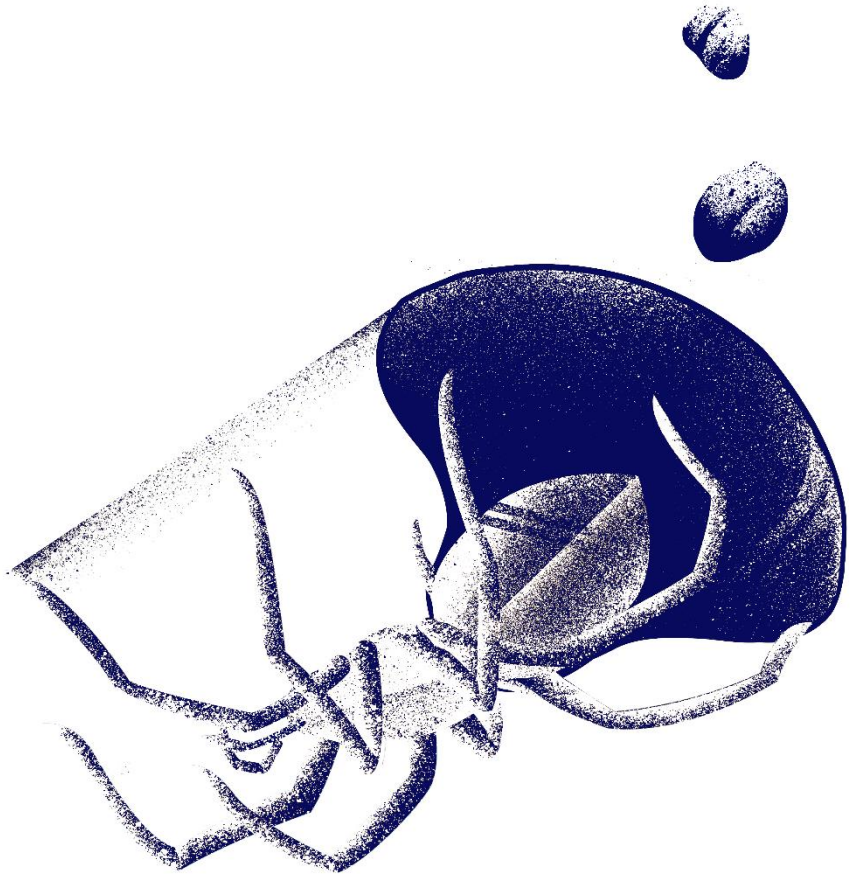
Molecular, genetic and morphological studies of *Argyroneta aquatica* silk, and analysis of a range of silk composites



Michelle Strickland MSci (Hons)

Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy

May 2019

Dedication

A huge number of people have contributed their time and expertise in discussion, advice and assistance during this project and I am indebted to all of them. First and foremost, I would like to thank Sara Goodacre for her ever-enthusiastic supervision and promoting a warm and supportive environment. I would also like to thank members of the SpiderLab, in particular Antje Hundertmark, Alastair Gibbons and Ella Deutsch, who have all made my time a much more fun and interesting experience!

My thanks also go to Louise Johnson and Chris Wade for critical evaluation of this thesis and post-viva comments. Also to John Brookfield and John Owers-Bradley for sharing their expertise, support and assistance.

This thesis is dedicated, in part, to the memory of Dr Victor Tudorica; colleague, friend and a gentleman, who originally took the lead in our *Argyroneta* research. Last, and most certainly not least: my project likely would not have started without James Fowkes, whose encouragement, unwavering support (and quiet tolerance of spiders) means the world to me.

Declaration

I declare that this thesis is the result of my own work, unless otherwise stated, and that it has not been presented as part of any other degree.

Michelle Strickland

Abstract

Argyroneta aquatica are unique in their ability to build an underwater sheet web that is used to hold a bubble of air drawn from the surface. The diving bell allows dissolved oxygen to diffuse into the air bubble, enabling the spiders to remain submersed for extended periods of time. With recent advances in synthetic silk research, this property makes the diving bell of particular interest as a future biomaterial.

This thesis describes molecular, genetic and morphological studies of *A. aquatica* and their silk and shows the phylogenetic relationship between *A. aquatica, Cybaeus angustiarum* and *Pholcus phalangioides* silks and a range of silks from across the Araneae. The conservation of a pH-dependent salt bridge in the C-terminal region of *A. aquatica* silk, which extends across all silk types in the Araneae, suggests *A. aquatica* have retained the same silk fibre formation process as terrestrial spiders, a finding relevant to the future synthetic production of *A. aquatica* silk. Morphological observations of the diving bell web and *A. aquatica*'s spinning apparatus show a layer of hydrogel on both the haphazardly woven web and the spinnerets.

Lastly, this thesis describes an investigation into combinations of dragline silk and carbon nanomaterials (CNM), concluding that spiders exposed to CNM dispersions do not add CNM to their silk, but that CNM can be directly applied to silk fibres. Silk fibres coated in CNM show a small decrease in their tensile properties which may be a result of supercontraction due to the use of water in the CNM dispersions.

Contents

Dedication III

Abstract V

List of Figures XI

List of Tables XIII

1. General Introduction 1

1.1. Spider taxonomy and key species in this thesis 1

1.1.1. Spider taxonomy 1

1.1.2. *Argyroneta aquatica* 3

1.1.3. *Cybaeus angustiarum* 7

1.1.4. *Pholcus phalangioides* 7

1.1.5. Rationale for choice of species for study 8

1.2. Evolution of spider silk 9

1.3. Applications of silk by the Araneae 11

1.4. Non-Araneae silk producers 12

1.5. Spider silk protein macrostructure 13

1.6. Genetic structure of spider silk genes 15

1.7. Silk glands 16

1.8. Spider silk production 17

1.8.1. Extrusion of silk dope 17

1.8.2. Function of the N- and C-terminal domains during silk fibre formation 18

1.9. Human uses for spider silk 19

1.9.1 Traditional applications and folk medicine 19

1.9.2. Modern applications and modification of native spider silk 20

1.9.3. Synthetic silk 20

1.10. A summary of the aims of, and work contained in, this thesis 21

2. General Methodology 25

2.1. Spider collection and husbandry 25

2.1.1. Collection of specimens 25

2.1.2. Storage and transportation 25
2.1.3. Husbandry techniques 26
2.2. RNA sequencing 27
2.2.1. Dissection, storage and transportation of silk glands 27
2.2.2. RNA extraction 28
2.2.3. RNA sequencing, quality control and transcriptome assembly 28
2.3. Collection of dragline silk 28
3. Identification of novel spidroin genes in *Argyroneta aquatica*, *Cybaeus angustiarum* and *Pholcus phalangioides* 31
3.1. Introduction 31
3.1.1. Methods of identifying novel genes 31
3.2. Methods 33
3.2.1. Spidroin identification 33
3.3. Results 34
3.4. Discussion 38
3.4.1. Use of transcriptomics to sequence and identify novel spidroins 38
3.4.2. Use of sample tissue for transcriptome sequencing 39
3.5. Conclusion 41
4. Analysis of the terminal domains of spider silk, including novel spidroins from *Argyroneta aquatica*, *Cybaeus angustiarum* and *Pholcus phalangioides* 43
4.1. Abstract 43
4.2. Introduction 43
4.3. Materials and Methods 45
4.3.1. RNA extraction and transcriptome sequencing 45
4.3.2. Transcriptome assembly and identification of silk genes 46
4.3.3. Phylogenetic analysis 46
4.4. Results 47
4.4.1. Transcriptome analysis 47
4.4.2. Analysis of silk terminal domains 48
4.4.4. Conservation of a salt bridge structure across spider silk genes 58

4.5. Discussion **60**
4.5.1. *A. aquatica* transcriptome 60
4.5.2. *C. angustiarum* transcriptome 62
4.5.3. *P. phalangioides* transcriptome 63
4.5.4. Analysis of N- and C-terminal domains 63
4.5.5. Presence and conservation of salt bridges within the terminal domains 65

5. Morphology of the web and spinning apparatus of the water spider, *Argyroneta aquatica* and comparison with *Cybaeus angustiarum* **67**

5.1. Introduction **67**

5.3. Materials and methods **68**
5.3.1. Study animals 68
5.3.2. Preparation of diving bell and web specimens for environmental Scanning Electron Microscope (eSEM) imaging 68
5.3.3. Preparation of spider specimens for Scanning Electron Microscope (SEM) imaging 68
5.3.4. Imaging of silk glands 69

5.4. Results **69**
5.4.2. Morphology of the diving bell 69
5.4.3. Structure of *C. angustiarum* web 71
5.4.4. Spinnerets 72
5.4.5. Silk gland morphology 76

5.5. Discussion **78**
5.5.2. Structure of the diving bell of *A. aquatica* and web of *C. angustiarum* 78
5.5.3. SEM of spinnerets 79

5.6. Conclusions **80**

6. Modifying the physical properties of spider silk by combining with dispersed carbon nanomaterials **81**

6.1. Introduction **81**
6.1.1. Properties of carbon nanomaterials 82
6.1.2. Modification 84
6.1.3. Overview 85

6.2. Methods **86**
6.2.1. Carbon nanomaterial preparation and Raman spectroscopy 86
6.2.2. Application of carbon nanomaterials 89
6.2.3. Tensile testing of silk samples 90
6.2.4. Measurement of sample diameter 90
6.2.5. Calculations of tensile strength, stress and Young's Modulus 91
6.3. Results **92**
6.3.1. Measurements of silk fibres under SEM 92
6.3.2. "Spray on spider" experimentation 93
6.3.3. Tensile test of silks incorporated with CNM by spraying 96
6.4. Discussion **101**
6.4.1. Integration of CNM into spider silks 102
6.4.2. Discussion of techniques and suggested improvements 107
6.5. Conclusions **108**
7. General discussion **111**
7.1. Views, perspective and future directions **111**
7.1.1. Further sequencing of *Argyroneta aquatica*, *Cybaeus angustiarum* and *Pholcus phalangioides* spidroins 111
7.1.2. *A. aquatica* silk - carbon nanomaterial composites 115
7.1.3. Further research into silk of semi-aquatic spiders 116
7.1.4. Synthetic *A. aquatica* silk 117
8. Bibliography **119**
9. Appendix **137**
9.1. Local BLAST database spidroin sequence list **137**
9.2. Tensile test data **148**
10. Acknowledgements **153**
11. PIP Reflective Statement **155**
12. Publications **159**

List of Figures

Figure 1.1. Timeline depicting first appearance of spiders (Araneae) and first evidence of silk. 1
Figure 1.2. Adult female *A. aquatica*. 3
Figure 1.3. Image of an adult *A. aquatica* resting in a diving bell. 4
Figure 1.4. Section of updated spider phylogenetic tree. 6
Figure 1.5. Image of an adult female *C. angustiarum.* 7
Figure 1.6. Image of an adult female *P. phalangioides.* 8
Figure 1.7. Images of spider spinnerets and spigots. 10
Figure 1.8. Schematic overview of the different products of the silk glands of a female orb weaver. 14
Figure 1.9. Overview of the different regions of a silk gene/protein. 15
Figure 1.10. Diagram showing the different regions of a silk gland. 17
Figure 1.11. Overview of the silk spinning process. 18
Figure 2.1. Diagram depicting collection of dragline silk. 29
Figure 3.1. Sample alignment from a BLASTp result file. 37
Figure 4.1. Familial origins of silk sequences used in this study. 49
Figure 4.2. Translated alignment of all silk C-termini nucleotide sequences used in this study. 52
Figure 4.3. Translated alignment of all silk N-termini nucleotide sequences used in this study. 52
Figure 4.4 Maximum-likelihood tree of all silk C-terminal nucleotide sequences used in this study 53
Figure 4.5. Alignment of *Nephila clavipes* C-terminal sequences showing individual JPred secondary structure predictions (Jnet) beneath. 56
Figure 4.6. Alignment of *Nephila clavipes* N-terminal sequences showing individual JPred secondary structure predictions (Jnet) beneath. 57
Figure 4.7. Colourised alignment of *Nephila clavipes* C-terminal sequences. 59
Figure 5.1. Photographic and eSEM images of the diving bell formed by one adult female *A. aquatica.* 70
Figure 5.2. eSEM images of the web formed by one adult female *C. angustiarum.* 71
Figure 5.3. SEM images of the spinnerets of an adult female *A. aquatica.* 72
Figure 5.4. SEM images of the spinnerets of an adult female *C. angustiarum.* 74
Figure 5.5. Optical microscope images of the dissected silk glands of *A. aquatica*. 76

