipes*
Major Ampullate Spidroin 2	AAT75314.1	*Nephila clavipes*
Major Ampullate Spidroin 2	AAT75315.1	*Nephila clavipes*
Major Ampullate Spidroin 2	AAT75316.1	*Nephila clavipes*
Major Ampullate Spidroin 2	AAT75317.1	*Nephila clavipes*
Major Ampullate Spidroin 2	ACF19413.1	*Nephila clavipes*
Major Ampullate Spidroin 2	AAK30607.1	*Nephila inaurata madagascariensis*
Major Ampullate Spidroin 2	AAK30609.1	*Nephila senegalensis*
Major Ampullate Spidroin 2	ABD61599.1	*Uloborus diversus*
Major Ampullate Spidroin 2	ABD61600.1	*Uloborus diversus*
Major Ampullate Spidroin 2-1	AAT08434.1	*Kukulcania hibernalis*
Major Ampullate Spidroin 2-2	AAT08435.1	*Kukulcania hibernalis*
Major Ampullate Spidroin 2-Like	AAK30597.1	*Argiope trifasciata*
Major Ampullate Spidroin 2-Like	AAZ15322.1	*Nephila inaurata madagascariensis*
Major Ampullate Spidroin 2-Like	AAK30605.1	*Nephila inaurata madagascariensis*
Major Ampullate Spidroin 2-Like	ADG57597.1	*Parawixia bistriata*
Major Ampullate Spidroin 2a	ABD61593.1	*Deinopis spinosa*
Major Ampullate Spidroin 2b	ABD61594.1	*Deinopis spinosa*
Major Ampullate Spidroin 3	AAT08432.1	*Kukulcania hibernalis*
Major Ampullate Spidroin-Like	ADM14317.1	*Diguetia canities*
Major Ampullate Spidroin-Like	ADM14318.1	*Diguetia canities*
Major Ampullate Spidroin-Like	CAM32250.1	*Euprosthenops australis*
Major Ampullate Spidroin-Like	AAV91960.1	*Latrodectus geometricus*
Major Ampullate Spidroin-Like	ABR37275.1	*Nephilengys cruentata*
Minor Ampullate Silk Protein	AFV31615.1	*Araneus ventricosus*

Minor Ampullate Silk Protein	AFV31613.1	*Araneus ventricosus*
Minor Ampullate Silk Protein	AFV31614.1	*Araneus ventricosus*
Minor Ampullate Silk Protein	AFM29835.1	*Argiope argentata*
Minor Ampullate Silk Protein	AFM29836.1	*Argiope argentata*
Minor Ampullate Silk Protein	ABD61589.1	*Deinopis spinosa*
Minor Ampullate Silk Protein	ADM14321.1	*Latrodectus hesperus*
Minor Ampullate Silk Protein	ADM14322.1	*Latrodectus hesperus*
Minor Ampullate Silk Protein	ADM14320.1	*Metepeira grandiosa*
Minor Ampullate Silk Protein	ADM14328.1	*Metepeira grandiosa*
Minor Ampullate Silk Protein	ADM14329.1	*Metepeira grandiosa*
Minor Ampullate Silk Protein	AAC14590.1	*Nephila clavipes*
Minor Ampullate Silk Protein	AF027736.1	*Nephila clavipes*
Minor Ampullate Silk Protein	ADM14326.1	*Uloborus diversus*
Minor Ampullate Silk Protein	ADM14327.1	*Uloborus diversus*
Minor Ampullate Silk Protein	ABD61597.1	*Uloborus diversus*
Minor Ampullate Silk Protein 1-Like	ACB29694.1	*Latrodectus hesperus*
Minor Ampullate Silk Protein-Like	ABR37276.1	*Nephilengys cruentata*
Minor Ampullate Silk Protein-Like	ABR37277.1	*Nephilengys cruentata*
Minor Ampullate Silk Protein-Like	ABR37278.1	*Nephilengys cruentata*
Minor Ampullate Silk Protein-Like	ADG57595.1	*Parawixia bistriata*
Minor Ampullate Silk Protein-Putative	AZAQ01030745.1	*Stegodyphus mimosarum*
Minor Ampullate Spidroin 1	ABC72645.1	*Nephila antipodiana*
Minor Ampullate Spidroin 1	AAC14589.1	*Nephila clavipes*
Minor Ampullate Spidroin 1	AF027735.1	*Nephila clavipes*
Minor Ampullate Spidroin 2	AF027737.1	*Nephila clavipes*
Piriform Spidroin	AEP25627.1	*Araneus gemmoides*
Piriform Spidroin	ADN39425.1	*Argiope trifasciata*
Piriform Spidroin	ADN39427.1	*Nephila clavipes*
Piriform Spidroin	ADN39426.1	*Nephila clavipes*
Piriform Spidroin-Like	ADK56477.1	*Nephilengys cruentata*
Piriform Spidroin-Putative	AZAQ01087893.1	*Stegodyphus mimosarum*
Pyriform Spidroin 1	ACV41934.1	*Latrodectus hesperus*
Pyriform Spidroin 1	ADV40087.1	*Latrodectus hesperus*
Pyriform Spidroin 2	ADK92884.1	*Nephila clavipes*
Silk Gland Protein 1	AAR21194.1	*Argiope amoena*
Spidroin 1	AAC38957.1	*Nephila clavipes*
Spidroin 1	AZAQ01020238.1	*Stegodyphus mimosarum*
Spidroin 1a	ACF71407.1	*Avicularia juruensis*

Spidroin 1b	ACF71408.1	*Avicularia juruensis*
Spidroin 1c	ACF71409.1	*Avicularia juruensis*
Spidroin 2	AAC04503.1	*Araneus bicentenarius*
Spidroin 2	ACF71410.1	*Avicularia juruensis*
Spidroin 2a	AZAQ01082375.1	*Stegodyphus mimosarum*
Spidroin 2b	AZAQ01067030.1	*Stegodyphus mimosarum*
Spidroin 2c	AZAQ01099004.1	*Stegodyphus mimosarum*
Tubuliform Spidroin	AAX45293.1	*Araneus gemmoides*
Tubuliform Spidroin	AAX45294.1	*Araneus gemmoides*
Tubuliform Spidroin	AAX45292.1	*Argiope aurantia*
Tubuliform Spidroin	AAX45291.1	*Argiope aurantia*
Tubuliform Spidroin	AAX45295.1	*Nephila clavipes*
Tubuliform Spidroin-1	ADM14323.1	*Agelenopsis aperta*
Tubuliform Spidroin-1	ADM14330.1	*Agelenopsis aperta*
Tubuliform Spidroin-1	ADM14331.1	*Agelenopsis aperta*
Tubuliform Spidroin-1	AFA43480.1	*Argiope amoena*
Tubuliform Spidroin-1	AAY28932.1	*Argiope argentata*
Tubuliform Spidroin-1	AAY28945.1	*Argiope argentata*
Tubuliform Spidroin-1	AAY28952.1	*Argiope argentata*
Tubuliform Spidroin-1	ADM14332.1	*Argiope argentata*
Tubuliform Spidroin-1	ADM14333.1	*Argiope argentata*
Tubuliform Spidroin-1	AAY28942.1	*Argiope aurantia*
Tubuliform Spidroin-1	AAY28953.1	*Argiope aurantia*
Tubuliform Spidroin-1	AAY28944.1	*Cyrtophora moluccensis*
Tubuliform Spidroin-1	AAY28934.1	*Deinopis spinosa*
Tubuliform Spidroin-1	AAY28943.1	*Gea heptagon*
Tubuliform Spidroin-1	AAY28954.1	*Gea heptagon*
Tubuliform Spidroin-1	AAY28940.1	*Latrodectus geometricus*
Tubuliform Spidroin-1	AAY28950.1	*Latrodectus geometricus*
Tubuliform Spidroin-1	AAY28941.1	*Latrodectus hasseltii*
Tubuliform Spidroin-1	AAY28949.1	*Latrodectus hasseltii*
Tubuliform Spidroin-1	AAY28931.1	*Latrodectus hesperus*
Tubuliform Spidroin-1	AAY28937.1	*Latrodectus hesperus*
Tubuliform Spidroin-1	AAY28947.1	*Latrodectus hesperus*
Tubuliform Spidroin-1	ADV40185.1	*Latrodectus hesperus*
Tubuliform Spidroin-1	ABD24296.1	*Latrodectus hesperus*
Tubuliform Spidroin-1	AAY28938.1	*Latrodectus mactans*
Tubuliform Spidroin-1	AAY28946.1	*Latrodectus mactans*
Tubuliform Spidroin-1	AAY28939.1	*Latrodectus tredecimguttatus*

Tubuliform Spidroin-1	AAY28948.1	*Latrodectus tredecimguttatus*
Tubuliform Spidroin-1	AAY90151.1	*Nephila antipodiana*
Tubuliform Spidroin-1	AAY28951.1	*Steatoda grossa*
Tubuliform Spidroin-1	AAY28933.1	*Uloborus diversus*
Tubuliform Spidroin-Like	ABR37274.1	*Nephilengys cruentata*
Tubuliform Spidroin-Putative	AZAQ01117603.1	*Stegodyphus mimosarum*

Table S1.3 Identified silk genes in male Theridiidae spiders. Abbreviations: *AcSp1* (*aciniform spidroin 1*), *AgSF2* (*aggregate silk factor 2*), *AgSp1* (*aggregate spidroin 1*), *Flag* (*flagelliform spidroin*), *MaSp1* (*major ampullate spidroin 1*), *MaSp2* (*major ampullate spidroin 2*), *MaSp3* (*major ampullate spidroin 3*; shown in Chaw et al., 2015 as *MaSp'*), *MiSp* (*minor ampullate spidroin*), *PySp1* (*pyriform spidroin 1*), and *SCPs* (*spider coating peptide 1* and *2*), *TuSp1* (*tubuliform spidroin 1*), Nterm (amino terminal domain), and Cterm (carboxyl terminal domain).

Species	**Contig name**	**Gene Name**	**Gene Region**
Latrodectus geometricus	LgML_comp32173_c0	*AcSp1*	Cterm
Latrodectus geometricus	LgML_comp23156_c0	*AgSF2*	Cterm
Latrodectus geometricus	LgML_comp16474_c0	*AgSp1*	Cterm
Latrodectus geometricus	LgML_comp149066_c0	*Flag*	Cterm
Latrodectus geometricus	LgML_comp22311_c0	*MaSp1*	Cterm
Latrodectus geometricus	LgML_comp241_c0	*MaSp1*	Nterm
Latrodectus geometricus	LgML_comp17309_c0	*MaSp2*	Cterm
Latrodectus geometricus	LgML_comp14684_c0	*MaSp3*	Cterm
Latrodectus geometricus	LgML_comp31776_c1	*MiSp*	Cterm
Latrodectus geometricus	LgML_comp34462_c0	*MiSp*	Cterm
Latrodectus geometricus	LgML_comp1856_c0	*PySp1*	Cterm
Latrodectus geometricus	LgML_comp24742_c0	*PySp1*	Cterm
Latrodectus geometricus	LgML_comp21261_c0	*TuSp1*	Cterm
Latrodectus hesperus	LhML_comp18769_c0	*AcSp1*	Cterm
Latrodectus hesperus	LhML_comp15040_c0	*AgSF2*	Cterm
Latrodectus hesperus	LhML_comp12654_c0	*AgSp1*	Cterm
Latrodectus hesperus	LhML_comp18901_c0	*MaSp'*	Cterm
Latrodectus hesperus	LhML_comp13613_c0	*MaSp1*	Cterm
Latrodectus hesperus	LhML_comp9989_c0	*MaSp1*	Nterm
Latrodectus hesperus	LhML_comp5014_c0	*MaSp2*	Cterm
Latrodectus hesperus	LhML_comp17353_c1	*MiSp*	Cterm
Latrodectus hesperus	LhML_comp13758_c0	*PySp1*	Cterm
Latrodectus hesperus	LhML_comp17561_c0	SCP1	
Latrodectus hesperus	LhML_comp8101_c0	SCP1	
Latrodectus hesperus	LhML_comp12637_c0	*TuSp1*	Cterm
Steatoda grossa	SgML_comp14032_c0	*AcSp1*	Cterm
Steatoda grossa	SgML_comp50464_c0	*MaSp1*	Cterm
Steatoda grossa	SgML_comp68394_c0	*MaSp1*	Cterm
Steatoda grossa	SgML_comp61844_c0	*MaSp2*	Nterm
Steatoda grossa	SgML_comp68866_c0	*MiSp*	Cterm
Steatoda grossa	SgML_comp72094_c1	*MiSp*	Cterm

Steatoda grossa	SgML_comp72094_c1	*MiSp*	Cterm
Steatoda grossa	SgML_comp68599_c0	*PySp1*	Cterm
Steatoda grossa	SgML_comp60457_c0	*TuSp1*	Cterm

Table S1.4. Expression of silk genes not expected to be found in male spiders based on the absence of functional spigots. Average expression of male reads mapped to species-specific male transcriptomes and total silk gland female reads mapped to species-specific global transcriptomes shown as reads per kilobase of transcript per million mapped reads (RPKM). Silk genes not recovered in this study indicated by ---. Silk genes with expression levels RPKM < 1 are indicated by < 1.

	RPKM						
	TuSp1	***ECP-1***	***ECP-2***	***AgSp1***	***AgSF2***	***SCP-1***	***SCP-2***
***Latrodectus hesperus* global**	4670.37	379.01	5.07	1371.48	9156.74	6113.33	1088.30
***Latrodectus hesperus* male**	1.54	< 1	---	1.96	203.74	2.82	---
***Latrodectus geometricus* global**	3633.22	594.92	1844.35	1840.47	1084.42	< 1	< 1
***Latrodectus geometricus* male**	11.16	---	---	1.70	1.35	---	---
***Steatoda grossa* global**	5164.97	72.58	310.35	1828.57	< 1	---	< 1
***Steatoda grossa* male**	34.99	---	---	---	---	---	---

Table S1.5. Number of venom genes shared across focal spider species. Male venom genes with average expression levels RPKM < 1 of male reads mapped to species-specific male transcriptomes were not counted. *Data for female venom genes taken from Haney et al. (2016). Abbreviations: ICK = inhibitory cysteine knot toxins, CRISP = cysteine-rich secretory proteins.

	Latrotoxins	Latrodectins	ICK	CRISP
Latrodectus hesperus* female	14	5	4	6
***Latrodectus hesperus* male**	2	2	1	1
Latrodectus geometricus* female	17	4	2	5
***Latrodectus geometricus* male**	1	1	1	3
Steatoda grossa* female	12	6	3	1
***Steatoda grossa* male**	0	2	2	1

Table S1.6. Expression of venom genes found in male spiders. Average expression of male reads mapped to species-specific male transcriptomes shown in reads per kilobase of transcript per million mapped reads (RPKM). Venom genes with average expression RPKM < 1 are indicated by < 1. Abbreviations: ICK = inhibitory cysteine knot toxins, CRISP = cysteine-rich secretory proteins.

	RPKM		
♂	*L. hesperus*	*L. geometricus*	*S. grossa*
Latrotoxins	38.71	1.02	<1
Latrodectins	41.63	2.74	13.48
ICK	8.33	2.79	5.66
CRISP	6.27	10.97	138.41

Chapter 2

Semi-Aquatic Spider Silks: Transcripts, Proteins, and Silk fibers of the Fishing Spider, *Dolomedes triton* (Pisauridae)

Abstract

To survive in terrestrial and aquatic environments, spiders often rely heavily on their silk. The vast majority of silks that have been studied are from orb-weaving or cob-web weaving species, leaving the silks of water-associated spiders largely undescribed. We characterize silks from the semi-aquatic spider *Dolomedes triton*. From silk gland RNA-Seq libraries, we report 18 silk transcripts representing four categories of known silk protein types: aciniform, ampullate, pyriform, and tubuliform. Proteomic and structural analyses (SEM, EDS, contact angle) of *D. triton's* submersible egg sac reveal similarities to silks from aquatic caddisfly larvae. We identified two layers in *D. triton* egg sacs, notably a highly hydrophobic outer layer with a different elemental composition compared to egg sacs of terrestrial spiders. These features may provide *D. triton* egg sacs with their water repellent properties.

Introduction

Spiders (Araneae) are distributed worldwide with nearly all of the over 47,000 species (World Spider Catalog, 2018) occurring in terrestrial habitats. Yet, a few species are associated with aquatic and semi-aquatic environments. Fishing spiders (*Dolomedes*, Pisauridae) are commonly found around ponds and lakes. They are active hunters and do not build webs to catch prey, but instead, they either ambush or stalk aquatic invertebrates or small fish. Nonetheless, silk use in fishing spiders is crucial to their survival. Both male and female *Dolomedes triton* (Walckenaer, 1837) spiders use silk as their trailing safety line (dragline) to move from one place to another. However, once *D. triton* spiders become sexually mature, each sex also uses silk for different purposes. For example, males use silk to construct sperm webs and to wrap nuptial gifts (Lang, 1996). Females deposit pheromones on their dragline silk to attract males (Gaskett, 2007; Roland and Rovner, 1983) and encase their eggs in sphere-shaped, silken egg sacs that they carry with their chelicerae (McAlister, 1960). Female *D. triton* spiders also build elaborate silken structures called nursery webs, where spiderlings are housed for a period of time after hatching.

As exemplified by *D. triton*, spiders use silk for an array of essential tasks related to their survival including prey capture, reproduction, locomotion, and protection of progeny (Foelix, 2011). Here, we describe the silk genes expressed by *D. triton* females and males and the characteristics of *D. triton* egg sacs. One predictor of the types of silk a spider is capable of producing is the complement of silk spigots on the spider's spinnerets. Each spigot is connected to its own silk gland, and spigots vary greatly in size,

shape, and location on a spinneret. Thus, studies of spigot morphology can be indicative of the types of silk and silk genes expressed by a spider (Correa-Garhwal et al., 2017). Based on previous studies of spigot morphology, we expect *D. triton* spiders to express genes associated with aciniform, major ampullate, minor ampullate, pyriform, and tubuliform silk glands (Griswold, 2005; Moon, 2008; Murphy and Roberts, 2015; Zhang et al., 2004). The silk genes expected are members of the spidroin (spider silk fibroin) gene family (Blasingame et al., 2009; Guerette et al., 1996; Hayashi et al., 2004; Hinman and Lewis, 1992). Spidroins are named for the gland type in which they were initially identified. Thus, the expected genes are for AcSp (a contraction of aciniform spidroin), MaSp (major ampullate spidroin), MiSp (minor ampullate spidroin), PySp (pyriform spidroin), and TuSp (tubuliform spidroin). Only females have tubuliform (egg case) spigots. Hence, we expect only females to express TuSp genes.

To investigate any molecular changes in the silk genes, we compare *D. triton* silk sequences to published spidroins sequences from other araneomorph spider species. Moreover, we compare *D. triton* spidroin gene characteristics to those of the velvet spider *Stegodyphus mimosarum* Pavesi, 1883 (Sanggaard et al., 2014), the most closely related species with published genomic and transcriptomic data. By comparing *D. triton* to *S. mimosarum*, we will investigate whether the silk sequences of *D. triton* have modifications that could be linked to survival in aquatic environments compared to the silks of a terrestrial species. Alternatively, spider silks in general might not need any specializations to function in water. In this scenario, *D. triton* silks could be considered pre-adapted to semi-aquatic environments. Furthermore, we investigate the composition

of *D. triton* submersible egg sacs. *D. triton* females carry their silken egg sacs over land, across the surface of water, and sometimes under water. We use a proteomic approach to identify the main proteins that contribute to the egg sac's waterproof properties.

Materials and Methods

Silk Gland Collection and RNA Isolation

Adult female and male *D. triton* were obtained from Todd Gearheart Enterprises (taranturaspiders.com; Florida, USA). Immediately after euthanization, the total set of silk glands was dissected from each spider and flash frozen in liquid nitrogen. Separate RNA extractions were done for the total set of silk glands from each individual spider. Briefly, total RNA was isolated from silk glands with TRIzol (Invitrogen, Carlsbad, CA, USA) and further purified with an RNeasy mini kit (Qiagen, Germantown, MD, USA). Residual genomic DNA was removed with Turbo DNAse (Thermo Fisher Scientific, Waltham, MA, USA). RNA extractions were quantified using a Nanodrop ND-1000 (Thermo Fisher Scientific). RNA quality was estimated by visualizing rRNA with a Bioanalyzer (Agilent 2100, Santa Clara, CA, USA).

RNA-Seq Library Construction and Sequencing

RNA-Seq libraries were made from *D. triton* silk glands using the Ovation Universal RNA-Seq System (NuGen, San Carlos, CA, USA). Four libraries were constructed, two from silk glands of females and two from the silk glands of males. Depletion of unwanted ribosomal RNA transcripts was done using Insert Dependent

Adaptor Cleavage (In-DAC) technology, which is part of the Ovation Universal RNA-Seq System. Libraries were indexed, combined in equimolar concentrations, and sequenced (paired end, 150 cycles) on a MiSeq System (Illumina, San Diego, CA, USA) at the University of California, Riverside Genomics Core Facility.

De novo Assembly

Raw sequencing reads from each FASTQ file were processed by clipping the adaptors and removing low quality reads with Trimmomatic (Bolger et al., 2014). For each library, the quality of the resulting filtered reads was assessed using FastQC (Babraham Bioinformatics FastQC Package). No k-mer normalization was performed following (Clarke et al., 2014). Two *D. triton* sex-specific silk gland transcriptomes were assembled with Trinity v2.1.1 using default parameters (Grabherr et al., 2011), one with the combined reads from the replicate male libraries, and the other with the reads from the female libraries. The two assemblies were combined into a single, comprehensive transcriptome with CAP3 (Huang and Madan, 1999). Quality of *D. triton* assembly was approximated using N50. Assembly completeness was determined by comparison to the arthropod set of Universal Single-Copy Orthologs (BUSCO v 1.2; Simão et al., 2015) and to a core eukaryotic gene dataset using CEGMA (Parra et al., 2007). Transcripts were automatically annotated using BLASTX searches (e-value < 1e-5) to both NCBI NR and UniProt_KB (Altschul et al., 1990). Putative chimeric and contaminants were removed from the resulting assemblies following Clarke et al. (2015). Transcripts were functionally annotated with Gene Ontology (GO) terms associated with the UniProt best

matches. To generate predicted proteins, assembled contigs were translated in the frame of its the best BLASTX hit to nr by e-value using a cutoff of 1e-5. If a BLASTX hit was not available, the longest open reading frame (ORF) was used to predict the amino acid sequence following Clarke et al. (2014).

Annotation of Spidroin Paralogs

To categorize spidroin family members, *D. triton* assembly was subject to BLASTX searches (e-value < 1 e-5) against a protein database with spidroin genes downloaded from NCBI nr proteins and UniProtKB/Swiss-Prot databases (September 2016) in Geneious v8.1.8 (Kearse et al., 2012). Each BLAST identified contig was visually inspected to confirm the presence of typical characteristics of spidroin genes, such as coding regions for conserved N- and C- terminal domains and repetitive regions (Figure 2.1A). Transcripts with a nucleotide identity >95% were manually collapsed into a single contig.

From the *D. triton* spidroin contigs, the N- and C-terminal encoding regions were translated and combined with published spidroin sequences from other araneomorph (true spider) species that represent a broad diversity of the spidroin family (Table S2.3). Spidroin termini sequences from a non-araneomorph spider were included as outgroups (*Bothriocyrtum californicum* fibroin 1, GenBank accessions EU117162 and HM752562). N-terminal regions (and separately, C-terminal regions) were aligned with ClustalW (Larkin et al., 2007) implemented in Geneious and refined by eye. Amino acid model test and maximum likelihood gene tree construction with 10,000 bootstrap replicates were

done in RAxML v8.2.8 (Stamatakis, 2014). LG and WAG likelihood amino acid substitution models were used for N- and C- terminal region alignments, respectively. Resulting trees were visualized with FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/Fig.tree/).

D. triton contigs were manually assigned to a spidroin type by their repetitive region characteristics, such as the presence of amino acid motifs that are diagnostic for specific spidroins (Gatesy et al., 2001), and phylogenetic position in the spidroin gene trees. We associated N-terminal region contigs with C-terminal region contigs as representing parts of the same spidroin only if the adjacent repetitive regions were nearly identical, as done in previous studies (Garb et al., 2010; Motriuk-Smith et al., 2005).

Read Mapping and Gene Expression

Filtered reads (see above) were mapped to species-specific *D. triton* assembly using TopHat2 v2.1.1 with default parameters (Kim et al., 2013). Reads Per Kilobase per Million mapped reads (RPKM) values were calculated for each of the spidroin terminal domains and compared across sexes and across species. Only spidroins with more than ten reads mapped and at least one RPKM were kept for further analysis. Differences in gene expression for *D. triton* transcripts were analyzed from read counts using the R package DEseq2 with default parameters (Love et al., 2014). Genes were considered differentially expressed if their adjusted p-values were <0.01.

Proteomic Analyses of D. triton Silk Fibers and Silk Glands

Egg sacs were harvested for proteomic analyses of the silk fibers that envelop the egg bundle. Egg sacs produced by female *D. triton* were cut open using clean micro scissors and the eggs were removed. The egg sacs were then examined under a stereomicroscope and two distinct layers were separated, a brown outer layer and a white inner layer (Figure 2.7). The two layers were collected and silk samples were stored at room temperature in clean, sealed containers. Two egg sacs from different *D. triton* females were collected for each egg sac layer.

Silk glands were also harvested for proteomic analyses. Tubuliform- and ampullate-shaped silk glands were individually dissected from two, mature *D. triton* females, placed in separate microfuge tubes, flash frozen in liquid nitrogen and stored at -80^{o}C. Two biological replicates were collected for each silk gland type.

Protein extractions were done following Chaw et al. (2015). Briefly, frozen silk glands were submerged in protein extraction buffer (10% glycerol, 50 nM Tris, pH 7.5, 5 nM $MgCl_2$, 2% SDS, 150 nM NaCl, 0.2% β-mercaptoethanol, 0.005 M EDTA) supplemented with Halt protease inhibitor cocktail (Thermo Fisher Scientific) and macerated with a pestle. For silk fibers, egg sac samples were minced with micro scissors, submerged in protein extraction buffer, and incubated overnight at room temperature. The solubilized protein samples (glands and fibers) were analyzed by SDS/PAGE on gels stained with Bio-Safe Coomassie blue (Bio-Rad, Hercules, CA, USA). Each protein sample was distributed among multiple lanes (~20 μL/lane) and protein bands were combined for in-gel digest. Trypsin/chymotrypsin in-gel digest was

done following the protocol from Arizona Proteomics Consortium (http://proteomics.arizona.edu/protocols). The digested peptides were extracted from the gel slices with sonication (Bioruptor standard sonication system, Diagenode, Denville, NJ, USA). After 30 minutes of sonication, peptides were purified from solution with Ziptips C_{18} (Millipore, Billerica, MA, USA), dried, and stored at -20^{o}C. The peptide samples were subjected to LC-MS/MS analysis on an LTQ Orbitrap Velos mass spectrometer (Thermo Fisher Scientific) equipped with a nanomate ESI source (Advion, Ithaca, NY, USA) at the University of Arizona's Arizona Proteomics Consortium.

The resulting tandem mass spectra were searched against the non-redundant longest open reading frame translation of our *D. triton* transcriptome (see above), Chelicerata proteins downloaded from NCBI (on October 17, 2013), and common contaminant proteins using Thermo Proteome Discoverer 1.3 (Thermo Fisher Scientific). Tryptic/chymotryptic peptides with up to two-missed cleavage sites were taken into consideration and iodoacetamide derivatives of cysteine and oxidation of methionines were specified as variable modifications. Protein and peptide identification results were visualized with Scaffold v4.7.3 (Proteome Software Inc., Portland, OR, USA). Proteins passing a minimum of two peptides identified at 95% protein confidence and 50% peptide confidence by the peptide and protein profile were accepted. Sample reports were exported from Scaffold. Contaminants were removed and biological replicates consolidated by taking the higher protein identification probability of the two replicates.

Annotation of Peroxidase Paralogs

Dolomedes triton peroxidases were combined with published sequences representing three different heme-containing peroxidase families: peroxidasin, peroxinectin, and myeloperoxidase (Table S2.4). The umbrella liverwort peroxidase was included as outgroup (*Marchantia polymorpha* peroxidase 1, GenBank accession: BAB97197). Protein sequences were aligned with ClustalW (Larkin et al., 2007) implemented in Geneious and refined by eye. Amino acid model test and maximum likelihood gene tree construction with 10,000 bootstrap replicates were done in RAxML v8.2.8 (Stamatakis, 2014). WAG likelihood amino acid substitution model was used for peroxidase alignments. Resulting tree was visualized with FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/Fig.tree/).

Elemental and Hydrophobicity Characterizations of D. triton Egg Sacs

Small sections of *D. triton* egg sacs were cut with clean micro scissors, mounted to aluminum pin stubs with carbon tape, and coated either with a thin layer of platinum-palladium for scanning electron microscopy (SEM) or carbon for Energy Dispersive X-Ray Spectroscopy (EDS; to avoid overlapping spectral lines from the coating). Electron micrographs were collected with a 5-10 kV accelerating voltage using a Mira3 (Tescan, Czech Republic). EDS spectra were taken using a Phoenix/Genesis (AMETEK Inc., Berwyn, PA, USA) and QUANTAX 400 6|60 mm^2 XFlash (Bruker, Billerica, MA, USA) system.

The contact angle of water with *D. triton* egg sacs was determined by pipetting 10 μL of deionized water onto intact egg sacs. Backlit images were taken using a Canon EDS 5DSR with macro lens, and advancing contact angle was measured using ImageJ (Rasband, National Institute of Health, https://imagej.nih.gov/ij/). For comparison to egg sacs from a fully terrestrial spider, *Latrodectus hesperus* (Western black widow) egg sacs were analyzed using the same methods.

Results and Discussion

De novo Transcriptome Assemblies

As genomic resources do not exist for *D. triton*, the first step in investigating gene expression levels was to construct a *de novo* assembled transcriptome from our four *D. triton* silk gland RNAseq libraries. Sequencing and assembly statistics are summarized in Table S2.1. A total of 45,703 contigs were assembled. Assembly quality was assessed with BUSCO (Simão et al., 2015) using *Ixodes scapularis* (deer tick) as a reference. About 73.4% of the tick BUSCOs were identified as complete in the *D. triton* assembly. The completeness score of the *D. triton* assembly is lower than previously published assemblies of *Nephila clavipes* 99.1% (Babb et al., 2017), *Argiope argentata* 92.5% (Chaw et al., 2016), *Latrodectus hesperus* 93.1%, *L. geometricus* 83.4%, and *Steatoda grossa* 90.5% (Babb et al., 2017; Chaw et al., 2016; Clarke et al., 2015). A likely explanation for the difference BUSCO completeness is that the *D. triton* assembly included only silk gland tissues while the other assemblies included multiple types of tissues.

Dolomedes triton Spidroins

D. triton transcriptome assembly was screened for spidroin and other silk-associated sequences (spider coating peptides SCP-1 and SCP2, egg case silk proteins ECP-1 and ECP-2, and aggregate silk factors AgSF1 and AgSF2), but only spidroins were found. Spidroins have non-repetitive amino (N)-terminal and carboxyl (C)-terminal regions, which flank tandemly arrayed repeats in a lengthy repetitive region (Figure 2.1A). We identified a total of 18 contigs that contained partial-length spidroin transcripts (Table S2.2). Ten of the contigs possessed coding sequences for amino (N)-terminal regions and the other eight had coding sequences for carboxyl (C)-terminal regions (Figure 2.1). Most (16; 89%) contigs also included adjacent repetitive sequence (Figure 2.2). We identified both N- and C-terminal domains for all of our identified *D. triton* spidroins with the exception of the spidroin that we refer to as "Sp" (only N- terminal domain known; Figure 2.1).

Our *D. triton* spidroin contigs expand the diversity of spidroins from non-araneoid spiders in terms of repeat composition and number of novel termini. Over the years, spidroins have been identified in cDNA studies (e.g., Correa-Garhwal and Garb, 2014; Garb et al., 2010; Gatesy et al., 2001; Rising et al., 2007; Tian et al., 2004). More recently, next generation sequencing has contributed tremendously to spidroin diversity not only in number but also in sequence completeness and repeat structure. For example, the genome of the non-araneoid velvet spider *Stegodyphus mimosarum* yielded 19 contigs (Sanggaard et al., 2014), and the genome of the golden orb-web spider *Nephila clavipes* identified 28 spidroins (Babb et al., 2017; Sanggaard et al., 2014).

Aciniform, Pyriform, and Tubuliform Spidroins

We found that *D. triton* have aciniform, pyriform, and tubuliform spidroin contigs (AcSp, PySp, and TuSp; Figure 2.1). The AcSp, PySp, and TuSp contigs are long enough to contain complete or near complete repeat units (Figure 2.2). AcSp, PySp, and TuSp repeats were aligned to the corresponding spidroin repeats from *S. mimosarum* (e.g., an alignment of only AcSp repeats; Figure S2.1). We found high sequence similarity and conservation among the aligned repeats, corroborating our annotations. Also, maximum likelihood analyses of the C- and N-terminal region sequences show that *D. triton* AcSp, PySp, and TuSp sequences group together with spidroins of the same respective type from the comparison species (Figures 2.3 and 2.4).

Aciniform spidroins are the primary component of aciniform silk, which is used in prey-wrapping, egg case construction, and web decoration (Chaw et al., 2014; Herberstein et al., 2000; Vasanthavada et al., 2007). Consistent with previously described aciniform spidroins from other species, *D. triton* AcSps are composed of long (~180 or ~208 amino acids), complex repeat units (Ayoub et al., 2013; Chaw et al., 2014; Vasanthavada et al., 2007). *D. triton* has at least two AcSp variants (vA, vB), which were found to contain two distinct types of repeat units. Substantial differences in repeat composition as well as N- and C- terminal regions of the variants suggest that they represent separate loci rather than alleles. Thus, we posit that *D. triton* has at least two AcSp loci (Figure 2.2 and Figure S2.1). The three most abundant amino acids in *D. triton* AcSp repeats are serine, glycine, and alanine, the combination of which account for 45% of vA, and 59% of vB. Despite the prevalence of alanine and glycine, AcSp repetitive

sequences are depleted of some of the amino acid sequence motifs (alternating glycine and alanine couplets, glycine-glycine-X) that are common in spidroins such as MaSp1 (Gatesy et al., 2001; Figure 2.2 and Figure S2.1). There are, however, poly-alanine amino acid motifs in our *D. triton* AcSp vB, but not vA. Poly-alanine amino acid motifs are common to a subset of spidroins, such as MaSp1 and MaSp2, but are rare or absent in AcSp spidroins (Ayoub et al., 2013, 2007; Chaw et al., 2014; Hayashi et al., 2004; Sanggaard et al., 2014). Long repeats of alanine are hypothesized to be involved in β-sheet formation and play an important role in fiber strength and toughness (Hinman and Lewis, 1992; Lawrence et al., 2004; Simmons et al., 1996; Trancik et al., 2005; van Beek et al., 2002). Similarly, poly-alanine motifs found in *D. triton* AcSp could contribute to the formation of nanocrystalline regions, in turn increasing the toughness of the fiber. Given that aciniform silk from araneoid spiders has been found to be extraordinarily tough largely due to high extensibility (Blackledge and Hayashi, 2006; Hayashi et al., 2004), tensile tests of non-araneoid aciniform are needed to investigate whether high toughness is general to aciniform silk or unique to araneoid aciniform silk.

As with the AcSp spidroins, *D. triton* PySp sequences show sequence similarities to PySp sequences from other species (Figure S2.1). PySp is the main protein found in pyriform silk, which is excreted as a composite fiber and glue that anchors other silk types to substrates (Blasingame et al., 2009; Kovoor & Zylberberg, 1980, 1982). Predicted amino acid compositions of *D. triton* PySp repeats show high levels of the polar amino acids serine (28%) and glutamine (10%), as well as very low levels of glycine (2%). PySp compositions of orb- and cob-weaving spiders are similarly rich in

serine and glutamine and poor in glycine (Blasingame et al., 2009; Chaw et al., 2017; Geurts et al., 2010; Perry et al., 2010). More generally, the co-occurrence of high serine with low glycine (and vice-versa) has been broadly observed across the spidroin family (Starrett et al., 2012). *D. triton* PySp sequences were also found to contain runs of proline alternating with another amino acid (PX; Figure S2.1). This PX amino acid motif is also found in PySp1 repeats from orb-weaving spiders, the house spider *Parasteatoda tepidariorum* PySp2*,* as well as *S. mimosarum* PySp and is hypothesized to provide extensibility to pyriform silk fibers in orb-weaving spiders (Chaw et al., 2017; Geurts et al., 2010; Perry et al., 2010).

Similar to AcSp and PySp, *D. triton* TuSp repetitive sequence is fairly well conserved (Figure S2.1). TuSp is associated with tubuliform silk, which is used to wrap egg cases and, in most spiders, is produced in specialized silk glands called tubuliform glands. *D. triton* TuSp sequence contains repeats that are largely composed of serine and alanine. Moreover, *D. triton* TuSp terminal regions form a monophyletic clade with other tubuliform spidroins in maximum likelihood analyses (Figures 3-4). Our phylogenetic analysis also recovered the tubuliform clade as sister to aciniform clade, a relationship supportive of previous hypothesis of spidroin evolution (Ayoub et al., 2013; Garb et al., 2010).

Ampullate Spidroins

Unlike aciniform, pyriform, or tubuliform spidroins, which had at most two variants per type, our *D. triton* ampullate spidroin category encompasses more variants.

D. triton was predicted to have silk genes associated with major and minor ampullate silk glands based on the presence of major and minor ampullate silk spigots on their spinnerets (Zhang et al., 2004; Moon, 2008; Murphy & Roberts, 2015). Major ampullate and minor ampullate silks are used for draglines and in the construction of prey-capture webs. In orb web and cobweb weaving spiders, the major and minor ampullate silk glands are dramatically different in size (Chaw et al., 2015; Jeffery et al., 2011; Mullen, 1969; Peters, 1955; Vollrath, 2000). However, in our silk gland dissections, the ampullate shaped glands of *D. triton* were not dramatically different in size or shape within an individual; instead they were similar to each other. Furthermore, maximum likelihood analyses of N- and C-terminal regions showed that our Amp sequences group with previously described MaSp (e.g., MaSp1, MaSp2, MaSp3) and MiSp from other species in a large, diverse clade that we termed "Amp" (Figures 3-4). We did not recover a monophyletic MaSp clade, nor a monophyletic MiSp clade. Thus, we were unable to definitively categorize our *D. triton* ampullate-like spidroin contigs as major ampullate spidroin (MaSp) or minor ampullate spidroin (MiSp). Instead, we refer to our ampullate transcripts with the more general abbreviation, "AmSp", then "N" or "C" for whether the transcript contains an amino or carboxy-terminal region, followed by a letter for each variant (v) type (e.g., *D. tri*_AmSp_C_vA, *D. tri*_AmSp_C_vB; Figure 2.2).

The repetitive regions of our Amp sequences (e.g. *D. tri*_AmSp_N_vB), have similarity to previously reported major ampullate spidroin 1 (MaSp1) sequences, such as the presence of the amino acid motifs poly-A, $(GA)_n$, and $(GGX)_n$ (Ayoub and Hayashi, 2008; Xu and Lewis, 1990; Zhang et al., 2013). However, given that our contigs

represent partial-length transcripts, the sequence differences among our variants, and that our spidroin termini trees did not recover a definitive MaSp1 clade, we kept the conservative Amp annotation rather than rename our contigs as MaSp1.

We also searched for minor ampullate spidroins, which are thus far only known from araneoid species and are characterized by long consensus repeats primarily made of $(GX)_n$, A_n, and GGX interrupted by regions called spacers (Chen et al., 2012; Colgin and Lewis, 1998; Vienneau-Hathaway et al., 2017). These MiSp spacers are rich in threonine, serine, and valine, which is distinct from the amino acid composition of the MiSp repetitive region. MiSp spacers also are conserved in sequence and length across species (Chen et al., 2012; Vienneau-Hathaway et al., 2017). We searched our Amp spidroins for MiSp-like repetitive motifs and spacers and although we found a few Amp sequences with threonine and serine rich motifs, none of the *D. triton* Amp sequences contained all of the araneoid MiSp features (Figure 2.2). Furthermore, terminal region phylogenetic analyses show that our Amp sequences group with MaSp sequences instead of the araneoid MiSp clade (Figures 3-4). Future studies could examine whether the various ampullate-shaped silk glands vary in which *D. triton* spidroins they express, despite their morphological homogeneity.