Figure 6.1. Schematic representation of some of the available carbon nanomaterials. 81
Figure 6.2. Chart showing the Raman spectra of the four CNM used in this study. 88
Figure 6.3. Representation of "spray on spider" (A) and "spray on silk" (B) application and collection process. 89
Figure 6.4. Area of toughness on a stress-strain chart. 92
Figure 6.5. Raman spectra of original graphite and nanotube samples compared to those measured in *A. diadematus* silk after "spray on silk" experiments. 95
Figure 6.6. Images of *A. diadematus* silk after "spray on silk" experiments. 96
Figure 6.7. Results of tensile tests of *P. phalangioides* silk. 98
Figure 6.8. Results of tensile tests of *A. diadematus* 1 and 2 silk. 99
Figure 6.9. Results of tensile tests of *A. diadematus* 3 silk 100

List of Tables

Table 3.1. Summary of silk spidroin sequences used in the custom BLAST database. 35
Table 3.2. Summary of transcriptome data. 35
Table 3.3. Summary of putative spidroin sequences identified. 36
Table 4.1. Summary of unique novel N- and C-terminal silk spidroin sequences identified from the combined *A. aquatica* transcriptomes. 47
Table 4.2. Examples of translated repeat regions from *Argyroneta aquatica* and *Argiope trifasciata*. 48
Table 5.1. Description of the spigots found in adult *A. aquatica* and *C. angustiarum.* 75
Table 6.1. Average values of tensile properties of CNM, dragline silks, steel and Kevlar. 83
Table 6.2. Commercial suppliers of CNM and their dimensions. 87
Table 6.3. Average measurements of diameter and subsequently-calculated cross-sectional area of silk fibre samples. 93
Table 6.4. Record of spiders after application of CNM or control substances. 93
Table 6.5. Number of silk samples collected and tested. 97
Table 6.6. Record of silk samples damaged after treatment and before tensile testing could be completed. 97
Table 9.1. Full list of all spidroin sequences included in the custom local BLAST database. 137
Table 9.2 Transformed tensile data from the three groups of tensile-tested silk analysed in this study. 148

1. General Introduction

1.1. Spider taxonomy and key species in this thesis

Spiders (Order: Araneae) are eight-legged arthropods, found on every continent with the single exception of Antarctica. As of May 2019, there are 48,165 formally identified species of spider in 117 families (Catalog, 2019), making it the largest known order of arachnids and one of the most diverse groups within the animal kingdom. One of the defining features of the Araneae is their creative and varied use of silk and ability to utilise silk throughout their lifetime. A number of insects make silks, but these animals typically only produce silk for a single purpose at a single time in development (e.g. building a cocoon to enable the transition from larvae to adult).

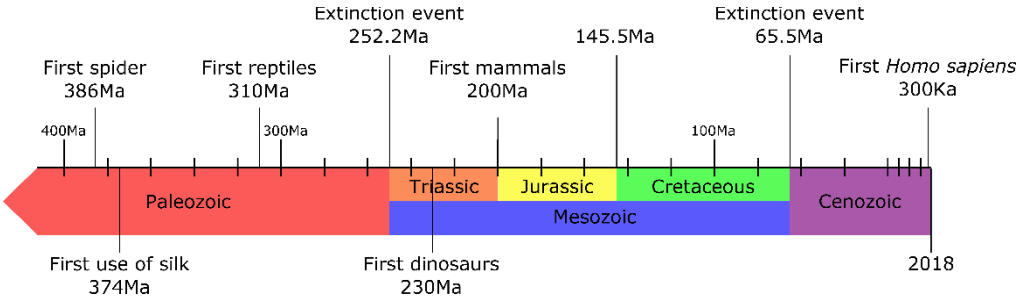


Figure 1.1. Timeline depicting first appearance of spiders (Araneae) and first evidence of silk.

1.1.1. Spider taxonomy

The first spiders occurred some 386 million years ago, with the first fossilised evidence of silk use some 12 million years later (Figure 1.1). All of the Araneae retain two features considered to be significant in spider evolution and as such have been the main focus of spider systematics; silk produced from a pair of spinnerets located at the posterior end of the body, and exteriorly-located male copulatory organs (pedipalps) which have to be charged with sperm prior to copulation. Recent studies of a number of mitochondrial and nuclear marker genes have attempted to clarify existing taxonomic relationships and web-building behaviour, resulting in a number of updates to established relationships (Wheeler et al., 2017, Fernández et al., 2018).

A greatly reduced evolutionary tree can be found in section 4.4, Figure 4.1; the latest phylogenetic data representing 932 species from 115 of the 117 families,

along with a more complete tree, is available in (Wheeler et al., 2017). Some genetic information is known about silk spidroins in some parts of the Araneae, and the level of specialisation at the gene level appears to reflect the level of specialisation to which some spiders apply their silk.

The properties and applications of silk by spiders has evolved considerably since its first use. Silk genes found in the basal Liphistiidae and Mygalomorphae are typically unspecialised and the proteins are produced in one or two large silk glands (Starrett et al., 2012). After the Mygalomorphae come the Araneomorphae, split into the (former) haplogyne and entelegyne. The haplogyne were originally grouped based on gross morphology of the external female genitalia (epigyne); haplogyne females lack the hardened (sclerotized) genital plate found in the entelegyne, the morphology of which is unique to each species and used for species identification (Coddington and Levi, 1991). However, the recent study by Wheeler *et al.* refutes the monophyly of the haplogyne group, instead breaking it into a number of clades basal to the entelegyne, including the Synspermiata and Palpimanoidea.

The entelegyne are separated into the Orbiculariae and retrolateral tibial apophysis (RTA) clade, named for a backwards-facing projection on the tibia of the male pedipalps (Coddington and Levi, 1991). The Orbiculariae, forming approximately 25% of all spider species, are named for the use of orb webs (characterised as a flat web featuring an elastic capture spiral) and were originally considered to be monophyletic based on morphological data, however DNA analyses are beginning to show this is not the case (discussed by Fernández and Hormiga (Fernández et al., 2018, Hormiga and Griswold, 2014)). The orb weavers are believed to have the greatest diversification and specialisation of silks and silk glands, with the greatest number of silk glands (up to seven) each producing a specific type of silk intended to fulfil a specific function (for details see section 1.2 – 1.7). As a result, the majority of silk-related data have been obtained from this group. Little research has been undertaken into silks of the RTA clade and web structures are less complex than those of the Orbiculariae, but these spiders have been found adapted to environments which pose particular challenges, such as the marine spiders *Desis* whose silk is exposed to saltwater (Baehr et al., 2017) and the diving bell spider *Argyroneta aquatica* (Braun, 1931).

1.1.2. *Argyroneta aquatica*



Figure 1.2. Adult female *A. aquatica*.
Out of water, showing the matt grey coating of hydrophobic hairs on the abdomen and sparsely coating the rear two pairs of legs. Adult males are visually similar, with a thinner body, longer legs and enlarged pedipalps. Photo by Radek Šich, www.arachnology.cz.

Argyroneta aquatica (Family: Dictynidae, RTA clade), colloquially known as the underwater or diving bell spider, are the only spider species that has adapted to spend almost its entire life underwater (Catalog, 2019). The latest phylogenetic analysis moves *A. aquatica* from the Cybaeidae to the Dictynidae, a family that is part of the RTA clade (Wheeler et al., 2017). Due to the absence of other aquatic species, and the genetic similarity to other members of this family, it is fair to assume that *A. aquatica* are descended from a terrestrial ancestor.

These spiders are widespread across Europe and Asia up to 62° latitude north and can be found in clean, densely-vegetated areas of fresh water ponds and lakes with low water flow (Masumoto et al., 1998a, Kang et al., 2014, Schütz and Taborsky, 2003, De Bakker et al., 2006, Neumann and Kureck, 2013, Crome, 1951, Seyyar and Demir, 2009). The southernmost boundary of their dispersion is undescribed, although at time of writing no records appear to exist south of the Mediterranean Sea (in the west) or China (east).

A. aquatica is one of the few spider species where the adult male is considered to be larger than the female. Males have a body length of 9-12mm and females 8-15mm, but males have a greater leg span (Schütz and Taborsky, 2003, Schütz and Taborsky, 2005, Bee et al., 2017). All *A. aquatica* are brown in colour with the

exception of the abdomen and underside of the cephalothorax, which are covered in a dark grey, dense layer of hydrophobic hairs, which also extend along two pairs of legs (Figure 1.2). The hairs on the body trap a layer of air whilst underwater, allowing the spider to continue to breathe whilst submersed. The hairs on the rearmost pair of legs enable the legs to be used as a supporting frame, expanding the bubble and allowing an excess of air to be dragged underwater and released into the diving bell (Neumann and Woermann, 2013, Schütz and Taborsky, 2003). The air bubble encapsulating the abdomen of the spider is not sufficient for long-term submersion and individuals regularly surface or return to the diving bell in order to replenish the bubble. This typically involves pausing just beneath the surface of the water or diving bell, turning and pushing the abdomen through the water meniscus into the air and pausing before diving again.



Figure 1.3. Image of an adult *A. aquatica* resting in a diving bell.
Spider in image is kept in a 0.5l plastic cup submersed in a tank of water and has spun a diving bell connected to the sides of the cup. Photo by Ella Deutsch.

Both sexes spin underwater sheet webs, connected to plants and stone underwater by anchor lines, which are used to hold the air which has been carried down from the surface (Figure 1.3, (Seymour and Hetz, 2011)). The resulting structure, known as a diving bell, acts as an air reservoir and functions as a physical gill. The diving bell allows of carbon dioxide produced by the spider to pass through the silk into the surrounding water, and dissolved oxygen to pass in the opposite direction into the air bubble (Seymour and Hetz, 2011, Messner and Adis, 1995, Braun, 1931). Diving bells differ in size depending on the sex of

individuals and their needs (Schutz et al., 2007); females typically build larger bells in which they may rest out of water, whilst males build smaller bells which may only accommodate the abdomen (Schütz and Taborsky, 2003, Masumoto et al., 1998b). These bells function as an underwater venue for mating, resting, eating and as the foundation for an egg sac (Schütz and Taborsky, 2003).

Female *A. aquatica* typically spend the majority of their time sitting inside their diving bell, consuming tiny fish larvae, underwater crustaceans or insects (Nyffeler and Pusey, 2014) that bump into the numerous anchor and walking lines surrounding the diving bell. Males are more mobile, actively hunting for prey and searching for females with which to mate.

Intra-species variation: potential subspeciation of A. aquatica

A study from 2002 by Ono highlighted morphological differences in the male palps between *A. aquatica* specimens found in Japan compared with those found in Europe, concluding that the two represent different *A. aquatica* subspecies (*A. a. japonica* and *A. a. aquatica* respectively) (Ono, 2002). Unlike the transfer of *A. aquatica* from Cybaeidae to Dictynidae, the proposed subspecies have not been officially accepted by the World Spider Catalogue (Catalog, 2019).

Reclassification of *A. aquatica* from Cybaeidae to Dictynidae

Until recently, *A. aquatica* were classified as a member of the Cybaeidae family, but have since been transferred to the Dictynidae after the publication of a phylogenetic study of 115 spider families (Figure 1.4, following page) (Wheeler et al., 2017, Fernández et al., 2018). This does not necessitate reinvestigation of the outcomes described in this thesis, although it will be noted that one of the justifications for using *Cybaeus angustiarum* in section 4 was that the two spiders were of the same family and therefore more closely related than is now apparent.