We did not to recover contigs that closely match previously published *Dolomedes tenebrosus* spidroins (named fibroins 1 and 2; Gatesy et al., 2001). But we did find that our *D. tri*_AmSp_C_vA, *D. tri*_AmSp_C_vB, and *D. tri*_AmSp_C_vD cluster with *D. tenebrosus* fibroins 1 and 2 with moderate support (Figure 2.4). The repetitive region of *D. tri*_AmSp_C_vA and *D. tri*_AmSp_C_vB have similarities with the repetitive region

of *D. tenebrosus* fibroin 2. Furthermore, *D. tri*_AmSp_C_vA and *D. tenebrosus* fibroin 2 C-terminal regions have 88% and 83% pairwise identity at the nucleotide and amino acid levels, respectively, suggesting they likely represent closely related paralogs. *D. tri*_AmSp_C_vD has no repetitive region and thus we cannot assess its similarity to the published *D. tenebrosus* fibroin 1 repeat region.

Orphan Spidroin

One of our spidroin transcripts did not cluster with any of the traditionally named spidroin clades (Figure 2.3) and is likely to represent a new spidroin type (Figure 2.2). The novel sequence is a *D. triton* N-terminal region transcript that, in the phylogenetic analysis, is placed as the sister-group of the PySp clade (*D. tri*_Sp_N; Figure 2.3). However, the support for this placement is low (57 %) and the repetitive sequence of this transcript is divergent from that of PySp. Thus, we were unable to assign a spidroin type to this sequence. Hence, we gave it the general name of "spidroin" only (abbreviated as "Sp" in *D. tri*_Sp_N).

Differential gene expression of D. triton

D. triton sex-specific, total silk gland RNAseq libraries (combinations of all silk glands in individual spiders; note that males lack tubuliform silk glands) were each mapped to our *D. triton* transcriptome. To avoid possible over- or under-counting due to the incompletely known spidroin repetitive regions (Chaw et al., 2016), spidroin expression levels were estimated from reads mapped to the C-terminal regions only. We

found that ampullate spidroins, particularly *D. tri_AmSp_C_vA* and *D. tri _AmSp_C_vB*, were highly expressed in both sexes compared to the other spidroin genes (Figure 2.5). This finding is consistent with *D. triton* ampullate spidroins being components of their draglines, which are needed to move from one location to another throughout a fishing spider's life.

To investigate sex-biased gene expression of other genes besides spidroins, we compared transcript abundance between silk gland tissues of males and females. We found 203 transcripts to be significantly differentially expressed (DE) in *D. triton* spiders (Figure 2.6; Supplementary File 1; adjusted p-values < 0.1). Of these, 138 (68%) were overexpressed in females and 65 (32%) were overexpressed in males. Using sequence similarity to UniProt full proteins to predict gene function, we found multiple differentially expressed genes (88 out of 203) that were assigned GO terms related to cellular upkeep and function (e.g., cellular component, molecular function, and biological process; Supplementary File 1).

Overexpressed transcripts in female spiders included the *TuSp* genes, as expected, because *D. triton* females actively produce TuSp proteins for egg case production (males do not). Male *D. triton* spiders lack tubuliform glands, the silk glands that produce fibers used in egg sac construction by female spiders (Foelix, 2011; Murphy & Roberts, 2015; Zhang et al., 2004). However, *TuSp* silk genes were found to be expressed in *D. triton* males, albeit 2.5 times lower than in conspecific females (Figure 2.5). Marginal expression of *TuSp* transcripts has also been previously documented in male Theridiidae, explained as possibly due to incomplete transcriptional shutdown and silk gene co-

expression at low levels (Correa-Garhwal et al., 2017). Overexpressed transcripts in *D. triton* males included the *Amp* genes, suggesting that, compared to females, males are producing a greater proportion of silks composed of ampullate spidroins. This could be related to the males having a roving lifestyle as they search for females.

Our DE results show unevenness in favor of female sequences (138 female vs. 65 male sequences). Difference in number of DE genes between sexes has also been reported in crabs (Liu et al., 2015), salmon (Farlora et al., 2014), freshwater fish (Hale et al., 2010), and red abalone (Valenzuela-Muñoz et al., 2014), and is thought to be associated to reproductive processes such as egg production. Future studies of sex-biased gene expression should include reproductive organs and other tissue types to identify which of the genes that are differentially expressed in silk glands might be more generally implicated in sex-dependent expression.

Proteomics of D. triton silk glands and egg cases

Peptide sequencing of the silk dope from tubuliform and ampullate silk glands of *D. triton* females provided us with empirical evidence for 342 proteins predicted from our transcriptome assembly. Seventy-five percent of the identified proteins were present in both gland types, with the majority (219 out of 255) annotated with GO terms (Supplementary File 2). The functions of these shared proteins suggest heavy involvement in core cellular functions and metabolism (e.g., vesicles and biological macromolecule synthesis). We also identified proteins that were unique to each silk gland type: 49 ampullate-specific and 38 tubuliform-specific proteins. 61% and 79% of the

gland specific proteins, ampullate and tubuliform gland respectively, were annotated with GO terms associated with cell functioning and protein biosynthesis (Supplementary File 2). Following our expectations based on sequence motifs and gene tree analyses (Figures 2-4), spidroins annotated as Amp were detected in ampullate glands, and spidroins annotated as TuSp were detected in tubuliform glands. Surprisingly, AcSp was also observed in tubuliform glands. Recent proteomic studies have detected AcSp proteins in the major ampullate glands of the cob-web weaver *Latrodectus hesperus* but not in tubuliform glands (Chaw et al., 2015; Larracas et al., 2016). *AcSp* gene expression has also been detected in the aggregate, flagelliform, and minor ampullate silk glands of the orb-web *Nephila clavipes*, but not the tubuliform glands (Babb et al., 2017). Our study provides further evidence for species-specific variation in co-expression patterns of spidroin types. Altering gene expression levels could be a way for spiders to modify silk physical properties.

D. triton spiders are one of the only spider species known to submerge themselves in water while carrying egg sacs (McAlister, 1960). Microscopy of *D. triton* egg sacs shows the complexity of the submersible egg sac (Figure 2.7). We identified two distinct layers present in *D. triton* egg sacs: an inner layer that is loosely woven with two different sized diameter fibers and an outer layer that is tightly woven, also with two different sized diameter fibers as well as an unknown but prevalent coating. To investigate the uniqueness of these layers, we compared micrographs of the egg sac layers of *D. triton* with the corresponding layers of egg sacs from the golden orb-weaver *Nephila clavipes*, the Western black widow *L. hesperus*, and the silver garden spider

Argiope argentata. Scanning electron microscope (SEM) micrographs reveal considerable variation in the composition and architecture of each egg sac (Figure S2.2). Egg sacs from all species show the presence of large diameter fibers in the inner and outer layers; most likely, these fibers emerge from tubuliform glands. We found *D. triton*, *A. argentata*, and *L. hesperus* egg sacs to have a second fiber type that is much narrower in diameter.

To identify the protein components of *D. triton* egg sac layers, we analyzed the two distinct layers seen in the SEM micrographs (Figure 2.7). In total, we found evidence for 40 proteins predicted from our transcriptome (Supplementary File 3). The majority of the identified proteins, such as tubuliform spidroin (TuSp), were detected on both the outer and the inner layers. Finding tubuliform spidroin (TuSp) proteins between the two layers is consistent with both layers having large diameter, tubuliform, silk fibers (Figure 2.7). Among the proteins unique to the outer layer, we found aciniform spidroin (*D. tri*_AcSp_N_vB and *D. tri*_AcSp_C_vB). *D. tri*_AcSp_N_vB is the same spidroin found in the tubuliform gland proteome (Supplementary File 2). Aciniform spidroin is also expected to be abundant in aciniform silk, and aciniform silk has been described as being incorporated in egg sacs of *L. hesperus* (Vasanthavada et al., 2007). The presence of small diameter silk fibers observed in the microscopy analyses suggest that aciniform silk is used in the egg sacs of *D. triton*. Although the bulk of *D. triton* egg sac is mostly composed of tubuliform silk fibers, our study shows that egg sacs are a mixture of large and small diameter fibers (likely tubuliform and aciniform).

Three predicted proteins identified in peptide sequences from *D. triton* egg sac are homologous to peroxidases. Peroxidases belong to a family of heme-containing enzymes that are involved in diverse biological roles such as peroxidase-catalyzed protein crosslinking in barnacle cement (So et al., 2017, 2016) and mosquito egg chorion layer (Li et al., 1996; Li and Li, 2006). A peroxidase has also been identified in the major and minor ampullate silk glands of the orb-weaver *Nephila senegalensis* and this enzyme (NsPox) is thought to contribute to the initiation of disulfide links in outer layer silk proteins during fiber processing (Pouchkina et al., 2003; Vollrath and Knight, 1999). Maximum likelihood analysis of *D. triton,* other spider peroxidases, and non-spider representatives of heme-containing peroxidase families (peroxidasin, peroxinectin, and myeloperoxidases) show that they cluster with other spider peroxidases (Supplementary Figure S2.3). Since the spider peroxidases do not fall into any of the named heme-containing peroxidase families, we simply refer to each of our transcripts as a "peroxidase" (Pox), followed by a letter for each variant (v) type (e.g., *D. tri*_Pox_vA, *D. tri*_ Pox_vB, *D. tri*_ Pox_vC).

*D. tri*_Pox_vA was the only *D. triton* peroxidase also identified by peptide sequencing of protein extracts of the ampullate shaped silk glands (Supplementary File 2). In the gene tree analysis, this *D. triton* Pox belongs to a clade of only spider peroxidases, including *S. mimosarum* peroxidasin and NsPox (Supplementary Figure S2.3). Thus, *D. tri*_Pox_vA may be an ampullate gland specific peroxidase (Pouchkina et al., 2003; Supplementary Files 2 and 3). By contrast, *D. tri*_Pox_vB and *D. tri*_Pox_vC were not detected in the peptide sequences from ampullate glands, nor from tubuliform

silk glands, suggesting that these peroxidases may be expressed in the other silk glands (aciniform or pyriform silk glands). Hence, the different spider peroxidases appear to have a silk gland-type specific expression.

D. triton egg sacs and potential aquatic specializations

Energy dispersive X-ray spectrometry (EDS) was used to obtain information regarding the elemental composition of *D. triton* egg cases by comparing the outer and inner layer of each egg sac (Figure 2.7; Figures S4-S5). We found calcium and phosphorus to be present in coated the outer layer of *D. triton* egg sacs. Since these elements have not been previously described from other spider silks, we also analyzed egg sacs from other spider species by EDS (Figure S2.5). We found no significant traces of calcium or phosphorus in fully terrestrial spider egg sacs (*N. clavipes*, *L. hesperus*, *A. argentata*), suggesting that these elements may be utilized in the coating of *D. triton* egg sacs. Intriguingly, aquatic caddisfly larvae silk fibers, which do not have a coating, have high levels of these elements (*Brachycentrus echo*; Stewart and Wang, 2010). Thus, EDS of *D. triton* egg sacs revealed differences with the egg sacs from other spider species (all non-aquatic) and similarities to silks from an aquatic caddisworm (Trichoptera).

To further investigate *D. triton* egg sac specializations to aquatic environments, its water-repellent properties were investigated. Hydrophobicity, the property of a surface to repel water, was quantified using contact angle (Callies and Quere, 2005; Morris et al., 2012). Contact angle is the angle formed by the intersection of the liquid-solid interface (Zisman, 1964). Surfaces are considered hydrophobic if the contact angle is 90-120° or

super-hydrophobic if the contact angle is >120° (Callies and Quere, 2005; Feng et al., 2002). We found that the *D. triton* egg sac is demonstrably hydrophobic, with a contact angle of 116.7±14.5° (N=5, Figure S2.6). To establish whether hydrophobicity was specific to *D. triton* egg sacs or a general property of spider egg sacs, *L. hesperus* egg sacs were also tested using the same methods. The contact angle of *L. hesperus* egg sacs was 95.6±15.4° (N=6), twenty degrees less than that of *D. triton* egg sacs, but still categorized as hydrophobic (Vetter et al., 2016). This indicates that hydrophobicity is not unique to *D. triton* egg sacs, and is instead a more general feature of spider egg sacs. The higher contact angle found in the *D. triton* egg sac compared to *L.hesperus* could be attributed to the coating found on the *D. triton* egg sac (Figure 2.7C; Figure S2.4). The coating could function as a semi-permeable membrane, keeping egg sacs buoyant and the eggs dry when submerged, yet allowing for gas exchange by the developing eggs when on land.

Conclusions

We identified 18 new spidroin contigs from the fishing spider *Dolomedes triton* (Figures 2.1-2.2). We confirmed our prediction based on spigot morphology that *D. triton* spiders of both sexes should express genes associated with aciniform, ampullate, and pyriform silk glands, and females should also express genes associated with tubuliform (egg case) silk glands (Figure 2.5). Although *D. triton* spidroins show diversity in repeat composition (Figure 2.2 and Figure S2.1), sequence comparison shows no obvious evidence that *D. triton* spidroins have unique modifications for semi-aquatic

environments. It appears that ancestral spider silks already had molecular and physical properties that supported survival in aquatic and semi-aquatic environments, prior to the evolution of fishing spiders. Thus, *D. triton* spidroins could be considered pre-adapted to semi-aquatic environments. However, we found that *D. triton* egg sacs contained calcium and phosphorus. Calcium and phosphorus have also been noted as present in aquatic insect silk fibers but not the egg sacs of non-aquatic spiders. These elements may contribute to the water-repellant properties of fishing spider egg sacs (Figure S2.5).

Figures

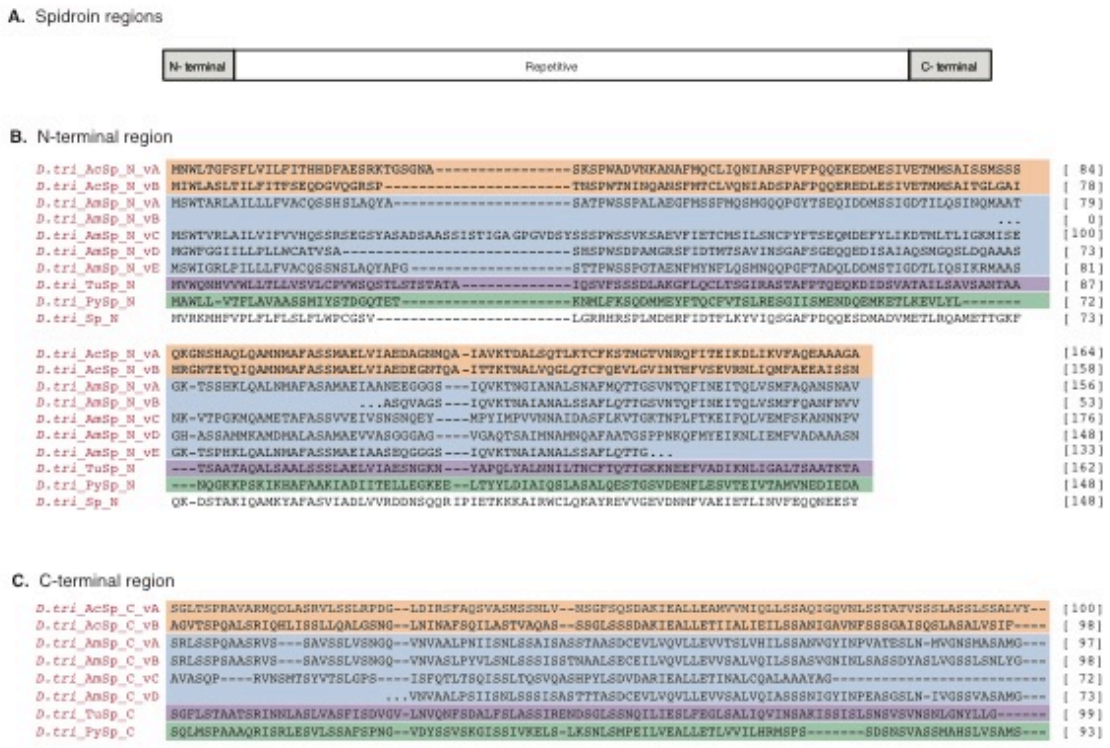


Figure 2.1. Conserved regions of spidroins. **A.** Schematic organization of a spidroin primary structure. **B.** Multiple sequence alignment of amino (N)-terminal regions. **C.** Multiple sequence alignment of carboxyl (C)-terminal regions. Spidroin names abbreviated as in Table S2.2. Shaded boxes indicate spidroin type sequences annotated as AcSp (orange), Amp (blue), TuSp (purple), and PySp (green). Ellipses indicate missing upstream or downstream sequence Dashes are alignment gaps. Amino acid positions for each sequence are numbered on the right.

Figure 2.2. Spidroin repetitive sequences. **A.** Repetitive sequence adjacent to N-terminal region. **B.** Repetitive sequence adjacent to C-terminal region. Shaded boxes indicate spidroins annotated as in Figure 2.1. Spidroin names abbreviated as in Table S2.2. Amino acids abundant in silks are highlighted: alanine (red), serine (blue), and glycine (green).

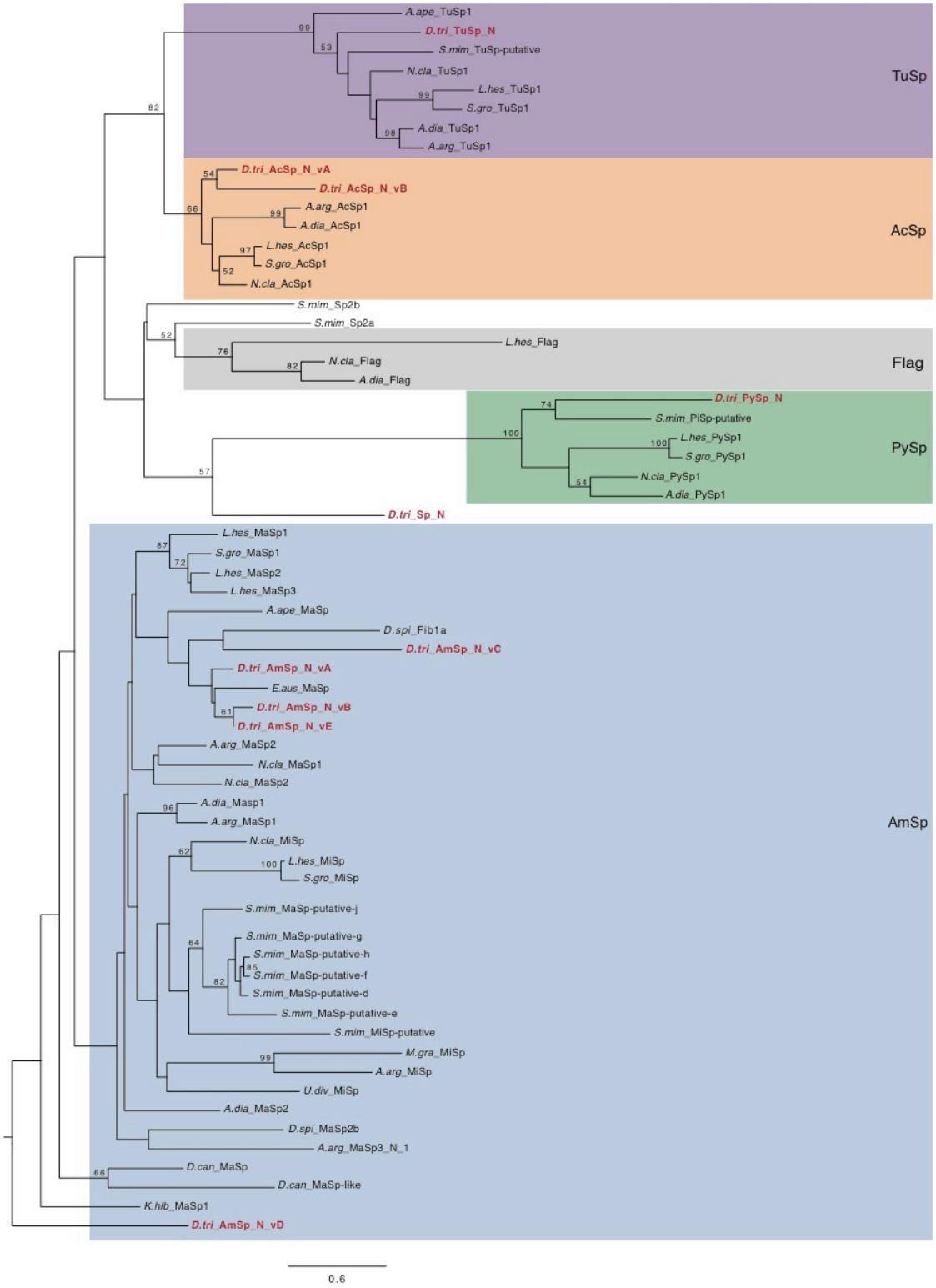

A.ape_TuSp1
D.tri_TuSp_N
S.mim_TuSp-putative
N.cla_TuSp1
L.hes_TuSp1
S.gro_TuSp1
A.dia_TuSp1
A.arg_TuSp1
TuSp
D.tri_AcSp_N_vA
D.tri_AcSp_N_vB
A.arg_AcSp1
A.dia_AcSp1
L.hes_AcSp1
S.gro_AcSp1
N.cla_AcSp1
AcSp
S.mim_Sp2b
S.mim_Sp2a
L.hes_Flag
N.cla_Flag
A.dia_Flag
Flag
D.tri_PySp_N
S.mim_PiSp-putative
L.hes_PySp1
S.gro_PySp1
N.cla_PySp1
A.dia_PySp1
PySp
D.tri_Sp_N
L.hes_MaSp1
S.gro_MaSp1
L.hes_MaSp2
L.hes_MaSp3
A.ape_MaSp
D.spi_Fib1a
D.tri_AmSp_N_vC
D.tri_AmSp_N_vA
E.aus_MaSp
D.tri_AmSp_N_vB
D.tri_AmSp_N_vE
A.arg_MaSp2
N.cla_MaSp1
N.cla_MaSp2
A.dia_Masp1
A.arg_MaSp1
N.cla_MiSp
L.hes_MiSp
S.gro_MiSp
S.mim_MaSp-putative-j
S.mim_MaSp-putative-g
S.mim_MaSp-putative-h
S.mim_MaSp-putative-f
S.mim_MaSp-putative-d
S.mim_MaSp-putative-e
S.mim_MiSp-putative
M.gra_MiSp
A.arg_MiSp
U.div_MiSp
A.dia_MaSp2
D.spi_MaSp2b
A.arg_MaSp3_N_1
D.can_MaSp
D.can_MaSp-like
K.hib_MaSp1
D.tri_AmSp_N_vD
AmSp
0.6

Figure 2.3. Maximum likelihood tree of spidroin N-terminal regions. *Dolomedes triton* spidroin paralogs highlighted in red. Tree rooted with the California trapdoor spider *Bothriocyrtum californicum* fibroin 1 (not shown). Shared boxes indicate spidroin types as in Figure 2.1, with the addition of flag (gray). Names abbreviated as in Tables S2.2-S2.3. Bootstrap percentages >50% are shown. Scale bar represents substitutions per site.

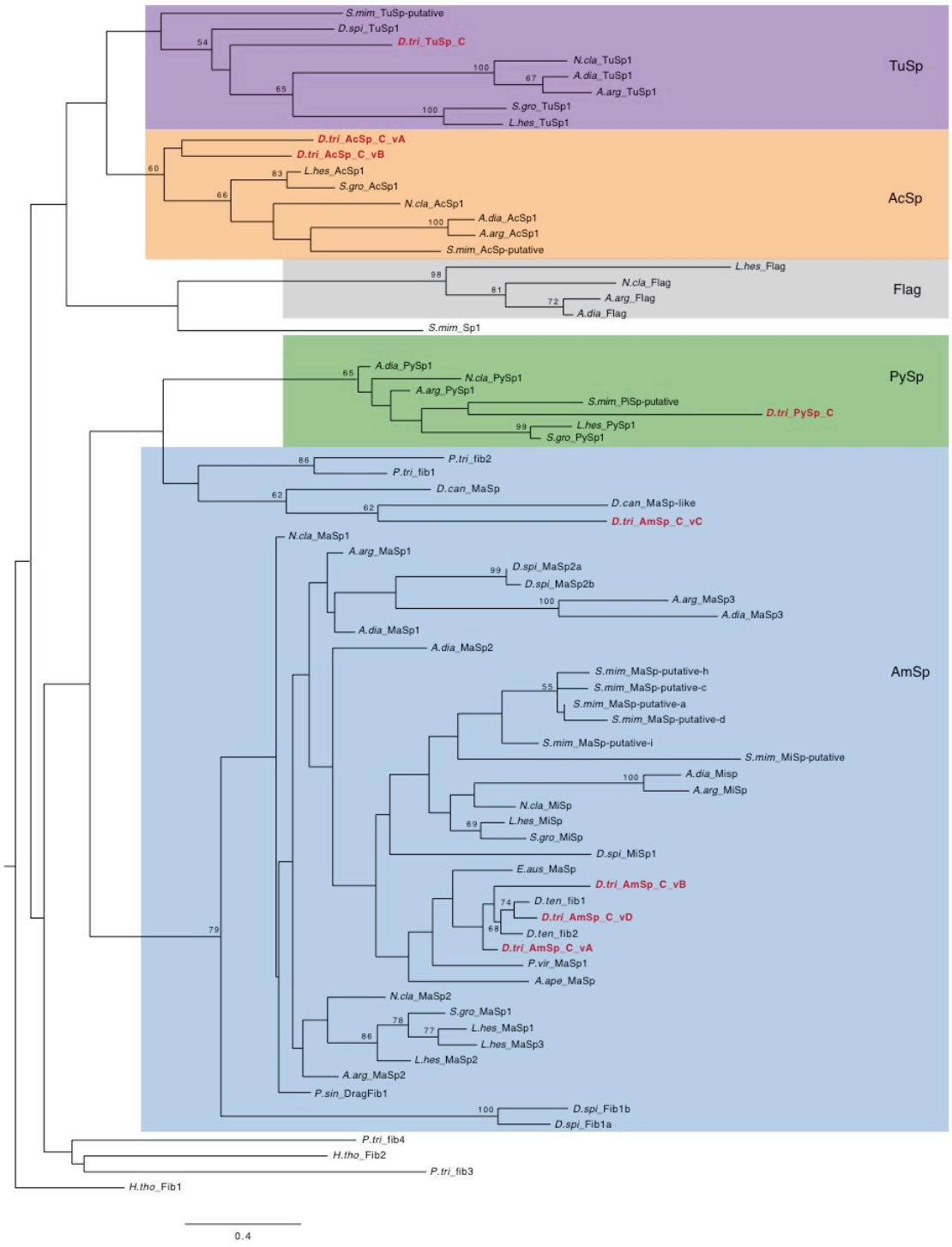

TuSp
AcSp
Flag
PySp
AmSp
S.mim_TuSp-putative
D.spi_TuSp1
D.tri_TuSp_C
N.cla_TuSp1
A.dia_TuSp1
A.arg_TuSp1
S.gro_TuSp1
L.hes_TuSp1
D.tri_AcSp_C_vA
D.tri_AcSp_C_vB
L.hes_AcSp1
S.gro_AcSp1
N.cla_AcSp1
A.dia_AcSp1
A.arg_AcSp1
S.mim_AcSp-putative
L.hes_Flag
N.cla_Flag
A.arg_Flag
A.dia_Flag
S.mim_Sp1
A.dia_PySp1
N.cla_PySp1
A.arg_PySp1
S.mim_PiSp-putative
D.tri_PySp_C
L.hes_PySp1
S.gro_PySp1
P.tri_fib2
P.tri_fib1
D.can_MaSp
D.can_MaSp-like
D.tri_AmSp_C_vC
N.cla_MaSp1
A.arg_MaSp1
D.spi_MaSp2a
D.spi_MaSp2b
A.arg_MaSp3
A.dia_MaSp3
A.dia_MaSp1
A.dia_MaSp2
S.mim_MaSp-putative-h
S.mim_MaSp-putative-c
S.mim_MaSp-putative-a
S.mim_MaSp-putative-d
S.mim_MaSp-putative-i
S.mim_MiSp-putative
A.dia_Misp
A.arg_MiSp
N.cla_MiSp
L.hes_MiSp
S.gro_MiSp
D.spi_MiSp1
E.aus_MaSp
D.tri_AmSp_C_vB
D.ten_fib1
D.tri_AmSp_C_vD
D.ten_fib2
D.tri_AmSp_C_vA
P.vir_MaSp1
A.ape_MaSp
N.cla_MaSp2
S.gro_MaSp1
L.hes_MaSp1
L.hes_MaSp3
L.hes_MaSp2
A.arg_MaSp2
P.sin_DragFib1
D.spi_Fib1b
D.spi_Fib1a
P.tri_fib4
H.tho_Fib2
P.tri_fib3
H.tho_Fib1
0.4

Figure 2.4. Maximum likelihood tree of spidroin C-terminal regions. *Dolomedes triton* spidroin paralogs highlighted in red. Tree rooted with the California trapdoor spider *Bothriocyrtum californicum* fibroin 1 (not shown). Shared boxes indicate spidroin types as in Figure 2.1, with the addition of flag (gray). Names abbreviated as in Tables S2.2-S2.3. Bootstrap percentages >50% are shown. Scale bar represents substitutions per site.

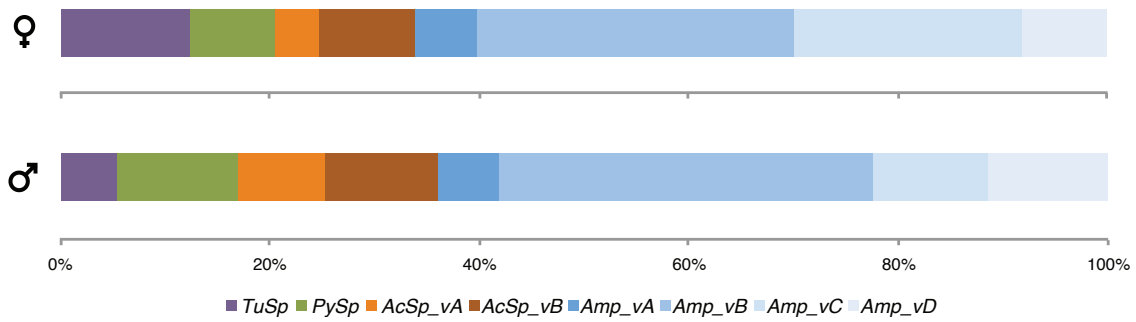


Figure 2.5. Stacked bar graphs of relative spidroin gene expression levels, based on C-terminal coding regions. Total silk gland tissue from *Dolomedes triton* females (top) and males (bottom). Spidroin genes abbreviated as in Table S2.2, except that the designation "_C" for carboxy-terminal region removed from all names. Percentages show average expression of normalized reads per kilobase of transcript per million mapped reads (RPKM).

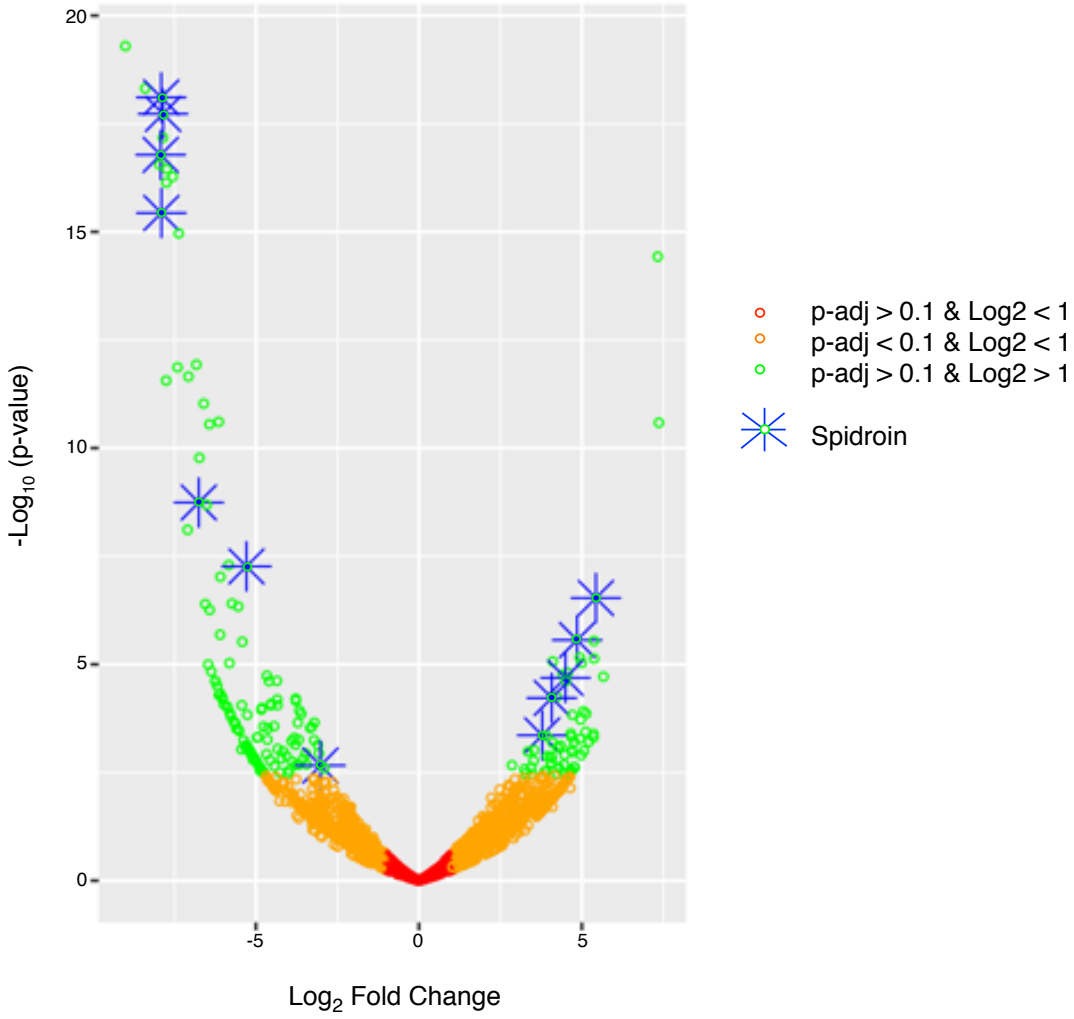


Figure 2.6. Volcano plot of the significance level of gene expression differences between the silk glands of *Dolomedes triton* females and males. Each point represents an individual gene. Differentially expressed genes are highlighted in green (genes with $\log_2$FC >1 and p-adjusted value < 0.1). Differentially expressed spidroin genes are highlighted with blue stars with green centers

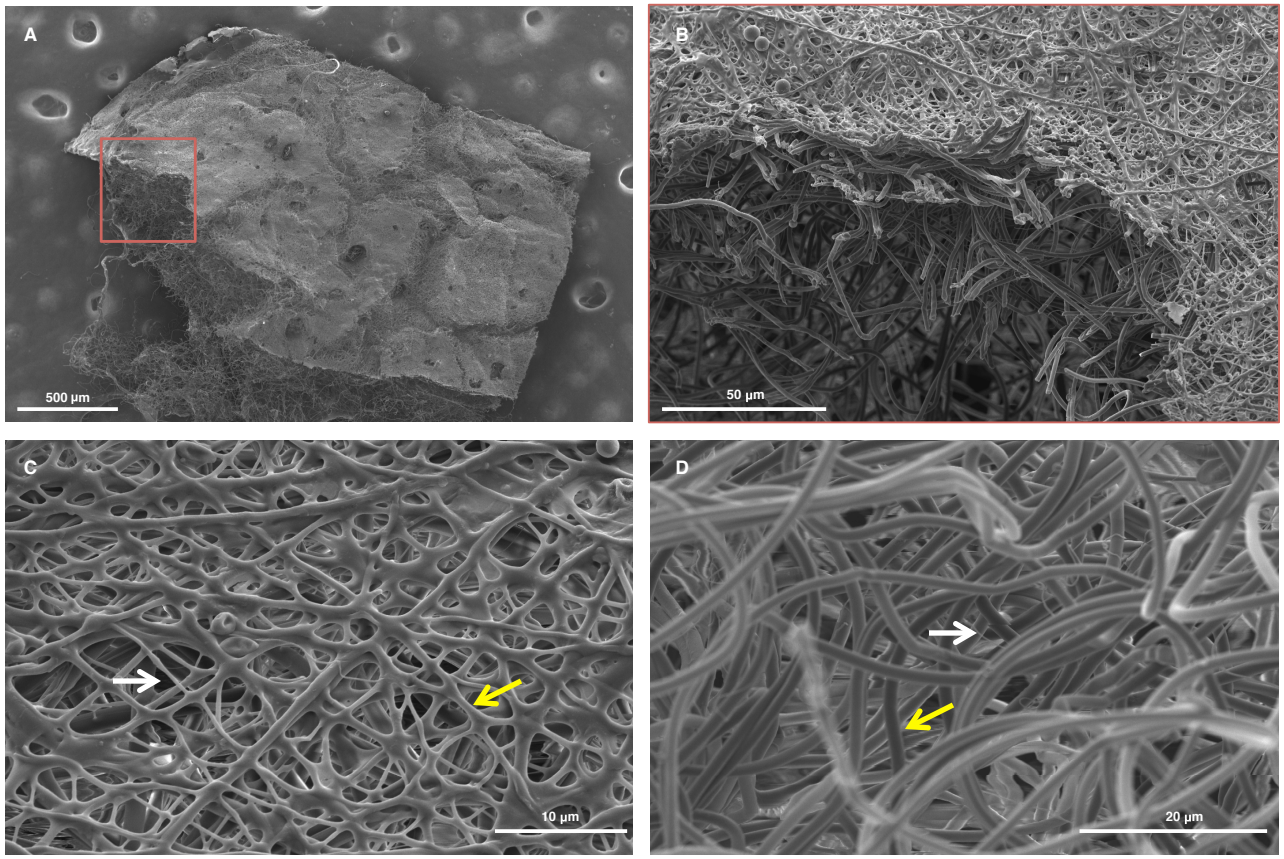


Figure 2.7. Scanning electron micrographs of a *Dolomedes triton* egg sac. **A.** Large area micrograph of *D. triton* egg sac, red square indicates the area shown in **B.** **B.** Outer and inner layers identified in *D. triton* egg sac. **C.** Egg sac outer layer showing large (yellow arrow) and small (white arrow) diameter fibers with copious amounts of coating. **D.** Egg sac inner layer showing predominantly large diameter, tubuliform, silk fibers.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Ayoub, N.A., Garb, J.E., Kuelbs, A., Hayashi, C.Y., 2013. Ancient properties of spider silks revealed by the complete gene sequence of the prey-wrapping silk protein (AcSp1). Mol. Biol. Evol. 30, 589–601. https://doi.org/10.1093/molbev/mss254

Ayoub, N.A., Garb, J.E., Tinghitella, R.M., Collin, M.A., Hayashi, C.Y., 2007. Blueprint for a high-performance biomaterial: full-length spider dragline silk genes. PloS One 2, e514. https://doi.org/10.1371/journal.pone.0000514

Ayoub, N.A., Hayashi, C.Y., 2008. Multiple recombining loci encode MaSp1, the primary constituent of dragline silk, in widow spiders (*Latrodectus*: Theridiidae). Mol. Biol. Evol. 25, 277–286. https://doi.org/10.1093/molbev/msm246

Babb, P.L., Lahens, N.F., Correa-Garhwal, S.M., Nicholson, D.N., Kim, E.J., Hogenesch, J.B., Kuntner, M., Higgins, L., Hayashi, C.Y., Agnarsson, I., Voight, B.F., 2017. The *Nephila clavipes* genome highlights the diversity of spider silk genes and their complex expression. Nat. Genet. 49, 895–903.

Blackledge, T.A., Hayashi, C.Y., 2006. Silken toolkits: biomechanics of silk fibers spun by the orb web spider Argiope argentata (Fabricius 1775). J. Exp. Biol. 209, 2452–2461. https://doi.org/10.1242/jeb.02275

Blasingame, E., Tuton-Blasingame, T., Larkin, L., Falick, A.M., Zhao, L., Fong, J., Vaidyanathan, V., Visperas, A., Geurts, P., Hu, X., others, 2009. Pyriform spidroin 1, a novel member of the silk gene family that anchors dragline silk fibers in attachment discs of the black widow spider, *Latrodectus hesperus*. J. Biol. Chem. 284, 29097–29108.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Callies, M., Quere, D., 2005. On water repellency. Soft Matter 55–61.

Chaw, R.C., Arensburger, P., Clarke, T.H., Ayoub, N.A., Hayashi, C.Y., 2016. Candidate egg case silk genes for the spider *Argiope argentata* from differential gene expression analyses. Insect Mol. Biol. n/a-n/a. https://doi.org/10.1111/imb.12260

Chaw, R.C., Correa-Garhwal, S.M., Clarke, T.H., Ayoub, N.A., Hayashi, C.Y., 2015. Proteomic evidence for components of spider silk synthesis from black widow silk glands and fibers. J. Proteome Res. 14, 4223–4231. https://doi.org/10.1021/acs.jproteome.5b00353

Chaw, R.C., Saski, C.A., Hayashi, C.Y., 2017. Complete gene sequence of spider attachment silk protein (PySp1) reveals novel linker regions and extreme repeat homogenization. Insect Biochem. Mol. Biol. 81, 80–90. https://doi.org/10.1016/j.ibmb.2017.01.002

Chaw, R.C., Zhao, Y., Wei, J., Ayoub, N.A., Allen, R., Atrushi, K., Hayashi, C.Y., 2014. Intragenic homogenization and multiple copies of prey-wrapping silk genes in *Argiope* garden spiders. BMC Evol. Biol. 14, 31. https://doi.org/10.1186/1471-2148-14-31

Chen, G., Liu, X., Zhang, Y., Lin, S., Yang, Z., Johansson, J., Rising, A., Meng, Q., 2012. Full-length minor ampullate spidroin gene sequence. PLoS ONE 7, e52293. https://doi.org/10.1371/journal.pone.0052293

Clarke, T.H., Garb, J.E., Hayashi, C.Y., Arensburger, P., Ayoub, N.A., 2015. Spider transcriptomes identify ancient large-scale gene duplication event potentially important in silk gland evolution. Genome Biol. Evol. 7, 1856–1870.

Clarke, T.H., Garb, J.E., Hayashi, C.Y., Haney, R.A., Lancaster, A.K., Corbett, S., Ayoub, N.A., 2014. Multi-tissue transcriptomics of the black widow spider reveals expansions, co-options, and functional processes of the silk gland gene toolkit. BMC Genomics 15. https://doi.org/10.1186/1471-2164-15-365

Colgin, M.A., Lewis, R.V., 1998. Spider minor ampullate silk proteins contain new repetitive sequences and highly conserved non-silk-like “spacer regions”. Protein Sci. Publ. Protein Soc. 7, 667–672.

Correa-Garhwal, S.M., Chaw, R.C., Clarke, T.H., Ayoub, N.A., Hayashi, C.Y., 2017. Silk gene expression of theridiid spiders: implications for male-specific silk use. Zoology 122, 107–114. https://doi.org/10.1016/j.zool.2017.04.003

Correa-Garhwal, S.M., Garb, J.E., 2014. Diverse formulas for spider dragline fibers demonstrated by molecular and mechanical characterization of spitting spider silk. Biomacromolecules 15, 4598–4605.

Farlora, R., Araya-Garay, J., Gallardo-Escárate, C., 2014. Discovery of sex-related genes through high-throughput transcriptome sequencing from the salmon louse Caligus rogercresseyi. Mar. Genomics 15, 85–93. https://doi.org/10.1016/j.margen.2014.02.005

Feng, L., Li, S., Li, Y., Li, H., Zhang, L., Zhai, J., Song, Y., Liu, B., Jiang, L., Zhu, D., 2002. Super-Hydrophobic Surfaces: From Natural to Artificial. Adv. Mater. 14, 1857–1860. https://doi.org/10.1002/adma.200290020

Foelix, R., 2011. Biology of Spiders. Oxford University Press.

Garb, J.E., Ayoub, N.A., Hayashi, C.Y., 2010. Untangling spider silk evolution with spidroin terminal domains. BMC Evol. Biol. 10, 243. https://doi.org/10.1186/1471-2148-10-243

Gaskett, A.C., 2007. Spider sex pheromones: emission, reception, structures, and functions. Biol. Rev. 82, 27–48. https://doi.org/10.1111/j.1469-185X.2006.00002.x

Gatesy, J., Hayashi, C., Motriuk, D., Woods, J., Lewis, R., 2001. Extreme diversity, conservation, and convergence of spider silk fibroin sequences. Science 291, 2603–2605. https://doi.org/10.1126/science.1057561

Geurts, P., Zhao, L., Hsia, Y., Gnesa, E., Tang, S., Jeffery, F., Mattina, C.L., Franz, A., Larkin, L., Vierra, C., 2010. Synthetic spider silk fibers spun from pyriform spidroin 2, a glue silk protein discovered in orb-weaving spider attachment discs. Biomacromolecules 11, 3495–3503. https://doi.org/10.1021/bm101002w

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., 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

Griswold, C.E., 2005. Atlas of Phylogenetic Data for Entelegyne Spider (Araneae: Araneomorphae: Entelegynae) with Comments on Their Phylogeny.

Guerette, P.A., Ginzinger, D.G., Weber, B.H.F., Gosline, J.M., 1996. Silk properties determined by gland-specific expression of a spider fibroin gene family. Science 272, 112–115. https://doi.org/10.1126/science.272.5258.112

Hale, M.C., Jackson, J.R., DeWoody, J.A., 2010. Discovery and evaluation of candidate sex-determining genes and xenobiotics in the gonads of lake sturgeon (*Acipenser fulvescens*). Genetica 138, 745–756. https://doi.org/10.1007/s10709-010-9455-y

Hayashi, C.Y., Blackledge, T.A., Lewis, R.V., 2004. Molecular and mechanical characterization of aciniform silk: uniformity of iterated sequence modules in a

novel member of the spider silk fibroin gene family. Mol. Biol. Evol. 21, 1950–1959. https://doi.org/10.1093/molbev/msh204

Herberstein, M.E., Craig, C.L., Coddington, J.A., Elgar, M.A., 2000. The functional significance of silk decorations of orb-web spiders: a critical review of the empirical evidence. Biol. Rev. 75, 649–669.

Hinman, M.B., Lewis, R.V., 1992. Isolation of a clone encoding a second dragline silk fibroin. *Nephila clavipes* dragline silk is a two-protein fiber. J. Biol. Chem. 267, 19320–19324.

Huang, X., Madan, A., 1999. CAP3: A DNA sequence assembly program. Genome Res. 9, 868–877.

Jeffery, F., La Mattina, C., Tuton-Blasingame, T., Hsia, Y., Gnesa, E., Zhao, L., Franz, A., Vierra, C., 2011. Microdissection of black widow spider silk-producing glands. J. Vis. Exp. JoVE.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., Salzberg, S.L., 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14, R36.

Kovoor, J., Zylberberg, L., 1982. Fine structural aspects of silk secretion in a spider. II. Conduction in the pyriform glands. Tissue Cell 14, 519–530. https://doi.org/10.1016/0040-8166(82)90044-1

Kovoor, J., Zylberberg, L., 1980. Fine structural aspects of silk secretion in a spider (*Araneus diadematus*). I. Elaboration in the pyriform glands. Tissue Cell 12, 547–556. https://doi.org/10.1016/0040-8166(80)90044-0

Lang, A., 1996. Silk Investment in Gifts By Males of the Nuptial Feeding Spider *Pisaura Mirabilis* (Araneae: Pisauridae). Behaviour 133, 697–716. https://doi.org/10.1163/156853996X00431

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948. https://doi.org/10.1093/bioinformatics/btm404

Larracas, C., Hekman, R., Dyrness, S., Arata, A., Williams, C., Crawford, T., Vierra, C.A., 2016. Comprehensive proteomic analysis of spider dragline silk from black widows: a recipe to build synthetic silk fibers. Int. J. Mol. Sci. 17, 1537. https://doi.org/10.3390/ijms17091537

Lawrence, B.A., Vierra, C.A., Moore, A.M., 2004. Molecular and mechanical properties of major ampullate silk of the black widow spider, *Latrodectus hesperus*. Biomacromolecules 5, 689–695.

Li, J., Hodgeman, B.A., Christensen, B.M., 1996. Involvement of peroxidase in chorion hardening in *Aedes aegypti*. Insect Biochem. Mol. Biol. 26, 309–317. https://doi.org/10.1016/0965-1748(95)00099-2

Li, J.S., Li, J., 2006. Major chorion proteins and their crosslinking during chorion hardening in *Aedes aegypti mosquitoes*. Insect Biochem. Mol. Biol. 36, 954–964. https://doi.org/10.1016/j.ibmb.2006.09.006

Liu, Y., Hui, M., Cui, Z., Luo, D., Song, C., Li, Y., Liu, L., 2015. Comparative transcriptome analysis reveals sex-biased gene expression in juvenile chinese mitten crab *Eriocheir sinensis*. PLOS ONE 10, e0133068. https://doi.org/10.1371/journal.pone.0133068

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. https://doi.org/10.1186/s13059-014-0550-8

McAlister, W.H., 1960. The maintenance of the sternal egg sac position by pisaurid spiders. Am. Midl. Nat. 63, 149–151.

Moon, M.-J., 2008. Fine structure of the silk spigots in the spider *Dolomedes sulfureus* (Araneae: Pisauridae). Appl. Microsc. 38, 89–96.

Morris, R.H., Atherton, S., Shirtcliffe, N.J., McHale, G., Dias, T., Newton, M.I., 2012. Hydrophobic smart material for water transport and collection, in: Breedon, P. (Ed.), Smart Design. Springer London, pp. 49–55.

Motriuk-Smith, D., Smith, A., Hayashi, C.Y., Lewis, R.V., 2005. Analysis of the conserved n-terminal domains in major ampullate spider silk proteins. Biomacromolecules 6, 3152–3159. https://doi.org/10.1021/bm050472b

Mullen, G.R., 1969. Morphology and histology of the silk glands in *Araneus sericatus* cl. Trans. Am. Microsc. Soc. 88, 232–240. https://doi.org/10.2307/3224495

Murphy, J.A., Roberts, M.J., 2015. Spider families of the world and their spinnerets. British arachnological society.

Parra, G., Bradnam, K., Korf, I., 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics 23, 1061–1067. https://doi.org/10.1093/bioinformatics/btm071

Perry, D.J., Bittencourt, D., Siltberg-Liberles, J., Rech, E.L., Lewis, R.V., 2010. Piriform spider silk sequences reveal unique repetitive elements. Biomacromolecules 11, 3000–3006.

Peters, H.M., 1955. Über den Spinnapparat von *Nephila madagascariensis* (Radnetzspinnen, Fam. Argiopidae). Z. Für Naturforschung B 10, 395–404.

Pouchkina, N.N., Stanchev, B.S., McQueen-Mason, S.J., 2003. From EST sequence to spider silk spinning: identification and molecular characterisation of *Nephila senegalensis* major ampullate gland peroxidase NsPox. Insect Biochem. Mol. Biol. 33, 229–238. https://doi.org/10.1016/S0965-1748(02)00207-2

Rising, A., Johansson, J., Larson, G., Bongcam-Rudloff, E., Engström, W., Hjälm, G., 2007. Major ampullate spidroins from Euprosthenops australis: multiplicity at protein, mRNA and gene levels. Insect Mol. Biol. 16, 551–561. https://doi.org/10.1111/j.1365-2583.2007.00749.x

Roland, C., Rovner, J.S., 1983. Chemical and vibratory communication in the aquatic pisaurid spider *Dolomedes Triton*. J. Arachnol. 11, 77–85.

Sanggaard, K.W., Bechsgaard, J.S., Fang, X., Duan, J., Dyrlund, T.F., Gupta, V., Jiang, X., Cheng, L., Fan, D., Feng, Y., Han, L., Huang, Z., Wu, Z., Liao, L., Settepani, V., Thøgersen, I.B., Vanthournout, B., Wang, T., Zhu, Y., Funch, P., Enghild, J.J., Schauser, L., Andersen, S.U., Villesen, P., Schierup, M.H., Bilde, T., Wang, J., 2014. Spider genomes provide insight into composition and evolution of venom and silk. Nat. Commun. 5, 3765. https://doi.org/10.1038/ncomms4765

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Simmons, A.H., Michal, C.A., Jelinski, L.W., 1996. Molecular Orientation and Two-Component Nature of the Crystalline Fraction of Spider Dragline Silk. Science 271, 84–87.

So, C.R., Fears, K.P., Leary, D.H., Scancella, J.M., Wang, Z., Liu, J.L., Orihuela, B., Rittschof, D., Spillmann, C.M., Wahl, K.J., 2016. Sequence basis of Barnacle Cement Nanostructure is Defined by Proteins with Silk Homology. Sci. Rep. 6.

So, C.R., Scancella, J.M., Fears, K.P., Essock-Burns, T., Haynes, S.E., Leary, D.H., Diana, Z., Wang, C., North, S., Oh, C.S., Wang, Z., Orihuela, B., Rittschof, D., Spillmann, C.M., Wahl, K.J., 2017. Oxidase activity of the barnacle adhesive interface involves peroxide-dependent catechol oxidase and lysyl oxidase enzymes. ACS Appl. Mater. Interfaces 9, 11493–11505.

Stamatakis, A., 2014. Raxml version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Starrett, J., Garb, J.E., Kuelbs, A., Azubuike, U.O., Hayashi, C.Y., 2012. Early events in the evolution of spider silk genes. PloS One 7, e38084. https://doi.org/10.1371/journal.pone.0038084

Stewart, R.J., Wang, C.S., 2010. Adaptation of Caddisfly Larval Silks to Aquatic Habitats by Phosphorylation of H-Fibroin Serines. Biomacromolecules 11, 969–974. https://doi.org/10.1021/bm901426d

Tian, M., Liu, C., Lewis, R., 2004. Analysis of major ampullate silk cDNAs from two non-orb-weaving spiders. Biomacromolecules 5, 657–660. https://doi.org/10.1021/bm034391w

Trancik, J.E., Czernuszka, J.T., Cockayne, D.J.H., Viney, C., 2005. Nanostructural physical and chemical information derived from the unit cell scattering amplitudes of a spider dragline silk. Polymer 46, 5225–5231.

Valenzuela-Muñoz, V., Bueno-Ibarra, M.A., Escárate, C.G., 2014. Characterization of the transcriptomes of Haliotis rufescens reproductive tissues. Aquac. Res. 45, 1026–1040. https://doi.org/10.1111/are.12044

van Beek, J.D., Hess, S., Vollrath, F., Meier, B.H., 2002. The molecular structure of spider dragline silk: Folding and orientation of the protein backbone. Proc. Natl. Acad. Sci. 99, 10266–10271. https://doi.org/10.1073/pnas.152162299

Vasanthavada, K., Hu, X., Falick, A.M., La Mattina, C., Moore, A.M., Jones, P.R., Yee, R., Reza, R., Tuton, T., Vierra, C., 2007. Aciniform spidroin, a constituent of egg case sacs and wrapping silk fibers from the black widow spider *Latrodectus hesperus*. J. Biol. Chem. 282, 35088–35097.

Vetter, R.S., Tarango, J., Campbell, K.A., Tham, C., Hayashi, C.Y., Choe, D.-H., 2016. Efficacy of several pesticide products on brown widow spider (Araneae: Theridiidae) egg sacs and their penetration through the egg sac silk. J. Econ. Entomol. 109, 267–272. https://doi.org/10.1093/jee/tov288

Vienneau-Hathaway, J.M., Brassfield, E.R., Lane, A.K., Collin, M.A., Correa-Garhwal, S.M., Clarke, T.H., Schwager, E.E., Garb, J.E., Hayashi, C.Y., Ayoub, N.A., 2017. Duplication and concerted evolution of MiSp-encoding genes underlie the material properties of minor ampullate silks of cobweb weaving spiders. BMC Evol. Biol. 17, 78. https://doi.org/10.1186/s12862-017-0927-x

Vollrath, F., 2000. Strength and structure of spiders' silks. Rev. Mol. Biotechnol. 74, 67–83. https://doi.org/10.1016/S1389-0352(00)00006-4

Vollrath, F., Knight, D.P., 1999. Structure and function of the silk production pathway in the Spider *Nephila edulis*. Int. J. Biol. Macromol. 24, 243–249. https://doi.org/10.1016/S0141-8130(98)00095-6

Xu, M., Lewis, R.V., 1990. Structure of a protein superfiber: spider dragline silk. Proc. Natl. Acad. Sci. 87, 7120–7124.

Zhang, J.-X., Zhu, M.-S., Song, D.-X., 2004. A review of the chinese nursery-web spiders (Araneae, Pisauridae). J. Arachnol. 32, 353–417.

Zhang, Y., Zhao, A.-C., Sima, Y.-H., Lu, C., Xiang, Z.-H., Nakagaki, M., 2013. The molecular structures of major ampullate silk proteins of the wasp spider, *Argiope bruennichi*: A second blueprint for synthesizing de novo silk. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 164, 151–158.

Zisman, W.A., 1964. Relation of the equilibrium contact angle to liquid and solid constitution, in: Contact Angle, Wettability, and Adhesion, Advances in Chemistry. American Chemical Society, pp. 1–51.

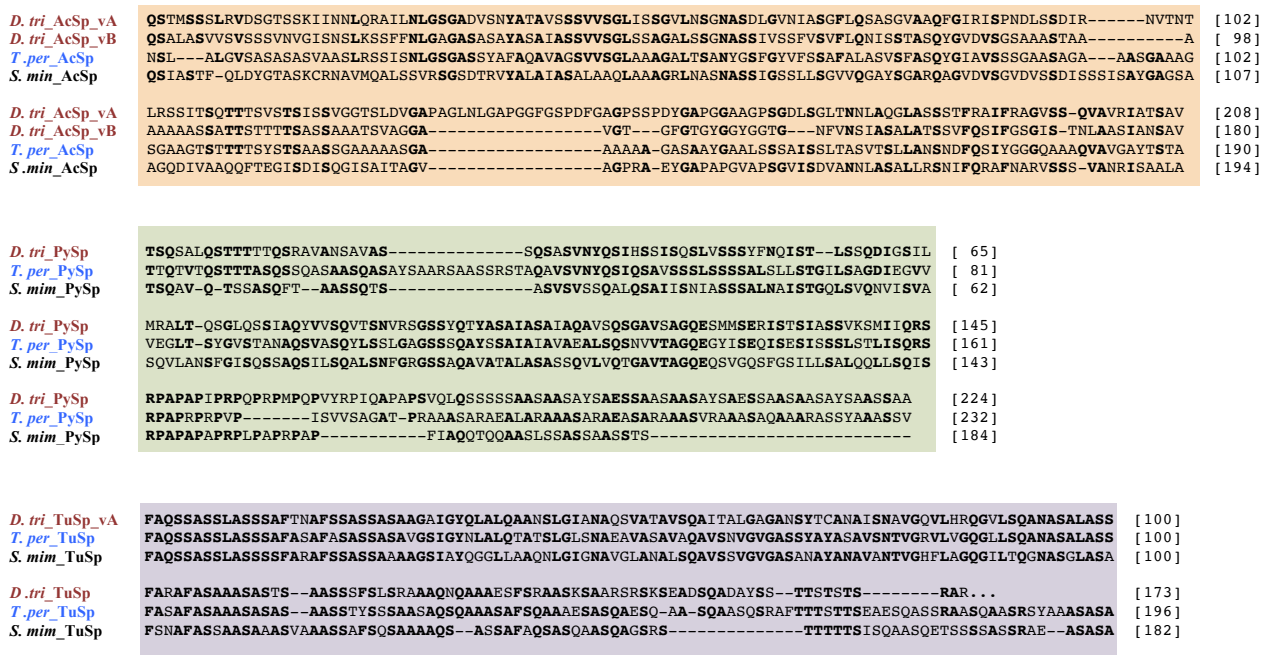


Figure S2.1. Exemplar repeat units for three spidroin paralog groups. Multiple alignments comparing spidroin paralogs encoding aciniform (AcSp), tubuliform (TuSp) and pyriform (PySp) spidroins of *Dolomedes triton* (red), *Tengella perfuga* (blue) and *Stegodyphus mimosarum* (AcSp, GenBank: KK121179.1; TuSp GenBank: KK121684.1; PySp GenBank: KK116737.1). Names abbreviated as in Table S2.2. Amino acids conserved > 50% across all sequences are indicated in bold. Gaps inserted into the alignment are indicated by dashes and missing sequence by periods. Amino acid positions for each sequence are numbered on the right.

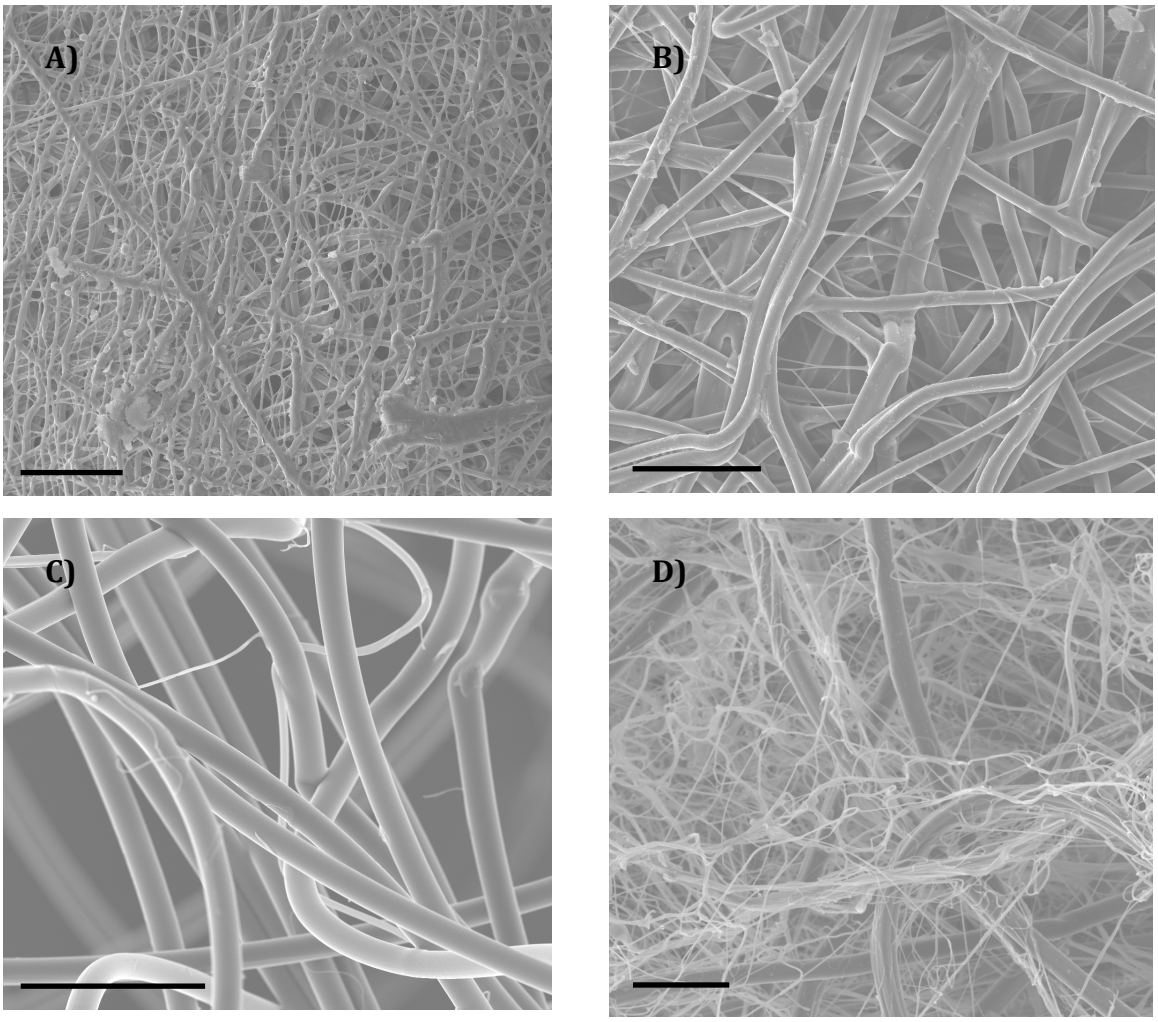
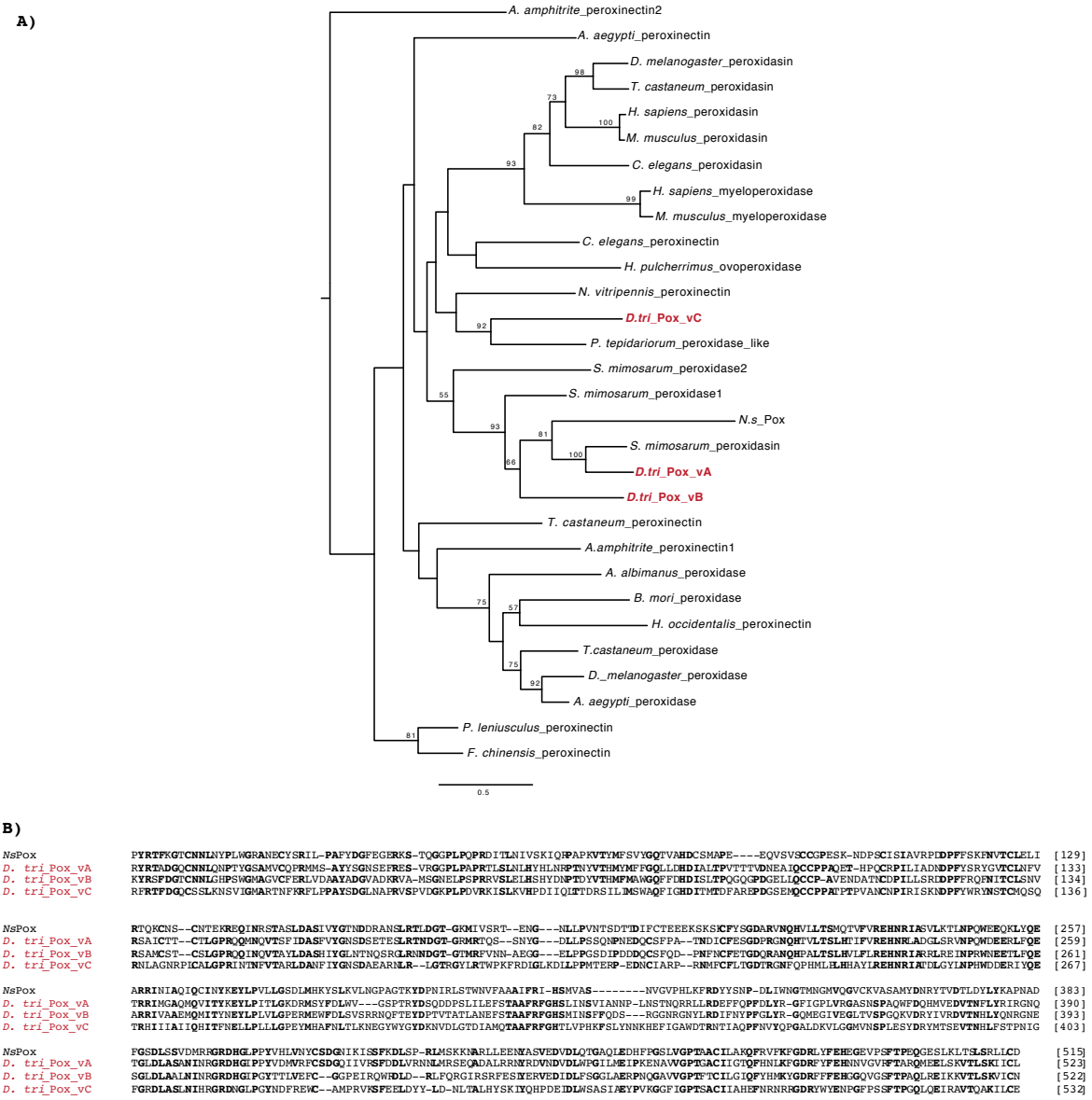


Figure S2.2. Scanning electron micrographs of eggs sacs from **A)** *Dolomedes triton*, **B)** *Latrodectus hesperus*, **C)** *Nephila clavipes*, and **D)** *Argiope argentata*. Scale bar 20 μm.



B)

```
NsPox       PYRTFKGTCNNLNYPLWGRANECYSRIL-PAFYDGFEGERKS-TQGGPLPQPRDITLNIVSKIQHPAPKVTYMFSVYGQTVAHDCSMAPE----EQVSVSCCGPESK-NDPSCISIAVRPIDPFFSKFNVTCLELI  [129]
D. tri_Pox_vA  RYRTADGQCNNLQNPTYGSAMVCQPRMMS-AYYSGNSEFRES-VRGGPLPAPRTLSLNLHYHLNRPTNYVTHMYMFPGQLLDHDIALTPVTTTVDNEAIQCCPPAQET-HPQCRPILIADNDPFYSRYGVTCLNFV  [133]
D. tri_Pox_vB  KYRSFDGTCNNLGHPSWGMAGVCFERLVDAAYADGVADKRVA-MSGNELPSPRRVSLELHSHYDNPTDYVTHMFMAWGQFFDHDISLTPQGQGPDGELLQCCP-AVENDATNCDPILLSRIDDPFFRQFNITCLSNV  [134]
D. tri_Pox_vC  RFRTFDGQCSSLKNSVIGMARTNFNRFLPPAYSDGLNAPRVSPVDGKPLPDVRKISLKVHPDIIQLTTDRSILLMSWAQFIGHDITMTDFARPPDGSEMQCCPPATPTPVANCNPIRISKNDPFYWRINSTCMQSQ  [136]

NsPox       RTQKCNS--CNTEKREQINRSTASLDASIVYGTNDDRANSLRTLDGT-GKMIVSRT--ENG---NLLPVNTSDTTDIFCTEEEKSKSKCFYSGDARVNQHVLLTSMQTVFVREHNRIASVLKTLNPQWEEQKLYQE  [257]
D. tri_Pox_vA  RSAICTT--CTLGPRQQMNQVTSFIDASFVYGNSDSETESLRTNDGT-GRMRTQS--SNYG---DLLPSSQNPNEDQCSFPA--TNDICFESGDPRGNQHTVLTSLHTIFVREHNRLADGLSRVNPQWDEERLFQE  [259]
D. tri_Pox_vB  RSAMCST--CSLGPRQQINQVTAYLDASHIYGLNTNQSRGLRNNDGT-GTMRFVNN-AEGG---ELPPGSDIPDDDQCSFQD--PNFNCFETGDQRANQHPALTSLHVLFLREHNRIARRLREINPRWNEETLFQE  [261]
D. tri_Pox_vC  RNLAGNRPICALGPRINTNFVTARLDANFIYGNSDAEARNLR--LGTRGYLRTWPKFRDIGLKDILPPMTERP-EDNCIARP--RNMFCFLTGDTRGNFQPHMLHLHHAYLREHNRIATDLGYINPHWDDERIYQE  [267]

NsPox       ARRINIAQIQCINYKEYLPVLLGSDLMHKYSLKVLNGPAGTKYDPNIRLSTWNVFAAAIFRI-HSMVAS--------NVGVPHLKFRDYYSNP-DLIWNGTMNGMVQGVCKVASAMYDNRYTVDTLDYLYKAPNAD  [383]
D. tri_Pox_vA  TRRIMGAQMQVITYKEYLPITLGKDRMSYFDLWV---GSPTRYDSQDDPSLILEFSTAAFRFGHSLINSVIANNP-LNSTNQRRLLRDEFFQPPFDLYR-GFIGPLVRGASNSPAQWFDQHMVEDVTNFLYRIRGNQ  [390]
D. tri_Pox_vB  ARRIVAAEMQMITYNEYLPLVLGPERMEWFDLSVSRRNQFTEYDPTVTATLANEFSTAAFRFGHSMINSFFQDS---RGGNRGNYLRDIFNYPFGLYR-GQMEGIVEGLTVSPGQKVDRYIVRDVTNHLYQNRGNE  [393]
D. tri_Pox_vC  TRHIIIAIQHITFNELLPLLLGPEYMHAFNLTLKNEGYWYGYDKNVDLGTDIAMQTAAFRFGHTLVPHKFSLYNNKHEFIGAWDTRNTIAQPFNVYQPGALDKVLGGMVNSPLESYDRYMTSEVTNHLFSTPNIG  [403]

NsPox       PGSDLSSVDMRRGRDHGLPPYVHLVNYCSDGNIKISSFKDLSP-RLMSKKNARLLEENASVEDVDLQTGAQLEDHFPGSLVGPTAACILAKQFRVFKFGDRLYFEHEGEVPSFTPEQGESLKLTSLSRLLCD  [515]
D. tri_Pox_vA  TGLDLASANINRGRDHGIPPYVDMVRFCSDGQIIVRSFDDLVRNNLMRSEQADALRRNYRDVNDVDLWFGILMEIPKENAVVGPTGACIIGTQFHNLKFGDRFYFEHNNVGVNFTARQMEELSKVTLSKIICL  [523]
D. tri_Pox_vB  SGLDLAALNINRGRDHGIPGYTTLVEFC--GGPEIRQWHDLD--RLFQRGIRSRFESIYERVEDIDLFSGGLAERPNQGAVVGPTFTCILGIQFYHMKYGDRFFFEHGGQVGSFTPAQLREIKKVTLSKVICN  [522]
D. tri_Pox_vC  PGRDLASLNIHRGRDNGLPGYNDFREWC--AMPRVKSFEELDYY-LD-NLTALHYSKIYQHPDEIDLWSASIAEYPVKGGFIGPTSACIIAHEFNRNRRGDRYWYENPGFPSSFTPGQLQEIRAVTQAKILCE  [532]
```

Figure S2.3. Phylogenetic analysis and sequence alignment of *Dolomedes triton* peroxidases. **A**) Maximum likelihood tree of *Dolomedes triton* (red) peroxidases and some representatives from the heme-containing peroxidase families (peroxidasin, peroxinectin, myeolperoxidases). Tree rooted with the umbrella liverwort *Marchantia polymorpha* peroxidase. Names abbreviated as in Table S2.4. Bootstrap percentages >50% are show. Scale bar represents substitutions per site. **B**) Multiple alignments of *Dolomedes triton* (red) peroxidases (*D. tri*_Pox_vA, *D. tri*_Pox_vB, and *D. tri*_Pox_vC) and *Nephila senegalensis* major ampullate gland peroxidase (NsPox). Amino acids conserved > 50% across all sequences are indicated in bold. Gaps inserted into the alignment are indicated by dashes and missing sequence by periods. Amino acid positions for each sequence are numbered on the right. Names abbreviated as in Table S2.4.

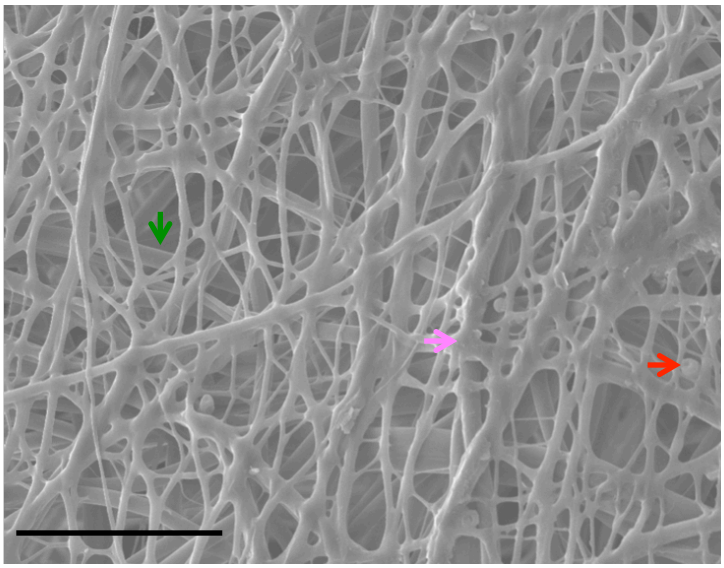


Figure S2.4. Scanning electron micrograph of outer layer in *D. triton* egg sac. Arrows corresponds to EDS spectra for *Dolomedes triton* outer layer fiber (green), *D. triton* outer layer coating (pink), and *D. triton* outer layer aggregation (red). Scale bar 20 μm.

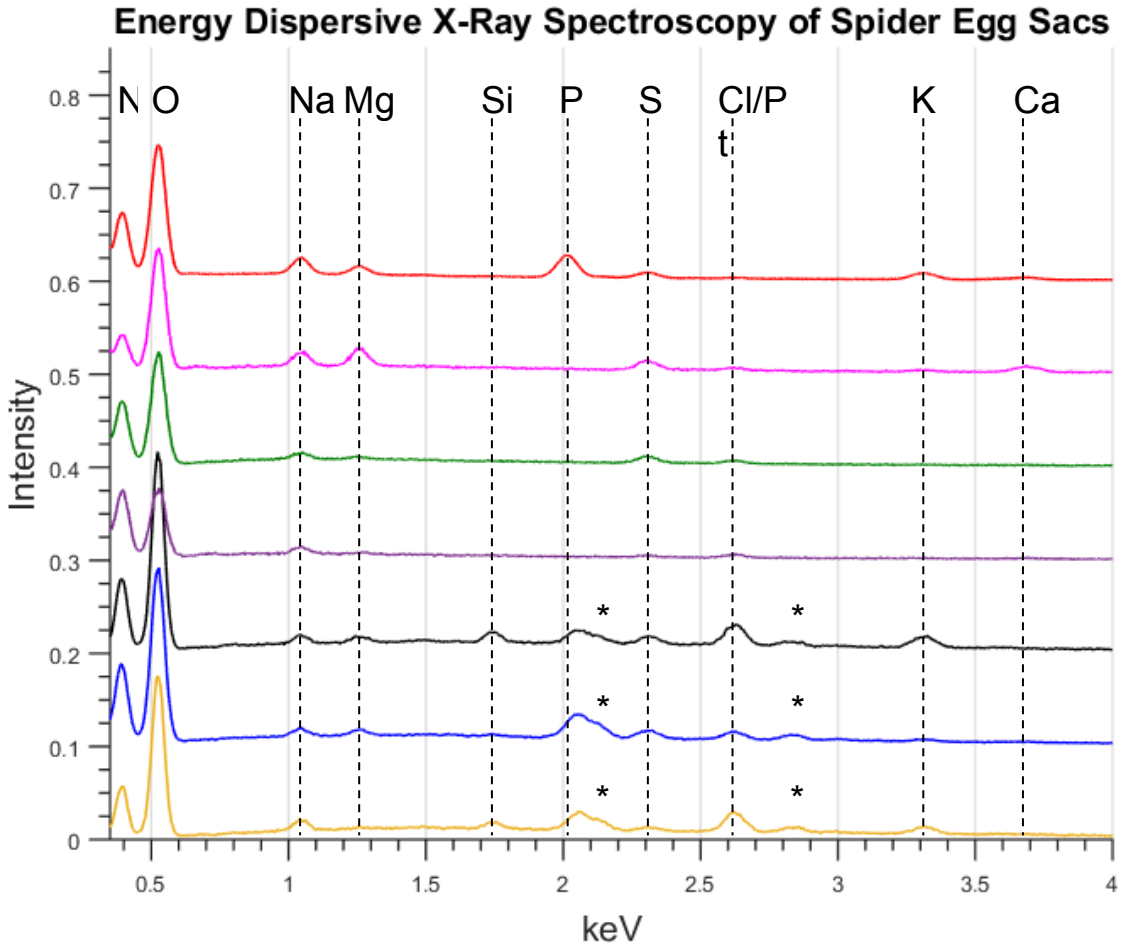


Figure S2.5. Elemental composition of multiple egg sacs. EDS spectra showing the elemental composition of the egg sac of (from bottom to top) *Argiope argentata* (yellow), *Nephila clavipes* (blue), *Latrodectus hesperus* (black), *D. triton* inner layer fiber (purple), *Dolomedes triton* outer layer fiber (green), *D. triton* outer layer coating (pink), *D. triton* outer layer aggregation (red). Elements indicated as N: nitrogen, O: oxygen, Na: sodium, Mg: magnesium, Si: Silicon (likely a contaminant), P: phosphorous, S: sulfur, Cl: chlorine, Pt: platinum, K: potassium, and Ca: calcium. * Corresponds to a coating element (Pt/Pd).

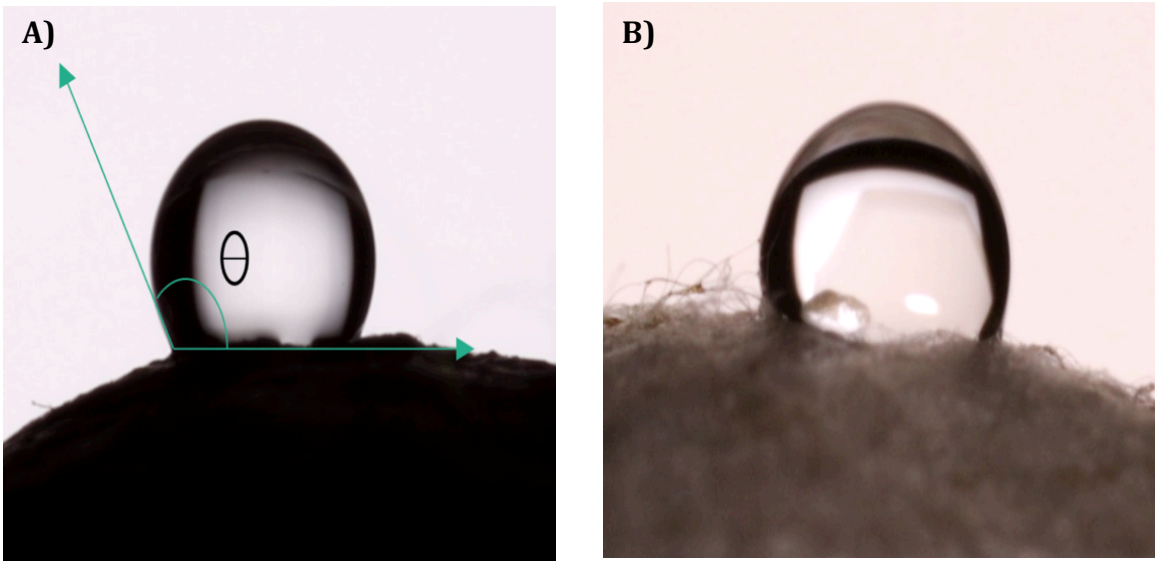


Figure S2.6. Static contact angle measurement of a water droplet on egg sac surface of **A)** *Dolomedes triton*; the contact is 107.7^{o} and **B)** *Latrodectus hesperus*; the contact angle 94.5^{o}. Scale bar 5 mm.