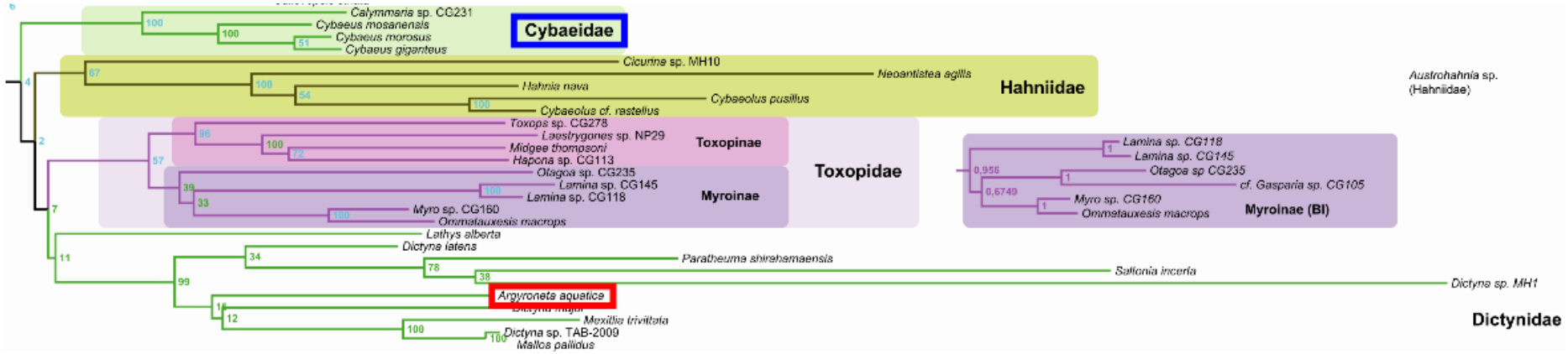


Figure 1.4. Section of updated spider phylogenetic tree.
Families Cybaeidae (top), Dictynidae (bottom) are shown, with *A. aquatica* in their new location in the Dictynidae (red box). Modified from Figure 5, Wheeler et al. 2017 [19].

1.1.3. *Cybaeus angustiarum*



Figure 1.5. Image of an adult female *C. angustiarum.* Showing heavy dark markings on the abdomen and striped legs in the usual pattern. Males have similar markings (Kurka et al., 2015).

Small terrestrial spiders with a body length of 6-10mm, *C. angustiarum* (Family: Cybaeidae, RTA clade) have a golden-coloured abdomen with three dark lines radiating across a medial line running anterior-posterior. The cephalothorax and legs are golden-brown, with dark brown rings around the legs. *C. angustiarum* live exclusively in habitats with areas close to 100% humidity including under stones and fallen wood, in decaying tree trunks and dense vegetation. Their reach extends across mountainous areas in the northernmost parts of Europe to Azerbaijan (Kurka et al., 2015).

1.1.4. *Pholcus phalangioides*

Known as the cellar spider or daddy long-legs, *P. phalangioides* (Family: Pholicdae, Synspermiata clade) have a worldwide distribution and are often found hanging from fine webs spun in residential properties, in which they will vibrate aggressively when disturbed (Foelix, 2011, Jackson and Brassington, 1987, Jackson, 1990, Huber, 2011).

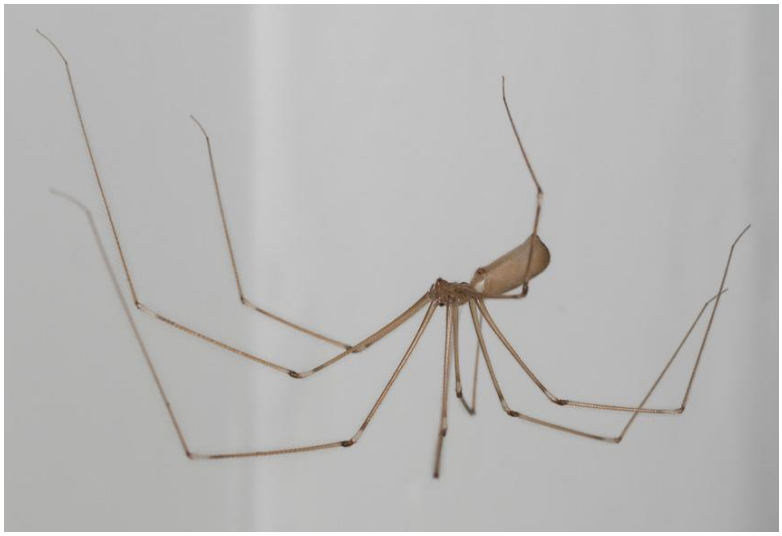


Figure 1.6. Image of an adult female *P. phalangioides.*
Spider is demonstrating its typical pose, suspended from a web.
Photo taken from (Seppo Koponen, 2015).

Pholcids have six eyes and are immediately recognisable by the long length of their legs in relation to their body (Figure 1.6). *P. phalangioides* can be distinguished from a similar-looking Pholcid, *Holocnemus pluchei,* by the colouring of the abdomen; *P. phalangioides* have a skull-like marking on the dorsal side of the abdomen whereas *H. pluchei* have a dark, jagged marking outlined in white. Males have a body length of 7-10mm and females 8-10mm, and the legspan can reach 7cm (Bee et al., 2017). *P. phalangioides* mate throughout the year and females carry their egg sac in their chelicerae (Foelix, 2011). *P. phalangioides* will consume other arachnids, including jumping spiders, redback spiders (*Latrodectus hasseltii*) and the physically larger house spider *Tegenaria domestica* (UK)/*Eratigena atrica* (rest of world), which they capture by wrapping with silk using their long legs before injecting their prey with venom (Jackson, 1990, Jackson and Brassington, 1987).

1.1.5. Rationale for choice of species for study

Whilst there is a substantial amount of silk-related genetic research available, the majority of it focuses on a small number of families within the Mygalomorphae and Orbiculariae. This leaves significant knowledge gaps within the Araneae both between and outside of these groups – little is known about the evolutionary steps taken between the Mygalomorphae and the Orbiculariae, and the RTA clade which

is a sister group to the Orbiculariae are almost completely unstudied, despite containing species with interesting and varied uses of silk.

Argyroneta aquatica (Clerck, 1757) use their silk to construct a web in which the spider resides and performs a number of actions, from feeding to mating and storing eggs (Schutz et al., 2007, Seymour and Hetz, 2011, Schütz and Taborsky, 2003, Schütz and Taborsky, 2011, Schütz and Taborsky, 2005). As descried above, it is the only known species capable of spinning silk whilst submersed in water which has an unusual and interesting property of permitting gas diffusion, apparently at a rate at which the spider may remain submersed for extended periods of time (Seymour and Hetz, 2011). *Cybaeus angustiarum* is a terrestrial species, closely related to *A. aquatica*, used here as a closely-related terrestrial species for comparison.

Pholcus phalangioides is a member of the Synspermiata, a recognisable and understudied group, knowledge of which may provide information of the evolutionary process by which undifferentiated silks of the Mygalomorphae evolved to the highly specialised silks of the Orbiculariae. *P. phalangioides* is also the only species with a Y chromosome (Král et al., 2006) and knowledge of the silk genes would demonstrate whether large changes can occur within the spider genome without affecting the universal nature of the silks.

1.2. Evolution of spider silk

Silk spidroins (spider fibroins) are of great interest scientifically due to the range of biophysical properties demonstrated in a variety of uses and functions, including silk-based glues (Wolff et al., 2015) and silken fibres with strength and toughness outperforming that of a number of man-made materials (Agnarsson et al., 2010). These may be used by the spider in combination with as-yet unidentified protein-based materials which are also produced in the same glands as the silk proteins (Seymour and Hetz, 2011). Arachnologists worldwide have been studying spiders for over 300 years, with studies covering their ability to travel huge distances, the use of individual silks and the origin of silk genes.

Found in fossils dating 374Ma, the earliest arachnid known to spin silk is *Attercopus fimbriunguis*, with fibres emerging from spigots mounted on ventral abdominal plates (Selden et al., 2008). The precursors to spinnerets, as arachnids evolved to better use this novel resource, the modified ventral plates evolved into

the prominent spinnerets we see today, adapted in multiple ways depending on the needs of individual species. Originally possessing 4 pairs, the majority of today's spiders retain three pairs of spinnerets (Figure 1.7) whilst a few have just two pairs. The four pairs of spinnerets are named as follows: the posterior lateral, posterior medial, anterior lateral and anterior medial spinnerets. In species with three pairs of spinnerets the "absent" fourth pair are the anterior medial spinnerets; these are retained by some Mesothelae (Glatz, 1972, Glatz, 1973) in (Foelix, 2011). Each of these spinnerets is mounted with multiple spigots, which are connected to one of up to seven glands found in the abdomen of the spider (see section 1.5, Figure 1.8).



Figure 1.7. Images of spider spinnerets and spigots.
Left: Photographic image of the spinnerets of the Australian garden orb weaver *Eriophora transmarina* extruding silk fibres. Visible are the posterior lateral spinnerets (top), anterior lateral spinnerets (bottom) and anal tubercle (top centre). The posterior medial spinnerets can be found between the two, larger, pairs of spinnerets. (Credit: Wikimedia user Jason7825, licensed under the Creative Commons Share Alike 4.0 International license Wikimedia commons.wikimedia.org).

Right: Post-production coloured SEM image of *Gasteracantha* sp. showing the spigots (green and gold) extruding silk fibres (purple) (right) (Credit: Dennis Kunkel Microscopy/Science Photo Library sciencephoto.com).

The equipment to spin silk has evolved in tandem with the silken material itself. Spider silk can be classed as belonging to one of several different types, of which seven have been characterised to a greater or lesser extent (see section 1.5). These seven types of silk are considered to be the most derived silks, because they have been specified for a unique collection of properties and used together to build a

range of structures from a simple sheet to a complex orb web. The most extensively studied silk, major ampullate silk, is well known for its strength and size, as well as being the most readily collected by those wanting to use it or who are interested in studying its properties.

1.3. Applications of silk by the Araneae

Spiders are perhaps best known as creators of orb webs (2D structure supporting a capture spiral) and dusty cobwebs. Different species have different functions for silk, although most spiders use specific silks for particular applications, e.g. egg cases, sperm webs and dragline silk. Egg sacks or cases are exclusively produced by female spiders and include delicate yet sturdy baubles hung in caves featuring 100% humidity (*Hickmania troglodytes*) (Doran et al., 1999), fluffy hemispheres secured in a sheltered location (*Araneus diadematus*) and simple bundle of eggs wrapped in a few strands of silk and carried (*Pholcus phalangioides*) (Foelix, 2011). Most egg sacks are built with the intention of protecting the eggs whilst the embryos develop. They are typically formed of multiple layers of silk and may be simple, multi-layered or complex and come in a range of colours (*Ero furcata* egg sacs appear reddish brown (Lissner and Scharff); *A. diadematus* are gold and *Tetragnatha nigrita* spin grey silk (Lissner and Scharff)). Sperm webs are solely used by male spiders; sexually mature males build a small sheet of silk that they use as an intermediary to transfer sperm from their genital opening to their palps prior to mating with a female (Foelix, 2011). Lastly, draglines are a commonly observed use of silk; smaller spiders such as jumping spiders use this as a safety line, producing a line of major ampullate silk that is occasionally glued to their surroundings in to arrest an accidental fall, whereas larger species such as those found in the Mygalomorphae leave a trail of silk along their path that is not strong enough to support their weight but can be used to build or cover up a burrow or as trip lines to detect prey (Coyle, 1986, Coddington and Levi, 1991).

Whilst some silks are used singly for specific purposes, combinations of silk are adapted by single or small numbers of species. One prominent use of silk is to assist with the capture of prey; overlooking the well-known orb web with its sticky capture spiral, examples include: bolas spiders (*Mastophora* sp. amongst others), which swing a length of silk with a droplet of glue and pheromones on the free end that attaches to flying prey that can then be reeled in and eaten (Yeargan, 1994, Eberhard, 1977). *Pasilobilus* sp. build a simple triangular web across which a series of loose threads are spun, coated in sticky glue. When an insect collides

with these threads they break, leaving the insect attached but dangling beneath the web (Robinson and Robinson, 1975). *Deinopis* sp. spin a miniature modified orb web between their front legs, which they use to scoop up unsuspecting prey which are trapped not by a sticky glue but by becoming entangled in fluffy cribellate silk (Coddington Jonathan, 1986). Trap door spiders (including the family *Ctenizidae*) build a silk-lined burrow, with a solid silken "lid", from which a number of silk trip lines are extended; the spider waits in its burrow for unsuspecting prey to brush against these lines before attacking (Coddington and Sobrevila, 1987).

Other common applications of silk include the use of silk to create a cocoon or shelter in which the spider may hide from the elements or remain unseen. As mentioned above, a number of species build silken tunnels (such as *Agelena labyrinthica*) or burrows that are lined with silk (including spiders from the *Ctenizidae* and *Antrodiaetidae* families). Intertidal spiders (*Desis* sp.) use silk to seal themselves into an air chamber, keeping the spiders dry in an air pocket as their burrows are submerged during the high tides. Only the underwater spider, *Argyroneta aquatica*, is capable of spinning silk whilst submersed. These spiders live in fresh water, building a sheet web that is used to capture a bubble of air, subsequently used as a temporary underwater retreat (Masumoto et al., 1998a, Seymour and Hetz, 2011) (see sections 1.1.2 and 5).