Table S2.1. Summary of *Dolomedes triton de novo* transcriptome assembly

No. Raw Paired Reads	13,264,529
No. Cleaned Paired Reads	13,008,603
No. Contigs	45,373
Total Length (bp)	15,671,409
N50 (bp)	444
BUSCO % complete (*Ixodes* reference)	73.4

Table S2.2. Spidroins from this study for *Dolomedes triton*.

Spidroin Name[a,b]	Library type	Contig name	Top BLAST hit Accesion	Top BLAST hit Description	Gene tree Clade
*D. tri*_AcSp_C_vA	RNA-seq of silk glands combined (Fem, Male)	*D. tri*_AcSp_C_vA_contig	gi|422900768|gb|AFX83561.1|	Aciniform spidroin 1, partial [*Latrodectus hesperus*]	AcSp
*D. tri*_AcSp_C_vB	RNA-seq of silk glands combined (Fem, Male)	*D. tri*_AcSp_C_vB_contig	gi|422900768|gb|AFX83561.1|	Aciniform spidroin 1, partial [*Latrodectus hesperus*]	AcSp
*D. tri*_AcSp_N_vA	RNA-seq of silk glands combined (Fem, Male)	*D. tri*_AcSp_N_vA_contig	gi|675387023|gb|KFM79920.1|	Hypothetical protein X975_02929, partial [*Stegodyphus mimosarum*]	AcSp
*D. tri*_AcSp_N_vB	RNA-seq of silk glands combined (Fem, Male)	*D. tri*_AcSp_N_vB_contig	gi|675387023|gb|KFM79920.1|	Hypothetical protein X975_02929, partial [*Stegodyphus mimosarum*]	AcSp
*D. tri*_AmSp_C_vA	RNA-seq of silk glands combined (Fem, Male)	*D. tri*_Amp_C_vA_contig	gb|AAK30599.1|AF350270_1|	Fibroin 2 [*Dolomedes tenebrosus*]	AmSp
*D. tri*_AmSp_C_vB	RNA-seq of silk glands combined (Fem, Male)	*D. tri*_Amp_C_vB_contig	gb|AAK30599.1|AF350270_1|	Fibroin 2 [*Dolomedes tenebrosus*]	AmSp
*D. tri*_AmSp_C_vC	RNA-seq of silk glands combined (Fem, Male)	12534_1_Amp_C_vC	gb|AFM97615.1|AFM97615.1|	Fibroin 1, partial [*Hypochilus thorelli*]	AmSp
*D. tri*_AmSp_C_vD	RNA-seq of silk glands combined (Fem, Male)	3010_1_Amp_C_vD	gb|AAK30599.1|AF350270_1|	Fibroin 2 [*Dolomedes tenebrosus*]	AmSp
*D. tri*_AmSp_N_vA	RNA-seq of silk glands combined (Fem, Male)	*D. tri*_Amp_N_vA_contig	gi|115635734|emb|CAJ90517.1|	Major ampullate spidroin 1 precursor [*Euprosthenops australis*]	AmSp
*D. tri*_AmSp_N_vB	RNA-seq of silk glands combined (Fem, Male)	DolMalecomp6778_c4-1_Amp_N_vB	gi|295982418|pdb|3LRD|	Chain A, Self-Assembly Of Spider Silk Proteins Controlled By A Ph-Sensitive Relay	AmSp
*D. tri*_AmSp_N_vC	RNA-seq of silk glands combined (Fem, Male)	196_1_Amp_N_vC	gi|115635734|emb|CAJ90517.1|	Major ampullate spidroin 1 precursor [*Euprosthenops australis*]	AmSp
*D. tri*_AmSp_N_vD	RNA-seq of silk glands combined (Fem, Male)	8251_1_Amp_N_vD	gi|303307772|gb|ADM14324.1|	Major ampullate spidroin, partial [*Agelenopsis aperta*]	AmSp
*D. tri*_AmSp_N_vE	RNA-seq of silk glands combined (Fem, Male)	DolFecomp3105_c0-1_Amp_N_vE	gi|115635734|emb|CAJ90517.1|	Major ampullate spidroin 1 precursor [*Euprosthenops australis*]	AmSp
*D. tri*_PySp_C	RNA-seq of silk glands combined (Fem)	8234_1_*D. tri*_PySp_C	gi|257124471|gb|ACV41934.1|	Pyriform spidroin 1 [*Latrodectus hesperus*]	PySp
*D. tri*_PySp_N	RNA-seq of silk glands combined (Fem)	*D. tri*_PySp_N	gi|JX112872|gb|AFN54363.1|	Major ampullate silk protein 2 [*Argiope bruennichi*]	PySp
*D. tri*_Sp_N	RNA-seq of silk glands combined (Fem, Male)	7528_1_Sp_N	gi|EU177661.1|gb|ABY67421.1|	Major ampullate spidroin 1 locus 1 [*Latrodectus hesperus*]	
*D. tri*_TuSp_C	RNA-seq of silk glands combined (Fem, Male)	*D. tri*_TuSp_C_contig	gi|303307770|gb|ADM14323.1|	Tubuliform spidroin 1, partial [*Agelenopsis aperta*]	TuSp
*D. tri*_TuSp_N	RNA-seq of silk glands combined (Fem, Male)	5425_1_*D. tri*_TuSp_N	gi|303307784|gb|ADM14332.1|	Tubuliform spidroin 1, partial [*Agelenopsis aperta*]	TuSp

[a] N or C in the spidroin names indicates whether a contig contains the N- or C- terminal region coding sequence.
[b] Variant name (e.g. _vA) does not indicate association of N- terminal transcript with C- terminal transcript.

Table S2.3. GenBank accession numbers for spidroins sequences used in phylogenetic analyses.

Sequence Name	Species	N-terminal region	C-terminal region
***A.ape*_MaSp**	*Agelenopsis aperta*	HM752573	AAT08436
***A.ape*_TuSp1**	*Agelenopsis aperta*	HM752576	--
***A.arg*_AcSp**	*Argiope argentata*	AHK09813	AHK09813
***A.arg*_Flag**	*Argiope argentata*	--	MF955778
***A.arg*_MaSp1**	*Argiope argentata*	MF955677	MF955761
***A.arg*_MaSp2**	*Argiope argentata*	MF955700	MF955804
***A.arg*_MaSp3**	*Argiope argentata*	MF955785	MF955690
***A.arg*_MiSp**	*Argiope argentata*	MF955726	MF955717
***A.arg*_PySp1**	*Argiope argentata*	AQR58363	AQR58363
***A.arg*_TuSp1**	*Argiope argentata*	ATW75951	ATW75951
***A.dia*_AcSp**	*Araneus diadematus*	MF955743	MF955754
***A.dia*_Flag**	*Araneus diadematus*	MF955789	MF955779
***A.dia*_Masp1**	*Araneus diadematus*	MF955789	MF955789
***A.dia*_MaSp2**	*Araneus diadematus*	MF955703	MF955809
***A.dia*_MaSp3**	*Araneus diadematus*	--	MF955691
***A.dia*_MiSp**	*Araneus diadematus*	--	MF955718
***A.dia*_PySp1**	*Araneus diadematus*	MF955713	MF955772
***A.dia*_TuSp1**	*Araneus diadematus*	MF955696	MF955799
***B.cal*_fibroin1**	*Bothriocyrtum californicum*	HM752562	EU117162
***D.can*_MaSp**	*Diguetia canities*	HM752564	HM752565
***D.can*_MaSp-like**	*Diguetia canities*	HM752566	HM752567
***D.spi*_Fib1a**	*Deinopis spinosa*	JX978170	DQ399326
***D.spi*_Fib1b**	*Deinopis spinosa*	--	DQ399327
***D.spi*_MaSp2a**	*Deinopis spinosa*	--	DQ399329
***D.spi*_MaSp2b**	*Deinopis spinosa*	HM752568	DQ399328
***D.spi*_MiSp1**	*Deinopis spinosa*	--	DQ399324
***D.spi*_Sp2b**	*Deinopis spinosa*	--	DQ399323
***D.spi*_TuSp1**	*Deinopis spinosa*	--	AY953073
***D.ten*_fib1**	*Dolomedes tenebrosus*	--	AF350269
***D.ten*_fib2**	*Dolomedes tenebrosus*	--	AF350270
***E.aus*_MaSp**	*Euprosthenops australis*	AM259067	AJ973155
***H.tho*_Fib1**	*Hypochilus thorelli*	--	JX102555
***H.tho*_Fib2**	*Hypochilus thorelli*	--	JX102556
***K.hib*_MaSp1**	*Kukulkania hibernalis*	HM752563	--
***L.hes* _Flag**	*Latrodectus hesperus*	MF955792	MF955781
***L.hes* _MaSp1**	*Latrodectus hesperus*	ABR68856	ABR68856
***L.hes* _MaSp2**	*Latrodectus hesperus*	ABR68855	ABR68855
***L.hes* _MaSp3**	*Latrodectus hesperus*	MF955786	MF955692
***L.hes* _MiSp**	*Latrodectus hesperus*	ARA91152	ARA91152
***L.hes* _PySp1**	*Latrodectus hesperus*	MF955714	MF955773
***L.hes* _TuSp1**	*Latrodectus hesperus*	MF955697	MF955801
***L.hes*_AcSp1**	*Latrodectus hesperus*	MF955746	MF955757
***M.gra*_MiSp**	*Metepeira grandiosa*	HM752575	--
***N.cla*_AcSp1**	*Nephila clavipes*	MF955747	MF955758
***N.cla*_Flag**	*Nephila clavipes*	MF955793	MF955782
***N.cla*_MaSp1**	*Nephila clavipes*	MF955682	MF955765
***N.cla*_MaSp2**	*Nephila clavipes*	MF955708	MF955815
***N.cla*_MiSp**	*Nephila clavipes*	MF955734	MF955722
***N.cla*_PySp1**	*Nephila clavipes*	MF955715	MF955774
***N.cla*_TuSp1**	*Nephila clavipes*	MF955698	MF955802
***N.cla*_TuSp1**	*Nephila clavipes*	MF955698	MF955802
***P.sin*_DragFib1**	*Psechrus sinensis*	--	AY666067
***P.tri* _fib2**	*Plectreurys tristis*	--	AF350282

***P.tri* _fib3**	*Plectreurys tristis*	--	AF350283
***P.tri* _fib4**	*Plectreurys tristis*	--	AF350284
***P.tri*_fib1**	*Plectreurys tristis*	--	AF350281
***P.vir*_MaSp1**	*Peucetia viridans*	--	GU306168
***S.gro*_AcSp**	*Stetaoda grossa*	MF955749	MF955760
***S.gro*_MaSp1**	*Stetaoda grossa*	MF955688	MF955770
***S.gro*_MiSp**	*Stetaoda grossa*	MF955737	MF955725
***S.gro*_PySp1**	*Stetaoda grossa*	MF955716	MF955775
***S.gro*_TuSp1**	*Stetaoda grossa*	MF955699	MF955803
***S.mim*_AcSp-putative**	*Stegodyphus mimosarum*	KFM79920	KFM79920
***S.mim*_MaSp-putative-a**	*Stegodyphus mimosarum*	--	KFM83271
***S.mim*_MaSp-putative-c**	*Stegodyphus mimosarum*	--	JT038023
***S.mim*_MaSp-putative-d**	*Stegodyphus mimosarum*	KFM59474	KFM59474
***S.mim*_MaSp-putative-e**	*Stegodyphus mimosarum*	KFM74936	--
***S.mim*_MaSp-putative-f**	*Stegodyphus mimosarum*	KFM61798	--
***S.mim*_MaSp-putative-g**	*Stegodyphus mimosarum*	KFM57717	--
***S.mim*_MaSp-putative-h**	*Stegodyphus mimosarum*	KFM61802	KFM61800
***S.mim*_MaSp-putative-i**	*Stegodyphus mimosarum*	KFM79313	KFM79313
***S.mim*_Misp-putative**	*Stegodyphus mimosarum*	KFM62627	KFM62627
***S.mim*_PiSp-putative**	*Stegodyphus mimosarum*	KFM75168	KFM68615
***S.mim*_Sp1**	*Stegodyphus mimosarum*	--	KFM60634
***S.mim*_Sp2a**	*Stegodyphus mimosarum*	KFM73910	--
***S.mim*_Sp2b**	*Stegodyphus mimosarum*	KFM70693	--
***S.mim*_TuSp-putative**	*Stegodyphus mimosarum*	KFM79920	KFM79920
U.div_MiSp	*Uloborus diversus*	HM752574	--

Table S2.4. Accession numbers from GeneBank and Peroxibase of peroxidase sequences used in phylogenetic analysis.

Peroxidase name	Accession number	Protein	Organism
*A. aegypti*_peroxidase	AAC97504.1	Peroxidase	*Aedes aegypti*
*A. aegypti*_peroxinectin	AaePxt01	Peroxinectin	*Aedes aegypti*
*A. albimanus*_peroxidase	AAD22196.1	Salivary peroxidase	*Anopheles albimanus*
*A. amphitrite*_peroxinectin1	AQY78510.1	Peroxinectin 2, partial	*Amphibalanus amphitrite*
*A. amphitrite*_peroxinectin2	AQY78509.1	Peroxinectin 1, partial	*Amphibalanus amphitrite*
*B mori*_peroxidase	XP_004930615.1	Peroxidase	*Bombyx mori*
*C. elegans*_peroxidasin	NP_505188.3	Peroxidasin homolog	*Caenorhabditis elegans*
*C. elegans*_peroxinectin	CelPxt01	Peroxinectin	*Caenorhabditis elegans*
*D. melanogaster*_peroxidase	Q01603.2	Peroxidase, chorion peroxidase	*Drosophila melanogaster*
*D. melanogaster*_peroxidasin	NP_523891.2	Peroxidasin isoform A	*Drosophila melanogaster*
*F. chinensis*_peroxinectin	FchPxt01	Peroxinectin	*Fenneropenaeus chinensis*
*H. occidentalis*_peroxinectin	KM384736.1	Peroxinectin	*Hesperophylax occidentalis*
*H. pulcherrimus*_ ovoperoxidase	BAA19738.1	Ovoperoxidase	*Hemicentrotus pulcherrimus*
*H. sapiens*_myeloperoxidase	NP_000241.1_	Myeloperoxidase precursor	*Homo sapiens*
*H. sapiens*_peroxidasin	NP_036425.1	Peroxidasin homolog precursor	*Homo sapiens*
*M. musculus*_myeloperoxidase	NP_034954.2	Myeloperoxidase precursor	*Mus musculus*
*M. musculus*_peroxidasin	NP_852060.2	Peroxidasin homolog precursor	*Mus musculus*
M. polymorpha peroxidase	BAB97197	Peroxidase 1	*Marchantia polymorpha*
*N. vitripennis*_peroxinectin	NviPxt01	Peroxinectin	*Nasonia vitripennis*
*Ns*Pox	AAO33164.1	Major ampullate gland peroxidase	*Nephila senegalensis*
*P. leniusculus*_peroxinectin	Q26059	Peroxinectin precursor	*Pacifastacus leniusculus*
*P. tepidariorum*_peroxidase	XP_015911676.1	Peroxidase-like	*Parasteatoda tepidariorum*
*S. mimosarum*_peroxidase1	KFM61817.1	Peroxidase, partial 1	*Stegodyphus mimosarum*
*S. mimosarum*_peroxidase2	KFM73005.1	Peroxidase, partial 2	*Stegodyphus mimosarum*
*S. mimosarum*_peroxidasin	KFM76656.1	Peroxidasin, partial	*Stegodyphus mimosarum*
*T. castaneum*_peroxidase	XP_967241.1	Predicted: Peroxidase	*Tribolium castaneum*
*T. castaneum*_peroxidasin	XP_968570.1	Predicted: Peroxidasin isoform X1	*Tribolium castaneum*
*T. castaneum*_peroxinectin	TcasPxt02	Peroxinectin	*Tribolium castaneum*

Chapter 3

Silk Genes and Silk Gene Expression in the Spider *Tengella Perfuga* (Zoropsidae), Including a Potential Cribellar Spidroin (Crsp)

Abstract

Most spiders spin multiple types of silk, including silks for reproduction, prey capture, and draglines. Spiders are a megadiverse group and the majority of spider silks remain uncharacterized. For example, nothing is known about the silk molecules of *Tengella perfuga*, a spider that spins sheet webs lined with cribellar silk. Cribellar silk is a type of adhesive capture thread composed of numerous fibrils that originate from a specialized plate-like spinning organ called the cribellum. The predominant components of spider silks are spidroins, members of a protein family synthesized in silk glands. Here, we use silk gland RNA-Seq and cDNA libraries to infer *T. perfuga* silks at the protein level. We show that *T. perfuga* spiders express 13 silk transcripts representing at least five categories of spider silk proteins (spidroins). One category is a candidate for cribellar silk and is thus named cribellar spidroin (CrSp). Studies of ontogenetic changes in web construction and spigot morphology in *T. perfuga* have documented that after sexual maturation, *T. perfuga* females continue to make capture webs but males halt web maintenance and cease spinning cribellar silk. Consistent with these observations, our candidate CrSp was expressed only in females. The other four spidroin categories correspond to paralogs of aciniform, ampullate, pyriform, and tubuliform spidroins. These spidroins are associated with egg sac and web construction. Except for the tubuliform spidroin, the spidroins from *T. perfuga* contain novel combinations of amino acid sequence motifs that have not been observed before in these spidroin types. Characterization of *T. perfuga* silk genes, particularly CrSp, expand the diversity of the spidroin family and inspire new structure/function hypotheses.

Introduction

Spiders are widely distributed and abundant in most terrestrial communities, and their evolutionary success is partly associated with diversification of silk usage (Blackledge et al., 2009; Bond and Opell, 1998; Craig, 2003). Silk is an important feature of spider biology, and all spiders produce silk for an array of essential, fitness-related tasks including prey capture, reproduction, locomotion, and protection of progeny (Foelix, 2011). Most studies on spider silk use and molecular composition have been heavily focused on orb-web and cob-web weaving spiders, but there are many other types of spiders (Ayoub et al., 2007; Ayoub and Hayashi, 2008; Babb et al., 2017; Blackledge and Hayashi, 2006; Blasingame et al., 2009; Eberhard, 1990; Hinman and Lewis, 1992; Kovoor and Zylberberg, 1980; Lawrence et al., 2004; Vasanthavada et al., 2007; Vienneau-Hathaway et al., 2017). For example, *Tengella perfuga*, Dahl 1901 (Zoropsidae) uses copious amounts of silk to build sheet webs with deep retreats in high elevation remnant cloud forest habitats in Nicaragua (Leister et al., 2013; Mallis and Miller, 2017). *T. perfuga* spiders belong to the grate shaped tapetum (GST) clade and are cribellate spiders (Polotow et al., 2015). Cribellate spiders have one pair of silk spinnerets modified into a cribellum, a plate-like spinning organ that is dotted with numerous miniscule spigots. From this dense field of spigots, thousands of ultrafine fibrils are produced; this silk type is referred to as cribellar silk (Eberhard and Pereira, 1993; Peters, 1992, 1987, 1984). Cribellar silk has adhesive properties and is an important functional element of capture webs spun by cribellate spiders (Hawthorn and Opell, 2003, 2002; Opell, 1994).

Silk production in spiders involves a combination of highly specialized structures and behaviors. Spider silk proteins are secreted from silk glands, which are located in the abdomen. Each silk gland has a duct that leads to its own spigot located on the spinnerets (Coddington, 1989; Gosline et al., 1986). Spider silk spigots are morphologically distinctive and are named according to the silk gland connected to them. From each spigot type, a functionally specific silk type emerges. For instance, pyriform silk fibers are produced from pyriform spigots, which can be visually distinguished from other types of spigots and are connected to pyriform silk glands (Coddington, 1989; Griswold et al., 2005).

Ontogeny of silk usage and silk spigots in *T. perfuga* has recently been examined (Alfaro et al., 2018a; Mallis and Miller, 2017). Adult female *T. perfuga* spiders use silk for foraging, building retreats, and constructing egg sacs. Silk use starts with spiderlings making small sheet webs without cribellar silk and as they mature, their webs become more complex with the addition of cribellar silk. Cribellar silk fibers fill the sheet of adult female webs, lining the retreat and knockdown lines that extend from the substrate to the sheet. After becoming sexually mature, males abandon their webs and adopt a wandering life style. Based on scanning electron microscopy, spigots corresponding to aciniform, cribellate, major ampullate, minor ampullate, pyriform, and tubuliform silk glands have been imaged for *T. perfuga* (Alfaro et al., 2018a). Additionally, there is a pair of spigot triads that are connected to unidentified gland types. Each triad is composed of a large spigot called the “modified spigot” and two smaller, flanking spigots. The most dramatic changes in *T. perfuga* spigot ontogeny involve the cribellum. With successive molts, the

number of cribellar spigots and size of the cribellum increases as the spiders molt to adulthood, except that males lose their cribellar spigots in the final molt (Alfaro et al., 2018a).

Here, we use expression libraries to characterize the silk genes of *T. perfuga*. Based on studies of their silk usage and silk spigot ontogeny (Alfaro et al., 2018a; Griswold et al., 2005; Mallis and Miller, 2017), we expect *T. perfuga* to express genes associated with aciniform, cribellate, major ampullate, minor ampullate, pyriform, and tubuliform silk glands. The predicted silk genes encode proteins known as spidroins (spidroin is a contraction of spider fibroins (Hinman and Lewis, 1992)), and the silk genes are members of the spidroin gene family (Blasingame et al., 2009; Guerette et al., 1996; Hayashi et al., 2004; Hinman and Lewis, 1992). Characteristics of aciniform, major ampullate, minor ampullate, pyriform, and tubuliform spidroins are largely known from araneoid orb-weaving spiders (Ayoub et al., 2013, 2007; Chaw et al., 2017; Garb and Hayashi, 2005; Hinman and Lewis, 1992; Vienneau-Hathaway et al., 2017; Zhao et al., 2006), which diverged ~214 million years ago from *T. perfuga* (Garrison et al., 2016). Additionally, because *T. perfuga* uses extensive amounts of cribellar silk, we anticipate discovering a spidroin that is expressed in mature females but not mature males. Such a spidroin will be a candidate constituent of cribellar silk, which to our knowledge, has yet to be described at the molecular level. Finally, we predict that silk genes associated with web construction will be highly expressed compared to other spidroin genes because *T. perfuga* spiders use copious amounts of silk in their capture webs.

Materials and Methods

cDNA Library Construction and Sequencing

T. perfuga reared by R. Alfaro were used for all the silk gland dissections. Spiders were anesthetized and immediately after euthanization, silk glands were dissected from each individual, flash frozen in liquid nitrogen, and stored at -80°C. Ampullate-shaped silk glands and smaller silk glands from mature male and females were used to make silk gland type-specific cDNA libraries; tubuliform silk glands from mature females were used to make a tubuliform gland cDNA library. Additional dissections were done to collect the full set of silk glands present in female and male spiders. The total silk gland complement in a spider included ampullate shaped, tubuliform-shaped (present in females only), and an assortment of small silk glands, including cribellar glands, which were close and still attached to the spinnerets. Tissues were used to construct and screen plasmid-based cDNA libraries following the methods described in Garb et al. (2007). The libraries were screened with γ-^{32}P-labeled oligonucleotide probes designed from previously characterized spidroins (Garb and Hayashi, 2005; Gatesy et al., 2001). To discover novel spidroins, about one third of each library was screened for size and clones with inserts > 600 base pairs were sequenced using T7 and SP6 universal primers. BLASTX searches revealed that the sequenced cDNAs included 30 spidroin clones. One clone, a tubuliform spidroin (*T. per*_TuSp_C), was fully sequenced (2,971 base pairs) using the transposon-based EZ-Tn5 <TET-1> insertion kit (Epicentre). *T. perfuga* cDNA clones were Sanger sequenced at the University of California Riverside (UCR) Genomics Core Facility.

RNA-Seq Library Construction, Sequencing, and Assembly

The total set of silk glands was dissected from each of two *T. perfuga* females raised by R. Alfaro. The glands were flash frozen in liquid nitrogen and stored at -80°C. Separate RNA extractions were done for the total set of silk glands from each individual spider following the methods of Starrett et al., (2012). Two RNA-Seq libraries were then made from cDNA prepared using the method described in Starrett et al., (2012) with the modification that first strand cDNA was primed with both oligo-d(T) and random hexamers. Indexed libraries were constructed from the cDNA with the Encore NGS Library System (NuGen). Sequencing (paired end, 100 cycles each) was done on a HiSeq System (Illumina) at the UCR Genomics Core Facility.

Raw sequencing reads from each FASTQ file were processed by clipping the adaptors and removing low quality reads with Trimmomatic (Bolger et al., 2014). Quality of resulting filtered reads was assessed using FastQC (Babraham Bioinformatics FastQC Package**)**. All *T. perfuga* reads were combined to assemble a *de novo* female silk gland transcriptome with Trinity v2.1.1 using default parameters (Grabherr et al., 2011). See Table S3.1 for assembly statistics. Quality of the *T. perfuga* assembly was approximated using N50 and completeness determined by comparison to the arthropod set of Universal Single-Copy Orthologs (BUSCO v 1.2; Simão et al., 2015). 97.7% of the *Ixodes* BUSCOs were identified as complete in the *T. perfuga* assembly.

Annotation

BLASTX searches (e-value < 1e-5) to both NCBI nr and UniProtKB were used to automatically annotate transcripts (Altschul et al., 1990). Putative chimeric and contaminant sequences were removed from the resulting assemblies following Clarke et al. (2015). Functional annotation was done with Gene Ontology (GO) terms associated with the best UniProt matches. Translation of assembled contigs based on the frame of the best BLASTX hit to nr by e-value was used to generate predicted proteins. If a transcript had no BLASTX hit, amino acid sequence was predicted using the longest open reading frame (ORF) following Clarke et al. (2014).

Spidroin gene family members identified from the automatic annotation were further examined with additional BLASTX searches (e-value < 1 e-5) against a protein database with spidroin genes downloaded from NCBI nr proteins and UniProtKB/Swiss-Prot databases (September 2016) in Geneious v8.1.8 (Kearse et al., 2012). Visual inspection confirmed the presence of known characteristics of spidroin genes, such as repetitive regions and coding regions for conserved N- and C- terminal domains (Table S3.2). To be conservative in reporting the number of new spidroins, transcripts with pairwise nucleotide identities >95% were considered to represent the same variant and were collapsed into single contigs using majority rule.

Phylogenetic and Expression Analyses of Spidroin Family Members

Phylogenetic analyses of spidroin family members were done by aligning the N- and C-terminal region translations of *T. perfuga* spidroin contigs with published spidroin sequences from araneomorph (true spider) spiders (Table S3.3). A spidroin terminal

region from a non-araneomorph spider, *Bothriocyrtum californicum*, was used to root each analysis (GenBank accessions EU117162 and HM752562). Amino acid alignments were done with ClustalW (Larkin et al., 2007) as implemented in Geneious and refined by eye. Amino acid model test and maximum likelihood gene tree construction with 10,000 bootstrap replicates were done in RAxML v8.2.8 (Stamatakis, 2014). JTT and LG likelihood amino acid substitution models were used for N- and C- terminal region alignments, respectively. Resulting trees were visualized with FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/Fig.tree/).

The relative levels of spidroin gene expression in *T. perfuga* silk glands were quantified by mapping filtered sequencing reads from *T. perfuga* RNA-Seq libraries (combination of all silk glands within individual mature females, i.e. two biological replicates) to our *T. perfuga* transcriptome using TopHat2 v2.1.1 with default parameters (Kim et al., 2013). Reads Per Kilobase per Million mapped read (RPKM) values were calculated for each spidroin transcript. Spidroins with at least ten mapped reads and one RPKM were kept for further analysis.

Results and Discussion

Tengella perfuga Spidroins

Spidroins are structural proteins composed of a large repetitive region bounded by conserved non-repetitive amino and carboxyl terminal regions (Garb et al., 2010; Gatesy et al., 2001). We identified 13 spidroin contigs from *T. perfuga* spiders that contain N- or C- terminal coding regions and the adjacent repetitive regions (Table S3.2, Figure S3.2).

These spidroin contigs are associated with ampullate, aciniform, pyriform, tubuliform, and cribellate silk glands (Table S3.2). Maximum likelihood analyses of the C- and N-terminal region sequences show that *T. perfuga* AmSp (ampullate), AcSp (aciniform), PySp (pyriform), and TuSp (tubuliform) sequences group together with spidroins of the same respective type from the comparison species (Figure 3.1 and Figure S3.1).

T. perfuga has multiple ampullate spidroin variants. Six transcripts were identified as ampullate spidroins, three with the N-terminal region and the other three with the C-terminal region. Phylogenetic analyses of the terminal regions show that our ampullate sequences cluster within a diverse clade of major and minor ampullate spidroins (Figure 3.1 and Figure S3.1). Relationships among C-terminal encoding sequences indicate that all three *T. perfuga* ampullate spidroin variants cluster in their own clade with moderate support (Figure 3.1; 73%). Similarly, all the *Stegodyphus mimosarum* termini in the Ampullate group also form their own clade (Figure 3.1; 65%). As has been observed in other analyses of spidroin C-termini, the relationships among ampullate spidroins within and across species are complicated, suggesting turnover (birth, death) and/or sequence conversion (Ayoub et al., 2007; Ayoub and Hayashi, 2008; Garb and Hayashi, 2005; Gatesy et al., 2001; Hayashi et al., 1999).

The repetitive region of *T. perfuga* ampullate spidroins share amino acid sequence motifs with the minor ampullate (minor ampullate spidroin-MiSp) and major ampullate (major ampullate spidroin1 and 2-MaSp1 and MaSp2) spidroins of orb-web and cob-web weaving spiders (Figure S3.2). These motifs, such as poly-alanine (A_n), glycine-alanine $(GA)_n$, and glycine-glycine-X $(GGX)_n$, where X is a subset of all amino acids, have been

related to differences in tensile properties between silk types that are primarily composed of MaSp1, MaSp2, or MiSp (Simmons et al., 1996a; Sponner et al., 2005; Vienneau-Hathaway et al., 2017).

Although the repetitive regions of the *T. perfuga* ampullate spidroins have these motifs, the repeat sequences do not correspond to those of MiSp, MaSp1, or MaSp2. One explanation could be that major and minor ampullate silks have similar functions in *T. perfuga*. In orb-web weavers, MiSp is the primary component of minor ampullate silk, which is used in the temporary spiral during orb-web construction (e.g. Chen et al., 2012; Colgin and Lewis, 1998). In contrast, MaSp1 and MaSp2 are the main components of major ampullate silk, which is the primary silk type in draglines and the frame and spokes of the orb-web (Ayoub et al., 2007; Hinman and Lewis, 1992; Zhang et al., 2013). However, *T. perfuga* spiders do not build orb webs, and thus, the primary function of both their major and minor ampullate silks is likely to be as components of the dragline. Because the repeat sequences of *T. perfuga* ampullate spidroins do not obviously correspond to MiSp, MaSp1, or MaSp2 of orb-weavers, we annotated our *T. perfuga* ampullate-type spidroins with the general name of Ampullate Spidroin (AmSp), with a version letter to distinguish them from each other (Table S3.2).

Unlike the multiple *T. perfuga* ampullate spidroins, we found evidence for only a single locus each for aciniform, pyriform, and tubuliform spidroins. *T. perfuga* aciniform spidroin (AcSp), the presumed main component of aciniform silk, has a similar repetitive region to the AcSp from other species (Ayoub et al., 2013; Chaw et al., 2014; Hayashi et al., 2004; Vasanthavada et al., 2007). *T. perfuga* AcSp has a repeat length of 190 amino

acids (aa), which is similar to that of orb-web weavers in the genus *Argiope* (200-204 aa Chaw et al., 2014; Hayashi et al., 2004) and to the two sections that compose the 375 aa aciniform spidroin repeat of the cob-web weaver *Latrodectus hesperus* (the two ~190 aa sections are alignable to each other and to the AcSp from other species, Ayoub et al., 2013). The *T. perfuga* AcSp repeat also has a substantial serine content (24%) and presence of poly-serine motifs (Figure S3.2). Unlike previously known aciniform spidroins, the repetitive region of *T. perfuga* AcSp has poly-alanine amino acid motifs (Fig S2). Indeed, poly-alanine is more prevalent and in longer stretches than poly-serine in *T. perfuga* AcSp. Poly-alanine motifs are common in other spidroins such as MaSp1 and are thought to contribute to fiber tensile strength (Ayoub et al., 2007; Becker et al., 1994; Hinman and Lewis, 1992; Jenkins et al., 2013; Lawrence et al., 2004; Simmons et al., 1996b; Trancik et al., 2005). It is possible that the poly-alanine in *T. perfuga* AcSp sequences could also increase the strength of *T. perfuga* aciniform silk fibers.

T. perfuga PySp contains a novel combination of known PySp amino acid sequence motifs. In other species, PySp is the main component of pyriform silk, which is used to anchor silk fibers to a substrate (Blasingame et al., 2009; Kovoor and Zylberberg, 1982, 1980). *T. perfuga* PySp sequence contains one short (10 aa) stretch of alternating prolines (PX). This amino acid motif has also been identified in PySp from the cribellate spider *S. mimosarum*, orb-web weaving species, and one cob-web weaving spider (*Parasteatoda tepidariorum*). The PX motif is hypothesized to provide extensibility to pyriform silk fibers in orb-web weaving spiders (Chaw et al., 2017; Geurts et al., 2010; Perry et al., 2010; Sanggaard et al., 2014). In addition to PX, *T. perfuga* PySp also

contains motifs with short runs of alanines (AAASARAEAXAR, AAASXRAA; black boxes in Figure S3.2), which are similar to motifs that thus far were only found in PySp from the cob-web weaver *L. hesperus* (AAARAQAQAERAKAE, AAARAQAQAE; Blasingame et al., 2009; Correa-Garhwal et al., 2017). *L. hesperus* lacks PX, while *S. mimosarum*, *P. tepidariorum*, and orb-web weaving species lack these alanine-rich motifs (Babb et al., 2017; Chaw et al., 2017; Geurts et al., 2010; Perry et al., 2010; Sanggaard et al., 2014). Thus, *T. perfuga* PySp is noteworthy for containing both motifs. The conservation of PX motifs and alanine-rich motifs in *T. perfuga* PySp sequences suggests similar functional constraints on *T. perfuga* PySp and PySp from orb-web and cob-web weaving spiders.

T. perfuga TuSp has extraordinary sequence conservation among the four repeat units represented in our contig, which despite being nearly 3 kb is still a partial transcript (*T. per_TuSp_C*). The tandem arrayed, 194 aa repeat units in *T. perfuga* TuSp had >94% average pairwise identity at the amino acid and nucleotide levels. This high sequence similarity among tandem repeats has been observed in TuSp from other species, where it is likely explained by intragenic concerted evolution (Garb et al., 2007; Garb and Hayashi, 2005; Tian and Lewis, 2005). Additionally, *T. perfuga* TuSp repeats are similar in length and amino acid composition, largely composed of serine and alanine, to TuSp repeats from orb-web and cob-web weaving spiders.

Cribellar Candidate, CrSp

T. perfuga had two novel spidroin transcripts, one containing a N-terminal region and the other a C-terminal region (*T. per*_CrSp_C in Figure 3.1 and *T. per*_Sp_N in Figure S3.1). These *T. perfuga* transcripts had different top BLASTX hits, both of which were spidroins from the same species, the cribellate spider, *S. mimosarum*. Because the two *S. mimosarum* spidroins are located on separate genome assembly scaffolds and have dissimilar repetitive region sequences, we considered our two *T. perfuga* spidroin transcripts as also representing separate loci.

The *T. perfuga* transcript containing the N-terminal region was given the name "Sp" to indicate that it is a spidroin family member, but cannot be assigned to a known category. This transcript was placed as sister to the flagelliform clade with 50% support in the phylogenetic analysis (*T. per*_Sp_N; Figure S3.1). This affinity of a *T. perfuga* spidroin with flagelliform spidroins is unexpected because flagelliform spigots are associated only with ecribellate orb-weavers and their kin (Araneoidea) (Griswold et al., 2005, 1999, 1998; Townley and Tillinghast, 2009). Additionally, *T. per*_Sp_N spidroin repetitive sequence lacks the motifs that are characteristic of flagelliform spidroins (proline-rich motifs, intervening spacers; (Hayashi and Lewis, 2001, 1998)). Flagelliform spigots, however, have been hypothesized to be homologous to pseudoflagelliform spigots, which are unique to some cribellate taxa (Coddington, 1986; Eberhard and Pereira, 1993; Hajer, 1991). Recently, Alfaro et al. (Alfaro et al., 2018b) proposed that the modified spigot of *Tengella* is homologous to the modified/pseudoflagelliform silk spigot in other cribellate species. Thus, *T. per*_Sp_N may be associated with

pseudoflagelliform glands. More specific annotation of this *T. perfuga* spidroin beyond "Sp" (e.g., as a pseudoflagelliform spidroin) requires future work to obtain more complete sequence and more closely related spidroins.

We were able to be more definitive with out annotation of the novel *T. perfuga* transcript that contained the spidroin C-terminal region; we associated our *T. perfuga* sequence with cribellar silk and thus named it *T. per*_CrSp_C. Recent studies describing the web-building ontogeny of *T. perfuga* found that females deploy vast amounts of cribellar silk during web and retreat construction (Mallis and Miller, 2017). By contrast, *T. perfuga* males were found to use cribellar silk as juveniles and then lose the spigots associated with cribellar silks at their final molt (Alfaro et al., 2018a; Mallis and Miller, 2017). Consistent with this observation, *T. per*_CrSp_C was present in the female tissue cDNA library constructed from the small glands attached to the spinnerets, which is where cribellar glands are expected to be located. This transcript was not present in our male (mature) tissue cDNA libraries (Table S3.2).

Phylogenetic analysis of C-terminal regions provides further support for our annotation of this *T. perfuga* transcript as a CrSp. *T.per*_CrSp_C formed a clade with *S. mim*_Sp1 and the two C-termini shared 55% aa identity (Figure 3.1A, 1B). Additionally, the repetitive sequence of *T. per*_CrSp_C and *S. mim*_Sp1 lack motifs that are characteristic of other spidroin types, but share a novel 158 aa long repeat unit (56% identity at the aa level; Figure 3.1C). The identification of *T. per*_CrSp_C only in females, its distinct repeat sequence, and the placement of *T. per*_CrSp_C in a separate clade from the previously known spidroin types, all support that *T. per*_CrSp_C is a

cribellar silk spidroin. In our analysis (Fig. 1), the presence of CrSp orthologs only in the cribellate spiders *S. mimosarum* and *T. perfuga* suggests that CrSp has been lost in spiders that are secondarily ecribellate (without a cribellum).

Spidroin Gene Expression in T. perfuga Spiders

To investigate the relationship of silk gene expression and silk use in *T. perfuga*, we examined spidroin gene expression levels by comparing RPKM of contigs containing C-terminal regions. A similar pattern of expression was observed using contigs containing N-terminal regions. We found spidroin gene expression in *T. perfuga* spiders to be dominated (91%) by genes associated with egg sac construction (*T. per_TuSp_C*) and web construction (*T. per_CrSp_C*, *T. per_AmSp_C_vA*, *T. per_AmSp_C_vB*, and *T. per_AmSp_C_vC*). Female spiders wrap their egg cases mostly with tubuliform silk fibers to protect the developing embryos (Garb and Hayashi, 2005; Tian and Lewis, 2005). Thus, it was expected that the expression of *T. per_TuSp_C* would be one of the highest expressed spidroins in *T. perfuga* females (second highest, Figure 3.2), and absent in our male (mature) silk gland cDNA libraries given that males do not make egg cases (Table S3.2).