1.4. Non-Araneae silk producers

At least 23 groups of insects (including the *Hymenoptera*, *Trichoptera*, *Hemiptera* and *Lepidoptera*), a small number of arachnids (including *Pseudoscorpiones* and *Acari*), *Crustacea*, and *Myriapoda* (Hazan et al., 1975, Del-Claro and Tizo-Pedroso, 2009, Craig, 2003, Kronenberger et al., 2012, Yonemura et al., 2006, Chang et al., 2005, Zhou et al., 2000, Sutherland et al., 2007) are known to make forms of silk. Most use silk in order to build a protective structure of some form, typically at a single point in their lifecycle (e.g. pupal case), whilst others utilise silk as support or protection for eggs (Sutherland et al., 2010). Only one family of insects are able to spin silk throughout their lifecycle (Orthoptera: Gryllacrididae), which they use to line burrows or fold together leaves to form cocoons in which they reside (Morton and Rentz, 1983), although webspinners (Order: Embidiina) produce a form of silk during their larval and adult stages.

Perhaps the most well-known silk-producing animal is the silk worm, the larval form of the moth *Bombyx mori* (Order: Lepidoptera), which spins a cocoon in which it undergoes the metamorphic transition to its adult stage. Identified and cultivated by the Chinese some 5,000 years ago, silk cloth woven from fibres unravelled from the cocoons made by these larvae have become a widely traded commodity.

The similarities and differences between spider and silkworm silk and silk production have been compared in detail (Andersson et al., 2016). Spiders produce silks throughout their lifetime; silkworms only produce silk during their fifth instar, prior to metamorphosis. Where spiders have several types of gland and produce silks with different properties, silk worms produce a single type of silk, formed of two fibres bound together with sericin, from a pair of modified salivary glands (Andersson et al., 2016, Asakura et al., 2007). These fibres are formed of a number of different proteins including "heavy" and "light" fibroins linked by disulphide bonds and fibroin p25 (bound by hydrophobic interactions).

Whilst spider silk (e.g. major ampullate silk) and silkworm silk share an overall structure (two terminal domains and a highly repetitive region consisting significantly of glycine and alanine amino acid residues), the amino acid sequences of the terminal domains and the repetitive regions differ and both form different primary and secondary structures (reviewed in (Andersson et al., 2016)). Perhaps most importantly, spider major ampullate silk has a much greater tensile strength.

1.5. Spider silk protein macrostructure

Silk spidroins are proteinaceous fibres, either used exclusively as a dry fibre or in combination with other materials (such as glycoproteins and inorganic salts) to form sticky glues or silk cements (Grawe et al., 2014, Wolff et al., 2015, Vollrath and Edmonds, 1989, Vollrath et al., 1990, Kovoor and Zylberberg, 1980). Spider silks are typically categorised as one of seven types (Figure 1.3), as an unspecialised or uncategorised silk (Strickland et al., 2018, Starrett et al., 2012) or cribellate silk. The categories of silk described are named after the gland from which the spidroins are produced, or in the case of cribellate silk, after the post-production modifying apparatus (the cribellum).

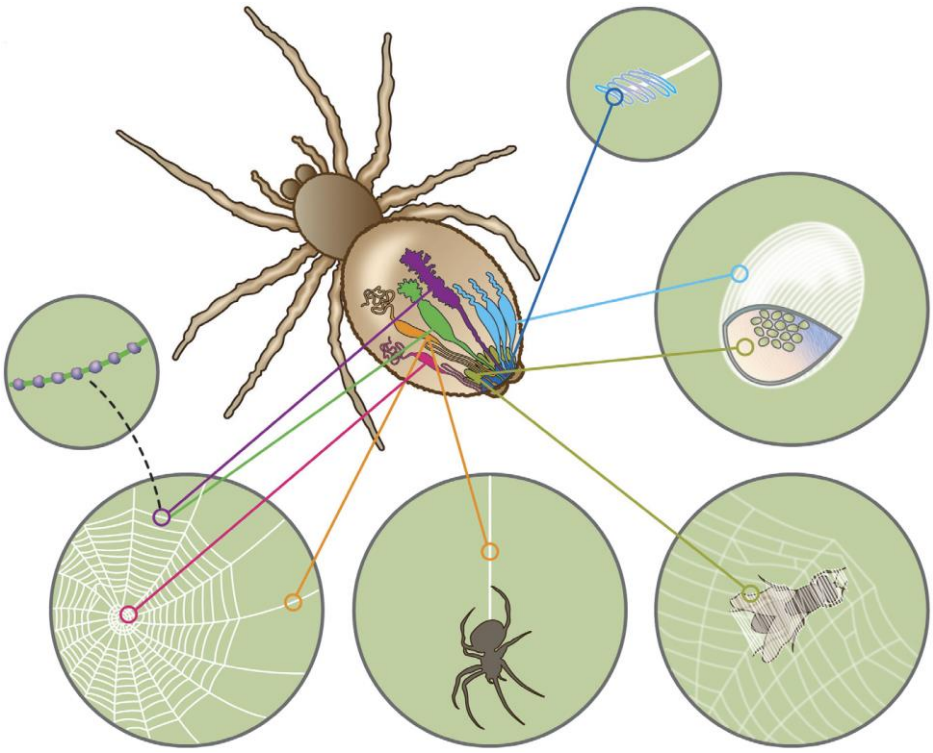


Figure 1.8. Schematic overview of the different products of the silk glands of a female orb weaver.
Clockwise from top right: piriform glands (attachment discs), tubuliform (also known as cylindriform) glands (outer egg case), aciniform glands (inner egg case and prey-wrapping), major ampullate glands (dragline and radial/anchor lines), minor ampullate (scaffolding spiral), flagelliform glands (capture spiral), aggregate glands (sticky coating on capture spiral). Note: males do not possess tubuliform silk glands. Modified from Figure 1, Rising & Johanssen (2015) (Rising and Johansson, 2015).

Genetic data suggests that a genome duplication event very early in the history of the Araneae (estimated around 380MA and concurrent with the division of the Mygalomorphae and the Araneomorphae) has driven the evolution of multiple silk types (Clarke et al., 2015). Each of the characterised silks has evolved different biophysical properties in order to fulfil the purpose for which they are required (Yarger et al., 2018). Major ampullate (dragline) silk has a high tensile strength useful for supportive threads (Agnarsson et al., 2010); flagelliform silk is highly elastic and spun into a shock-absorbing capture spiral in Araneoidea webs (Hayashi and Lewis, 1998); piriform silk is used in combination with piriform glue to form cement-like attachment discs (Wolff et al., 2015); aciniform is the toughest silk fibre, used as a wrapping material to secure prey (Hayashi et al., 2004);

aggregate glue is composed of nanofibrils and glue, which attracts water from the air, forming droplets which bead on flagelliform fibres and act as windlasses to maintain the tension of the capture spiral (Vollrath et al., 1990, Vollrath and Edmonds, 1989).

Within each silk fibre, spidroins orientate along the length of the fibre, forming β-sheets and amorphous regions (van Beek et al., 2002, Hayashi et al., 1999). Each of the different spidroins demonstrates different biophysical properties, attributed to particular molecular structures formed by combinations of amino acids interacting as individual spidroins fold together to form the fibre. For example, tensile strength is attributed to β-sheet formations containing a number of poly-alanine repeats, whereas elasticity is conferred by glycine-proline-glycine-X-X (where X is a small subset of amino acids) (Ayoub et al., 2007).

1.6. Genetic structure of spider silk genes



Figure 1.9. Overview of the different regions of a silk gene/protein.
Non-repetitive amino-terminal (green) and carboxyl-terminal (red) domains encompass a region formed of consecutive repetitive motifs. The secondary structure of the terminal domains is conserved between silk types/species, whereas the repetitive region varies between silk types.

Spider silks all likely share a common evolutionary origin. Specialisation of silk genes begins early in spider evolution, where it is thought gene duplication, recombination and diversification initiated the formation of the diverse collection existing today (Clarke et al., 2015, Starrett et al., 2012, Ayoub et al., 2013). Spider silk genes are modular in form, consisting of three regions; a non-repetitive amino-terminal (N-terminal) domain (around 130 amino acid residues), a carboxyl-terminal (C-terminal) domain (approx. 110 amino acid residues) and an extended repetitive region (Figure 1.4). The terminal domains are conserved evolutionarily (Challis et al., 2006, Strickland et al., 2018, Collin et al., 2011) and are essential for the correct formation of both natural and synthetic silk fibres (Andersson et al., 2017, Stark et al., 2007). The repetitive region dictates the biophysical properties of the final silk product (Beckwitt et al., 1998). The

repetitive regions form the greatest diversity between silk types, with particular motifs conferring specific attributes. Major ampullate spidroins (MaSp), in particular, contain a number of short motifs repeated inconsistently which result in the formation of beta-sheets in the secondary structure, conferring toughness and irreversible extensibility [37, 39, 40]. There are two variants of MaSp spidroins; whilst both MaSp1 and MaSp2 exhibit poly-A, poly-GA and GGX motifs (where X is a subset of amino acids) MaSp1 and MaSp2 can be distinguished by the presence or absence of proline residues as only MaSp2 contains the motifs GPG and GPGX [41, 166]. Other silks, such as tubuliform spidroins and the aciniform spidroins found in Araneus trifasciata, have much longer repetitive blocks of up to 200 amino acids [38].

The repetitive regions form the greatest diversity between silk types, with particular motifs conferring specific attributes. Major ampullate spidroins (MaSp), in particular, contain a number of short motifs repeated inconsistently, which result in the formation of beta-sheets in the secondary structure, conferring toughness and irreversible extensibility (van Beek et al., 2002, Hayashi and Lewis, 1998, Hayashi et al., 1999, Xu et al., 2015). Other silks, such as tubuliform spidroins and the aciniform spidroins found in *Araneus trifasciata*, have much longer repetitive blocks of up to 200 amino acids (Hayashi et al., 2004). Piriform spidroins are likely the most diverged of the silks; both glycine and proline residues are absent from the repeating region (Blasingame et al., 2009), although the amino acid composition may not be indicative of the properties of the silk; the secondary structure is more indicative of the final properties of the silk (Blamires et al., 2016).

Due to the repetitive nature of the repeating region, silk genes have traditionally been difficult to sequence and assemble in full. However, sequencing technologies and assembly programs are improving; the first full-length gene sequence was published in 2007 (Ayoub et al., 2007) and more recently, spider genomes have become available (Sanggaard et al., 2014, Schwager et al., 2017), with the genome of *Nephila clavipes* being the first to be >99% annotated (Babb et al., 2017).

1.7. Silk glands

The vast majority of spiders have multiple silk glands, with members of the Orbiculariae known to contain seven individual, highly specialised glands, each of which produces a single, highly specialised spidroin. The simplest silk glands can

be found in the most basal species, as found in the Mygalomorphae (also known as the Orthognatha), where only one or two simple glands may be found (Palmer Jacqueline, 1985, Palmer et al., 1982).

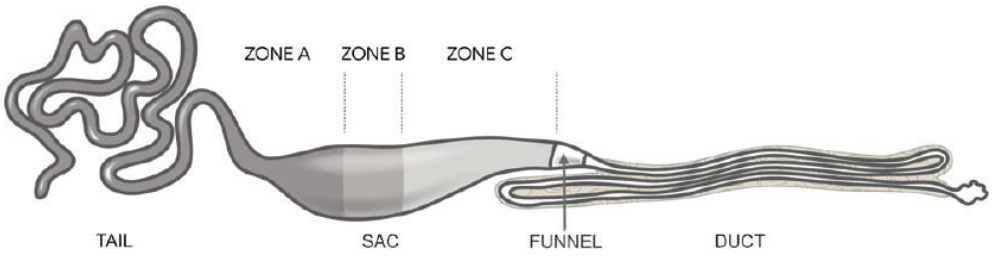


Figure 1.10. Diagram showing the different regions of a silk gland.
Spidroins are produced by epithelial cells in Zones A and B, whilst carbonic anhydrase is produced in Zone C and along the duct. Taken from Figure 3, Andersson et al. (2013) (Andersson et al., 2013).

Silk glands are, with two known exceptions (the venom glands of the spitting spider *Scytodes thorascica* (Correa-Garhwal and Garb, 2014) and *Nephila clavipes* (Babb et al., 2017)), the only organs in which spidroins are produced. Silk glands come in a range of sizes and shapes; typically the major ampullate gland is the largest and is only found once; tubuliform (or cylindriform) glands are long and cylindrical in shape, whereas aciniform glands are small and can be found in clusters close to the spinnerets. Overall, however, the glands are simple in form and composed of the collecting sack, preceded by single or bifurcated production tubules that lead to the extrusion duct and ultimately terminate in the spinnerets. The walls of the glands are composed of a single layer of three types of columnar epithelial cell: two (in zones A and B) produce silk fibroin, whilst the third (zone C) is involved in the physiological regulation of silk production (Kovoor and Zylberberg, 1980, Vollrath and Knight, 2001, Andersson et al., 2013, Kovoor, 1987) (Figure 1.10).

1.8. Spider silk production

1.8.1. Extrusion of silk dope

Spidroins (silk proteins) are produced in the uppermost part of the silk gland (Figure 1.11, top), in columnar cells containing large numbers of secretory vesicles (Bell and Peakall, 1969, Plazaola and C. Candelas, 1991).

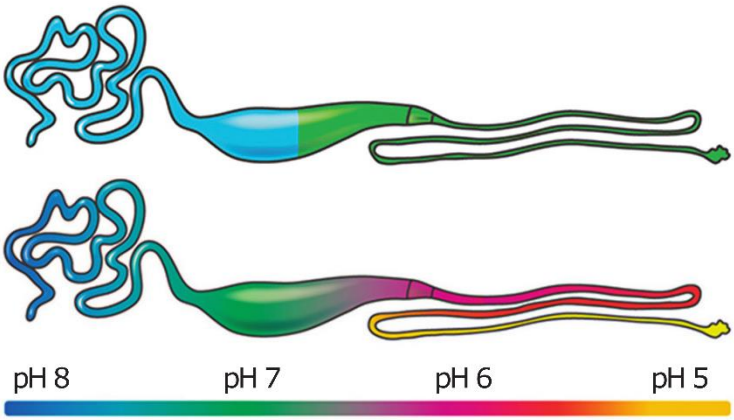


Figure 1.11. Overview of the silk spinning process.
(Top) Spidroins are produced at the end of the silk gland (blue) and stored in the central collecting sack. Carbonic anhydrase is produced in the area coloured green, in order to maintain the pH gradient (bottom). Modified from Figure 2, Rising & Johansson (2015) (Rising and Johansson, 2015).

The protein is stored as a liquid dope in the sack portion of the gland, where the spidroins form either individual micelles in solution (Vollrath and Knight, 1999, Heim et al., 2009) or a liquid crystalline feedstock (Vollrath and Knight, 2001). The spinning duct utilises carbonic anhydrase to maintain a pH-gradient that begins at pH7 in the gland and decreases to pH5.7 by the end of the duct (Andersson et al., 2014) (Figure 1.6, bottom); this is important for the transition of liquid silk dope to a solid fibre. When silk is required, the silk dope travels down the collecting duct, where in addition to the change in pH it experiences a combination of sheer and stress forces, dehydration, and removal of ions (Vollrath et al., 1998, Tillinghast et al., 1984). At a fixed point in the duct, the liquid dope converts to a solid fibre, although the diameter may not be entirely fixed until the fibre is finally extruded past valves just inside the spigots (Wilson, 1962b). Final dehydration occurs as the fibre is placed under tension by the back legs of the spider.

1.8.2. Function of the N- and C-terminal domains during silk fibre formation

A number of studies have examined the structures and roles of the terminal domains in silk fibre formation. In short, the C-terminal domain drives spontaneous fibre formation (Stark et al., 2007) whereas the N-terminal domain acts as a molecular brake, restricting fibre formation by preventing the spidroins from agglutinating until the spidroin reaches a point in the duct where the pH

gradient falls below pH5.5 (Kronqvist et al., 2014, Askarieh et al., 2010). Fibre formation is assisted by reversible pre-alignment of the N-terminal domains between pH6.5 – 5.5 (Kronqvist et al., 2014). Additionally, the C-terminal domain is a highly stable dimer above pH6.5 (Andersson et al., 2014, Gao et al., 2013), however the decreasing pH of the silk duct disrupts a salt bridge within the domain (Andersson et al., 2014), destabilising the structure and facilitating the unfolding of the structure and formation of β-sheet fibrils (Rising and Johansson, 2015, Martel et al., 2008). This also likely induces the formation of β-sheets in the repetitive region of the spidroins (Andersson et al., 2014, Rising and Johansson, 2015).

1.9. Human uses for spider silk

1.9.1 Traditional applications and folk medicine

Spider silk has been used in traditional medicine and practise for hundreds of years (Bon, 1710). Examples include applying freshly-spun webs to open wounds to assist with healing; silk appears to bind to the flesh and assist healing (Vollrath et al., 2002), perhaps by acting as an antimicrobial agent (Wright and Goodacre, 2012) or by acting as a source of vitamin K (Heimer, 1988).

Individual threads of silk have been used in instruments of measurement such as the transit instrument at the Greenwich Observatory, London, where a single strand of *Araneus diadematus* silk is used to mark 0° longitude, preferred over metallic or keratin strands due to silk's exceptional thinness and consistency in diameter and inertia when exposed to changing temperature (Morgan, 2016). Dragline silk has also been used to form the crosshairs in traditional weapons and optical instruments such as the reticulum in micrometres (Morgan, 2016, Bedini, 2005, Turner, 1992). More recently, studies using spider webs as a natural collector of airborne particles examine how long viral particles can remain contagious in air (Usborne, 2016).

Birds are also known to collect and utilise spider silk in their nest building; for example the long-tailed tit (*Aegithalos caudatus*) and some species of hummingbird use silk as a binding agent or for decoration or camouflage during the construction of their nests (Morgan, 2016, Hansell, 1996).

1.9.2. Modern applications and modification of native spider silk

In the field of biomaterials research, a number of studies have investigated the biocompatibility aspect of spider silk with thoughts leaning towards future biomedical applications. In regenerative medicine, tissues are replaced by a degenerative material that acts as a scaffold, being overgrown by the surrounding cells. In one study, the major ampullate (dragline) silk of *Nephila* sp. was used as the basis of a simple cell scaffold for nerve regeneration; human Schwann cells would bind to the silk thread in 15 minutes and after 24 hours had proliferated and formed multiple layers encompassing the fibre completely (Allmeling et al., 2006). Other studies have shown similar potential, including in cartilage (Gellynck et al., 2006), urology (Steins et al., 2015) and skin regeneration (Wendt et al., 2011).

Combining silk and materials with interesting properties is uncommon and a few studies exist specifically with respect to spider silk, although investigations into the result of combining silks with unusual compounds prove interesting. One study exposes spiders to dispersions of carbon nanomaterials, including graphene, resulting in fibres with significant increases in tensile strength and toughness (Lepore et al., 2017), whereas others apply the same structures directly to dragline silk, although without the same positive effects (Steven et al., 2013). Other groups use atomic layer deposition, in which a sample is heated *en vacuo* to high temperatures before atomised compounds are introduced, which form a layer on the sample as it cools. This has been used to incorporate metal ions into spider silk in order to increase the toughness of dragline silk (Lee et al., 2009) and in one other case samples of silk have been immersed in chloroauric acid, which allows the deposition of gold particles, resulting in the formation of an electrically conductive material which shrinks or expands relative to humidity (Singh et al., 2007).

1.9.3. Synthetic silk

As described earlier, spider silk is a biological material with a diverse range of physical properties. Due to difficulties encountered in farming spiders for their silks (with the exception of the social spiders, all spiders are cannibalistic and will often consume smaller or weaker individuals), spider silks have not been a widely collected material, unlike the readily harvested cocoons of the silk worm (*Bombyx mori*). The only known mass-collection of spider silk was an eight-year project to collect golden orb weaver (*Nephila madagascariensis*) silk, leading to the

production of an elaborately embroidered cape and shawl later displayed at the Victoria & Albert museum in London (UK) in 2012 (Peers and Godley, 2011).

Synthetic silk research has begun to demonstrate potential ways of mass-producing synthetic silk. Early experiments with transgenic plants such as potato and tobacco (Scheller et al., 2001) and bovine epithelial cells (Lazaris et al., 2002) have not made significant progress, although silk proteins have been manufactured and purified. Other modified organisms include transgenic silkworms (Teulé et al., 2012), yeast (Arthur, 2017) and goats, in which recombinant silk proteins are made in the mammary glands and can be collected in the milk then extracted and purified (Karatzas et al., 2007).

Recent developments utilising bacterial cultures have yielded promising results. In 2007, the first synthetic "mini spidroins" to be produced signalled a new wave of synthetic silk research (Stark et al., 2007); in 2018 another group further modified this mini spidroin to improve yields, indicating synthetic spider silk may yet become commercially viable (Andersson et al., 2017).

Meanwhile, other milestones were being achieved. The first report of the production of a full-sized spidroin was reported (Xia et al., 2010) whilst high-resolution structures of the amine and carboxyl-terminal (N- and C-terminal) regions of major ampullate silk were described (Hagn et al., 2010, Askarieh et al., 2010). These were followed by detailed information of how these domains interacted (Andersson et al., 2014), allowing a better understanding of how silk proteins folded and combined to form a fibre. Additionally, researchers generated a modified mini-spidroin by substituting the amino acid residue methionine for the artificial amino acid L-azidohomoalanine. Subsequent copper catalysed azide-alkyne cycloaddition (or "click-chemistry") allows for the addition of a number of compounds either pre-or post-fibre formation; the example in the study utilised fluorophores and an antibiotic, but the authors stress that these may be substituted for a number of antibiotics, antifungals or growth factors, used singly or combined into multifunctional fibres (Harvey et al., 2017).

1.10. A summary of the aims of, and work contained in, this thesis

This thesis aims to further scientific knowledge of spider silks, specifically of the species *Argyroneta aquatica* but also the evolutionary history of spider silk and

the relationship between its genetic sequence, structure and function. Additionally, some investigation into the effect of combining native spider silk with compounds displaying interesting physical properties is also discussed. Each study is a discrete piece of work and presented with its own introduction and discussion, however all chapters are relevant to ongoing research into synthetic silk and future applications of spider silk as a biomaterial. A complete bibliography can be found towards the end, in section 8.

The first chapter (section 3) assesses the effectiveness of RNA extraction and subsequent sequencing of the whole abdomen of a spider against the dissected, combined, silk glands of the same species, with respect to obtaining the greatest data regarding silk genes expressed by that species.

The second chapter (section 4) considers the whole of the genetic work undertaken here, examining evolutionary and structural conservation within spidroin N- and C-terminal domains. In particular, it builds on work by Challis, Goodacre and Hewett (Challis et al., 2006), keeping in mind secondary structure features described by Hagn (Hagn et al., 2010) and observed by others. Utilising the greatly-expanded database of silk sequences on GenBank, along with novel silk sequences described in the previous chapter, this study examines whether a salt bridge, a pH-sensitive bond between two residues within the C-terminus, is evolutionarily conserved, and investigates whether similar features may be present/conserved in the N-terminal domain. The phylogenetic relationships between terminal domains of different spider species are also examined.

Section 5 delves into the morphology of the silk and spinning apparatus of *Argyroneta aquatica*, the only known truly-aquatic spider species. The silken structure ("diving bell") is the main focal point for this chapter: whilst the expressed spidroins were sequenced and described in sections three and four, little is known about the adaptations this spider has made in order to enable the underwater production of silk and determine whether it is truly unique in its ability. This chapter examines the morphology of the silk glands, spinnerets and spigots, and the diving bell itself, in order to expand our understanding of the unique behaviour *A. aquatica* demonstrate.

The final chapter (section 6) investigates and attempts to replicate work by Lepore (Lepore et al., 2015, Lepore et al., 2017), who exposed spiders to three carbon nanomaterials (CNM) and presented results showing a large increase in

tensile strength. Here, we exposed both spiders and their pristine dragline silk to the same CNM, included a fourth CNM and examine the Raman spectra and tensile properties of subsequently produced, and coated silk, respectively.

2. General Methodology

2.1. Spider collection and husbandry

2.1.1. Collection of specimens

Argyroneta aquatica were collected from the Máchovo jezero Lake, Doksy, Czechia in October 2014, June 2015 and August 2016. Spiders were found around the edges of small ponds, which were separated from the main lake by reeds. Fine nets could be used to "scoop" submersed plant matter and debris from around the base of reeds near to the shoreline; individuals would be caught in the debris and extract themselves, allowing easy collection.

Cybaeus angustiarum were collected from the Krkonoše region of Czechia in June 2015 and Autumn 2016. Spiders were found in pine forest, underneath logs and large stones, and were difficult to catch due to their speed and lack of a dragline (individuals would drop to the ground from heights up to a metre without attaching a dragline).