The webs of *T. perfuga* spiders are sheet-like, with deep retreats and knockdown lines extending from overhanging substrate to the sheet below (Mallis and Miller, 2017). These structures are composed of at least two different silk types, with the primary silk type corresponding to dragline (ampullate) silk, and the secondary type corresponding to cribellar silk (Mallis and Miller, 2017). We found ampullate spidroin genes (*T.*

per_AmSp_C_vA, *T. per_AmSp_C_vB*, and *T. per_AmSp_C_vC*) to have the highest combined relative expression compared to other spidroin genes in *T. perfuga* female silk glands (Figure 3.2). This suggests that the ampullate spidroins are the most abundant proteins produced, which is consistent with ampullate silk fibers being the primary component of the sheet web, retreat, and knockdown lines. Differences in expression levels among ampullate spidroins were also observed. One ampullate spidroin (*T. per_AmSp_C_vA*) had the highest relative expression, when compared to other spidroins (Figure 3.2). We also found *T. per_CrSp_C*, the gene for the putative cribellar spidroin (CrSp), to account for ~ 7% of total *T. perfuga* silk gene expression.

Conclusions

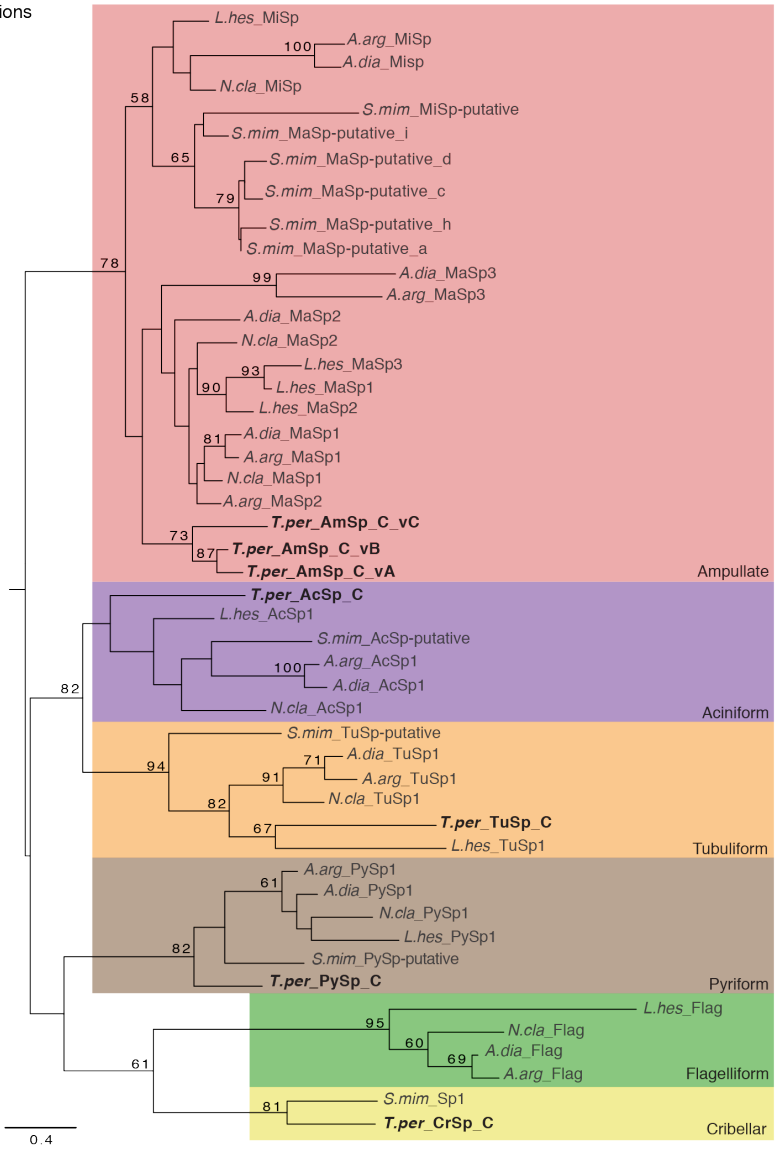
We identified 13 new spidroin contigs from the cribellate spider *T. perfuga*. All are partial length, seven of which are N-terminal region fragments and six C-terminal region fragments (Table S3.2). This means that there are at least seven spidroin genes present in *T. perfuga* genome. As predicted from spigot morphology, we found *T. perfuga* spiders to express genes associated with previously described aciniform, ampullate, tubuliform, and pyriform silks.

T. perfuga use multiple variants of ampullate spidroins and ampullate spidroins account for most of the spidroin expression in a female (Figure 3.2). The second most highly expressed spidroin in females is TuSp, which is involved in egg case production. The *T. perfuga* TuSp is well conserved compared to the TuSp from other species. However, the *T. perfuga* aciniform and pyriform spidroins have novel combinations of

motifs (Figure S3.2). The *T. perfuga* AcSp has poly-serine as is typical of AcSp, but surprisingly also has poly-alanine, which may enhance tensile strength. Similarly, the *T. perfuga* PySp possesses both the PX motif prevalent in the PySp of some species, and alanine rich motifs (AAASARAEAXAR, AAASXRAA) similar to the PySp from cob-web weavers. *T. perfuga* has the first PySp that we know of that combines the PX extensibility motif and the alanine-rich motifs in the same repeat, which is a combination that has structure/function implications.

We also documented expression of a candidate cribellar spidroin, CrSp. *T. perfuga* is a cribellate spider, although males lose the ability to spin cribellar silk when they mature. We show evidence that *T. perfuga* CrSp is expressed by *T. perfuga* mature females but not mature males. *T. perfuga* CrSp has distinctive repetitive and C-terminal region sequences and gene tree analysis and pairwise alignments show an affinity with a spidroin from *S. mimosarum*, another cribellate species (Figure 3.1). Discovery of a candidate cribellate spidroin is significant as it provides insights into our understanding of the composition of cribellar silk. Furthermore, we can now begin to relate CrSp sequence to the adhesive properties of cribellar silk and trace the evolution of CrSp across different cribellate and ecribellate spider lineages.

A. C-terminal regions



B. CrSp C-terminal regions

```
T.per_CrSp_C  TRISSLVSVIRSSLPRAGKKFDYLSFARGLSRMILDISIANPSYSSSDIL [50]
S.mim_Sp1     KRIASLISVIISSLPAPGGKFDYLTFARGLASLLSDIRAGNPSYSASDVI [50]

T.per_CrSp_C  VEGLLNALAATLQLLEQANLSEINVINSQYVTRTLSGALRVAF-SGQLA  [98]
S.mim_Sp1     TEGLLEALVAFIQMEEYITLSDRPIEYSDYVTKAISDSLNVAFKSQQLI  [99]
```

C. CrSp repeats

```
T.per_CrSp_C  AFGSHLYGTLLVNPRFVTVFGSDFSLERSRLFLSVLSSRIHSFPQFSSIPVQYLLNRYTDVVASIPFGSSEQIYARRIAQ [ 80]
S.mim_Sp1     AFGSHLYGTLLVNPRFSTLFGSEFSLEKVRPFLFALASHIHSFSQFSSISANDLFERYIEVVNALPLGSSVQAYALALSQ [ 80]

T.per_CrSp_C  ETASVLYKNNLLSWQILASEDAAVDKAAEDAGAVLSQFASLSDQSISLSSSTEDVAASMAASAVLSPSVLETLATAEA-  [158]
S.mim_Sp1     ATAELLYENNLLSWDALAKEDAEAAGAGEAQATVSSTLVS--------SSTVESAAAETAASAILSPSVLSILSSSESE  [151]
```

Figure 3.1. Phylogenetic analysis of *Tengella perfuga* spidroins and alignment of *T. perfuga* cribellar C-terminal and repeat regions with *Stegodyphus mimosarum* Spidroin 1. (**A)** C-terminal regions maximum likelihood tree. Shaded boxes indicate spidroin types, annotated as ampullate (pink), aciniform (purple), tubuliform (orange), pyriform (brown), flagelliform (green), and cribellar (yellow). Tree rooted with California trapdoor spider *Bothriocyrtum californicum* fibroin 1 (not shown). Names abbreviated as in Tables S3.2-S3.3. Bootstrap percentages ≥ 50% are shown. Scale bar represents substitutions per site. (**B**) C-terminal regions and (**C**) repeat regions of *T. perfuga* cribellar spidroin aligned with *Stegodyphus mimosarum* Spidroin 1 (*S.mim*_Sp1). Gaps inserted into the alignment are indicated by dashes. Total amino acids shown on the right.

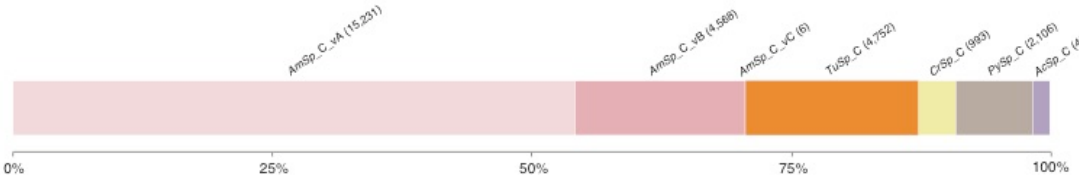


Figure 3.2. Relative silk gene expression in female *Tengella perfuga* silk glands. Silk transcripts containing the C-terminal domain are shown. Average expression from two biological replicates of *T. perfuga* total silk gland library reads mapped to our *de novo T. perfuga* transcriptome. Expression is shown as reads per kilobase of transcript per million mapped reads (RPKM, average total for each transcript shown in parentheses). Colors indicate spidroin types as in Figure 3.1. Names abbreviated as in Table S3.2. Total RPKM of silk genes 28,114.

References

Alfaro, R.E., Griswold, C.E., Miller, K.B., 2018a. The ontogeny of the spinning apparatus of *Tengella perfuga* Dahl (Araneae: Zoropsidae). Invertebrate Biology Forthcoming.

Alfaro, R.E., Griswold, C.E., Miller, K.B., 2018b. -Comparative spigot ontogeny across the spider tree of life. PeerJ 6, e4233. https://doi.org/10.7717/peerj.4233

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Ayoub, N.A., Garb, J.E., Kuelbs, A., Hayashi, C.Y., 2013. Ancient properties of spider silks revealed by the complete gene sequence of the prey-wrapping silk protein (AcSp1). Mol. Biol. Evol. 30, 589–601. https://doi.org/10.1093/molbev/mss254

Ayoub, N.A., Garb, J.E., Tinghitella, R.M., Collin, M.A., Hayashi, C.Y., 2007. Blueprint for a high-performance biomaterial: full-length spider dragline silk genes. PloS One 2, e514. https://doi.org/10.1371/journal.pone.0000514

Ayoub, N.A., Hayashi, C.Y., 2008. Multiple recombining loci encode MaSp1, the primary constituent of dragline silk, in widow spiders (*Latrodectus*: Theridiidae). Mol. Biol. Evol. 25, 277–286. https://doi.org/10.1093/molbev/msm246

Babb, P.L., Lahens, N.F., Correa-Garhwal, S.M., Nicholson, D.N., Kim, E.J., Hogenesch, J.B., Kuntner, M., Higgins, L., Hayashi, C.Y., Agnarsson, I., Voight, B.F., 2017. The *Nephila clavipes* genome highlights the diversity of spider silk genes and their complex expression. Nat. Genet. 49, 895–903. https://doi.org/10.1038/ng.3852

Becker, M.A., Mahoney, D.V., Lenhert, P.G., Eby, R.K., Kaplan, D., Adams, W.W., 1994. X-ray Moduli of Silk Fibers from *Nephila clavipes* and *Bombyx mori*, in: Silk Polymers, ACS Symposium Series. American Chemical Society, Washington, DC, pp. 185–195. https://doi.org/10.1021/bk-1994-0544.ch017

Blackledge, T.A., Hayashi, C.Y., 2006. Silken toolkits: biomechanics of silk fibers spun by the orb web spider *Argiope argentata* (Fabricius 1775). J. Exp. Biol. 209, 2452–2461. https://doi.org/10.1242/jeb.02275

Blackledge, T.A., Scharff, N., Coddington, J.A., Szüts, T., Wenzel, J.W., Hayashi, C.Y., Agnarsson, I., 2009. Reconstructing web evolution and spider diversification in the molecular era. Proc. Natl. Acad. Sci. U. S. A. 106, 5229–5234. https://doi.org/10.1073/pnas.0901377106

Blasingame, E., Tuton-Blasingame, T., Larkin, L., Falick, A.M., Zhao, L., Fong, J., Vaidyanathan, V., Visperas, A., Geurts, P., Hu, X., others, 2009. Pyriform spidroin 1, a novel member of the silk gene family that anchors dragline silk fibers in attachment discs of the black widow spider, *Latrodectus hesperus*. J. Biol. Chem. 284, 29097–29108.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bond, J.E., Opell, B.D., 1998. Testing adaptive radiation and key innovation hypotheses in spiders. Evolution 52, 403–414. https://doi.org/10.2307/2411077

Chaw, R.C., Saski, C.A., Hayashi, C.Y., 2017. Complete gene sequence of spider attachment silk protein (PySp1) reveals novel linker regions and extreme repeat homogenization. Insect Biochem. Mol. Biol. 81, 80–90. https://doi.org/10.1016/j.ibmb.2017.01.002

Chaw, R.C., Zhao, Y., Wei, J., Ayoub, N.A., Allen, R., Atrushi, K., Hayashi, C.Y., 2014. Intragenic homogenization and multiple copies of prey-wrapping silk genes in *Argiope* garden spiders. BMC Evol. Biol. 14, 31. https://doi.org/10.1186/1471-2148-14-31

Chen, G., Liu, X., Zhang, Y., Lin, S., Yang, Z., Johansson, J., Rising, A., Meng, Q., 2012. Full-length minor ampullate spidroin gene sequence. PLoS ONE 7, e52293. https://doi.org/10.1371/journal.pone.0052293

Clarke, T.H., Garb, J.E., Hayashi, C.Y., Arensburger, P., Ayoub, N.A., 2015. Spider transcriptomes identify ancient large-scale gene duplication event potentially important in silk gland evolution. Genome Biol. Evol. 7, 1856–1870. https://doi.org/10.1093/gbe/evv110

Clarke, T.H., Garb, J.E., Hayashi, C.Y., Haney, R.A., Lancaster, A.K., Corbett, S., Ayoub, N.A., 2014. Multi-tissue transcriptomics of the black widow spider reveals expansions, co-options, and functional processes of the silk gland gene toolkit. BMC Genomics 15. https://doi.org/10.1186/1471-2164-15-365

Coddington, J.A., 1989. Spinneret silk spigot morphology: evidence for the monophyly of orbweaving spiders, Cyrtophorinae (Araneidae), and the group Theridiidae plus Nesticidae. J. Arachnol. 17, 71–95.

Coddington, J.A., 1986. The monophyletic origin of the orb web. Spiders Webs Behav. Evol. Shear Ed Stanf. Univ. Press Stanf. Calif. 319–363.

Colgin, M.A., Lewis, R.V., 1998. Spider minor ampullate silk proteins contain new repetitive sequences and highly conserved non-silk-like “spacer regions”. Protein Sci. Publ. Protein Soc. 7, 667–672.

Correa-Garhwal, S.M., Chaw, R.C., Clarke, T.H., Ayoub, N.A., Hayashi, C.Y., 2017. Silk gene expression of theridiid spiders: implications for male-specific silk use. Zoology 122, 107–114. https://doi.org/10.1016/j.zool.2017.04.003

Craig, C.L., 2003. Spiderwebs and silk: tracing evolution from molecules to genes to phenotypes. Oxford University Press.

Eberhard, W., Pereira, F., 1993. Ultrastructure of cribellate silk of nine species in eight families and possible taxonomic implications (Araneae: Amaurobiidae, Deinopidae, Desidae, Dictynidae, Filistatidae, Hypochilidae, Stiphidiidae, Tengellidae). J. Arachnol. 21, 161–174.

Eberhard, W.G., 1990. Function and phylogeny of spider webs. Annu. Rev. Ecol. Syst. 21, 341–372. https://doi.org/10.1146/annurev.es.21.110190.002013

Foelix, R., 2011. Biology of Spiders. Oxford University Press.

Garb, J.E., Ayoub, N.A., Hayashi, C.Y., 2010. Untangling spider silk evolution with spidroin terminal domains. BMC Evol. Biol. 10, 243.

Garb, J.E., DiMauro, T., Lewis, R.V., Hayashi, C.Y., 2007. Expansion and intragenic homogenization of spider silk genes since the Triassic: evidence from Mygalomorphae (tarantulas and their kin) spidroins. Mol. Biol. Evol. 24, 2454–2464. https://doi.org/10.1093/molbev/msm179

Garb, J.E., Hayashi, C.Y., 2005. Modular evolution of egg case silk genes across orb-weaving spider superfamilies. Proc. Natl. Acad. Sci. U. S. A. 102, 11379–11384. https://doi.org/10.1073/pnas.0502473102

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.

Gatesy, J., Hayashi, C., Motriuk, D., Woods, J., Lewis, R., 2001. Extreme diversity, conservation, and convergence of spider silk fibroin sequences. Science 291, 2603–2605. https://doi.org/10.1126/science.1057561

Geurts, P., Zhao, L., Hsia, Y., Gnesa, E., Tang, S., Jeffery, F., Mattina, C.L., Franz, A., Larkin, L., Vierra, C., 2010. Synthetic spider silk fibers spun from pyriform spidroin 2, a glue silk protein discovered in orb-weaving spider attachment discs. Biomacromolecules 11, 3495–3503. https://doi.org/10.1021/bm101002w

Gosline, J.M., DeMont, M.E., Denny, M.W., 1986. The structure and properties of spider silk. Endeavour 10, 37–43. https://doi.org/10.1016/0160-9327(86)90049-9

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., 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

Griswold, C.E., Coddington, J.A., Hormiga, G., Scharff, N., 1998. Phylogeny of the orb-web building spiders (Araneae, Orbiculariae: Deinopoidea, Araneoidea). Zool. J. Linn. Soc. 123, 1–99.

Griswold, C.E., Coddington, J.A., Platnick, N.I., Forster, R.R., 1999. Towards a phylogeny of entelegyne spiders (Araneae, Araneomorphae, Entelegynae). J. Arachnol. 53–63.

Griswold, C.E., Ramírez, M.J., Coddington, J.A., Platnick, N.I., 2005. Atlas of phylogenetic data for Entelegyne spiders (Araneae: Araneomorphae: Entelegynae), with comments on their phylogeny. Proc.-Calif. Acad. Sci. 56, 1.

Guerette, P.A., Ginzinger, D.G., Weber, B.H.F., Gosline, J.M., 1996. Silk properties determined by gland-specific expression of a spider fibroin gene family. Science 272, 112–115. https://doi.org/10.1126/science.272.5258.112

Hajer, J., 1991. Notes on the spinning of the spiders *Hyptiotes paradoxus* C.L.K., 1834, and *Uloborus wakkenaerius* Latr., 1806 (Araneae: Uloboridae). Bulletin de la Société des Science Naturelles de Neuchâtel 116, 99–103. https://doi.org/DOI 10.5169/seals-89371

Hawthorn, A.C., Opell, B.D., 2003. van der Waals and hygroscopic forces of adhesion generated by spider capture threads. J. Exp. Biol. 206, 3905–3911.

Hawthorn, A.C., Opell, B.D., 2002. Evolution of adhesive mechanisms in cribellar spider prey capture thread: evidence for van der waals and hygroscopic forces: Evolution of adhesive mechanisms. Biol. J. Linn. Soc. 77, 1–8.

Hayashi, C.Y., Blackledge, T.A., Lewis, R.V., 2004. Molecular and mechanical characterization of aciniform silk: uniformity of iterated sequence modules in a novel member of the spider silk fibroin gene family. Mol. Biol. Evol. 21, 1950–1959. https://doi.org/10.1093/molbev/msh204

Hayashi, C.Y., Lewis, R.V., 2001. Spider flagelliform silk: lessons in protein design, gene structure, and molecular evolution. Bioessays 23, 750–756.

Hayashi, C.Y., Lewis, R.V., 1998. Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks. J. Mol. Biol. 275, 773–784. https://doi.org/10.1006/jmbi.1997.1478

Hayashi, C.Y., Shipley, N.H., Lewis, R.V., 1999. Hypotheses that correlate the sequence, structure, and mechanical properties of spider silk proteins. Int. J. Biol. Macromol. 24, 271–275.

Hinman, M.B., Lewis, R.V., 1992. Isolation of a clone encoding a second dragline silk fibroin. *Nephila clavipes* dragline silk is a two-protein fiber. J. Biol. Chem. 267, 19320–19324.

Jenkins, J.E., Sampath, S., Butler, E., Kim, J., Henning, R.W., Holland, G.P., Yarger, J.L., 2013. Characterizing the secondary protein structure of black widow dragline silk using solid-state NMR and X-ray diffraction. Biomacromolecules 14, 3472–3483. https://doi.org/10.1021/bm400791u

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., Salzberg, S.L., 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14, R36. https://doi.org/10.1186/gb-2013-14-4-r36

Kovoor, J., Zylberberg, L., 1982. Fine structural aspects of silk secretion in a spider. II. Conduction in the pyriform glands. Tissue Cell 14, 519–530.

Kovoor, J., Zylberberg, L., 1980. Fine structural aspects of silk secretion in a spider (*Araneus diadematus*). I. Elaboration in the pyriform glands. Tissue Cell 12, 547–556. https://doi.org/10.1016/0040-8166(80)90044-0

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948. https://doi.org/10.1093/bioinformatics/btm404

Lawrence, B.A., Vierra, C.A., Moore, A.M., 2004. Molecular and mechanical properties of major ampullate silk of the black widow spider, *Latrodectus hesperus*. Biomacromolecules 5, 689–695.

Leister, M., Mallis, R., Miller, K., 2013. The male of *Tengella perfuga* Dahl, 1901 with re-description of the female and comparisons with *T. radiata* (Kulczynski, 1909) (Araneae: Tengellidae). Zootaxa 3709, 185–199.

Mallis, R., Miller, K., 2017. Natural history and courtship behavior in *Tengella perfuga* Dahl, 1901 (Araneae: Zoropsidae). BioOne 45, 166–176.

Opell, B.D., 1994. Factors governing the stickiness of cribellar prey capture threads in the spider family Uloboridae. J. Morphol. 221, 111–119. https://doi.org/10.1002/jmor.1052210109

Perry, D.J., Bittencourt, D., Siltberg-Liberles, J., Rech, E.L., Lewis, R.V., 2010. Piriform spider silk sequences reveal unique repetitive elements. Biomacromolecules 11, 3000–3006.

Peters, H.M., 1992. On the spinning apparatus and the structure of the capture threads of *Deinopis subrufus* (Araneae, Deinopidae). Zoomorphology 112, 27–37.

Peters, H.M., 1987. Fine structure and function of capture threads, in: Ecophysiology of Spiders. Springer, Berlin, Heidelberg, pp. 187–202. https://doi.org/10.1007/978-3-642-71552-5_13

Peters, H.M., 1984. The spinning apparatus of Uloboridae in relation to the structure and construction of capture threads (Arachnida, Araneida). Zoomorphology 104, 96–104.

Polotow, D., Carmichael, A., Griswold, C.E., 2015. Total evidence analysis of the phylogenetic relationships of Lycosoidea spiders (Araneae, Entelegynae). Invertebr. Syst. 29, 124–163. https://doi.org/10.1071/IS14041

Sanggaard, K.W., Bechsgaard, J.S., Fang, X., Duan, J., Dyrlund, T.F., Gupta, V., Jiang, X., Cheng, L., Fan, D., Feng, Y., Han, L., Huang, Z., Wu, Z., Liao, L., Settepani, V., Thøgersen, I.B., Vanthournout, B., Wang, T., Zhu, Y., Funch, P., Enghild, J.J., Schauser, L., Andersen, S.U., Villesen, P., Schierup, M.H., Bilde, T., Wang, J., 2014. Spider genomes provide insight into composition and evolution of venom and silk. Nat. Commun. 5, 3765. https://doi.org/10.1038/ncomms4765

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Simmons, A.H., Michal, C.A., Jelinski, L.W., 1996a. Molecular orientation and two-component nature of the crystalline fraction of spider dragline silk. Science 271, 84–87. https://doi.org/10.1126/science.271.5245.84

Simmons, A.H., Michal, C.A., Jelinski, L.W., 1996b. Molecular orientation and two-component nature of the crystalline fraction of spider dragline silk. Science 271, 84–87.

Sponner, A., Unger, E., Grosse, F., Weisshart, K., 2005. Differential polymerization of the two main protein components of dragline silk during fibre spinning. Nat. Mater. 4, 772–775. https://doi.org/10.1038/nmat1493

Stamatakis, A., 2014. Raxml version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Starrett, J., Garb, J.E., Kuelbs, A., Azubuike, U.O., Hayashi, C.Y., 2012. Early events in the evolution of spider silk genes. PloS One 7, e38084. https://doi.org/10.1371/journal.pone.0038084

Tian, M., Lewis, R.V., 2005. Molecular characterization and evolutionary study of spider tubuliform (eggcase) silk protein. Biochemistry (Mosc.) 44, 8006–8012. https://doi.org/10.1021/bi050366u

Townley, M.A., Tillinghast, E.K., 2009. Developmental changes in spider spinning fields: a comparison between *Mimetus* and *Araneus* (Araneae: Mimetidae, Araneidae). Biol. J. Linn. Soc. 98, 343–383. https://doi.org/10.1111/j.1095-8312.2009.01297.x

Trancik, J.E., Czernuszka, J.T., Cockayne, D.J.H., Viney, C., 2005. Nanostructural physical and chemical information derived from the unit cell scattering amplitudes of a spider dragline silk. Polymer 46, 5225–5231. https://doi.org/10.1016/j.polymer.2005.04.007

Vasanthavada, K., Hu, X., Falick, A.M., La Mattina, C., Moore, A.M., Jones, P.R., Yee, R., Reza, R., Tuton, T., Vierra, C., 2007. Aciniform spidroin, a constituent of egg case sacs and wrapping silk fibers from the black widow spider *Latrodectus hesperus*. J. Biol. Chem. 282, 35088–35097.

Vienneau-Hathaway, J.M., Brassfield, E.R., Lane, A.K., Collin, M.A., Correa-Garhwal, S.M., Clarke, T.H., Schwager, E.E., Garb, J.E., Hayashi, C.Y., Ayoub, N.A., 2017. Duplication and concerted evolution of MiSp-encoding genes underlie the material properties of minor ampullate silks of cobweb weaving spiders. BMC Evol. Biol. 17, 78. https://doi.org/10.1186/s12862-017-0927-x

Zhang, Y., Zhao, A.-C., Sima, Y.-H., Lu, C., Xiang, Z.-H., Nakagaki, M., 2013. The molecular structures of major ampullate silk proteins of the wasp spider, *Argiope bruennichi*: A second blueprint for synthesizing de novo silk. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 164, 151–158.

Zhao, A.-C., Zhao, T.-F., Nakagaki, K., Zhang, Y.-S., SiMa, Y.-H., Miao, Y.-G., Shiomi, K., Kajiura, Z., Nagata, Y., Takadera, M., Nakagaki, M., 2006. Novel molecular and mechanical properties of egg case silk from wasp spider, *Argiope bruennichi*. Biochemistry (Mosc.) 45, 3348–3356. https://doi.org/10.1021/bi052414g

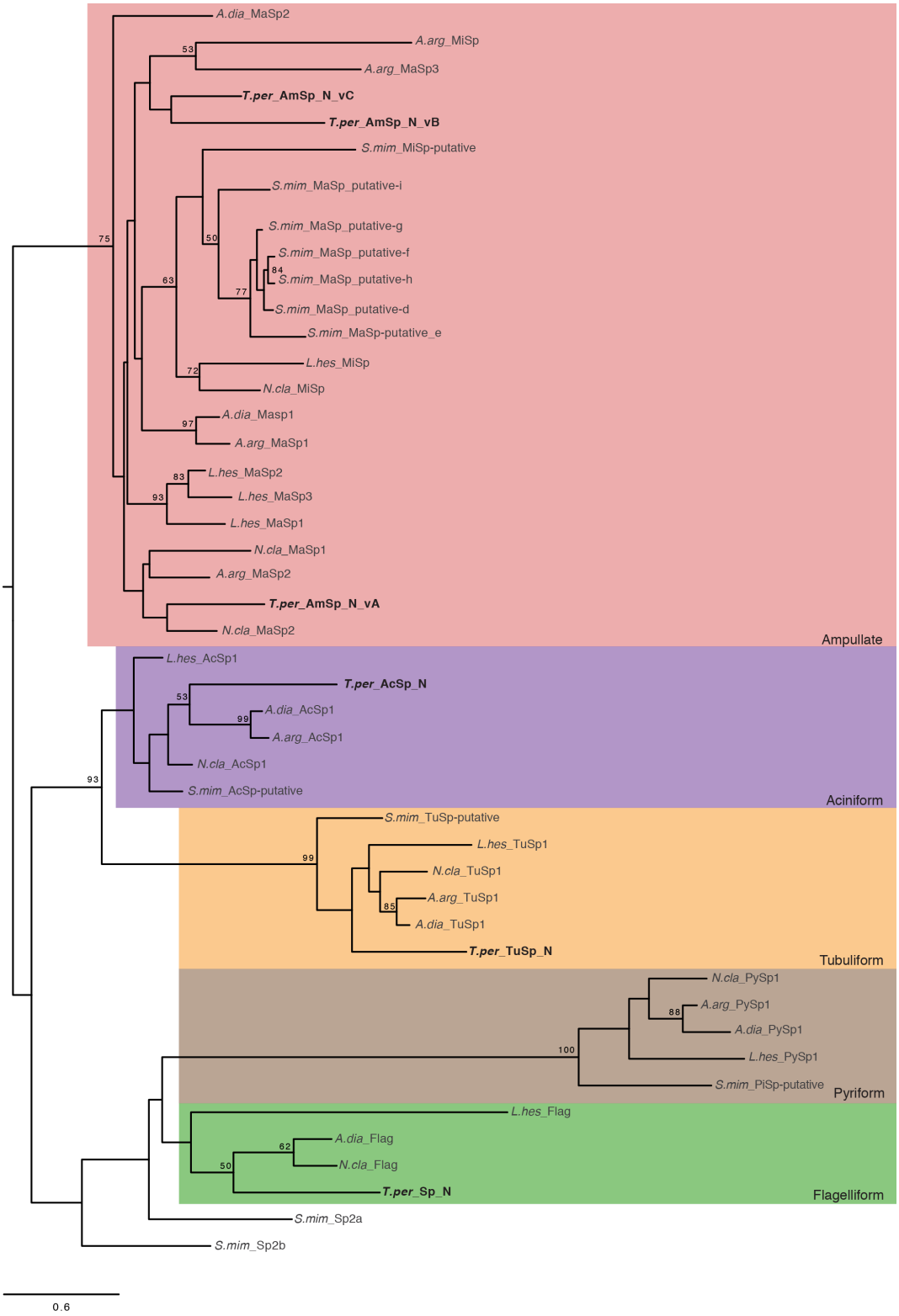

A.dia_MaSp2
A.arg_MiSp
A.arg_MaSp3
T.per_AmSp_N_vC
T.per_AmSp_N_vB
S.mim_MiSp-putative
S.mim_MaSp_putative-i
S.mim_MaSp_putative-g
S.mim_MaSp_putative-f
S.mim_MaSp_putative-h
S.mim_MaSp_putative-d
S.mim_MaSp-putative_e
L.hes_MiSp
N.cla_MiSp
A.dia_Masp1
A.arg_MaSp1
L.hes_MaSp2
L.hes_MaSp3
L.hes_MaSp1
N.cla_MaSp1
A.arg_MaSp2
T.per_AmSp_N_vA
N.cla_MaSp2
Ampullate
L.hes_AcSp1
T.per_AcSp_N
A.dia_AcSp1
A.arg_AcSp1
N.cla_AcSp1
S.mim_AcSp-putative
Aciniform
S.mim_TuSp-putative
L.hes_TuSp1
N.cla_TuSp1
A.arg_TuSp1
A.dia_TuSp1
T.per_TuSp_N
Tubuliform
N.cla_PySp1
A.arg_PySp1
A.dia_PySp1
L.hes_PySp1
S.mim_PiSp-putative
Pyriform
L.hes_Flag
A.dia_Flag
N.cla_Flag
T.per_Sp_N
Flagelliform
S.mim_Sp2a
S.mim_Sp2b
0.6

Figure S3.1. Maximum likelihood tree of spidroin N-terminal regions. Shaded boxes indicate spidroin types as in Figure 3.1. Tree rooted with California trapdoor spider *Bothriocyrtum californicum* fibroin 1 (not shown). Names abbreviated as in Tables S3.2-S3.3. Bootstrap percentages ≥ 50% are shown. Scale bar represents substitutions per site.

Figure S3.2. Spidroin repetitive sequences of *Tengella perfuga*. (**A**) Repetitive sequence adjacent to N-terminal region. (**B**) Repetitive sequence adjacent to C-terminal region. Spidroin names abbreviated as in Table S3.2. Amino acids abundant in silks are highlighted: alanine (red), serine (blue), and glycine (green). Pyriform amino acid motifs indicated in boxes. Total number of amino acids indicated in parentheses

Table S3.1. Summary of *Tengella perfuga de novo* transcriptome assembly.

No. Raw Paired Reads	81,814,325
No. Cleaned Paired Reads	79,352,198
No. Trinity Contigs	127,600
Total Length (bp)	87,246,581
N50 (bp)	1,052
BUSCO % complete (*Ixodes* reference)	97.7

Table S3.2. *Tengella perfuga* spidroins.

Spidroin Name[a,b]	Library type	Top BLASTx Hit Accession	Top BLASTx Hit Description	E-value
*T. per*_AcSp_C	RNA-Seq of silk glands combined, cDNA of tubuliform silk glands (Fem)	gi|422900768|gb|AFX83561.1|	Aciniform spidroin 1, partial [*Latrodectus hesperus*]	1.10e-14
*T. per*_AcSp_N	RNA-Seq of silk glands combined (Fem)	gi|675387023|gb|KFM79920.1|	Hypothetical protein X975_02929, partial [*Stegodyphus mimosarum*]	5.15e-51
*T. per*_AmSp_N_vA	RNA-Seq of silk glands combined (Fem)	gi|295982412|pdb|3LR2|	Chain A, Self-Assembly Of Spider Silk Proteins Is Controlled By A Ph-Sensitive Relay	2.84e-25
*T. per*_AmSp_N_vB	RNA-Seq of silk glands combined (Fem)	gi|164709230|gb|ABY67420.1|	Major ampullate spidroin 1 locus 3 [*Latrodectus geometricus*]	4.07e-26
*T. per*_AmSp_N_vC	RNA-Seq of silk glands combined (Fem)	gi|295982412|pdb|3LR2|	Chain A, Self-Assembly Of Spider Silk Proteins Is Controlled By A Ph-Sensitive Relay	5.95e-41
*T. per*_AmSp_C_vA	RNA-Seq of silk glands combined (Fem); cDNA of ampullate silk glands (Fem)	gi|38197751|gb|AAR13810.1|	Major ampullate -2 [*Argiope amoena*]	4.05e-21
*T. per*_AmSp_C_vB	RNA-Seq of silk glands combined (Fem)	gi|294440291|gb|ADE74592.1|	Major ampullate spidroin 1 [*Peucetia viridans*]	1.02e-24
*T. per*_AmSp_C_vC	RNA-Seq of silk glands combined (Fem); cDNA of ampullate silk glands (Male)	gi|294440291|gb|ADE74592.1|	Major ampullate spidroin 1 [*Peucetia viridans*]	5.81e-21
*T. per*_PySp_C	RNA-Seq of silk glands combined (Fem)	gi|1148301527|gbAQR58363.1|	Pyriform spidroin 1 [*Argiope argentata*]	4.78e-11
*T. per*_CrSp_C	RNA-Seq of silk glands combined, cDNA of small silk glands attached to spinneret (Fem)	gi|675367732|gb|KFM60634.1|	Hypothetical protein X975_22661 [*Stegodyphus mimosarum*]	6.19e-71
*T. per*_Sp_N	RNA-Seq of silk glands combined (Fem)	gi|675381008|gb|KFM73910.1|	Hypothetical protein X975_01894 [*Stegodyphus mimosarum*]	1.58e-26
*T. per*_TuSp_C	RNA-Seq of silk glands combined, cDNA of tubuliform silk glands (Fem)	gi|303307781|gb|ADM14330.1|	Tubuliform spidroin 1, partial [*Agelenopsis aperta*]	4.14e-42
*T. per*_TuSp_N	RNA-Seq of silk glands combined (Fem)	gi|303307781|gb|ADM14330.1|	Tubuliform spidroin 1, partial [*Agelenopsis aperta*]	2.39e-28

[a] N or C refer in the spidroin names indicate whether a contig contains the N- or C- terminal region coding sequence.

[b] Variant name (e.g. _vA) does not indicate association of N- term transcripts with C-terminal transcripts.

Table S3.3. Spidroin sequences from GenBank used in phylogenetic analyses.

Spidroin Name	Species	N-terminal region Accession	C-terminal region Accession
***A.arg*_AcSp1**	*Argiope argentata*	AHK09813	AHK09813
***A.arg*_Flag**	*Argiope argentata*	--	MF955778
***A.arg*_MaSp1**	*Argiope argentata*	MF955677	MF955761
***A.arg*_MaSp2**	*Argiope argentata*	MF955700	MF955804
***A.arg*_MaSp3**	*Argiope argentata*	MF955785	MF955690
***A.arg*_MiSp**	*Argiope argentata*	MF955726	MF955717
***A.arg*_PySp1**	*Argiope argentata*	AQR58363	AQR58363
***A.arg*_TuSp1**	*Argiope argentata*	ATW75951	ATW75951
***A.dia*_AcSp1**	*Araneus diadematus*	MF955743	MF955754
***A.dia*_Flag**	*Araneus diadematus*	MF955789	MF955779
***A.dia*_Masp1**	*Araneus diadematus*	MF955789	MF955789
***A.dia*_MaSp2**	*Araneus diadematus*	MF955703	MF955809
***A.dia*_MaSp3**	*Araneus diadematus*	--	MF955691
***A.dia*_MiSp**	*Araneus diadematus*	--	MF955718
***A.dia*_PySp1**	*Araneus diadematus*	MF955713	MF955772
***A.dia*_TuSp1**	*Araneus diadematus*	MF955696	MF955799
***B.cal*_Fibroin1**	*Bothriocyrtum californicum*	HM752562	EU117162
***L.hes*_AcSp1**	*Latrodectus hesperus*	AFX83557	AFX83557
***L.hes*_Flag**	*Latrodectus hesperus*	MF955792	MF955781
***L.hes*_MaSp1**	*Latrodectus hesperus*	F595246	F595246
***L.hes*_MaSp2**	*Latrodectus hesperus*	F595245	F595245
***L.hes*_MaSp3**	*Latrodectus hesperus*	MF955786	MF955692
***L.hes*_MiSp**	*Latrodectus hesperus*	ARA91152	ARA91152
***L.hes*_PySp1**	*Latrodectus hesperus*	MF955714	MF955773
***L.hes*_TuSp1**	*Latrodectus hesperus*	MF955697	MF955801
***N.cla*_AcSp1**	*Nephila clavipes*	MF955747	MF955758
***N.cla*_Flag**	*Nephila clavipes*	MF955793	MF955782
***N.cla*_MaSp1**	*Nephila clavipes*	MF955682	MF955765
***N.cla*_MaSp2**	*Nephila clavipes*	MF955708	MF955815
***N.cla*_MiSp**	*Nephila clavipes*	MF955734	MF955722
***N.cla*_PySp1**	*Nephila clavipes*	MF955715	MF955774
***N.cla*_TuSp1**	*Nephila clavipes*	MF955698	MF955802
***N.cla*_TuSp1**	*Nephila clavipes*	MF955698	MF955802
***S.mim*_AcSp-putative**	*Stegodyphus mimosarum*	KFM79920	KFM79920
***S.mim*_MaSp-putative-a**	*Stegodyphus mimosarum*	--	KFM83271
***S.mim*_MaSp-putative-c**	*Stegodyphus mimosarum*	--	JT038023
***S.mim*_MaSp-putative-d**	*Stegodyphus mimosarum*	KFM59474	KFM59474
***S.mim*_MaSp-putative-e**	*Stegodyphus mimosarum*	KFM74936	--
***S.mim*_MaSp-putative-f**	*Stegodyphus mimosarum*	KFM61798	--
***S.mim*_MaSp-putative-g**	*Stegodyphus mimosarum*	KFM57717	--
***S.mim*_MaSp-putative-h**	*Stegodyphus mimosarum*	KFM61802	KFM61800
***S.mim*_MaSp-putative-i**	*Stegodyphus mimosarum*	KFM79313	KFM79313
***S.mim*_Misp-putative**	*Stegodyphus mimosarum*	KFM62627	KFM62627
***S.mim*_PiSp-putative**	*Stegodyphus mimosarum*	KFM75168	KFM68615
***S.mim*_Sp1**	*Stegodyphus mimosarum*	--	KFM60634
***S.mim*_Sp2a**	*Stegodyphus mimosarum*	KFM73910	--
***S.mim*_Sp2b**	*Stegodyphus mimosarum*	KFM70693	--
***S.mim*_TuSp-putative**	*Stegodyphus mimosarum*	KFM79920	KFM79920

Chapter 4

Silks of Aquatic Spiders: Specializations for Life Underwater

Abstract

Spiders are commonly found in terrestrial environments. However, a few species occupy aquatic or semi-aquatic habitats. Spiders rely heavily on their silk for functions related to their survival such as reproduction and dispersal. Aquatic and semi-aquatic spiders have developed different silk-related adaptations to survive in aquatic environments. The vast majority of spider silks that have been studied are those of orb- and cob-web weaving species, leaving the silks of water-associated spiders as well as many other terrestrial spiders largely undescribed. Here, we characterize silks from spiders from different environments: aquatic spiders *Argyroneta aquatica* and *Desis marina* as well as the terrestrial *Badumna longinqua*. From silk gland RNA-seq libraries, we report a total of 47 transcripts representing different homologs of the spidroin (spider fibroin) gene family. Some of these spidroins correspond to known spidroin types (aciniform, ampullate, cribellar, pyriform, and tubuliform), while other spidroins represent novel branches of the spidroin gene family. We also report a hydrophobic amino acid motif (GV) that, to date, is only found in the spidroins of aquatic and semi-aquatic spiders. Comparison of spider silk sequences with silks from other arthropods that convergently evolved underwater lifestyles, shows that there is a diversity of strategies to function in aquatic environments.