Pholcus phalangioides were collected within the Crop Research Institute (Drnovská 507, CZ-16106 Prague 6 - Ruzyně, Czechia), Life Sciences building, University Park, University of Nottingham (UK), and from a residential property (Beeston, Nottinghamshire, UK) in June and August 2015 and September 2016.

Araneus diadematus were collected from University Park campus (University of Nottingham, UK) and surrounding area in August 2015, September 2016 and June 2018. Priority was given to collecting similarly-sized adult female spiders, where available, to maintain consistency in size and silk diameter. Collected *A. diadematus* were typically fully grown, gravid (egg-bound), adult females.

Steatoda nobilis were collected from two residential properties (Beeston, Nottinghamshire, UK; Bedford, Bedfordshire, UK) in August 2015.

2.1.2. Storage and transportation

For larger *A. aquatica* and *C. angustiarum*, 50ml falcon tubes containing a large piece of wet paper or moss (*Sphagnum* sp.) with a small hole in the lid were used for storage and transport. Smaller spiders could be kept in 1.5ml Eppendorf tubes with pierced caps, collectively stored in a none air-tight plastic container containing wet paper. *P. phalangioides*, *A. diadematus* and *S. nobilis* were collected in a variety of small plastic containers holding a small piece of wet paper, before

being immediately transported to the lab and transferred to more suitable accommodation.

A. aquatica and *C. angustiarum* could be stored in their original container at 4°C for up to two months without food; spiders would be transferred to pre-chilled containers and allowed to slowly acclimatise to room temperature. Other spiders were transported to the lab immediately after capture.

2.1.3. Husbandry techniques

Argyroneta aquatica

A. aquatica were housed individually in a variety of aquaria for up to two years, including: 0.5l plastic cups, which were pierced and placed in a large tank with a water depth of around 6cm; 3l plastic tanks measuring 25cm x 15cm x 20cm; or 60l glass tanks measuring 60cm x 40 cm x 35cm. Aquaria were initially covered with pierced plastic film held in place with elastic bands, or thin paper held by tape, until it was established that *A. aquatica* are incapable of climbing the plastic/glass walls. This practice was discontinued due to the rapid formation of black mould within the cup and on the film/paper (likely *Aspergillus niger*).

All aquaria were maintained with a minimum water depth of 6cm and with a minimum gap of 5cm between the water surface and lip of the containing structure (aquaria or plastic cup). Each tank was also equipped with a minimum of one air stone (not required for aeration, but instead to encourage a small amount of water movement and prevent a suffocating protein film forming on the surface of the water).

Cups and small tanks contained a polystyrene float embedded with two or three straws (as in De Bakker et al., 2006) and occasionally a small piece of moss (*Sphagnum* sp.); larger tanks were furnished with gravel substrate and two plastic plants. Occasionally empty water snail shells were offered and utilised as hiding spaces (typically anchored to the container with silk and containing a diving bell); spiderlings would also build miniature diving bells inside the straws. Tank type did not appear to influence survival; wild-caught and lab-born spiders could be kept for up to two years before perishing. In order to avoid cannibalism, all spiders were kept in individual containers except when newly hatched spiderlings or paired for mating experiments. A dry location (either inside an enlarged diving bell or on top of a polystyrene float) was essential to a successful moult.

All aquaria were kept at ambient room temperature (18 – 21˚C). Up to 25% water changes were performed weekly; tank water was taken from the mains water supply, pre-treated with Tetra Aqua safe according to manufacturer's instructions, and aerated for 24 hours prior to use. All *A. aquatica* were fed two to three times weekly with a small number of brine shrimp (*Artemia* sp.*),* rinsed in fresh water as these spiders are apparently sensitive to salt water. A population of *Daphnia* was also maintained in each tank as an alternate food source and algae control. Occasionally this diet was supplemented by live micro-crickets (< 5cm), bloodworm, glassworm and the larvae of opportunistic mosquitoes.

Cybaeus angustiarum

C. angustiarum were kept individually in small enclosures, measuring approximately 7cm x 7cm x 9cm, lined with fresh moss or wet cotton wool and containing a small piece of bark. Substrate was replaced every 6 weeks or as soon as mould appeared. Containers were kept at room temperature (18 – 21˚C) and the substrate lightly sprayed three times weekly to maintain 100% humidity. Each spider was fed a small number of laboratory-bred fruit flies (*Drosophila* sp.) two to three times weekly. Enclosures were cleaned as necessary.

Other terrestrial spiders

P. phalangioides, *A. diadematus* and *S. nobilis* spiders were kept in small individual enclosures, measuring approximately 7cm x 7cm x 9cm and containing a single straw placed diagonally inside the pot. Some enclosures also contained pieces of card folded into a tube, providing a hiding place. Ventilation was provided by cutting a hole in the lids of each pot, stoppered with cotton wool. Containers were kept at room temperature (18 – 21°C). Spiders were typically fed a small number of live fruit flies or micro crickets and watered with a small volume of distilled water sprayed against the side of the enclosure two to three times weekly. Enclosures were cleaned as necessary.

2.2. RNA sequencing

2.2.1. Dissection, storage and transportation of silk glands

Adult female specimens intended for sequencing were killed and the silk glands immediately dissected in RNAlater. Glands were transferred to 1.5ml Eppendorf tubes containing fresh RNAlater and stored at -80°C prior to RNA extraction and sequencing. This process was completed in approximately 15 minutes.

During transportation from the Crop Research Institute to the University of Nottingham the frozen silk glands were kept covered in dry ice inside an insulated carrier then transferred to -80°C freezer upon arrival.

2.2.2. RNA extraction

Dissected silk glands that had been stored in RNAlater at -80°C were defrosted and homogenised immediately prior to RNA extraction. For whole abdomen extractions, the spider was flash-frozen in liquid nitrogen and the abdomen removed and ground immediately prior to RNA extraction. The cephalothorax and limbs were discarded.

RNA was extracted using the Machery-Nagel NucleoSpin® RNA kit according to manufacturer's instructions; process completed in approximately 40 minutes. Purified RNA samples were stored at -80°C.

2.2.3. RNA sequencing, quality control and transcriptome assembly

RNA was sequenced using an Illumina NextSeq 500 and each transcriptome trimmed to high quality reads (removing artefacts such as sequencing adapters) using Scythe (Buffalo, 2014) and Sickle (Joshi and Fass, 2011) by the DeepSeq facility (University of Nottingham). Each transcriptome was *de novo* assembled using Trinity (version r20140413p1) (Grabherr et al., 2011) (at the time of assembly no suitable reference genomes were available), and assembled reads were screened for open reading frames and predicted protein sequences generated using TransDecoder (Haas et al., 2013).

2.3. Collection of dragline silk

Spiders naturally leave a trail of dragline silk that acts as a safety line when walking or running. In order to collect a line of pristine silk, test spiders were individually removed from their enclosure and encouraged to run across a corrugated arena, leaving a length of dragline silk suspended between the 'peaks' (Figure 2.1).

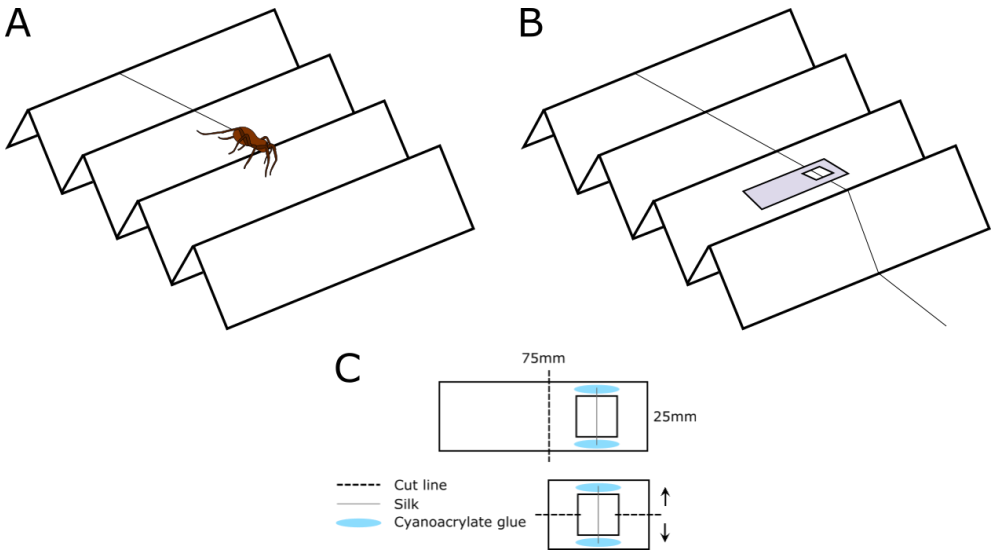


Figure 2.1. Diagram depicting collection of dragline silk. Spider walks across a corrugated arena, leaving a line of dragline silk (A). Corrugated arena after spider has been removed, with card frame applied (B). Diagram of card frame showing applied glue, attached silk and cut lines for tensile testing. Arrows show direction card frame pieces will be separated during tensile testing (C).

Cyanoacrylate glue was applied along two sides of a laser-cut card frame with a 15mm x 15mm or 8mm x 48mm window which was placed across the silk and left for 3-4 minutes to allow the glue to adhere (Figure 2.1). Next, the silk either side of the frame was cut with a sharp scalpel and samples were placed in a microscope slide box stored on the lab bench. The arena was brushed between spiders to remove any unobserved strands of silk.

3. Identification of novel spidroin genes in *Argyroneta aquatica*, *Cybaeus angustiarum* and *Pholcus phalangioides*

3.1. Introduction

3.1.1. Methods of identifying novel genes

Novel spidroin sequences are difficult to obtain: there is little conservation in the nucleotide sequence between silk types and between distantly related spider species, making primer design to allow direct amplification from genomic DNA without prior knowledge of the genetic sequence impossible. Furthermore, the length and the presence of long repetitive regions within spidroins makes obtaining novel sequences through PCR challenging (Ayoub et al., 2007, Babb et al., 2017, Casem et al., 2010, Hayashi et al., 2004, Hayashi and Lewis, 1998), even with prior knowledge of the genetic sequence to allow primer design. One alternative to PCR to isolate individual spidroins is to design probes to screen RNA, but this is time and resource expensive, risking the unintentional loss of genes when transforming and growing bacterial colonies (Hayashi and Lewis, 1998). For these reasons new approaches are needed to explore and characterise the full diversity of the silk gene family.

Generating a full genome requires significant effort, funding, and multiple specimens. The methods followed by Babb (Babb et al., 2017) describe the steps taken to sequence the genome of the golden orb weaver *Nephila clavipes*. Whilst >95% coverage was achieved, what is not detailed in the paper is the time and manpower involved, which would likely have been considerably greater than described here for a transcriptome.

Genomic sequencing by its very nature would cover all the silk genes the spider possesses but in addition to the cost of carrying out this work there is a further problem in that spidroins themselves may be difficult to assemble. This is because spidroins can be 4 – 16kb in length; for example the aciniform spidroin of *Argiope trifasciata* is 8618 nucleotides long and contains repeating regions of 600 nucleotides in length. Spidroins may also have introns, further increasing their length (Babb et al., 2017, Hayashi and Lewis, 2000, Motriuk-Smith et al., 2005). Additionally, evidence of genome duplication (Clarke et al., 2015, Schwager et al., 2017) and the presence of over 20 spidroins in the *Nephila clavipes* genome (Babb

et al., 2017), almost certainly including untranscribed sequences or pseudogenes, makes filtering out actively transcribed spidroins more difficult.

Transcriptomes are cheaper than genomes to sequence and provide a reduced dataset. As they contain RNA, not DNA, sequence, transcriptomes only include actively transcribed genes, immediately removing the need to discount pseudogenes and remove introns and other non-coding sequence. Additionally, by sequencing particular tissues only the genes transcribed in those tissues will be sequenced; for example, in a transcriptome of spider silk glands, silk mRNA will make up a comparatively greater percentage than if the transcriptome of a whole spider were to be sequenced. However, this method also means genes that are not currently being transcribed will be missed, or RNA that is present in low volumes unintentionally filtered out during the quality control stage. Additionally, sequences containing repetitive regions (such as the spidroin genes identified here) are more difficult to assemble correctly.