Introduction

Spiders use silk throughout their lives and most spider species are capable of producing multiple, functionally differentiated silks for diverse and essential purposes. The importance of silk to spiders is especially dramatic for water-associated spiders that rely on silk to survive immersion. Yet, the vast majority of silk molecular studies have been on terrestrial spiders (orb-weavers and cob-web weavers, e.g. Babb et al., 2017; Correa-Garhwal et al., 2017; Garb et al., 2010; Hayashi and Lewis, 2001). Spidroins (a contraction of "spider fibroins"; Hinman and Lewis, 1992) are the dominant components of spider silks. Spidroins have several distinctive characteristics. They tend to be very large proteins with high molecular weights >200 kiloDaltons (Ayoub et al., 2007, 2013a; Chen et al., 2012; Chaw et al., 2017a, 2017b). Also, the primary structure of a spidroin is mostly composed of a central repetitive region, consisting of repeating blocks of sequence that are enriched for the amino acids glycine, alanine, and serine (Garb et al., 2010; Gatesy et al., 2001). For example, the major ampullate proteins (MaSp1) proteins from the Western black widow (*Latrodectus hesperus*) are dominated by short glycine-rich regions and poly-alanine blocks, which form beta-sheets. The beta-sheets correspond to the crystalline domains that confer remarkable strength to dragline silk (Hayashi et al., 1999; Holland et al., 2008; Kümmerlen et al., 1996; Sponner et al., 2005). Furthermore, spidroins possess non-repetitive amino (N) and carboxy (C)-terminal regions that flank the central repetitive region and play an important role in fiber formation (Beckwitt and Arcidiacono, 1994; Gao et al., 2013; Huemmerich et al., 2004; Ittah et al., 2007; Sponner et al., 2004). For instance, N-terminal region regulates spidroin assembly by preventing

early aggregation and aiding self-assembly in a pH-dependent mechanism (Askarieh et al., 2010; Hedhammar et al., 2008). Contrary to the N-terminal region, major ampullate spidroin C-terminal region forms disulfide-connected dimers essential for protein-protein interaction and fiber formation regardless of pH (Ittah et al., 2007; Sponner et al., 2004).

To date, silk characteristics from aquatic spiders are entirely unknown and investigating silks of spiders from diverse habitats raises questions regarding silk-related specializations related to their ecology. In this study, we studied species in the superfamily Dictynoidea that use silk in three different environments (Figure 4.1; Forster, 1970). Within Dictynoidea, *Desis marina* and *Badumna longinqua*, members of the family Desidae, inhabit contrasting environments (Coddington and Levi, 1991; Spagna et al., 2010; Wheeler et al., 2017). The medium-sized, ecribellate spider *D. marina*, lives within cavities in kelp holdfasts or holes in rocks along the shorelines of New Zealand (Figure 4.1A; McQueen and McLay, 1983; McQueen et al., 1983; Vink et al., 2017). *D. marina* spiders are not known to make prey-catching webs; instead, they catch intertidal amphipods by ambush (Vink et al., 2017). Yet, silk is essential to these spiders for the construction of silk-lined retreats within kelp or rock cavities that protect them from tides and water pressure (Mclay and Hayward, 1987). *D. marina* spiders can remain submerged for up to 19 days by trapping air in their silken retreat, coupled with their lower respiration rate (McQueen and McLay, 1983).

To contrast with *D. marina*, we selected the confamilial *B. longinqua*, a medium-sized cribellate spider that is commonly found around buildings and small bushes in coastal urban and suburban areas (Main, 2001; World Spider Catalog, 2018). Unlike *D.*

marina, *B. longinqua* spiders build a lattice-like sheet web using cribellar silk that extends from tubular retreats in crevices such as dense foliage and small openings (Figure 4.1C; Adams and Manolis, 2014; Main, 2001). Cribellar silk is a dry glue-like matrix composed of small fibers that emerge from the cribellum, a plate-like spinning organ, that are combed by the spider as they are drawn out from the cribellar spigots (Eberhard and Pereira, 1993; Peters, 1984, 1987, 1992). *D. marina* spiders lack cribellums, due to secondary loss (the "ecribellate" condition). Because *B. longinqua* and *D. marina* are in the same family, yet one is terrestrial and the other is water-associated, they serve as a good model to investigate the role of spider silk in relation to their environment.

Our third focal species is *Argyroneta aquatica* (Dictynidae), another species in the Dictynoidea. While *D. marina* is found in marine habitats, *A. aquatic* is found in freshwater ponds and lakes in Northern Europe, spending all of its life underwater (McQueen et al., 1983). An *A. aquatica* spider builds a special underwater domed-shaped sheet-web, called a diving-bell, and use this unique web as an oxygen reservoir (Figure 4.1C; De Bakker et al., 2006; Seymour and Hetz, 2011). The spider transports oxygen from the water's surface down to the diving-bell using an air bubble kept in place by abdominal hydrophobic hairs, which are referred to as a plastron (Flynn and Bush, 2008; Marx and Messner, 2012). While *A. aquatica* underwater web can function as a physical gill, as it can exchange dissolved oxygen from water, periodic air renewal is needed at long intervals of time to avoid collapsing of the web (Pedersen and Colmer, 2012; Seymour and Hetz, 2011). In this study, we examine how spidroins, repeat sequence

composition, and silk gene expression compare among freshwater, marine, and non-aquatic spiders.

Spidroins are secreted by abdominal silk glands, which are connected to spigots located on spinnerets from where silk fibers are drawn. Silk spigots can be morphologically differentiated based on size and location. For example, major ampullate fibers are produced in major ampullate glands and major ampullate glands connect to major ampullate silk spigots, which can be distinguished on the basis of shape, size, and location from other silk spigots (Coddington, 1989; Griswold et al., 2005). Therefore the complement of silk spigots on a spider's spinnerets can be an indicator of the types of silks a spider can produce. Morphological studies of *A. aquatica, B. longinqua,* and *D. marina* have identified silk spigots presumed to be connected to aciniform, major ampullate, minor ampullate, tubuliform, and pyriform silk glands (Griswold et al., 2005; Wasowska, 1977).

Moreover, *B. longinqua* has an additional set of spigot triads that are connected to uncharacterized silk glands. Each triad is composed of a large spigot called the "modified spigot" and two smaller, paracribellar spigots (Griswold et al., 2005). *B. longinqua* also has spigots called the cribellar spigots on their cribellum that produces cribellar silk. *D. marina* also has a pair of "modified spigots" connected to unidentified silk glands (Griswold et al., 2005). Given that spidroins are named after the silk gland type in which they were first identified, we anticipate identifying the following spidroin genes: AcSp (a contraction of aciniform spidroin), CrSp (cribellar spidroin), MaSp (major ampullate spidroin), MiSp (minor ampullate spidroin), PySp (pyriform spidroin), TuSp (tubuliform

spidroin). Additionally, we expect to identify spidroins genes, Sp (a contraction of spidroin), associated with uncharacterized silk glands connected to modified and paracribellar spigots. Since females and not males have tubuliform (egg case) spigots, therefore, we expect only females to express TuSp genes.

B. longinqua and *D. marina* are both representatives of the family Desidae and are therefore more closely related to each other than either is to *A. aquatica*, which belongs to the family Dictynidae. Because of this familiar relationship, we would expect *B. longinqua* and *D. marina* to have a more similar complement of spidroin genes. Alternatively, if habitat is a major selective force shaping spidroin evolution, spidroins from the aquatic spiders *A. aquatica* and *D. marina* could share more similarities. We also compare the *A. aquatica* and *D. marina* spidroins to those of another aquatic spider, *Dolomedes triton* (Pisauridae; Chapter 2 of this dissertation).

Material and Methods

RNA-seq Library Construction and Sequencing

Adult female *B. longinqua* were collected in Vista (San Diego County), California, USA. The complete set of silk glands were extracted from each individual, flash frozen in liquid nitrogen, and stored at -80°C. Adult female and male *D. marina* were collected in Kauri Point Reserve, New Zealand. Adult female and male *A. aquatica* were collected in Neerpelt, Belgium. The cephalothoraxes (without venom glands) and complete set of silk glands were dissected from each individual *D. marina* and *A. aquatica* and the tissues were immediately submerged in RNALater (Sigma-Aldrich,

Milwaukee, WI, USA). Venom glands were removed from each cephalothorax to obtain a single type of non-silk gland control tissue. Separate RNA extractions were done for silk glands from each individual following the methods of Starrett et al. (2012). Twelve RNA-Seq libraries were made using the Ovation Universal RNA-Seq System (NuGen, San Carlos, CA). Libraries made from *A. aquatica* and *D. marina* tissues included: two sets of female silk glands, two sets of male silk glands, and two cephalothoraxes, for a total of six libraries per species. Libraries made from *B. longinqua* included two sets of female silk gland tissues. Bidirectional sequencing of the libraries (2x150 bp, mid-output) was done on a NextSeq 500 System (Illumina) at the University of California, Riverside Genomics Core Facility.

Transcriptome Assembly and Estimates of Expression Level

Low-quality reads and adaptors were removed from raw sequencing reads from each FASTQ file using Trimmomatic (Bolger et al., 2014). The quality of the resulting filtered reads was evaluated using FastQC (Babraham Bioinformatics FastQC Package**)**. All reads from the same species were combined for *de novo* assembly of species-specific transcriptomes with Trinity v2.1.1 using default parameters (Grabherr et al., 2011). Assembly statistics, including N50 as an approximation of assembly quality, are shown in Table 4.1. Filtered reads from each species were mapped to their corresponding species-specific assemblies using TopHat2 v2.1.1 with default parameters (Kim et al., 2013). Reads Per Kilobase per Million mapped reads (RPKM) values were calculated for each

transcript. A minimum of ten reads mapped and more than one RPKM in at least two assemblies was used as a cut-off for inclusion in the gene expression analyses.

Spidroin Annotation and Phylogenetic Analyses

A silk protein database composed of spidroin genes, spider coating peptides (SCP-1 and SCP-2), egg case silk proteins (ECP-1 and ECP-2), and aggregate silk factors (AgSF1 and AgSF2) was constructed from downloaded NCBI nr proteins and UniProtKB/Swiss-Prot databases (March 2017). This database was used to identify silk genes via BLASTX searches (e-value < 1 e-5) in Geneious v8.1.8. Only spidroin transcripts were found, there were no hits to spider coating peptides, egg case silk proteins, or aggregate silk factors. Spidroin transcripts were visually inspected to confirm the presence of typical characteristics of spidroin genes, such as coding regions for conserved N- and C- terminal domains and repetitive regions (Figure 4.2A). Transcripts with >90% nucleotide identity were considered to represent the same locus and were combined into a contig using a majority rule approach.

N- and C-terminal encoding regions from spidroin contigs were translated and combined with published spidroin sequences from other araneomorph (true spider) species representing a broad diversity of the spidroin family (Table 4.3). N- and C-terminal regions were aligned separately with ClustalW implemented in Geneious and refined by eye. JTT and WAG amino acid model tests were used for N- and C-terminal alignments respectively. Maximum likelihood gene trees were constructed with 10,000

bootstrap replicates using RAxMLv8.2.8 (Stamatakis, 2014). Resulting trees were visualized with FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/).

Proteomic Analysis and SEM of Argyroneta aquatica Diving Bells

A. aquatica were individually housed in aquariums. Fresh diving bells were harvested for proteomic analyses and scanning electron microscopy (SEM). Webs were taken out of the water and immediately processed. Protein extractions were done following Chaw et al. (2015). Peptide samples were subjected to LC-MS/MS analysis on an LTQ Orbitrap Velos mass spectrometer (Thermo Fisher Scientific) equipped with a nanomate ESI source (Advion, Ithaca, NY, USA) at the University of Arizona's Arizona Proteomics Consortium. Resulting tandem mass spectra were searched against the non-redundant longest open reading frame translation of our *A. aquatica* transcriptome, Chelicerata proteins downloaded from NCBI (on October 17, 2013), and common contaminant proteins using Thermo Proteome Discoverer 1.3 (Thermo Fisher Scientific). Protein and peptide identification results were visualized with Scaffold v4.7.3 (Proteome Software Inc., Portland, OR, USA). Proteins with at least one peptide identified at 95% protein confidence and 95% peptide confidence by peptide and protein profile were accepted.

From freshly harvested *A. aquatica* webs, 8x5x3 mm sections were cut with clean micro scissors, mounted to aluminum pin stubs with carbon tape, and coated with a thin layer of platinum-palladium for scanning electron microscopy (SEM). Electron

micrographs were collected with a 5-10 kV accelerating voltage using a Mira3 (Tescan, Czech Republic).

Results and Discussion

AcSp, TuSp, and PySp Sequences are Conserved in the Focal Species

We identified nine, 26, and 12 partial length spidroin contigs from the assemblies of *A. aquatica, B. longinqua*, and *D. marina*, respectively (Table 4.2). All 47 of these spidroin contigs were partial length transcripts containing the coding sequence for either the N- or C-terminal region. Most (44 of 47) transcripts included adjacent repetitive sequence. There were spidroins with significant sequence similarity to the repeat units of aciniform spidroin (AcSp), tubuliform spidroin (TuSp), and pyriform spidroin (PySp), which are associated with aciniform, tubuliform, and pyriform silk glands, respectively (Figure 4.2). Phylogenetic analyses of the N- and C-terminal encoding regions recover an AcSp clade, a TuSp clade, and a PySp clade (Figures 4.3 and 4.4). These clades are groupings of AcSp, TuSp, and PySp sequences of the focal species with the same spidroin types from the comparison species included in the analyses.

Aciniform spidroins have mostly been described from orb-web and cob-web weaving spiders, and are considered the main component of aciniform silk (Ayoub et al., 2013b; Chaw et al., 2014; Hayashi et al., 2004; Vasanthavada et al., 2007). Aciniform silk is used for web construction, web decoration and wrapping of prey. Similar to aciniform spidroins from other species, AcSp sequences from our three focal species have long repeats (ranging from 173-206 aa; Figure 4.2B; Ayoub et al., 2013b; Chaw et al., 2014; Chapters 2 and 3 of this dissertation). *B. longinqua* and *A. aquatica* each have

at least two AcSp variants, which we named variants A and B (AcSp_vA and AcSp_vB). The *B. longinqua* AcSp variants have different repeat units, sharing only 41% sequence similarity at the amino acid level. This suggests that *B. longinqua* has at least two AcSp loci. For the *A. aquatica* AcSp variants, only one transcript contained a complete repeat unit (*A. aqu*_AcSp_N_vB; Figure 4.2B). However, comparison of the available *A. aquatica* repetitive regions shows few similarities with only 25% sequence identity (over 80 aa), also consistent with two AcSp loci in the *A. aquatica* genome.

Female spiders protect their eggs by wrapping them in silk. The silken egg cases are mainly composed of tubuliform silk (Casem et al., 2010; Hu et al., 2005a, 2006). Each of our focal species contained transcripts that were long enough to have more than one complete repeat unit of TuSp (Figure 4.2C). The TuSp repeat unit of *A. aquatica* is 190 aa, 194 aa in *B. longinqua*, and 196 aa in *D. marina* is 196 aa. Repeat units from all three species are comparable in length to the 180-184 aa TuSp1 repeat units from the cob-web weaving *L. hesperus* (Hu et al., 2005b) and the orb-web weaving spiders, *Argiope bruennichi* (Zhao et al., 2006) and *Argiope argentata* (Chaw et al., 2017b).

Not only are the TuSp repeat units conserved in length across species, but they have conserved amino acid sequence motifs. The amino acid sequence motifs poly-serine, poly-alanine, and glycine-X (where X could be A, I, L, Q, S, T, V, or Y), are common in our focal species. Unlike the single, long poly-alanine motifs found in each major ampullate spidroin 1 repeat unit (MaSp1; Ayoub et al., 2007; Hinman and Lewis, 1992), poly-alanine motifs in TuSp are no more than four residues long and are dispersed across the sequence (Figure 4.2C). *A. aquatica* TuSp repeat also shares a five threonine

(boxed, Figure 4.2C) stretch with the orb-web weaving spiders *Argiope aurantia*, *Nephila clavipes*, and *Araneus gemmoides* (Tian and Lewis, 2005), the cob-web weaving spider *L. hesperus* (Hu et al., 2005b), and the velvet spider *Stegodyphus mimosarum*. As with previous hypotheses of spidroin evolution (e.g. Ayoub et al., 2013b; Chaw et al., 2014; Clarke et al., 2014; Prosdocimi et al., 2011; Starrett et al., 2012), we recovered a sister relationship of TuSp and AcSp clades in our maximum likelihood analysis using the C-terminal encoding region (Figure 4.4; 51% bootstrap support). This result contrasted with the more weakly supported relationship of tubuliform spidroins nested within aciniform sequences when using the N-terminal domain (Figure 4.3; 17% bootstrap support).

Pyriform silk is a composite of PySp-based fibers and a cement coating used to adhere a variety of silk fibers to a substrate or to each other (Kovoor and Zylberberg, 1980, 1982; Wolff et al., 2015). Pyriform spidroin, PySp, is the main protein found in pyriform silk (Blasingame et al., 2009; Chaw et al., 2017a; Geurts et al., 2010; Perry et al., 2010). Similar to AcSp and TuSp, PySp spidroins also show conservation of repeat length and composition. Although a full repeat unit was not obtained for *D. marina*, our *A. aquatica* and *B. longinqua* contigs included complete PySp repeat units (Figure 4.2D). PySp repeat units of *A. aquatica* (207 aa) and *B. longinqua* (198 aa) are very similar to each other, sharing 80% amino acid identity. As with PySp from orb-web and cob-web weaving spiders, PySp repeats from *A. aquatica* and *B. longinqua* are rich in serine and glutamine. They also share a proline-rich motif (boxed, Figure 4.2D) with PySp1 repeats from orb-web weaving spiders, the house spider *Parasteatoda tepidariorum* PySp2*,* the fishing spider *D. triton* PySp, and PySp from the cribellate spiders *S. mimosarum* and

Tengella perfuga (Chaw et al., 2017a; Geurts et al., 2010; Perry et al., 2010). The *B. longinqua* proline-rich motif has the same length as that of the orb-web weaver *A. argentata* (36 aa), and in orb-web weaving spiders, the proline-rich motif is thought to produce a random coil configuration that promotes elastomeric properties in pyriform silk (Perry et al., 2010). In contrast to these two terrestrial spiders, the fully aquatic *A. aquatica* has a proline-rich motif that is the same length as that of the semi-aquatic spider *D. triton* PySp (28 aa). Overall similarities in repeat unit composition and length of PySp-specific motifs suggests similar selective pressures acting on PySp from terrestrial and aquatic spiders. Chemical composition and nano-structure studies of pyriform silk are needed to understand the mechanism of anchoring silk fibers to wet surfaces, and whether there is a functional significance to the water-associated PySps having a shorter proline-rich motif.

While our sequences are partial length transcripts, the portion of the repetitive region immediately adjacent to the N- and C- terminal region of each *B. longinqua* AcSp variant, as well as AcSp, TuSp, and PySp from *D. marina,* were found to be nearly identical suggesting that each pairing represents two ends of the same locus (Figure 4.5). It is also possible that the similarity between the N- and C-terminal region transcripts may represent the ends of different gene copies with similar functions. Future work using long-read sequencing technologies could definitively associate the partial length contigs by characterizing complete spidroin mRNAs or genes.

B. longinqua has Putative Cribellar and Modified Spigot-Silk Specific Spidroins

Cribellate spiders use cribellar silk in their prey-capture web as a dry adhesive to secure freshly caught insects to their webs. Recent studies of silks from the cribellate spider *T. perfuga* identified a putative spidroin associated with cribellar silk, called CrSp (Cribellar Spidroin; Chapter 3). In maximum likelihood analysis of the C- terminal encoding regions, we found a *B. longinqua* transcript that grouped with *T. perfuga* CrSp and a *S. mimosarum* spidroin with moderate support (Figure 4.4; 67%). Comparison of the repetitive regions of the *B. longinqua* transcript to those of *T. perfuga* and *S. mimosarum* shows a well-conserved repeat unit (Figure 4.6; 51% aa identity). For these reasons, we annotated this *B. longinqua* transcript as Cribellar Spidroin (*B. lon*_C_CrSp; Table 4.2).

Unlike *B. longinqua*, *A. aquatica*, and *D. marina* do not have cribellums, do not produce cribellar silk, and CrSp was not found their transcriptome assemblies. This is further evidence associating *B. longinqua* CrSp specifically with cribellar silk production. Ancestral character state reconstruction studies of RTA spiders suggest that Dictynidae (the family of *A. aquatica*) and Desidae (the family of *D. marina*) are primitively cribellate (Alfaro et al., 2018a; Miller et al., 2010). This means that the absence of CrSp in *D. marina* and *A. aquatica* is due to gene loss because they are descended from a cribellate ancestor. Whether there are lingering CrSp pseudogenes or if CrSp loci are completely missing from *D. marina* and *A. aquatica* genomes warrant further investigation.

Another transcript containing the N- terminal region also showed affinities to spidroins from cribellate spiders. The transcript, *B. lon*_Sp_N_vA, grouped with a spidroin (*T .per*_Sp_N) from the cribellate spider *T. perfuga* (93% bootstrap support; Figure 4.3). The repetitive sequence from *B. longinqua* has high amounts of glycine mainly organized in couplets (GX, with X representing a subset of amino acids), similar to the repetitive sequence of *T .perfuga*. It has been hypothesized that *T. per*_Sp_N could be associated with the modified glands (Chapter 3 of this dissertation), based on homology of *T. perfuga* modified spigots to the modified/pseudoflagelliform silk spigots in other cribellate species (Alfaro et al., 2018a). The presence of this modified spigot has also been documented for *B. longinqua* (Griswold et al., 2005), suggesting that *B. lon*_Sp_N_vA and *T. per*_Sp_N represent the same spidroin type.

Ampullate Spidroin Repeat Sequences are Diverse in the Focal Species

Based on the presence of major and minor ampullate spigots on *A. aquatica, B. longinqua* and *D. marina* spinnerets (Griswold et al., 2005; Wasowska, 1977), we expected these species to express silk genes associated with major and minor ampullate silk glands. Major and minor ampullate silk glands produce major and minor ampullate silks, which are used by many species as draglines and for the construction of prey capture webs. We identified a total of 19 transcripts with similarities to previously described major and minor ampullate spidroins. Maximum likelihood analyses of the N- and C- terminal encoding regions show that these 19 sequences group with various MaSp (e.g., MaSp1, MaSp2, MaSp3) and MiSp sequences from other species, in a diverse

spidroin clade that we named Ampullate (Figures 4.3 and 4.4). Because we could not definitively categorize these transcripts as either MaSp or MiSp, we annotated them with the neutral name "AmSp", followed by an "N" or "C" for whether the transcript contains an amino or carboxy-terminal region. Additionally, we recovered more than one AmSp for two of the species and so we distinguish the different paralogs by a variant (v) letter (e.g., *D. mar*_AmSp_N_vA, *B. lon*_AmSp_C_vB; Table 4.2).

We found AmSp diversity to differ across species. Specifically, we found *B. longinqua* to have at least six AmSp variants in contrast to *A. aquatica* and *D. marina*, which appear to have only one or two variants, respectively (Figure 4.7). Regardless of species, all the AmSp repeat regions have a predominance of glycine, alanine, and serine residues (Figure 4.7). These amino acids appear in a variety of short sequence motifs. For example, *B. longinqua* AmSps have poly-glycine, poly-alanine, and poly-serine motifs (e.g., GGGG, AAAAA, SSS, respectively). Additionally, *B. longinqua* AmSps differ in their proportion of glycine-alanine motifs (40% of *B. lon*_AmSp_N_vA vs. 24% of *B. lon*_AmSp_C_vC; Figure 4.7). While the AmSp sequences vary extensively within and across species, we found one motif, GGYGQ, to be common in the repeats of *A. aquatica* and *B. longinqua* (shaded, Figure 4.7).

Terrestrial spider webs are susceptible to humidity; under these conditions, the major ampullate silk of some species have been observed to supercontract (Agnarsson et al., 2009; Boutry and Blackledge, 2010, 2013; Work, 1981). Supercontraction, the sudden reduction in silk fiber length when wetted with water, is thought to be due to the disruption of hydrogen bonds between proteins that allows for re-orientation and coiling

of silk molecules (van Beek et al., 1999; Savage and Gosline, 2008; Savage et al., 2004). Major ampullate sequence elements from orb-web and cob-web weaving spiders such as GPGXX (X is one of a small subset of amino acids) and YGGLGS(N)QGAGR amino acid motifs are thought to be associated with supercontraction (Boutry and Blackledge, 2010; Yang et al., 2000). Because silks from *A. aquatica* and *D. marina* are in constant contact with water, we would expect ampullate sequences to have a shortage of these supercontraction elements as supercontraction would not be beneficial in aquatic habitats. Consistent with this prediction, AmSp sequences from *A. aquatica* and *D. marina* spiders were found to lack the amino acid motifs associated with supercontraction.

Evidence for Additional Spidroin Types in the Focal Species

Five transcripts containing spidroin N-terminal regions do not show similarities to known spidroin types and are thus named "Sp" for "Spidroin" with no indicator of type, such as the Am, Cr, Py, or Tu of AmSp, CrSp, PySp, and TuSp. One of the Sp transcripts (*B. lon*_Sp_N_vB) groups with a *S. mimosarum* spidroin (*S. mim*_Sp2b) and their clade is positioned outside a diverse clade that includes pyriform, flagelliform, and cribellar spidroins (Figure 4.3). Although the repetitive region of *B. lon*_Sp_N_vB is very different from other spidroins it shares 53% amino acid similarity with the repeat region of *S. mim*_Sp2b suggesting that they probably represent the same spidroin type (Figure 4.8).

Another novel spidroin, *B. lon*_Sp_N_vC, is sister to a clade containing CrSp and Flag spidroins (Figure 4.3). The repetitive region of this spidroin has high amounts of the

amino acids glycine (36%) and serine (21%) that are largely present in glycine-serine couplets. However, glycine-serine is not present in CrSp or Flag spidroins (Figure 4.6; Hayashi and Lewis, 1998). Given its phylogenetic placement and distinctive repeat region characteristics, we were unable to assign *B. lon*_Sp_N_vC to any known spidroin category.

The last three novel spidroin transcripts, one from each focal species, have similar N-terminal regions to each other but different repeat units. *A. aqu*_Sp_N, *B. lon*_Sp_N_vD, and *D. mar*_N_Sp form a well-supported clade with other spidroins from *D. triton* and *S. mimosarum* (*D. tri*_Sp_N and *S .mim*_Sp2c; Figure 4.3; 85%). Given that members of this clade have an unclear relationship to known spidroin types, we refer to this clade as "Other Spidroins". Repeat units of spidroins within this clade do not have conserved sequences (Figure 9). The divergent repetitive regions could be due to different selection pressures compared to the stabilizing selection on the terminal domains. Repetitive regions from aquatic and semiaquatic spiders share some similarities. *A. aqu*_Sp_N, *D. mar*_Sp_N, and *D. tri*_Sp_N have repeated couplets of the hydrophobic amino acids glycine and valine (Figure 4.9). Our study of spidroins from semi-aquatic spiders did not find specific modifications associated with wet environments (Chapter 2 of this dissertation). But now with a comparison to spidroins of two aquatic spiders, it can be proposed that the higher concentration of hydrophobic amino acid motifs in Sp sequences in *A. aquatica*, *D. marina*, and *D. triton* (20-38% GV) compared to the terrestrial *B. longinqua* and *S. mimosarum* (2-4% GV) is associated with specializations for use in water.

Spidroin Gene Expression Levels in A. aquatica, B. longinqua, and D. marina

Relative levels of spidroin gene expression were quantified by mapping the reads from each silk gland RNAseq library to its respective transcriptome assembly. For all three species, there were silk gland libraries from females. For *A. aquatica* and *D. marina*, there were also silk gland libraries from males. We report the relative expression of spidroin transcripts containing the N-terminal region for each species. Similar expression patterns were observed when using transcripts containing the C-terminal encoding region. We did not combine the counts from the two regions because we could not definitively pair all the N-terminal region transcripts with the C-terminal transcripts. The main difference between the N- and C-terminal region counts is the expression of tubuliform spidroins in *A. aquatica* and *B. lonquingua*. For both species, tubuliform spidroins had the highest expression of all the spidroins according to the C-terminal region but not with the N-terminal region. The highly repetitive region of spidroins can lead to inaccurate estimation of gene expression when using partial-length transcripts, as are all our assembled spidroin transcripts (Chaw et al., 2016). Our TuSp transcripts with C-terminal regions include more repetitive sequence than the TuSp transcripts with N-terminal regions. Thus, the discrepancy in expression profiles could be due to the overestimation of expression levels using the TuSp transcripts with C-terminal regions given that they have a more complete repetitive region than N-terminal region transcripts.

AcSp genes have the highest relative expression of all spidroins in *A. aquatica* females and males. Specifically, one variant (*A. aqu*_AcSp_N_vA) accounted for 70% of total spidroin expression in females and 49% of males. The second highest expressed

spidroins are the AmSp group. Although we found both male and female *A. aquatica* spiders to have AmSp as the second highest expressed spidroin type, ampullate expression is higher in males (1,866 RPKM) than in conspecific females (364 RPKM). Observations of web construction by male and female *A. aquatica* spiders have documented that males use more walking threads during web construction when compared to females (De Bakker et al., 2006). This observation is consistent with our expression data, assuming that AmSp is a major constituent of walking threads (Figure 4.10).

While ampullate spidroin expression was second highest in *A. aquatica*, ampullate spidroin expression was by far the highest in *B. longinqua* (76%; Figure 4.10). The majority of this expression was driven by one AmSp variant, B. *lon_AmSp*_N_vE (37% of total spidroin expression). *B. longinqua* is a cribellate spider and prey-catching webs of other cribellate spiders have been observed to be composed of at least two different silk fibers, one fiber corresponding to ampullate silk and the other to cribellar silk (Alfaro et al., 2018b). This suggests that *B. lon_AmSp*_N_vE is the main protein being produced by *B. longinqua* spiders for dragline silk and their prey-catching webs.

Females and not males have tubuliform silk glands; thus, we expected females and not males to express tubuliform spidroins. Following our expectations, we found tubuliform spidroin genes in *D. marina* females and not males, to have the highest expression relative to other spidroin genes, while for males, ampullate spidroins were the most expressed (Figure 4.10). After spiders become sexually mature, female investment shifts from feeding to the production of egg cases, which are mainly composed of

tubuliform silk. Instead, mature males are thought to adapt a roving lifestyle in search of receptive females. The expression profiles of *D. marina* males and females are consistent with expression patterns described for cob-web weaving spiders (Correa-Garhwal et al., 2017). Surprisingly, we found that in *D. marina*, both sexes have a relatively high expression of the spidroin *D. mar*_Sp_N suggesting that this GV-rich spidroin may play an important role in silk-related functions associated to a semi-aquatic environment.

Composition of Retreats and Webs for Aquatic Environments

D. marina spiders live in the intertidal zone, where they construct silken retreats in which they store a bubble of air in order to withstand submersion by seawater for long periods of time (Figure 4.1A; McQueen and McLay, 1983). Both male and female *D. marina* spiders make these silk retreats. Consequently, we would expect silk genes that are expressed in both sexes at relatively equal proportions to be part of the retreat. It has previously been reported that *D. marina* silk retreats are made of fibers with two different diameters: a main fiber with a diameter of 1.54 μm and a second fiber with a diameter of 0.38 μm (Mclay and Hayward, 1987). Fiber diameter tends to correspond with spigot size; for example, major ampullate spigots are larger in size than aciniform spigots and major ampullate fibers are larger in diameter than aciniform fibers. Given that we found ampullate spidroins to be relatively highly expressed in female and male *D. marina* spiders, it is likely that *D. marina* spiders use ampullate silk as the main fiber and aciniform as the secondary fiber in their silken repeats. It is also possible that the primary fiber is composed of the "Sp" spidroin. If so, this implies that the Sp-producing silk gland

is connected to the modified spigot on the posterior lateral spinneret, which is similar in size to ampullate spigots (Griswold et al., 2005).

The underwater web of *A. aquatica* not only serves as shelter but is also an air reservoir where spiders spend most of their time (Figure 4.1D). Scanning electron microscope (SEM) micrographs of *A. aquatica* diving bell revealed a mixture of threads as well as a substantial amount of coating (Figure 4.11). Previous SEM studies of *A. aquatica* diving bells have shown different types of threads of variable thickness, flat bundles 200-400 μm across and a proteinaceous gel-like mass or hydrogel embedding all threads (Neumann and Kureck, 2013).

An approach to test the components of *A. aquatica* diving bell involved covering the opening of pyriform and major ampullate spigots on the anterior spinnerets with resin (Wasowska, 1977). It was shown that treated spiders were able to construct a normal, functional diving well but did not make the threads attaching the web to the surrounding vegetation. This suggests that the web is mainly produced by silk glands connected to spigots on the median and posterior spinnerets (minor ampullate and aciniform glands; Wasowska, 1977). Based on these observations, the different fiber morphologies observed in SEM micrographs (Figure 4.11), and the expression profile of *A. aquatica* with high expression of AcSp, AmSp, and Sp relative to all spidroins (Figure 4.10), we propose that the bell is composed of aciniform, pyriform, ampullate, and Sp silk. This hypothesis is partially supported by proteomic analysis of the diving bell. We found AcSp, PySp, and Sp spidroins to be part of *A. aquatica* diving bell (Table S4.1). This means that Sp is likely a component of minor ampullate silk and if so, is expressed in

minor ampullate glands. It also suggests that *A. aqu*_AmSp_N is likely a component of attaching threads and produced in major ampullate glands connected to major ampullate spigots on the anterior spinnerets (Wasowska, 1977). Ampullate silk has been observed to be used in web construction by *A. aquatica* spiders as the structural lines that attach the bell to aquatic plants (De Bakker et al., 2006). However, we were unable to collect all the structural lines of the diving bell and so it is possible that our collected sample did not include ampullate fibers, explaining why AmSp was not recovered in proteomic analyses.

The spidroins from aquatic and semi-aquatic spiders included in this study show little similarity in sequence to the silks of other arthropods that also live in aquatic habitats. This means that there are multiple solutions to evolving glues and silk fibers that function in wet environments. Silks from non-spider aquatic arthropods also have proteins with repetitive motifs that have been argued as playing an important role in survival and reproduction (Yang et al., 2014). For example, caddisfly larvae (Trichoptera) live in freshwater and spin protective cases and capture webs. The main structural protein in their silk is heavy chain fibroin, which has serine-rich motifs (SXSXSX) that interact with divalent ions (Yonemura et al., 2006). This interaction is thought to be essential for underwater adhesion. We found no evidence for similar serine-rich motifs in the spidroins of aquatic spiders (Figures 4.2, 4.7, and 4.9).

Similar to caddisfly larvae, amphipods and marine worms have molecular adaptations to marine environments. For example, the adhesive threads of the tube-building corophioid amphipod have a high proportion of basic residues, especially arginine (Kronenberger et al., 2012); the proteinaceous glue in the marine worm

(Polychaete) has proteins enriched with XGGYGYGGK repeat motifs and phosphorylated serine (Shao et al., 2009). We searched our spidroin sequences from aquatic and semi-aquatic spider species for features associated with marine adhesion but found no similarities to sequences from amphipods or marine worms. We did not find evidence of the convergent evolution of sequence elements used by other aquatic arthropods. Instead, we found that the spidroin sequences from water-associated spiders have high concentrations of the hydrophobic GV motif, which may therefore be a spider-specific strategy for using silk in aquatic environments.

Conclusion

We identified 47 spidroin transcripts in the assemblies from the focal species (Table 4.2). These transcripts correspond to aciniform, pyriform, tubuliform, and ampullate spidroins, plus a novel type of spidroin called "Sp". Aciniform, pyriform, and tubuliform spidroins from *A. aquatica*, *B. longinqua*, and *D. marina* are fairly conserved in repetitive and terminal regions across species and when compared to previously describes spidroins (Figures 4.2-4.4), suggesting similar selective pressures on spidroins independent of whether they are used in terrestrial only, semi-aquatic, or aquatic environments.

A spidroin recently identified to be associated with cribellar silk in the spider *Tengella perfuga* (CrSp; Chapter 3) was also discovered in the cribellate spider *B. longinqua.* The CrSp repeat units from both species are highly conserved in sequence (Figure 4.6). Another set of spidroins potentially associated with cribellate spiders was

also found. This set includes one Sp transcript from *B. longinqua* and one Sp transcript from *T. perfuga* that show similarities in the N-terminal and repetitive regions (*B. lon*_Sp_N_vA and *T. per*_Sp_N; Figure 4.3). We propose that these spidroins are produced in the modified spigot glands described by Alfaro et al. (2018).

The repetitive regions of our AmSp transcripts have similar amino acid composition to each other but show extensive variation in sequence (Figure 4.7). Although we found some amino acid motifs that are shared across species (e.g. poly-alanine, GGYGQ; Figure 4.7), we did not find any ampullate specific elements associated with supercontraction. Our finding suggests that ampullate silks from focal species do not undergo supercontraction as much as draglines from orb-web and cob-web weaving spiders.

Although our spidroins show diversity in repeat composition (Figure 4.2 and Figure S4.1), sequence comparison of aciniform, pyriform, tubuliform, and ampullate spidroins do not show unique modifications for semi-aquatic environments. However, we identified a highly hydrophobic amino acid motif, GV, that is shared among spider species associated with wet environments (Figure 4.9). This distinct amino acid motif is likely associated with the efficiency of underwater webs and retreats to withstand submersion, thus, increasing the fitness of aquatic and semi-aquatic spiders.

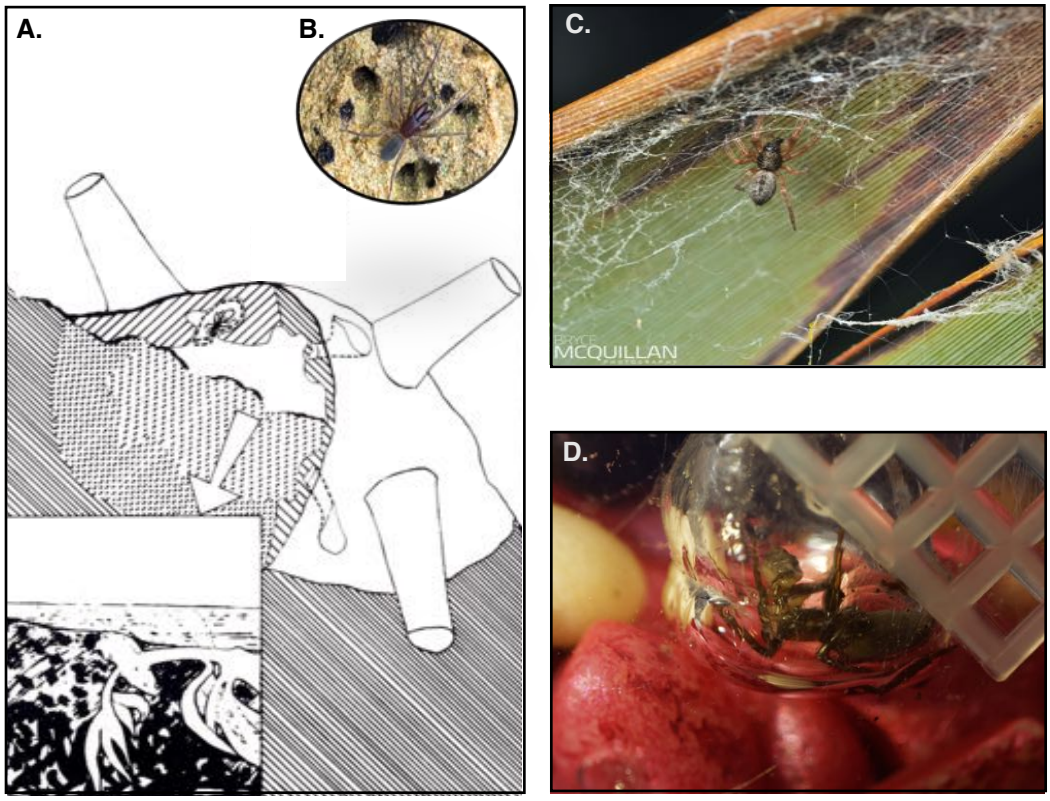


Figure 4.1. Focal spider species. Cross-section of holdfast positioned on a rock and location of the nest of the spider *Desis marina* (Desidae) (**A**); Image taken McQueen & McLay 1983. *Desis marina* on sandstone at katikati, New Zealand (**B**). Top view of *Badumna longinqua* (Desidae) sitting in its web (**C**). Side view of *Argyroneta aquatica* (Dictynidae) ventral side up inside its diving bell with a prey attached to its mouthparts (chelicerae) (**D**).

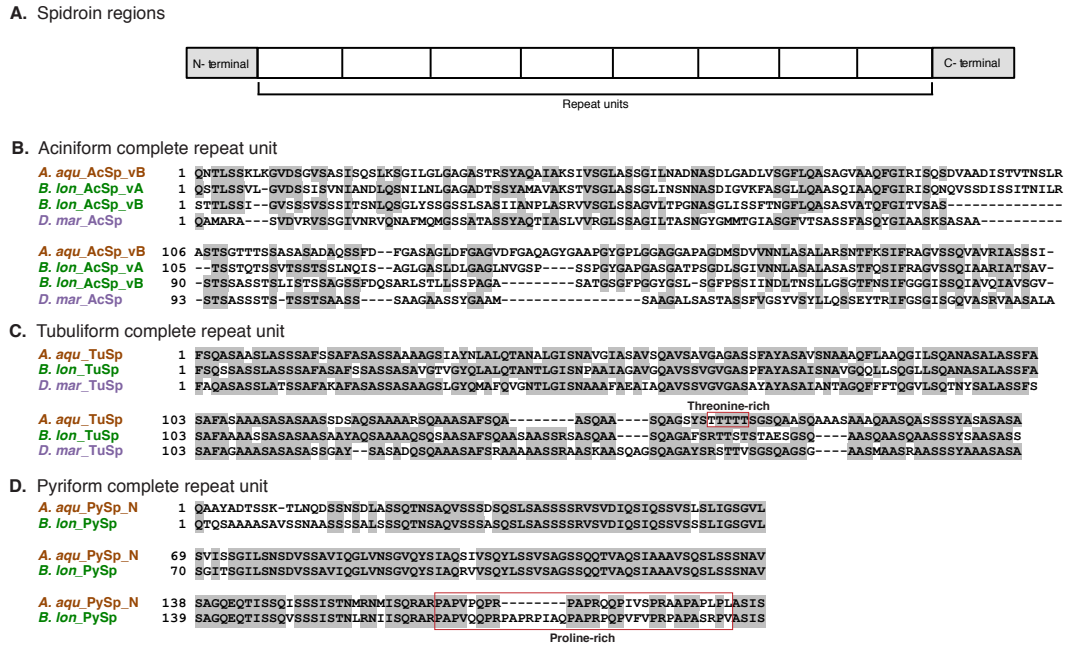


Figure 4.2. Exemplar repeat units for aciniform, tubuliform, and pyriform spidroins. (**A)** Schematic organization of a spidroin primary structure. Multiple alignments comparing spidroin paralogs encoding aciniform (AcSp; **B**), tubuliform (TuSp; **C**) and pyriform (PySp; **D**) spidroins of *Argyroneta aquatica* (brown), *Badumna longinqua* (green), *Desis marina* (purple). Names abbreviated as in Table 4.2. Amino acids conserved > 50% across all sequences are highlighted in grey. Gaps inserted into the alignment are indicated by dashes and missing sequence by periods. Amino acid positions for each sequence are numbered on the left. Proline and threonine rich regions shown in red boxes.

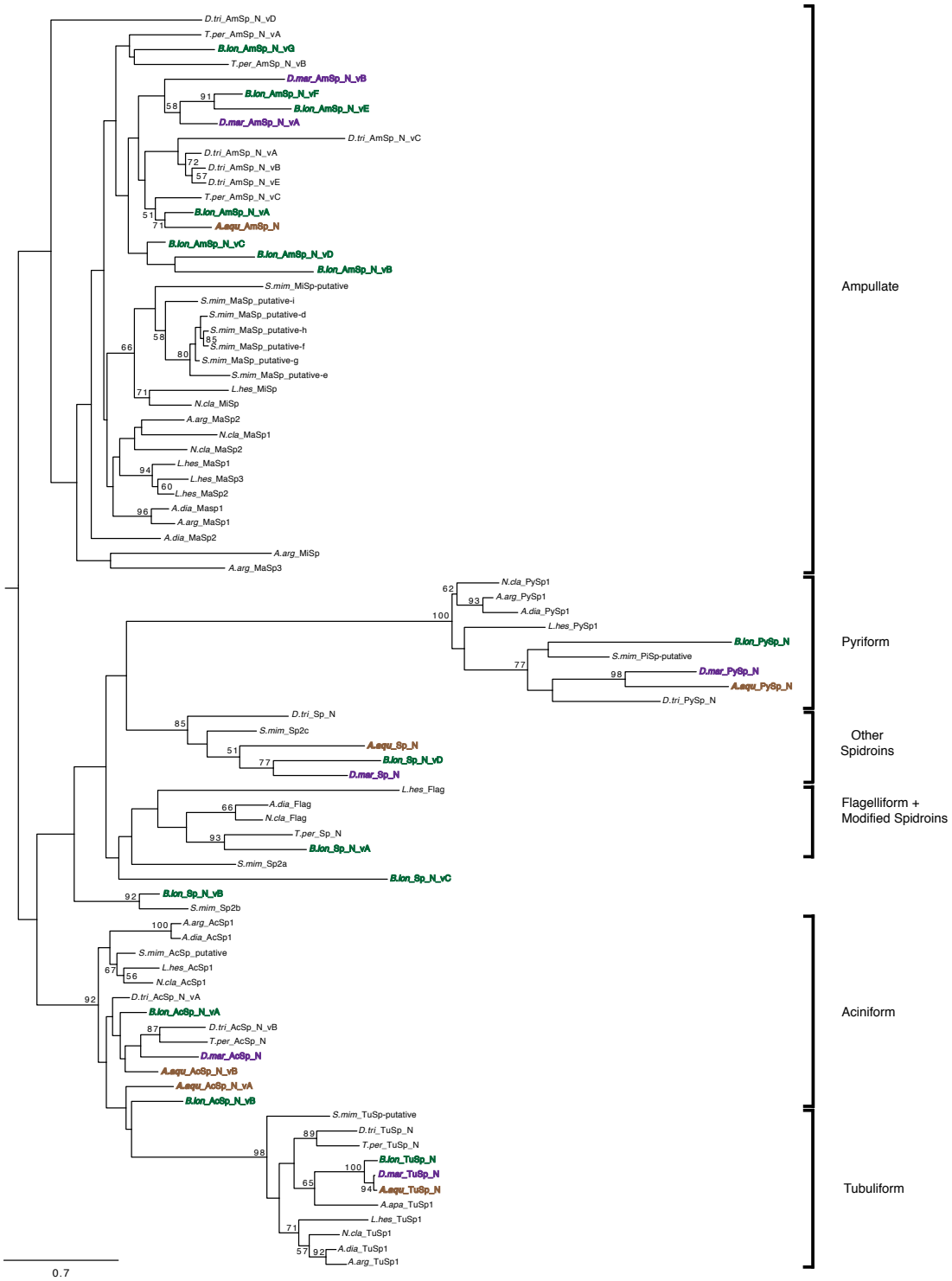

D.tri_AmSp_N_vD
T.per_AmSp_N_vA
B.lon_AmSp_N_vG
T.per_AmSp_N_vB
D.mar_AmSp_N_vB
B.lon_AmSp_N_vF
B.lon_AmSp_N_vE
D.mar_AmSp_N_vA
D.tri_AmSp_N_vC
D.tri_AmSp_N_vA
D.tri_AmSp_N_vB
D.tri_AmSp_N_vE
T.per_AmSp_N_vC
B.lon_AmSp_N_vA
A.aqu_AmSp_N
B.lon_AmSp_N_vC
B.lon_AmSp_N_vD
B.lon_AmSp_N_vB
S.mim_MiSp-putative
S.mim_MaSp_putative-i
S.mim_MaSp_putative-d
S.mim_MaSp_putative-h
S.mim_MaSp_putative-f
S.mim_MaSp_putative-g
S.mim_MaSp_putative-e
L.hes_MiSp
N.cla_MiSp
A.arg_MaSp2
N.cla_MaSp1
N.cla_MaSp2
L.hes_MaSp1
L.hes_MaSp3
L.hes_MaSp2
A.dia_Masp1
A.arg_MaSp1
A.dia_MaSp2
A.arg_MiSp
A.arg_MaSp3
N.cla_PySp1
A.arg_PySp1
A.dia_PySp1
L.hes_PySp1
B.lon_PySp_N
S.mim_PiSp-putative
D.mar_PySp_N
A.aqu_PySp_N
D.tri_PySp_N
D.tri_Sp_N
S.mim_Sp2c
A.aqu_Sp_N
B.lon_Sp_N_vD
D.mar_Sp_N
L.hes_Flag
A.dia_Flag
N.cla_Flag
T.per_Sp_N
B.lon_Sp_N_vA
S.mim_Sp2a
B.lon_Sp_N_vC
B.lon_Sp_N_vB
S.mim_Sp2b
A.arg_AcSp1
A.dia_AcSp1
S.mim_AcSp_putative
L.hes_AcSp1
N.cla_AcSp1
D.tri_AcSp_N_vA
B.lon_AcSp_N_vA
D.tri_AcSp_N_vB
T.per_AcSp_N
D.mar_AcSp_N
A.aqu_AcSp_N_vB
A.aqu_AcSp_N_vA
B.lon_AcSp_N_vB
S.mim_TuSp-putative
D.tri_TuSp_N
T.per_TuSp_N
B.lon_TuSp_N
D.mar_TuSp_N
A.aqu_TuSp_N
A.apa_TuSp1
L.hes_TuSp1
N.cla_TuSp1
A.dia_TuSp1
A.arg_TuSp1
0.7
Ampullate
Pyriform
Other Spidroins
Flagelliform + Modified Spidroins
Aciniform
Tubuliform

Figure 4.3. Maximum likelihood tree of spidroin N-terminal regions. *Argyroneta aquatica* (brown), *Badumna longinqua* (green), *Desis marina* (purple) spidroin paralogs highlighted in brown, green, and purple respectively. Tree rooted with the California trapdoor spider *Bothriocyrtum californicum* fibroin 1 (not shown). Names abbreviated as in Tables 4.2-4.3. Bootstrap percentages >50% are shown. . Scale bar represents substitutions per site.

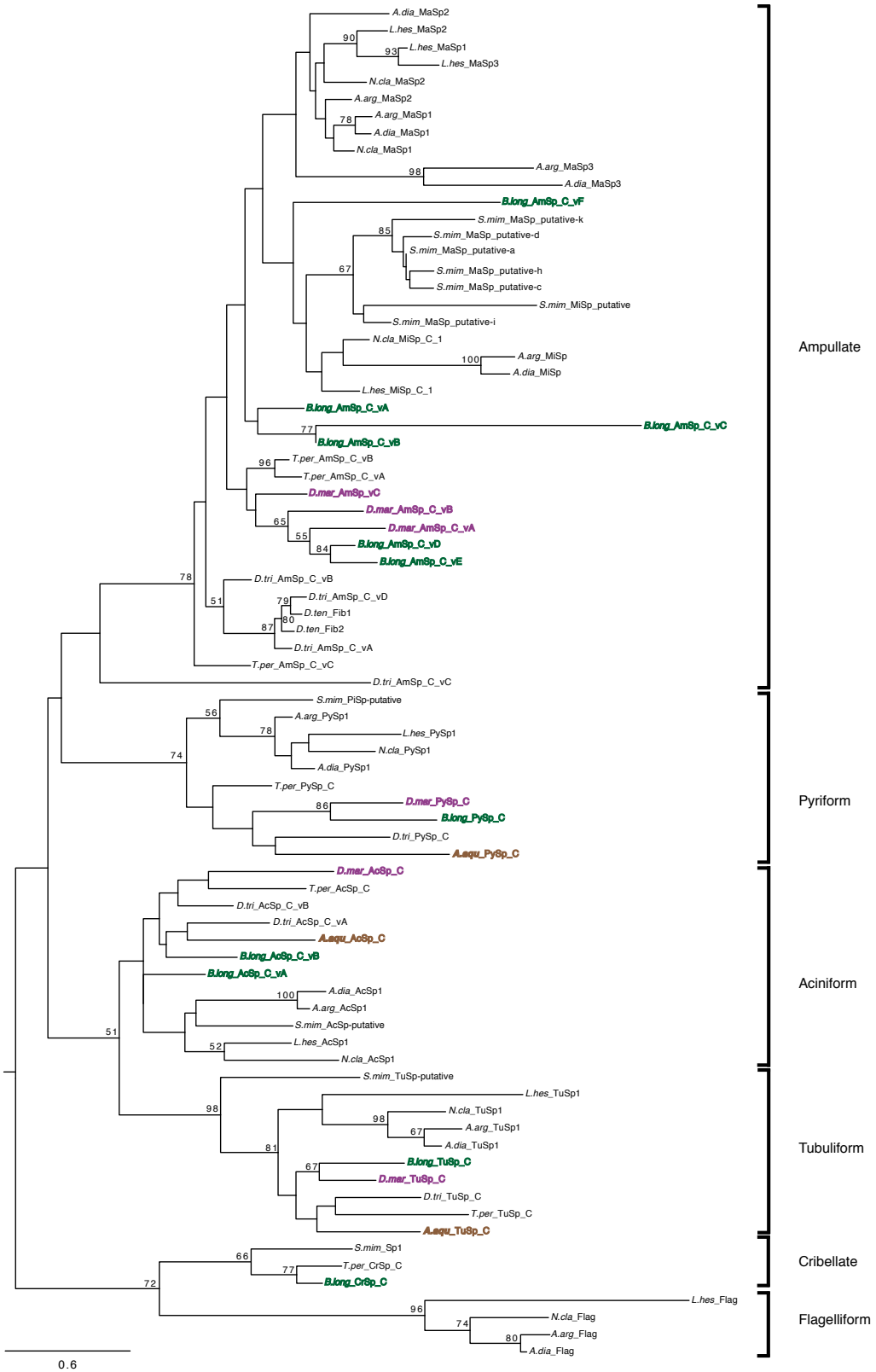

A.dia_MaSp2
L.hes_MaSp2
90
L.hes_MaSp1
93
L.hes_MaSp3
N.cla_MaSp2
A.arg_MaSp2
A.arg_MaSp1
78
A.dia_MaSp1
N.cla_MaSp1
A.arg_MaSp3
98
A.dia_MaSp3
B.long_AmSp_C_vF
S.mim_MaSp_putative-k
85
S.mim_MaSp_putative-d
S.mim_MaSp_putative-a
67
S.mim_MaSp_putative-h
S.mim_MaSp_putative-c
S.mim_MiSp_putative
S.mim_MaSp_putative-i
N.cla_MiSp_C_1
100
A.arg_MiSp
A.dia_MiSp
L.hes_MiSp_C_1
B.long_AmSp_C_vA
B.long_AmSp_C_vC
77
B.long_AmSp_C_vB
96
T.per_AmSp_C_vB
T.per_AmSp_C_vA
D.mar_AmSp_vC
D.mar_AmSp_C_vB
65
D.mar_AmSp_C_vA
55
84
B.long_AmSp_C_vD
B.long_AmSp_C_vE
78
D.tri_AmSp_C_vB
51
79
D.tri_AmSp_C_vD
80
D.ten_Fib1
87
D.ten_Fib2
D.tri_AmSp_C_vA
T.per_AmSp_C_vC
D.tri_AmSp_C_vC
S.mim_PiSp-putative
56
A.arg_PySp1
78
L.hes_PySp1
74
N.cla_PySp1
A.dia_PySp1
T.per_PySp_C
86
D.mar_PySp_C
B.long_PySp_C
D.tri_PySp_C
A.equ_PySp_C
D.mar_AcSp_C
T.per_AcSp_C
D.tri_AcSp_C_vB
D.tri_AcSp_C_vA
A.equ_AcSp_C
B.long_AcSp_C_vB
B.long_AcSp_C_vA
100
A.dia_AcSp1
A.arg_AcSp1
S.mim_AcSp-putative
51
52
L.hes_AcSp1
N.cla_AcSp1
S.mim_TuSp-putative
L.hes_TuSp1
98
98
N.cla_TuSp1
67
A.arg_TuSp1
81
A.dia_TuSp1
67
B.long_TuSp_C
D.mar_TuSp_C
D.tri_TuSp_C
T.per_TuSp_C
A.equ_TuSp_C
66
S.mim_Sp1
77
T.per_CrSp_C
72
B.long_CrSp_C
L.hes_Flag
96
N.cla_Flag
74
A.arg_Flag
80
A.dia_Flag
0.6
Ampullate
Pyriform
Aciniform
Tubuliform
Cribellate
Flagelliform

Figure 4.4. Maximum likelihood tree of spidroin C-terminal regions. *Argyroneta aquatica* (brown), *Badumna longinqua* (green), *Desis marina* (purple) spidroin paralogs highlighted in brown, green, and purple respectively. Tree rooted with the California trapdoor spider *Bothriocyrtum californicum* fibroin 1 (not shown). Names abbreviated as in Tables 4.2-4.3. Bootstrap percentages >50% are shown. . Scale bar represents substitutions per site.

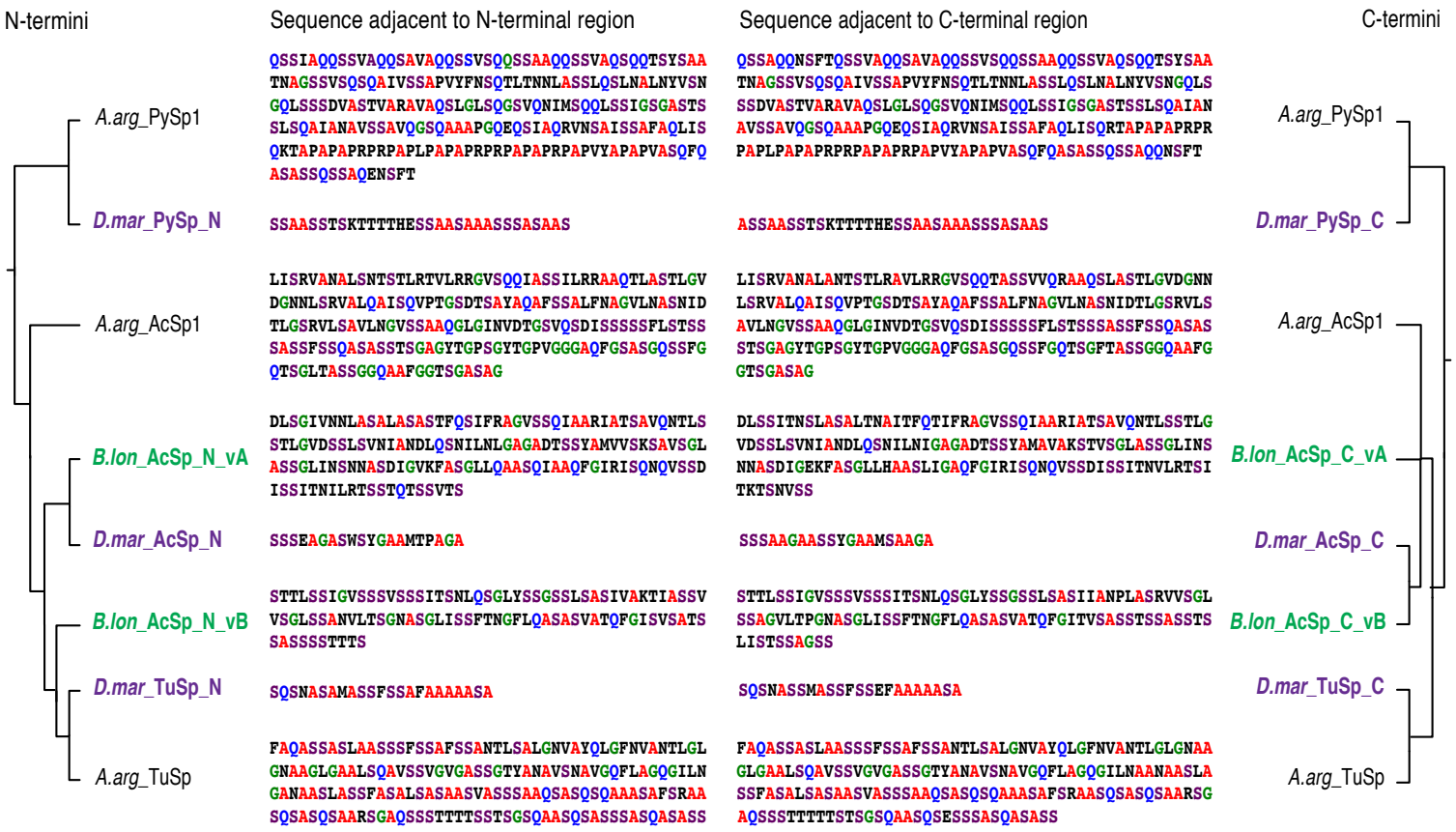


Figure 4.5. Comparison of repeats from transcript representing the same loci. *Argyroneta aquatica* (brown), *Badumna longinqua* (green), *Desis marina* (purple). Abundant amino acids highlighted: alanine (red), serine (purple), glycine (green), and glutamine (blue). Names abbreviated as in Tables 4.2-4.3.

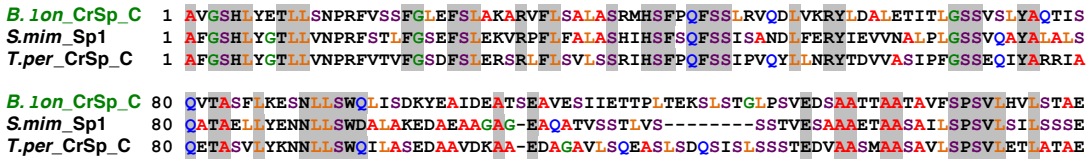


Figure 4.6. Cribellar spidroin (CrSp) repeat unit. Identified CrSp motif of *Badumna longinqua* (green) aligned to *Tengella perfuga* (*T.per*_CrSp), and *Stegodyphus mimosarum* (*S.mim*_Sp1; GenBank: KFM60634.1). Names abbreviated as in Tables 4.2-4.3. Amino acids conserved 100% across all sequences are shaded in grey. Abundant amino acids in silks are highlighted: alanine (red), serine (purple), glycine (green), leucine (orange), and glutamine (blue). Gaps inserted into the alignment are indicated by dashes. Amino acid positions for each sequence are numbered on the left.

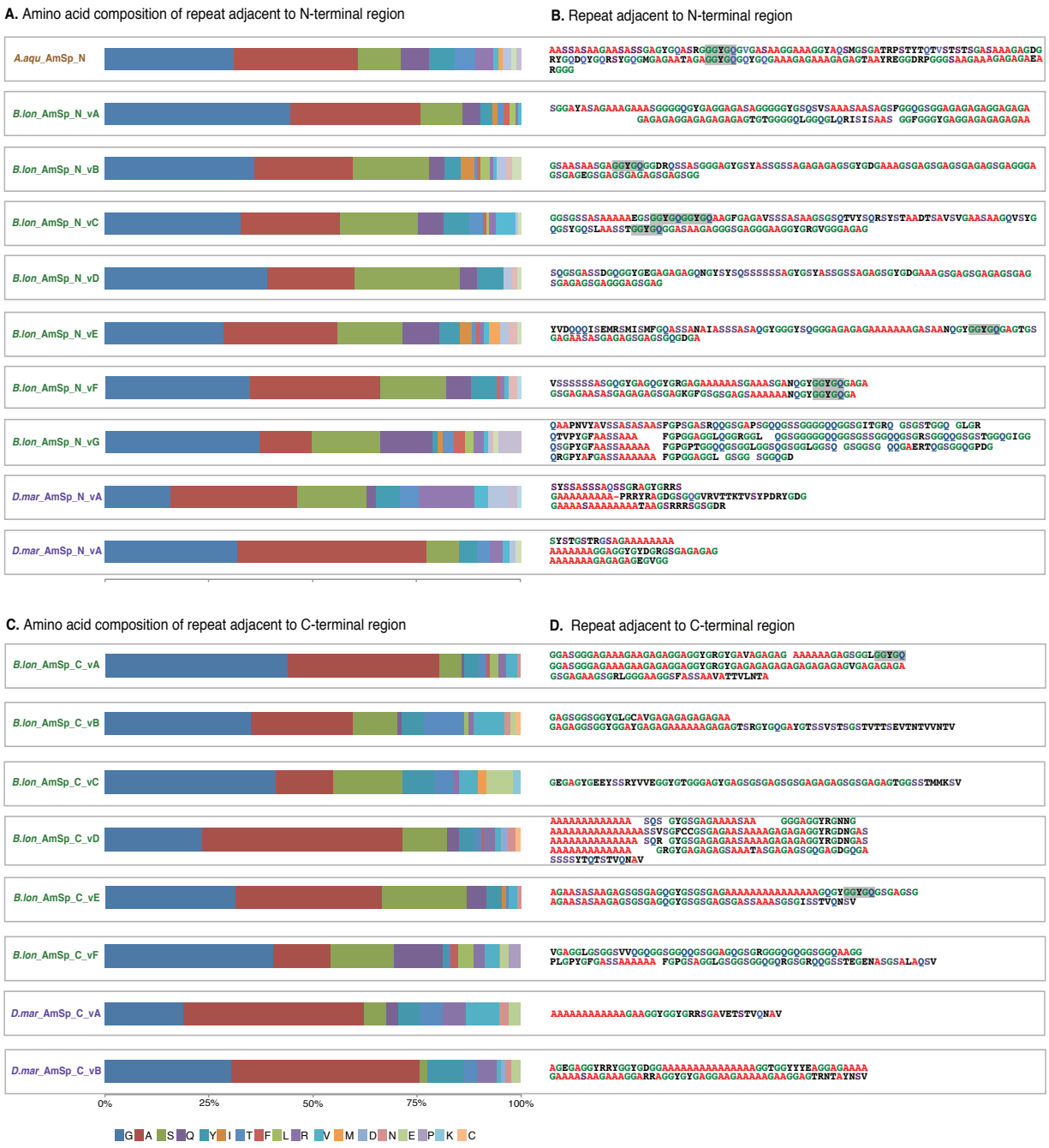


Figure 4.7. Repetitive regions of ampullate spidroins. Amino acid composition of AmSp repetitive regions (**A** and **C**) adjacent to the N-terminal (**B**) and C-terminal (**D**) regions of ampullate spidroins of *Argyroneta aquatica* (brown), *Badumna longinqua* (green), and *Desis marina* (purple). Names abbreviated as in Table 4.2. Amino acids abundant in ampullate silks are highlighted: alanine (red), serine (purple), glycine (green), and glutamine (blue). Amino acid motifs found in several sequences are highlighted in grey.

Figure 4.8. Pairwise alignment of *Badumna longinqua* Sp_N_vB (green) and *Stegodyphus mimosarum* (Sp2b, GenBank: KK117516.1) repetitive region. Names abbreviated as in Tables 4.2-4.3. Amino acids conserved 100% across all sequences are shaded in black. Gaps inserted into the alignment are indicated by dashes. Amino acid positions for each sequence are numbered on the left.

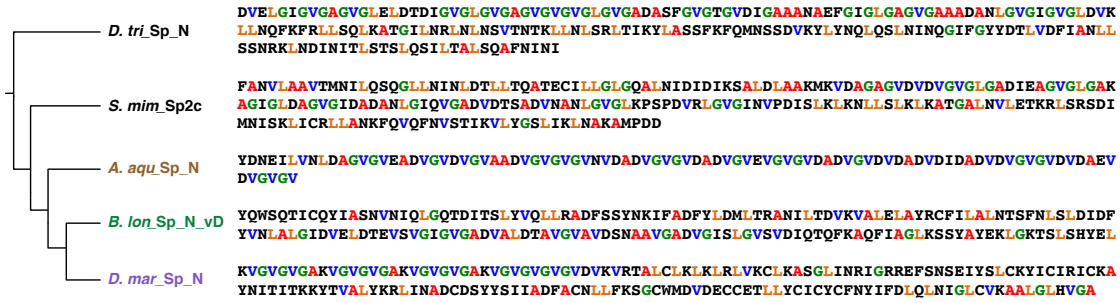


Figure 4.9. Comparison of exemplar repeats units of sequences from Sp clade. Sp spidroins for *Argyroneta aquatica* (brown), *Badumna longinqua* (green), *Desis marina* (purple) *Stegodyphus mimosarum* and *Dolomedes triton*. Names abbreviated as in Tables 4.2-4.3. Abundant amino acids colored in red (alanine), blue (valine), green (glycine) and orange (leucine).

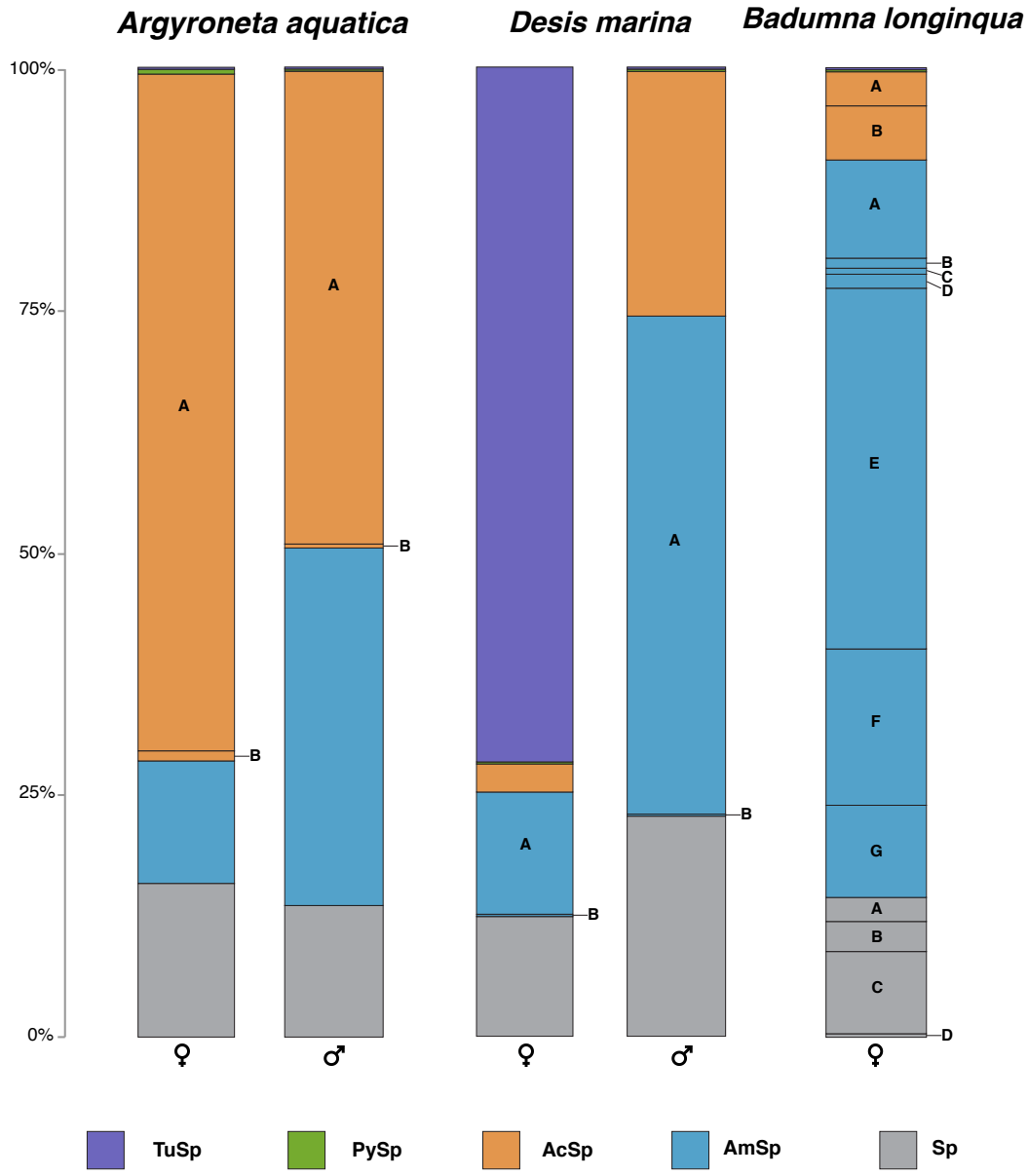


Figure 4.10. Relative expression levels of spidroin genes. Stacked bar graphs of gene expression levels in *Argyroneta aquatica* females (right) and males (right), *Desis marina* (right) and males (right), and *Badumna longinqua*. TuSp, PySp, AcSp, AmSp, and Sp are shown. Letters indicate different variants. Percentages of average expression of male and female reads mapped to species-specific transcriptomes. Percentages show reads per kilobase of transcript per million mapped reads (RPKM). Total RPKM of spidroins for *A. aquatica*: ♀ 2,883, ♂ 5,041; *D. marina*: ♀ 11,260, ♂ 8,155; and *B. longinqua* ♀ 43,284.

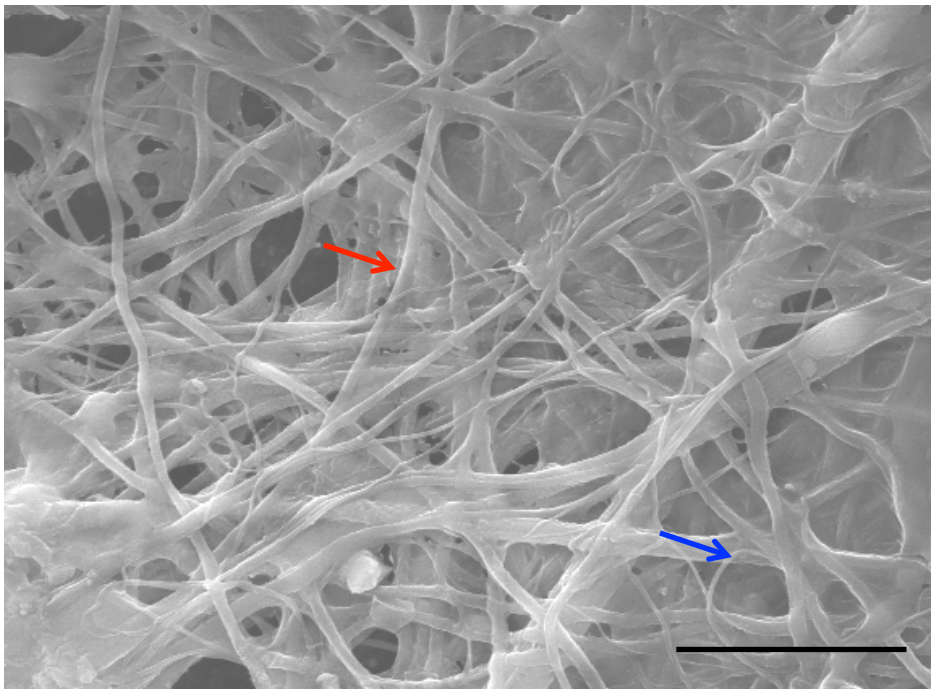


Figure 4.11. Scanning electron micrograph of *Argyroneta aquatica* diving bell. Large diameter fibers are indicated by a red arrow and smaller diameter fibers by a blue arrow. Scale bar 5 μm.