Identifying genes of interest in sequence data

High-throughput (formerly "next generation") sequencing methods such as Illumina (Solexa) or 454 pyrosequencing have been developed as a response to demand for low-cost-per-base sequencing services. The major disadvantage to these technologies however is their inability to sequence long "reads"; instead, maximum read lengths of hundreds of bases have been reported (Morozova and Marra, 2008). As such, any expected sequences longer than this require assembly, and software such as Trinity (Haas et al., 2013) has been developed to reassemble fragmented gene sequences. Whilst reference genomes are helpful in providing a scaffold upon which the fragmented transcriptome can be built, Trinity is capable of assembling transcriptomes *de novo* (without a reference genome). Once assembled, further software can be used to predict open reading frames and translate the transcriptome to amino acids (Haas et al., 2013). Finally, the BLAST algorithm (Basic Local Alignment Search Tool) (Altschul et al., 1990) can be used in conjunction with a custom made query database (such as one composed entirely of spidroin sequences) that searches the assembled transcriptome for specific sequences.

Non-sequencing variables to consider

Whilst generating transcriptomes has been chosen in this study to be the preferred method for sequencing novel spidroin sequences, a number of other

variables have to be considered. Firstly, there are costs and benefits to generating a transcriptome of the whole abdomen of the spider or the dissected silk glands. mRNA is a short-lived molecule, and cell machinery exists to break these molecules down as part of the regulation of gene regulation (Houseley and Tollervey, 2009). During RNA extraction this process is halted, and the RNA separated and purified. By sequencing RNA from the whole abdomen, we negate the need to dissect the silk glands and risk losing RNA due to degradation during the time required to perform the dissection and begin RNA extraction. However, by dissecting the silk glands we decrease the amount of RNA that will be sequenced, thereby increasing the relative amount of spidroin RNA in the transcriptome and potentially increasing the likelihood of more spidroin sequences being identified.

Further considerations that need to be made are the specimens from which the RNA is to be extracted from. Tubuliform spidroins (egg case silk) are exclusively produced by adult female spiders for wrapping and protecting their eggs, so adult female spiders should be used to ensure all spidroins may be present in the transcriptome. A related consideration is that adult females are likely to have mated with a male spider and started to produce eggs, leading to an increase in abdomen size without increasing the amount of silk gland tissue from which RNA can be extracted, so care may be required to select an adult female that is not yet gravid (egg carrying).

In this chapter we use three species of different size and ecology (section 1.1) and compare transcriptomes of the whole abdomen and combined dissected silk glands to examine whether the methods used are comparable in identifying novel spidroin genes. We assess whether sequencing the whole abdomen or glands provides more spidroin sequences and better data, and use this information to propose suitable methods for future *de novo* spidroin sequencing efforts.

3.2. Methods

3.2.1. Spidroin identification

Custom BLAST databases of spidroin nucleotide and protein sequences downloaded from GenBank and manually retrieved from the *Nephila clavipes* [36] and *Stegodyphus mimosarum* genomes [37] were generated to screen transcriptomes for silk sequences (see Table 3.1). Sequences identified by BLAST search were translated and manually examined for the “QALLE” motif (Challis et

al., 2006, Strickland et al., 2018) located approximately 40 amino acids before the "stop" codon, and an arginine located approximately 100 amino acids before the stop codon. Translation of the sequences allows for easier refinement by eye, as it appears there is greater sequence conservation at the codon level than at nucleotide level. Additionally, spidroins are predominantly formed of hydrophobic and hydrophilic residues and the identification of acidic or basic residues, such as the glycine found in the "QALLE" motif and the arginine found further along the sequence, is made easier by displaying the residues in colours reflecting as such – for an example, see section 4.4.2, Figure 4.2.

Identified partial spidroin sequences containing a C-terminal region were set aside and the remaining sequences aligned against a selection of trimmed N-terminal domains previously used in the BLAST databases; non-spidroin sequences which failed to align to existing N-terminal sequences were discarded; as with the C-terminal domain, a small number of acidic and basic residues are conserved at the codon level and readily identified by eye once aligned. None of the discarded sequences were manually identified as spidroin sequences without terminal domains (i.e. repetitive regions of silk spidroins).

Putative silk genes were tentatively named as one of the seven characterised spidroins or as an undefined spidroin sequence, depending on the greatest similarity to the top results in the BLAST searches against both the custom BLAST database and the GenBank database. Newly identified spidroins were included in the existing BLAST databases, which were used to repeat the BLAST search against the transcriptomes until no more novel spidroin sequences could be identified. Spidroin status was confirmed by BLAST-searching against the online NCBI GenBank database to identify any potential similarity to non-silk sequences, and confirm local BLAST result.

3.3. Results

Extensive GenBank searches resulted in a list of 340 spidroin sequences, including sequences representing the seven major silk types and a range of uncharacterised unspecialised spidroin sequences (see Table 3.1 and Appendix Table 9.2). These were used to generate nucleotide and protein BLAST databases.

Table 3.1. Summary of silk spidroin sequences used in the custom BLAST database.
Full list of sequences can be found in Appendix Table 9.2.

	Count
Sequences	340
Species	51
Families	21
Type-unknown spidroin sequences	53
Major ampullate sequences	123
Minor ampullate sequences	27
Tubuliform sequences	46
Aciniform sequences	61
Piriform sequences	11
Flagelliform sequences	11
Aggregate sequences	8

Table 3.2. Summary of transcriptome data.

	***A. aquatica* Dissected silk glands**	***A. aquatica* Whole abdomen**	***C. angustiarum* Whole abdomen**	***P. phalangioides* Dissected silk glands**	***P. phalangioides* Whole abdomen**
"Trimmed" RNA sequencing reads	23,086,892	20,935,042	22,140,818	75,245,460	72,488,374
Gene fragments assembled by Trinity	42,334	44,960	68,535	59,215	73,262
Protein sequences predicted by TransDecoder	13,895	13,349	16,418	27,503	34,832
BLASTn result	0	0	0	0	0
BLASTx result	76	23	52	29	39
BLASTp result	81	27	58	39	33
Putative spidroin sequences (N-term. only)	3	4	1	5	4
Putative spidroin sequences (C-term. only)	8	9	6	15	6

Five transcriptomes were generated, with three whole abdomen RNA extractions (one adult female *A. aquatica*, one *C. angustiarum* (sex and precise developmental stage undetermined) and one adult female *P. phalangioides*) and two from the combined silk glands of one adult female *A. aquatica* and one adult female *P. phalangioides* respectively. A number of novel silk spidroins were identified in the data by searching each assembled transcriptome for matches (e-value <0.00001) to the custom BLAST databases described above.

Silk spidroins show greater conservation at the amino acid level than at the nucleotide level and both terminal domains are conserved evolutionarily in all sequences (Strickland et al., 2018, Challis et al., 2006, Motriuk-Smith et al., 2005, Collin et al., 2018). BLASTp searches identified the greatest number of potential spidroin sequences, whilst BLASTn searches returned zero results. In total, 61 novel partial spidroin sequences, composed of a conserved terminal domain (Challis et al., 2006, Motriuk-Smith et al., 2005, Strickland et al., 2018) and short length of repetitive sequence, were identified in all of the five transcriptome assemblies (see Table 3.2, section 4.4 and Appendix 9.3). Within the retrieved *A. aquatica* and *P. phalangioides* transcriptome pairs (whole abdomen vs combined silk glands), a small number of sequences were found in each transcriptome pair whilst some were unique to each transcriptome.

Table 3.3. Summary of putative spidroin sequences identified.
(N-termini|C-termini) AcSp = Aciniform, AgSp = Aggregate, Flag = Flagelliform, MaSp = Major ampullate, MiSp = Minor ampullate, PiSp = Piriform, TuSp = Tubuliform/Cylindriform

Transcriptome	AcSp	AgSp	Flag	MaSp	MiSp	PiSp	TuSp	Uncharacterised spidroin
Argyroneta aquatica whole abdomen	2|2	-	-	2|3	-	0|3	0|1	-
Argyroneta aquatica combined silk glands	1|2	-	-	2|2	-	0|2	0|1	-
Cybaeus angustiarum whole abdomen	-	-	-	1|4	-	0|2	-	-
Pholcus phalangioides whole abdomen	-	-	-	3|0	-	-	-	1|6
Pholcus phalangioides combined silk glands	3|0	-	-	1|0	-	0|1	0|1	1|13

Due to the difficulty in assembling repeating sequences (Morozova and Marra, 2008), such as the repetitive region in the centre of all spidroin genes (section 1, Figure 1.4), we were only able to retrieve partial spidroin sequences from the transcriptomes, containing either the N- or C-terminal domain and a short length of repeating sequence. In all partial C-terminal spidroins the sequence ended with the terminal domain; all partial N-terminal spidroins with the exception of one (an *A. aquatica* aciniform sequence) included a methionine at the beginning and stretched the full length of the N-terminal domain.

```
Query= Gene.15693::c27065_g1_i1::g.15693::m.15693
Gene.15693::c27065_g1_i1::g.15693  ORF type:5prime_partial len:170
(-) c27065_g1_i1:105-614(-)

Length=170

> AY263390_translation Argiope amoena major ampullate spidroin-1
(MaSp1) mRNA, complete cds
Length=179

 Score = 83.6 bits (205),  Expect = 2e-022, Method: Compositional matrix adjust.
 Identities = 49/94 (52%), Positives = 70/94 (74%), Gaps = 4/94 (4%)

Query  78   LSSSASNARVSTAVSGLVGNG----VDALSNIVSSLSASVAATHPGLSGCEVLVQVLMEI  133
            LSS  + +RVS+AV+ LV +G      ALSN +S + + ++A++PGLSGC+VLVQ L+EI
Sbjct  44   LSSPGAASRVSSAVTSLVSSGGPTNSAALSNTISDVVSQISASNPGLSGCDVLVQALLEI  103

                 Query  134  ISALVHVLSSASIGQVNVNDAAYTAALVGDVISQ  167
            +SALVH+L SA+IGQVN + A  +A+LVG  + Q
Sbjct  104  VSALVHILGSANIGQVNSSSAGQSASLVGQSVYQ  137
```

Figure 3.1. Sample alignment from a BLASTp result file.
A novel *A. aquatica* C-terminal partial spidroin c27065_g1_i1 found in the whole abdomen transcriptome matched against 99 database sequences. The "QALLE" motif used to identify the C-terminal domain is visible in the subject sequence, represented as "QVLME" in the query.

Supported by the BLAST results, there is some evidence of sequence conservation at the codon level between spidroins of different species, particularly in the repeat regions of TuSp, PiSp and AcSp sequences, supported by the match locality in BLASTx and BLASTp searches. MaSp results were different; in one *A. aquatica* N-terminal sequence the top blast result found an 50% identity across amino acids in the terminal domain, but no significant similarity was found in any of the repeat/non-terminal amino acids in that sequence (the repeat region begins around the 167th amino acid). In an *A. aquatica* MaSp C-terminal result (query c27065_g1_i1, 170aa long, made significant matches with 99 database sequences) there was a 52% identity across amino acids 78 – 167 in the terminal domain, including the "QALLE" motif (in this case QVLME), but again no results covered the repeat/non-terminal region preceding the C-terminal domain (e.g. Figure 3.1).

3.4. Discussion

3.4.1. Use of transcriptomics to sequence and identify novel spidroins

A number of novel silk spidroins were identified using a transcriptome-based approach. As described in section 3.1, there are a number of methods to obtain silk sequences for study, although the most appropriate method will be a compromise between available time, funds and the intended use of the resulting data.

One of the benefits of transcriptomics is that a transcriptome is quicker to sequence and requires less preparation than a whole genome; RNA extraction from a live spider can be achieved within a couple of hours and only a single sample is needed per transcriptome. In this study, once the specimens for sequencing had been collected, each dissection took approximately 15 minutes and each RNA extraction was completed in under 45 minutes (other manufacturer's protocols may vary). Due to costs associated with running small numbers of samples, sequencing data in this study was not received for a number of months.

After sequencing and trimming, each assembly was processed overnight using the Trinity assembly software, and Transdecoder and initial BLAST searching was completed in a day. Manual annotation or identification of spidroin sequences was completed in approximately one week (exact time depended on the number of sequences being searched for and found in the transcriptome).

Future processing could be streamlined and automated using dedicated scripts and pipelines. In our results, transcriptome searches using the BLAST algorithm against GenBank and the custom spidroin database typically agreed in their results regarding spidroin type (perhaps unsurprising given the sequences in the database originated from GenBank); the use of GenBank to screen putative spidroin sequences served to protect against non-spidroin sequences being accidentally included in subsequent analyses.