Table 4.1. Summary of *de novo* transcriptome assemblies

Species	No. Raw Paired Reads	No. Cleaned Paired Reads	No. Contigs	Total Length (bp)	N50 (bp)
Argyroneta aquatica	57,211,947	52,277,995	343,077	209,100,116	805
Badumna longinqua	12,104,773	7,805,324	49,520	21,859,077	471
Desis marina	112,108,725	99,597,666	402,833	263,857,902	892

Table 4.2. Spidroins from this study for *Argyroneta aquatica*, *Badumna longinqua,* and *Desis marina*

Spidroin Name[a,b]	Contig name	Top BLASTx hit Accession	Top BLAST hit Description	E-value
***A. aqu*_AcSp_C**	DN138833_c0_g1_i1-1_AcSp_C	gi|587655300|gb|AHK09813.1|	Aciniform spidroin 1 [*Argiope argentata*]	5E-04
***A. aqu* _PySp_C**	26982_5_PySp_C	gi|257124471|gb|ACV41934.1|	Pyriform spidroin [*Latrodectus hesperus*]	9E-05
***A. aqu* _TuSp_C**	TuSp_C_contig	tr|A6YP77|A6YP77_9ARAC|	Fibroin 1 [*Hypochilus thorelli*]	1E-10
***A. aqu*_AcSp_N_vA**	212588_1_AcSp_N_vA	gi|587655228|gb|AHK09777.1|	Aciniform spidroin 1 N-terminal region variant 2, partial [*Argiope trifasciata*]	1E-45
***A. aqu*_AcSp_N_vB**	292563_1_AcSp_N_vB	gi|422900780|gb|AFX83566.1|	Aciniform spidroin 1, partial [*Latrodectus geometricus*]	1E-49
***A. aqu*_AmSp_N**	39346_1_AmSp_N	gi|193506893|gb|ACF19412.1|	Tubuliform spidroin 1, partial [*Agelenopsis aperta*]	9E-25
***A. aqu*_PySp_N**	110799_PySp_N	gi|675382271|KFM75168.1|	Hypothetical protein X975_11824, partial [*Stegodyphus mimosarum*]	2E-18
***A. aqu*_Sp_N**	Sp_N_contig	gi|675382271|KFM70693.1|	Hypothetical protein X975_03452, partial [*Stegodyphus mimosarum*]	9E-16
***A. aqu*_TuSp_N**	DN233571_c0_g1_i1-1_TuSp_N	gi|303307781|gb|ADM14330.1|	Egg case silk protein 1 [Argiope bruennichi]	5E-32
***B. lon*_AcSp_C_vA**	39029_2_AcSp_C_vA	gi|422900764|gb|AFX83559.1|	Aciniform spidroin 1, partial [*Latrodectus hesperus*]	6E-31
***B. lon*_AcSp_C_vB**	33606_1_AcSp_C_vB	gi|422900780|gb|AFX83566.1|	Aciniform spidroin 1, partial [*Latrodectus geometricus*]	2E-20
***B. lon*_AmSp_C_vA**	2322_1_AmSp_C_vA	gi|38197757|gb|AAR13813.1	Major ampullate spidroin-2, partial [*Argiope amoena*]	8E-10
***B. lon*_AmSp_C_vB**	28393_2_AmSp_C_vB	gi|55274114|gb|AAV48937.1|	Dragline silk spidroin 1, partial [*Psechrus sinensis*]	0.081
***B. lon*_AmSp_C_vC**	27482_1_AmSp_C_vC	gi|422900780|gb|AFX83566.1|	Aciniform spidroin 1, partial [Latrodectus geometricus]	2E-20
***B. lon*_AmSp_C_vD**	5222_4_AmSp_C_vD	gi|38197749|gb|AAR13809.1|	Major ampullate spidroin-2, partial [*Argiope amoena*]	0.076
***B. lon*_AmSp_C_vE**	14674_118_AmSp_C_vE	gi|55274104|gb|AAV48932.1|	Dragline silk spidroin 1, partial [*Cyrtophora moluccensis*]	5E-05
***B. lon*_AmSp_C_vF**	24990_2_AmSp_C_vF	gi|55274094|gb|AAV48927.1|	Dragline silk spidroin 1, partial [*Nephila pilipes*]	3E-07
***B. lon*_CrSp_C**	BlonF_DN25631_c1_g7_i1-1_CrSp_C	gi|392997864|gb|AFM97615.1|	Fibroin 1, partial [*Hypochilus thorelli*]	0.019
***B. lon*_PySp_C**	52759_3_PySp_C	gi|675382271|KFM75168.1|	Hypothetical protein X975_11824, partial [*Stegodyphus mimosarum*]	6E-07
***B. lon*_TuSp_C**	27600_2_TuSp_C	gi|303307770|gb|ADM14323.1|	Tubuliform spidroin 1, partial [*Agelenopsis aperta*]	7E-04
***B. lon*_AcSp_N_vA**	39029_1_AcSp_N_vA	gi|422900780|gb|AFX83566.1|	Aciniform spidroin 1, partial [*Latrodectus geometricus*]	1E-42
***B. lon*_AcSp_N_vB**	BlonF_DN25813_c1_g1_i1-1_AcSp_N_vB	gi|422900780|gb|AFX83566.1|	Aciniform spidroin 1, partial [*Latrodectus geometricus*]	2E-41
***B. lon*_AmSp_N_vA**	11418_9_AmSp_N_vA	gi|115635734|emb|CAJ90517.1	Major ampullate spidroin 1 precursor, partial	3E-30

		|	[*Euprosthenops australis*]	
***B. lon*_AmSp_N_vB**	BlonF_DN25728_c5_g3_i3-1_AmSp_N_vB	gi|87133239|gb|ABD24294.1|	Major ampullate spidroin 1, partial [*Latrodectus hesperus*]	8E-14
***B. lon*_AmSp_N_vC**	BlonF_DN25700_c5_g6_i2-1_AmSp_vC	gi|115635734|emb|CAJ90517.1|	Major ampullate spidroin 1 precursor, partial [*Euprosthenops australis*]	5E-15
***B. lon*_AmSp_N_vD**	BlonF_DN25728_c5_g7_i1-1_AmSp_N_vD	gi|193506893|gb|ACF19412.1|	Major ampullate spidroin 1B precursor, partial [*Nephila clavipes*]	9E-27
***B. lon*_AmSp_N_vE**	14674_166_AmSp_N_vE	gi|115635734|emb|CAJ90517.1|	Major ampullate spidroin 1 precursor, partial [*Euprosthenops australis*]	1E-26
***B. lon*_AmSp_N_vF**	25515_1_AmSp_N_vF	gi|193506893|gb|ACF19412.1|	Major ampullate spidroin 1B precursor, partial [*Nephila clavipes]*	6E-31
***B. lon*_AmSp_N_vG**	27565_1_AmSp_N_vG	gi|193506893|gb|ACF19412.1|	Major ampullate spidroin 1B precursor, partial [*Nephila clavipes*]	2E-49
***B. lon*_Sp_N_vA**	52650_1_Sp_N_vA	gi|587655204|gb|AHK09765.1|	Aciniform spidroin 1 N-terminal region variant 3, partial [*Argiope argentata*]	8E-15
***B. lon*_PySp_N**	BlonF_DN22192_c0_g1_i1-1_PySp_N	gi|675382271|gb|KFM75168.1|	Hypothetical protein X975_11824 [*Stegodyphus mimosarum*]	8E-22
***B. lon*_Sp_N_vC**	BlonF_DN25631_c2_g1_i1-1_Sp_N_vC	gi|422900780|gb|AFX83566.1|	Aciniform spidroin 1, partial [*Latrodectus geometricus*]	2e-23
***B. lon*_Sp_N_vB**	28432_1_Sp_N_vB	gi|303307752|gb|ADM14314.1|	Major ampullate spidroin 1, partial [*Kukulcania hibernalis*]	2e-10
***B. lon*_Sp_N_vD**	BlonF_DN24903_c1_g1_i1-1_Sp_N_vD	gi|303307750|gb|ADM14313.1|	Fibroin 1, partial [*Bothriocyrtum californicum*]	2e-11
***B. lon*_TuSp_N**	BlonF_DN12313_c0_g1_i1-1_TuSp_N	gi|303307781|gb|ADM14330.1|	Tubuliform spidroin 1, partial [*Agelenopsis aperta*]	1E-16
***D. mar*_AcSp_C**	47739_11896_AcSp_C	gi|587655300|gb|AHK09813.1|	Aciniform spidroin 1 [*Argiope argentata*]	2E-08
***D. mar*_AmSp_C_vA**	181647_1_AmSp_C_vA	gi|32815671|gb|AAP88232.1|	Major ampullate spidroin-1, partial [*Argiope amoena*]	9E-12
***D. mar*_AmSp_C_vB**	DN116420_c0_g1_i1-1_AmSp_C_vB	gi|32815671|gb|AAP88232.1|	Major ampullate spidroin-1, partial [*Argiope amoena*]	2E-14
***D. mar*_PySp_C**	DN126967_c4_g2_i1-1_PySp_C	gi|675382271|gb|KFM75168.1|	Hypothetical protein X975_11824 [*Stegodyphus mimosarum*]	1E-50
***D. mar*_TuSp_C**	DN125817_c4_g1_i1-1_TuSp_C	gi|303307770|gb|ADM14323.1|	Tubuliform spidroin 1, partial [*Agelenopsis aperta*]	7E-17
***D. mar*_Sp_C**	DN254754_c0_g1_i1-1_Sp_C	gi|1263285|gb|AAC47009.1|	Fibroin-2, partial [*Araneus diadematus*]	3E-20
***D. mar*_AcSp_N**	401495_1_AcSp_N	gi|422900780|gb|AFX83566.1|	Aciniform spidroin 1, partial [*Latrodectus geometricus*]	2E-47
***D. mar*_AmSp_N_vA**	339931_1_AmSp_N_vA	gi|193506891|gb|ACF19411.1|	Major ampullate spidroin 1A precursor, partial [*Nephila clavipes*]	2E-35
***D. mar*_AmSp_N_vB**	DN98982_c0_g1_i1-1_AmSp_N_vB	gi|303307772|gb|ADM14324.1|	Major ampullate spidroin, partial [*Agelenopsis aperta*]	1E-21
***D. mar*_PySp_N**	DN116815_c0_g1_i1-1_PySp_N	gi|675382271|gb|KFM75168.1|	Hypothetical protein X975_11824 [*Stegodyphus mimosarum*]	4E-19
***D. mar*_Sp_N**	357401_1_Sp_N	gi|164709230|gb|ABY67420.1|	Major ampullate spidroin 1 locus 3, partial [*Latrodectus geometricus*]	1E-04

***D. mar*_TuSp_N**	TuSp_N_vA	gi|303307781|gb|ADM14330.1|	Tubuliform spidroin 1, partial [*Agelenopsis aperta*]	2E-17

[a] N or C in spidroin names indicate whether a contig contains the N- or C- terminal region coding sequence.

[b] Variant name (e.g. _vA) does not indicate association of N- term transcripts with C-terminal transcripts.

Table 4.3. GenBank accession numbers for spidroins sequences used in phylogenetic analyses.

Sequence Name	Species	N-terminal region	C-terminal region
***A.ape*_MaSp**	*Agelenopsis aperta*	HM752573	AAT08436
***A.ape*_TuSp1**	*Agelenopsis aperta*	HM752576	--
***A.arg*_AcSp**	*Argiope argentata*	AHK09813	AHK09813
***A.arg*_Flag**	*Argiope argentata*	--	MF955778
***A.arg*_MaSp1**	*Argiope argentata*	MF955677	MF955761
***A.arg*_MaSp2**	*Argiope argentata*	MF955700	MF955804
***A.arg*_MaSp3**	*Argiope argentata*	MF955785	MF955690
***A.arg*_MiSp**	*Argiope argentata*	MF955726	MF955717
***A.arg*_PySp1**	*Argiope argentata*	AQR58363	AQR58363
***A.arg*_TuSp1**	*Argiope argentata*	ATW75951	ATW75951
***A.dia*_AcSp**	*Araneus diadematus*	MF955743	MF955754
***A.dia*_Flag**	*Araneus diadematus*	MF955789	MF955779
***A.dia*_Masp1**	*Araneus diadematus*	MF955789	MF955789
***A.dia*_MaSp2**	*Araneus diadematus*	MF955703	MF955809
***A.dia*_MaSp3**	*Araneus diadematus*	--	MF955691
***A.dia*_MiSp**	*Araneus diadematus*	--	MF955718
***A.dia*_PySp1**	*Araneus diadematus*	MF955713	MF955772
***A.dia*_TuSp1**	*Araneus diadematus*	MF955696	MF955799
***B.cal*_fibroin1**	*Bothriocyrtum californicum*	HM752562	EU117162
***D.ten*_fib1**	*Dolomedes tenebrosus*	--	AF350269
***D.ten*_fib2**	*Dolomedes tenebrosus*	--	AF350270
***L.hes* _Flag**	*Latrodectus hesperus*	MF955792	MF955781
***L.hes* _MaSp1**	*Latrodectus hesperus*	ABR68856	ABR68856
***L.hes* _MaSp2**	*Latrodectus hesperus*	ABR68855	ABR68855
***L.hes* _MaSp3**	*Latrodectus hesperus*	MF955786	MF955692
***L.hes* _MiSp**	*Latrodectus hesperus*	ARA91152	ARA91152
***L.hes* _PySp1**	*Latrodectus hesperus*	MF955714	MF955773
***L.hes* _TuSp1**	*Latrodectus hesperus*	MF955697	MF955801
***L.hes*_AcSp1**	*Latrodectus hesperus*	MF955746	MF955757
***N.cla*_AcSp1**	*Nephila clavipes*	MF955747	MF955758
***N.cla*_Flag**	*Nephila clavipes*	MF955793	MF955782
***N.cla*_MaSp1**	*Nephila clavipes*	MF955682	MF955765
***N.cla*_MaSp2**	*Nephila clavipes*	MF955708	MF955815
***N.cla*_MiSp**	*Nephila clavipes*	MF955734	MF955722
***N.cla*_PySp1**	*Nephila clavipes*	MF955715	MF955774
***N.cla*_TuSp1**	*Nephila clavipes*	MF955698	MF955802
***N.cla*_TuSp1**	*Nephila clavipes*	MF955698	MF955802
***S.mim*_AcSp-putative**	*Stegodyphus mimosarum*	KFM79920	KFM79920
***S.mim*_MaSp-putative-a**	*Stegodyphus mimosarum*	--	KFM83271
***S.mim*_MaSp-putative-c**	*Stegodyphus mimosarum*	--	JT038023
***S.mim*_MaSp-putative-d**	*Stegodyphus mimosarum*	KFM59474	KFM59474
***S.mim*_MaSp-putative-e**	*Stegodyphus mimosarum*	KFM74936	--
***S.mim*_MaSp-putative-f**	*Stegodyphus mimosarum*	KFM61798	--
***S.mim*_MaSp-putative-g**	*Stegodyphus mimosarum*	KFM57717	--
***S.mim*_MaSp-putative-h**	*Stegodyphus mimosarum*	KFM61802	KFM61800
***S.mim*_MaSp-putative-i**	*Stegodyphus mimosarum*	KFM79313	KFM79313
***S.mim*_Misp-putative**	*Stegodyphus mimosarum*	KFM62627	KFM62627
***S.mim*_PiSp-putative**	*Stegodyphus mimosarum*	KFM75168	KFM68615

***S.mim*_Sp1**	*Stegodyphus mimosarum*	--	KFM60634
***S.mim*_Sp2a**	*Stegodyphus mimosarum*	KFM73910	--
***S.mim*_Sp2b**	*Stegodyphus mimosarum*	KFM70693	--
***S.mim*_TuSp-putative**	*Stegodyphus mimosarum*	KFM79920	KFM79920

References

Adams, R.J., and Manolis, T.D. (2014). Field Guide to the Spiders of California and the Pacific Coast States (University of California Press).

Agnarsson, I., Boutry, C., Wong, S.-C., Baji, A., Dhinojwala, A., Sensenig, A.T., and Blackledge, T.A. (2009). Supercontraction forces in spider dragline silk depend on hydration rate. Zoology *112*, 325–331.

Alfaro, R.E., Griswold, C.E., and Miller, K.B. (2018a). -Comparative spigot ontogeny across the spider tree of life. PeerJ *6*, e4233.

Alfaro, R.E., Griswold, C.E., and Miller, K.B. (2018b). The ontogeny of the spinning apparatus of *Tengella perfuga* Dahl (Araneae: Zoropsidae). Invertebrate Biology Forthcoming.

Askarieh, G., Hedhammar, M., Nordling, K., Saenz, A., Casals, C., Rising, A., Johansson, J., and Knight, S.D. (2010). Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. Nature *465*, 236–238.

Ayoub, N.A., Garb, J.E., Tinghitella, R.M., Collin, M.A., and Hayashi, C.Y. (2007). Blueprint for a high-performance biomaterial: full-length spider dragline silk genes. PloS One *2*, e514.

Ayoub, N.A., Garb, J.E., Kuelbs, A., and Hayashi, C.Y. (2013a). Ancient properties of spider silks revealed by the complete gene sequence of the prey-wrapping silk protein (AcSp1). Mol. Biol. Evol. *30*, 589–601.

Ayoub, N.A., Garb, J.E., Kuelbs, A., and Hayashi, C.Y. (2013b). Ancient properties of spider silks revealed by the complete gene sequence of the prey-wrapping silk protein (AcSp1). Mol. Biol. Evol. *30*, 589–601.

Babb, P.L., Lahens, N.F., Correa-Garhwal, S.M., Nicholson, D.N., Kim, E.J., Hogenesch, J.B., Kuntner, M., Higgins, L., Hayashi, C.Y., Agnarsson, I., et al. (2017). The *Nephila clavipes* genome highlights the diversity of spider silk genes and their complex expression. Nat. Genet. *49*, 895–903.

Beckwitt, R., and Arcidiacono, S. (1994). Sequence conservation in the C-terminal region of spider silk proteins (Spidroin) from *Nephila clavipes* (Tetragnathidae) and *Araneus bicentenarius* (Araneidae). J. Biol. Chem. *269*, 6661–6663.

van Beek, J.D., Kümmerlen, J., Vollrath, F., and Meier, B.H. (1999). Supercontracted spider dragline silk: a solid-state NMR study of the local structure. Int. J. Biol. Macromol. *24*, 173–178.

Blasingame, E., Tuton-Blasingame, T., Larkin, L., Falick, A.M., Zhao, L., Fong, J., Vaidyanathan, V., Visperas, A., Geurts, P., Hu, X., et al. (2009). Pyriform spidroin 1, a novel member of the silk gene family that anchors dragline silk fibers in attachment discs of the black widow spider, *Latrodectus hesperus*. J. Biol. Chem. *284*, 29097–29108.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Boutry, C., and Blackledge, T.A. (2010). Evolution of supercontraction in spider silk: structure–function relationship from tarantulas to orb-weavers. J. Exp. Biol. *213*, 3505–3514.

Boutry, C., and Blackledge, T.A. (2013). Wet webs work better: humidity, supercontraction and the performance of spider orb webs. J. Exp. Biol. *216*, 3606–3610.

Casem, M.L., Collin, M.A., Ayoub, N.A., and Hayashi, C.Y. (2010). Silk gene transcripts in the developing tubuliform glands of the Western black widow, *Latrodectus hesperus*. J. Arachnol. *38*, 99–103.

Chaw, R.C., Zhao, Y., Wei, J., Ayoub, N.A., Allen, R., Atrushi, K., and Hayashi, C.Y. (2014). Intragenic homogenization and multiple copies of prey-wrapping silk genes in *Argiope* garden spiders. BMC Evol. Biol. *14*, 31.

Chaw, R.C., Correa-Garhwal, S.M., Clarke, T.H., Ayoub, N.A., and Hayashi, C.Y. (2015). Proteomic evidence for components of spider silk synthesis from black widow silk glands and fibers. J. Proteome Res. *14*, 4223–4231.

Chaw, R.C., Arensburger, P., Clarke, T.H., Ayoub, N.A., and Hayashi, C.Y. (2016). Candidate egg case silk genes for the spider *Argiope argentata* from differential gene expression analyses. Insect Mol. Biol. *25*, 757–768.

Chaw, R.C., Saski, C.A., and Hayashi, C.Y. (2017a). Complete gene sequence of spider attachment silk protein (PySp1) reveals novel linker regions and extreme repeat homogenization. Insect Biochem. Mol. Biol. *81*, 80–90.

Chaw, R.C., Collin, M., Wimmer, M., Helmrick, K.L., and Hayashi, C.Y. (2017b). Egg case silk gene sequences from *Argiope* spiders: evidence for multiple loci and a loss of function between paralogs. G3 GenesGenomesGenetics *8*, 231–238.

Chen, G., Liu, X., Zhang, Y., Lin, S., Yang, Z., Johansson, J., Rising, A., and Meng, Q. (2012). Full-length minor ampullate spidroin gene sequence. PLoS ONE *7*, e52293.

Clarke, T.H., Garb, J.E., Hayashi, C.Y., Haney, R.A., Lancaster, A.K., Corbett, S., and Ayoub, N.A. (2014). Multi-tissue transcriptomics of the black widow spider reveals expansions, co-options, and functional processes of the silk gland gene toolkit. BMC Genomics *15*.

Coddington, J.A. (1989). Spinneret silk spigot morphology: evidence for the monophyly of orbweaving spiders, Cyrtophorinae (Araneidae), and the group Theridiidae plus Nesticidae. J. Arachnol. *17*, 71–95.

Coddington, J.A., and Levi, H.W. (1991). Systematics and evolution of spiders (Araneae). Annu. Rev. Ecol. Syst. 565–592.

Correa-Garhwal, S.M., Chaw, R.C., Clarke, T.H., Ayoub, N.A., and Hayashi, C.Y. (2017). Silk gene expression of theridiid spiders: implications for male-specific silk use. Zoology *122*, 107–114.

De Bakker, D., Baetens, K., Van Nimmen, E., Gellynck, K., Mertens, J., Van Langenhove, L., and Kiekens, P. (2006). Description of the structure of different silk threads produced by the water spider *Argyroneta aquatica* (Clerck, 1757)(Araneae: Cybaeidae). Belg. J. Zool. *136*, 137–143.

Eberhard, W., and Pereira, F. (1993). Ultrastructure of cribellate silk of nine species in eight families and possible taxonomic implications (Araneae: Amaurobiidae, Deinopidae, Desidae, Dictynidae, Filistatidae, Hypochilidae, Stiphidiidae, Tengellidae). J. Arachnol. *21*, 161–174.

Flynn, M.R., and Bush, J.W.M. (2008). Underwater breathing: the mechanics of plastron respiration. J. Fluid Mech. *608*, 275–296.

Forster, R.R. (1970). The spiders of New Zealand. Part III. Desidae, Dictynidae, Hahniidae, Amaurobioididae, Nicodamidae. (Otago Museum Bulletin), pp. 1–184.

Gao, Z., Lin, Z., Huang, W., Lai, C.C., Fan, J., and Yang, D. (2013). Structural characterization of minor ampullate spidroin domains and their distinct roles in fibroin solubility and fiber formation. PLoS ONE *8*, e56142.

Garb, J.E., Ayoub, N.A., and Hayashi, C.Y. (2010). Untangling spider silk evolution with spidroin terminal domains. BMC Evol. Biol. *10*, 243.

Gatesy, J., Hayashi, C., Motriuk, D., Woods, J., and Lewis, R. (2001). Extreme diversity, conservation, and convergence of spider silk fibroin sequences. Science *291*, 2603–2605.

Geurts, P., Zhao, L., Hsia, Y., Gnesa, E., Tang, S., Jeffery, F., Mattina, C.L., Franz, A., Larkin, L., and Vierra, C. (2010). Synthetic spider silk fibers spun from pyriform spidroin 2, a glue silk protein discovered in orb-weaving spider attachment discs. Biomacromolecules *11*, 3495–3503.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from RNA-seq data without a reference genome. Nat. Biotechnol. *29*, 644–652.

Griswold, C.E., Ramírez, M.J., Coddington, J.A., and Platnick, N.I. (2005). Atlas of phylogenetic data for Entelegyne spiders (Araneae: Araneomorphae: Entelegynae), with comments on their phylogeny. Proc. Calif. Acad. Sci. *56*, 1.

Hayashi, C.Y., and Lewis, R.V. (1998). Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks. J. Mol. Biol. *275*, 773–784.

Hayashi, C.Y., and Lewis, R.V. (2001). Spider flagelliform silk: lessons in protein design, gene structure, and molecular evolution. Bioessays *23*, 750–756.

Hayashi, C.Y., Shipley, N.H., and Lewis, R.V. (1999). Hypotheses that correlate the sequence, structure, and mechanical properties of spider silk proteins. Int. J. Biol. Macromol. *24*, 271–275.

Hayashi, C.Y., Blackledge, T.A., and Lewis, R.V. (2004). Molecular and Mechanical Characterization of Aciniform Silk: Uniformity of Iterated Sequence Modules in a Novel Member of the Spider Silk Fibroin Gene Family. Mol. Biol. Evol. *21*, 1950–1959.

Hedhammar, M., Rising, A., Grip, S., Martinez, A.S., Nordling, K., Casals, C., Stark, M., and Johansson, J. (2008). Structural properties of recombinant nonrepetitive and repetitive parts of major ampullate spidroin 1 from *Euprosthenops australis*: implications for fiber formation. Biochemistry (Mosc.) *47*, 3407–3417.

Hinman, M.B., and Lewis, R.V. (1992). Isolation of a clone encoding a second dragline silk fibroin. *Nephila clavipes* dragline silk is a two-protein fiber. J. Biol. Chem. *267*, 19320–19324.

Holland, G.P., Creager, M.S., Jenkins, J.E., Lewis, R.V., and Yarger, J.L. (2008). Determining Secondary Structure in Spider Dragline Silk by Carbon−Carbon Correlation Solid-State NMR Spectroscopy. J. Am. Chem. Soc. *130*, 9871–9877.

Hu, X., Kohler, K., Falick, A.M., Moore, A.M., Jones, P.R., Sparkman, O.D., and Vierra, C. (2005a). Egg Case Protein-1 a new class of silk proteins with fibroin-like properties from the spider *Latrodectus Hesperus*. J. Biol. Chem. *280*, 21220–21230.

Hu, X., Lawrence, B., Kohler, K., Falick, A.M., Moore, A.M., McMullen, E., Jones, P.R., and Vierra, C. (2005b). Araneoid egg case silk: a fibroin with novel ensemble repeat units from the black widow spider, *Latrodectus hesperus*. Biochemistry (Mosc.) *44*, 10020–10027.

Hu, X., Kohler, K., Falick, A.M., Moore, A.M., Jones, P.R., and Vierra, C. (2006). Spider egg case core fibers: trimeric complexes assembled from TuSp1, ECP-1, and ECP-2. Biochemistry (Mosc.) *45*, 3506–3516.

Huemmerich, D., Helsen, C.W., Quedzuweit, S., Oschmann, J., Rudolph, R., and Scheibel, T. (2004). Primary structure elements of spider dragline silks and their contribution to protein solubility. Biochemistry (Mosc.) *43*, 13604–13612.

Ittah, S., Michaeli, A., Goldblum, A., and Gat, U. (2007). A model for the structure of the C-terminal domain of dragline spider silk and the role of its conserved cysteine. Biomacromolecules *8*, 2768–2773.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

Kovoor, J., and Zylberberg, L. (1980). Fine structural aspects of silk secretion in a spider (*Araneus diadematus*). I. Elaboration in the pyriform glands. Tissue Cell *12*, 547–556.

Kovoor, J., and Zylberberg, L. (1982). Fine structural aspects of silk secretion in a spider. II. Conduction in the pyriform glands. Tissue Cell *14*, 519–530.

Kronenberger, K., Dicko, C., and Vollrath, F. (2012). A novel marine silk. Naturwissenschaften *99*, 3–10.

Kümmerlen, J., van Beek, J.D., Vollrath, F., and Meier, B.H. (1996). Local Structure in Spider Dragline Silk Investigated by Two-Dimensional Spin-Diffusion Nuclear Magnetic Resonance. Macromolecules *29*, 2920–2928.

Main, B.Y. (2001). Historical ecology, responses to current ecological changes and conservation of Australian spiders. J. Insect Conserv. *5*, 9–25.

Marx, M.T., and Messner, B. (2012). A general definition of the term “plastron“ in terrestrial and aquatic arthropods. Org. Divers. Evol. *12*, 403–408.

Mclay, C.L., and Hayward, T.L. (1987). Reproductive biology of the intertidal spider *Desis marina* (Araneae: Desidae) on a New Zealand rocky shore. J. Zool. *211*, 357–372.

McQueen, D.J., and McLay, C.L. (1983). How does the intertidal spider Desis marina (Hector) remain under water for such a long time? N. Z. J. Zool. *10*, 383–391.

McQueen, D.J., Pannell, L.K., and McLay, C.L. (1983). Respiration rates for the intertidal spider Desis marina (Hector). N. Z. J. Zool. *10*, 393–399.

Miller, J.A., Carmichael, A., Ramírez, M.J., Spagna, J.C., Haddad, C.R., Řezáč, M., Johannesen, J., Král, J., Wang, X.-P., and Griswold, C.E. (2010). Phylogeny of entelegyne spiders: Affinities of the family Penestomidae (NEW RANK), generic phylogeny of Eresidae, and asymmetric rates of change in spinning organ evolution (Araneae, Araneoidea, Entelegynae). Mol. Phylogenet. Evol. *55*, 786–804.

Neumann, D., and Kureck, A. (2013). Composite structure of silken threads and a proteinaceous hydrogel which form the diving bell wall of the water spider *Agyroneta aquatica*. SpringerPlus *2*, 223.

Pedersen, O., and Colmer, T.D. (2012). Physical gills prevent drowning of many wetland insects, spiders and plants. J. Exp. Biol. *215*, 705–709.

Perry, D.J., Bittencourt, D., Siltberg-Liberles, J., Rech, E.L., and Lewis, R.V. (2010). Piriform spider silk sequences reveal unique repetitive elements. Biomacromolecules *11*, 3000–3006.

Peters, H.M. (1984). The spinning apparatus of Uloboridae in relation to the structure and construction of capture threads (Arachnida, Araneida). Zoomorphology *104*, 96–104.

Peters, H.M. (1987). Fine structure and function of capture threads. In Ecophysiology of Spiders, (Springer, Berlin, Heidelberg), pp. 187–202.

Peters, H.M. (1992). On the spinning apparatus and the structure of the capture threads of *Deinopis subrufus* (Araneae, Deinopidae). Zoomorphology *112*, 27–37.

Prosdocimi, F., Bittencourt, D., Silva, F.R. da, Kirst, M., Motta, P.C., and Rech, E.L. (2011). Spinning gland transcriptomics from two main clades of spiders (order: Araneae) - insights on their molecular, anatomical and behavioral evolution. PLOS ONE *6*, e21634.

Savage, K.N., and Gosline, J.M. (2008). The role of proline in the elastic mechanism of hydrated spider silks. J. Exp. Biol. *211*, 1948–1957.

Savage, K.N., Guerette, P.A., and Gosline, J.M. (2004). Supercontraction stress in spider webs. Biomacromolecules *5*, 675–679.

Seymour, R.S., and Hetz, S.K. (2011). The diving bell and the spider: the physical gill of *Argyroneta aquatica*. J. Exp. Biol. *214*, 2175–2181.

Shao, H., Bachus, K.N., and Stewart, R.J. (2009). A water-borne adhesive modeled after the sandcastle glue of *P. californica*. Macromol. Biosci. *9*, 464–471.

Spagna, J.C., Crews, S.C., and Gillespie, R.G. (2010). Patterns of habitat affinity and Austral/Holarctic parallelism in dictynoid spiders (Araneae : Entelegynae). Invertebr. Syst. *24*, 238–257.

Sponner, A., Unger, E., Grosse, F., and Weisshart, K. (2004). Conserved C-termini of spidroins are secreted by the major ampullate glands and retained in the silk thread. Biomacromolecules *5*, 840–845.

Sponner, A., Unger, E., Grosse, F., and Weisshart, K. (2005). Differential polymerization of the two main protein components of dragline silk during fibre spinning. Nat. Mater. *4*, 772–775.

Stamatakis, A. (2014). Raxml version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 1312–1313.

Starrett, J., Garb, J.E., Kuelbs, A., Azubuike, U.O., and Hayashi, C.Y. (2012). Early events in the evolution of spider silk genes. PloS One *7*, e38084.

Tian, M., and Lewis, R.V. (2005). Molecular characterization and evolutionary study of spider tubuliform (eggcase) silk protein. Biochemistry (Mosc.) *44*, 8006–8012.

Vasanthavada, K., Hu, X., Falick, A.M., La Mattina, C., Moore, A.M., Jones, P.R., Yee, R., Reza, R., Tuton, T., and Vierra, C. (2007). Aciniform spidroin, a constituent of egg case sacs and wrapping silk fibers from the black widow spider *Latrodectus hesperus*. J. Biol. Chem. *282*, 35088–35097.

Vink, C.J., McQuillan, B.N., Simpson, A., and Correa-Garhwal, S.M. (2017). The marine spider, *Desis marina* (Araneae: Desidae): New observations and localities. The Weta 71–79.

Wasowska, S. (1977). Studies of the spinning apparatus in spiders: Postembryonic morphogeny of the spinning aparatus. Zool. Pol. *26*, 355–405.

Wheeler, W.C., Coddington, J.A., Crowley, L.M., Dimitrov, D., Goloboff, P.A., Griswold, C.E., Hormiga, G., Prendini, L., Ramírez, M.J., Sierwald, P., et al. (2017). The spider tree of life: phylogeny of Araneae based on target-gene analyses from an extensive taxon sampling. Cladistics *33*, 574–616.

Wolff, J.O., Grawe, I., Wirth, M., Karstedt, A., and N. Gorb, S. (2015). Spider's super-glue: thread anchors are composite adhesives with synergistic hierarchical organization. Soft Matter *11*, 2394–2403.

Work, R.W. (1981). A Comparative Study of the Supercontraction of Major Ampullate Silk Fibers of Orb-Web-Building Spiders (Araneae). J. Arachnol. *9*, 299–308.

World Spider Catalog (2018). World Spider Catalog.

Yang, Y.J., Jung, D., Yang, B., Hwang, B.H., and Cha, H.J. (2014). Aquatic proteins with repetitive motifs provide insights to bioengineering of novel biomaterials. Biotechnol. J. *9*, 1493–1502.

Yang, Z., Liivak, O., Seidel, A., LaVerde, G., Zax, D.B., and Jelinski, L.W. (2000). Supercontraction and backbone dynamics in spider silk: 13C and 2H NMR studies. J. Am. Chem. Soc. *122*, 9019–9025.

Yonemura, N., Sehnal, F., Mita, K., and Tamura, T. (2006). Protein composition of silk filaments spun under water by caddisfly larvae. Biomacromolecules *7*, 3370–3378.

Zhao, A.-C., Zhao, T.-F., Nakagaki, K., Zhang, Y.-S., SiMa, Y.-H., Miao, Y.-G., Shiomi, K., Kajiura, Z., Nagata, Y., Takadera, M., et al. (2006). Novel molecular and mechanical properties of egg case silk from wasp spider, *Argiope bruennichi*. Biochemistry (Mosc.) *45*, 3348–3356.

Table S4.1. Proteins identified in *Argyroneta aquatica* diving bell. Information not available indicated by dashes.

Identified Protein	Molecular Weight (kDa)	Top BLAST Hit Accession	Top BLAST Hit Description	E-value
***A. aqu*_AcSp_N_vA**	26	gi|193506893|gb|ACF19412.1|	Tubuliform spidroin 1, partial [*Agelenopsis aperta*]	9.0E-25
***A. aqu*_PySp_C**	47	gi|257124471|gb|ACV41934.1|	Pyriform spidroin [*Latrodectus hesperus*]	9.0E-05
***A. aqu*_Sp_N**	28	gi|422900780|gb|AFX83566.1|	Aciniform spidroin 1, partial [*Latrodectus geometricus*]	2.0e-23
DN13205_c0_g1_i1-1	8	gi|33667938|gb|AAQ24546.1|	Blo t 9 allergen [*Blomia tropicalis*]	2.0E-09
253304_1	33	gi|40548521|gb|AAR87381.1|	Tropomyosin [*Neoscona nautica*]	2.0E-149
50840_1	29	gi|318087320|gb|ADV40252.1|	Hypothetical protein, partial [*Latrodectus hesperus*]	8.0E-07
DN141454_c0_g2_i1-1	12	gi|506965343|gb|AGM32062.1|	Hypothetical protein [*Coptotermes formosanus*]	6.4E-02
242175_1	49	gi|523712986|gb|AGQ56699.1|	Vitellogenin 2 [*Neoseiulus cucumeris*]	2.0E-14
DN139557_c0_g2_i1-1	25	gi|1009581138|ref|XP_015921887.1|	Apolipophorins-like [*Parasteatoda tepidariorum*]	5.0E-37
31706_7562	143	gi|2073373|dbj|BAA19844.1|	Alpha-2-macroglobulin [*Limulus sp.*]	0.0E+00
318196_1	38	gi|344178917|dbj|BAK64111.1|	Alpha-2 macroglobulin [*Hasarius adansoni*]	2.0E-49
332195_1	10	gi|671759229|gb|AII98023.1|	BLTX669 [*Nephila pilipes*]	6.0E-10
DN138176_c0_g2_i2-1	50	gi|332027696|gb|EGI67764.1|	Tubulin alpha-1 chain [*Acromyrmex echinatior*]	0.0E+00
DN126632_c0_g5_i1-1	24	gi|405952567|gb|EKC20363.1|	Hypothetical protein CGI_10006238 [*Crassostrea gigas*]	1.0E-35
370852_1	14	gi|556102150|gb|ESO90802.1|	Hypothetical protein LOTGIDRAFT_233588 [*Lottia gigantea*]	3.0E-05
26982_11	35	gi|431921509|gb|ELK18875.1|	Heat shock cognate 71 protein [*Pteropus alecto*]	4.4E+00
DN76940_c0_g1_i1-1	16	gi|492904312|ref|WP_006034718.1|	Elongation factor Tu [*Rickettsiella grylli*]	2.0E-94
DN75073_c0_g1_i1-1	13	gi|492904817|ref|WP_006035223.1|	Thioredoxin [*Rickettsiella grylli*]	7.0E-52
DN104795_c0_g1_i1-1	42	gi|499265039|ref|WP_010962432.1|	MULTISPECIES: ATP synthase subunit beta [*Wolbachia*]	0.0E+00
21585_7	74	gi|170063054|ref|XP_001866937.1|	Scavenger receptor cysteine-rich protein [*Culex quinquefasciatus*]	5.0E-67
265412_1	250	gi|241575657|ref|XP_002403224.1|	Trichohyalin, putative [*Ixodes scapularis*]	9.0E-106
345535_1	32	gi|241744740|ref|XP_002405466.1|	Alkyl hydroperoxide reductase, thiol specific antioxidant, putative [*Ixodes scapularis*]	5.0E-112

DN131058_c3_g1_i1-1	36	gi|241999408|ref|XP_002434347.1|	Glyceraldehyde 3-phosphate dehydrogenase, putative [*Ixodes scapularis*]	0.0E+00
DN143327_c0_g2_i1-1	14	gi|291225136|ref|XP_002732558.1|	Probable serine/threonine-protein kinase kinX [*Saccoglossus kowalevskii*]	1.0E-03
DN120520_c0_g2_i1-1	21	gi|291241282|ref|XP_002740542.1|	Apolipoprotein D-like [*Saccoglossus kowalevskii*]	1.0E-45
322416_3	17	gi|302783030|ref|XP_002973288.1|	Hypothetical protein SELMODRAFT_232026 [*Selaginella moellendorffii*]	1.0E-20
31706_6984	25	gi|339261586|ref|XP_003367832.1|	Putative trypsin Inhibitor like cysteine rich domain protein, partial [*Trichinella spiralis*]	1.0E-03
281967_1	54	gi|391326419|ref|XP_003737714.1|	PREDICTED: glutamate dehydrogenase, mitochondrial [*Galendromus occidentalis*]	0.0E+00
DN32275_c0_g2_i1-1	13	gi|391341738|ref|XP_003745184.1|	PREDICTED: synaptotagmin-15-like [*Metaseiulus occidentalis*]	1.0E-45
358067_1	72	gi|391342852|ref|XP_003745729.1|	PREDICTED: glucose dehydrogenase [FAD, quinone] [*Galendromus occidentalis*]	4.0E-143
31706_8269	71	gi|391342852|ref|XP_003745729.1|	PREDICTED: glucose dehydrogenase [FAD, quinone] [*Galendromus occidentalis*]	3.0E-145
282009_1	68	gi|391342852|ref|XP_003745729.1|	PREDICTED: glucose dehydrogenase [FAD, quinone] [*Galendromus occidentalis*]	7.0E-138
DN137736_c0_g1_i1-1	29	gi|499048261|ref|XP_004574680.1|	Polyserase-2-like [*Maylandia zebra*]	1.0E-30
280834_1	33	gi|556948298|ref|XP_005987015.1|	PREDICTED: low choriolytic enzyme-like isoform X2 [*Latimeria chalumnae*]	1.0E-39
DN143934_c0_g2_i2-1	29	gi|594066187|ref|XP_006057071.1|	PREDICTED: apolipoprotein L3-like [*Bubalus bubalis*]	4.0E-03
DN133244_c0_g1_i2-1	53	gi|573910962|ref|XP_006643199.1|	PREDICTED: transmembrane protease serine 9-like [*Lepisosteus oculatus*]	1.0E-43
31706_1002	14	--	--	--
DN119698_c1_g1_i1-1	5	--	--	--
DN134830_c112_g2_i2-1	6	--	--	--
DN197674_c0_g1_i1-1	5	--	--	--
DN21975_c0_g1_i1-1	9	--	--	--
44288_1	11	--	--	--
DN152011_c3_g1_i1-1	8	--	--	--
DN152897_c7_g1_i1-1	7	--	--	--
365077_1	8	--	--	--
DN150297_c1_g5_i1-1	9	--	--	--

DN17735_c0_g1_i1-1	6	--	--	--
DN129222_c0_g1_i1-1	8	--	--	--

Concluding Remarks

Over millions of years spiders have thrived in a verity of habitats and environments from land to aquatic. Spiders have acquired strategies to accomplish multiple tasks critical to their survival such as reproduction, dispersal, and prey capture. Studies have suggested that the ability of spiders to inhabit such varied environments is liked to their silk use. Despite vast morphological and structural studies of silk, not much emphasis has been given to molecular adaptations of spider silk pertaining to their habitat. With my research I made an attempt to understand the silk structure and diversity at molecular and genetic level. First, I studied sex-specific gene expression in three different Theridiidae species. Both male and female theridiid spiders expressed silk genes associated with aciniform, major ampullate, minor ampullate, and pyriform silk glands. Females were found to have higher expression of silk genes associated with egg production such as tubuliform silk genes. Whereas male spiders had higher expression of silk genes associated with displacement (e.g. dragline silk genes). Within males, a species-specific silk gene expression was also observed. To better understand the silk biology of male spiders, studies of developmental stages are needed to assess the temporal expression pattern of silk genes.

In the second chapter of my study, I examined the molecular composition of silk from aquatic spiders that enable them to survive, capture prey, and reproduce underwater The characterization of 18 spidroins transcripts from the semi-aquatic spider *Dolomedes triton* (Pisauridae) revealed diversity in their repeat composition and show differential spidroin gene expression between sexes. Despite spidroin repeat diversity, there was no

evidence for adaptive modification in *D. triton* spidroin sequences. This suggests that ancestral spider silks are pre-adapted to function in wet environments. Contrarily, proteomic and structural analyses of *D. triton* submersible egg sac identified distinct features that may contribute to the egg sac's waterproof properties.

I characterized the structural proteins of silks of *Tengella perfuga*. *T.perfuga* are terrestrial cribellate spider species that make prey-capture webs with vast amounts of cribellar silk. In this study, I identified the presence of at least seven spidroin genes in *T. perfuga* genome. I also described novel combinations of amino acid motifs for most *T. perfuga* spidroins and documented the expression of a potential candidate cribellar spidroin, CrSp. The discovery of a cribellar candidate provides insights into the composition of cribellar silk Future studies are needed to relate CrSp sequences to the physical properties of cribellar silk and to investigate the evolution of CrSp across different spider linages.

Within the superfamily Dictynoidea spider species are distributed in diverse environments from terrestrial to aquatic. I studied, the terrestrial *Badumna longinqua*, the aquatic fully aquatic spider *Argyroneta aquatica,* and the marine semi-aquatic spider *Desis marina*. I addition I also tried to establish the similarities and / or differences between male and female spiders in aquatic species. I identified 47 silk gene transcripts of spiders within Dictynoidea from three different environments. Silk genes corresponding to aciniform, pyriform, tubuliform, and ampullate spidroins, show diverse repetitive region but show no evidence of modification for changing environments. However, spider species associated with aquatic habitats share a highly hydrophobic

amino acid motif, GV, that is likely associated with a spider-specific strategy for using silk in aquatic environments. Examination of sex-biased expression in aquatic and semi-aquatic spiders shows similar patterns to in previous chapters. Namely, spidroin gene expression differs between males and females across species for silks related to changes in lifestyle after sexual maturation. To better understand the mechanisms of silks function underwater, chemical composition and nano-structure analyses of silks from aquatic spider species are needed.

Combinational approaches are needed to further investigate the role of silk in spiders that adapt to changing environments. My work describes silk genes that inspire new structure/function hypothesis. This field will greatly benefit from trascriptomic, genomic, proteomic, and structural studies to expand the opportunities for unique biomimetic materials.