The spidroin sequences obtained here are incomplete

Analysis of the spidroin sequences obtained suggested there were sequences unique to each transcriptome within the data analysed. This, coupled with unpaired N- and C-terminal domains (see Table 3.3), indicates that there are whole or partial spidroins missing from the data set. Missing data includes a

number of N-terminal and repetitive regions, which could be due to any or a combination of the reasons described earlier. Likely candidates include incomplete assembly due to the difficulty of assembling repetitive sequence, insufficient example N-terminal regions in the BLAST database or the possibility that N-terminal domains are more diverse than anticipated.

Some attempts have been made by other researchers in this group to isolate individual *A. aquatica* spidroin genes from cDNA, using PCR and primers designed from the transcriptomes obtained here. These attempts were unsuccessful, possibly due to difficulties connected with a group of genes which are large and highly repetitive (one aciniform spidroin is a single 19kb long exon) (Ayoub et al., 2013, Babb et al., 2017). Attempting to PCR spidroins after a DNA extraction is likely to have the additional challenge of a longer PRC product as some spidroins have been shown to contain intron sequence (Babb et al., 2017).

3.4.2. Use of sample tissue for transcriptome sequencing

Whole abdomen vs. combined silk glands

There appears to be no difference in whether greater numbers of partial spidroins were obtained from the dissected silk glands or the whole abdomen transcriptomes (see Table 3.2), which may reflect the limits of using transcriptomics to identify and analyse spidroin genes. Using whole abdomens in smaller spider species minimises the risk of RNA degradation during dissection. By using non-gravid females the proportion of silk glands (and their associated RNA) in specimen tissue was kept as high as possible. Larger species may be better suited to dissection to keep non-gland tissue to a minimum – although the size of the glands is not available, extremely large species such as the giant huntsman spider (*Heteropoda maxima*) likely possess silk glands which exceed the maximum tissue sample size of RNA extraction protocols, thereby requiring use of a part of the silk glands, in which case care would need to be taken to include only gland tissue from the production tubules and sac, where silk spidroins are transcribed.

One observation made during this study is the difficulty in dissecting silk glands whilst submersed in RNAlater. The glands and surrounding tissue were more difficult to separate compared with performing the same dissections in physiological solution (Řezáč, 2015). This may have increased the time to complete dissection of the silk glands which could have continued to degrade until

the glands were fully immersed in the solution. Our method of dissecting silk glands in RNA later and transfer to -80°C may have inadvertently damaged the stored samples as RNA later is intended to be used only at room temperature. Instead we could have dissected the glands in physiological solution in order to minimise the dissection time, immediately flash-freezing each sample in liquid nitrogen to halt RNase activity (Correa-Garhwal et al., 2018).

Combined vs individual glands

A further consideration to be made is whether to combine or separate silk glands prior to sequencing. Almost all of the spidroins identified in the *P. phalangioides* transcriptomes showed greatest similarity to unclassified spidroins in the BLAST searches (e.g. AF350283 *Plectreurys tristis* fibroin 3) and few were significantly similar to more established spidroin groups (e.g. aciniform, tubuliform). Spidroins are named after the gland in which they are produced – novel spidroin sequences in a transcriptome made from a whole abdomen or combined dissected silk glands cannot be labelled in the same way without further sequencing of gland tissue to identify the origin of such sequences. This is not necessarily inhibitive in the search for novel spidroins; knowledge of the sequence allows for the design of probes or PCR primers which can be used to search cDNA of individual glands. However, these methods assume that spidroin proteins are exclusively produced by a single gland; recent studies suggest this may not always be typical (Babb et al., 2017). If the silk glands are large, sufficient funding is available, and the spidroins are of particular interest, possibly the quickest process would be to re-sequence the transcriptomes of individual silk glands.

Presence of repeating sequence in spidroins

No complete (full-length) sequences were identified in the assembled transcriptome, likely due to difficulties assembling the repetitive region of the spidroin. Instead, the majority of sequences were incomplete, containing a length of repetitive region and either the N- or C-terminal domain. As such, it is difficult to pair sequences of the same spidroin type with N- and C-terminal domains where more than one of each have been identified, although a single pair association would also be tentative; multiple sequences may be present but not successfully sequenced or assembled post-sequencing as data may be lost through sequencing error or RNA degradation.

Repetitive proteins are difficult to assemble (Morozova and Marra, 2008). If repeat sequence fragments are identical the assembly software may discard them and it is also difficult to know which order repeats come in, relevant if there are single nucleotide polymorphisms between repeats. The repetitive regions of spidroins also vary; for example, major ampullate spidroins are composed of a selection of motifs repeated inconsistently throughout the spidroin, whereas the repetitive regions of aciniform spidroins are composed of tandem repeats of up to 200 amino acids in length (Hayashi et al., 2004).

It is possible there were some "repeat only" regions in the transcriptomes, but these are difficult to confirm manually. When identifying spidroins and analysing relationships the terminal regions are more useful, and typically include small amounts of repetitive region. However, repeat regions would tell us a bit more about the structure and properties of the final silk product so if identified in future should not be completely discounted.

An additional concern is that the TransDecoder software has an automatic minimum open reading frame of 300 nucleotides and will not translate anything shorter, although this setting can be modified. This is to minimise the occurrence of false positives, but risks the potential loss of partial spidroin sequence (such as short or fragmented repeat regions), relevant as the BLASTn searches seen here were ineffective in identifying spidroin sequences, whereas BLASTp (using a protein query to search against a protein database) was most effective. It is not known why in this case BLASTx searches did not return as many putative spidroin sequences as BLASTp, although the use of sequences trimmed using TransDecoder to only include identifiable open reading frames in BLASTp searches may have increased the e-value of individual queries that had insignificant results in the equivalent BLASTx search.

3.5. Conclusion

Our method of using transcriptomes to identify novel spidroins was successful; we were able to identify a number of partial spidroin sequences in all three species examined, including some which were not of characterised spidroins. By comparing whole abdomen and dissected silk glands we suggest that both are suitable starting material and the choice of material may be more relevant to the RNA extraction process than the final data set. The recently published transcriptome generated by Correa-Garhwal *et al.* (Correa-Garhwal et al., 2018)

came from dissected glands, immediately flash-frozen in liquid nitrogen, as the source of RNA for a transcriptome of *Dolomedes triton*, suggesting at least one research group favours the method of dissecting silk glands. Utilisation of emerging nanopore sequencing technologies may drive the publication of more complete spidroin sequences, allowing for greater study into the genetic and molecular structure of spider silks.

Transcriptomics can be a relatively quick way of retrieving spidroin sequence data from novel species without a reference genome. The identification of conserved terminal domains in our data allows for an amount of phylogenetic analysis to be undertaken in spidroin genes and this, coupled with the presence of small amounts of repeat sequence, will allow for further investigation into the evolutionary history, structure, and function of spider silk.

4. Analysis of the terminal domains of spider silk, including novel spidroins from *Argyroneta aquatica*, *Cybaeus angustiarum* and *Pholcus phalangioides*

4.1. Abstract

Spiders produce multiple silks with different physical properties that allow them to occupy a diverse range of ecological niches, including the underwater environment. Despite this functional diversity, past molecular analyses show a high degree of amino acid sequence similarity between C-terminal regions of silk genes that appear to be independent of the physical properties of the resulting silks; instead, this domain is crucial to the formation of silk fibres.

Here we present an analysis of the terminal domains of all known types of spider silk and include silk sequences from *Argyroneta aquatica*, *Cybaeus angustiarum* and *Pholcus phalangioides*. Our work indicates that spiders have retained a highly conserved mechanism of silk assembly, despite the extraordinary diversification of species, silk types and applications of silk over 350 million years. Sequence analysis of the silk C-terminal domain across the entire gene family shows the conservation of two uncommon amino acids that are implicated in the formation of a salt bridge, a functional bond essential to protein assembly. Analysis of the N-terminal domain shows conservation of a helical structure, supporting existing research. This conservation extends to the novel sequences isolated here.

This finding is relevant to research regarding the artificial synthesis of spider silk, suggesting that synthesis of all silk types will be possible using a single process.

4.2. Introduction

Spider silk proteins form multiple types of materials, including fibres and glues, each with specific mechanical properties and functions (Blamires et al., 2017, Garb, 2013, Hormiga and Griswold, 2014, Eisoldt et al., 2011). Gene duplication, recombination and diversification are thought to have generated this huge diversity of different silk types, each having a different ecological function (Clarke et al., 2015). These functions include safety lines, the structural frameworks of webs, external layers of protection around egg sacs, and securing prey (Hinman et al., 2000, Rising and Johansson, 2015).

In all spiders, silk proteins are produced and stored as liquid dope in silk glands (which may or may not be adapted to produce specific spidroins). When required,

the liquid silk dope is passed through elongated ducts and extruded through spigots mounted on articulate spinnerets (Rising and Johansson, 2015, Vollrath and Knight, 2001, Askarieh et al., 2010).

All silk genes have three components; a type-specific repetitive region flanked by conserved N-terminal (Motriuk-Smith et al., 2005, Rising et al., 2006) and C-terminal (Challis et al., 2006, Collin et al., 2016, Gnesa et al., 2012, Hagn et al., 2010) domains. Whilst the repetitive region determines the mechanical properties of each silk product (Hayashi and Lewis, 1998, Hayashi et al., 1999), the terminal domains work together ensuring that individual proteins assemble correctly and that the fibre forms at the correct stage of the spinning process (Rising and Johansson, 2015, Andersson et al., 2014, Andersson et al., 2017). The N-terminal domain restricts the formation of silk fibres to a precise point in the silk duct, preventing silk proteins stored in the silk gland from agglutinating (Askarieh et al., 2010). The C-terminal domain drives spontaneous fibre formation, likely through use of a pH-sensitive 'salt bridge' (Stark et al., 2007, Ittah et al., 2006), where noncovalent interactions between one basic and one acidic residue are disrupted at low pH because the acidic residue becomes protonated and is no longer charged.

Salt bridges, either individual or paired, have been proposed to explain the dimeric bundling of 4 or 5 alpha helices in the C-termini of one particular silk, the major ampullate (Hagn et al., 2010, Sponner et al., 2005, Sponner et al., 2004). Whilst the miniature spidroin 4RepCT, formed of 4 copies of a MaSp repetitive region and one C-terminus, has been sufficient to produce self-assembling silk fibres (Stark et al., 2007), it has recently been shown in the minispidroin NT2RepCT that including the N-terminal region achieves greater efficiency in the production of synthetic fibres (Andersson et al., 2017). However, studies have shown that the level of solubility demonstrated by the terminal regions differs between spider species, leading to further questions around their function and the suitability of individual terminal regions for use in silk protein synthesis (Askarieh et al., 2010, Andersson et al., 2014, Andersson et al., 2017).

A degree of amino acid sequence conservation has been observed from studies of small numbers of different silks (Challis et al., 2006, Hagn et al., 2010, Beckwitt and Arcidiacono, 1994, Gnesa et al., 2012, Sponner et al., 2004, Sponner et al., 2005, Collin et al., 2016). What is not clear is the extent to which this conservation

is maintained, particularly where silk proteins exhibit vastly different properties (e.g. the glue-like aggregate and piriform silk versus the superior strength of major ampullate and aciniform silks). Additionally, the diving bell spider *Argyroneta aquatica* spins silk whilst completely submersed in fresh water; given how silk proteins are dehydrated as part of the spinning process, does this "extreme" environment necessitate variation within the silk protein and spinning process?

Genetic analysis carried out thus far has been performed largely on spiders within the Mygalomorphae and particular groups within the Araneomorphae, such as the highly specialised Orbiculariae. This has left many groups understudied, including the Haplogyne and aquatic species. Whilst any data retrieved from *P. phalangioides* will inform on the evolutionary processes undertaken between the Mygalomorphae and Araneomorphae, one highly important question is whether the silk spidroins produced by *A. aquatica* are fundamentally different to those of terrestrial spiders. Whilst there is evidence of similar applications (egg sac, anchor threads, sheet web), previous studies have shown an unusual feature seemingly novel to *A. aquatica* webs. De Bakker et al. observed a film-like coating on the walls of the diving bell, covering a mesh of haphazardly woven silk (De Bakker et al., 2006). This film has been proposed to be a proteinaceous, hydrophilic hydrogel (Neumann and Kureck, 2013), thought to be linked to the gas-permeable properties of the bell. At the time of writing, there has been no speculation as to the composition of this hydrogel, including whether it could be a modified form of silk.

Here, we analyse the genetic sequences silks of all known types from spiders of as many groups as are available in GenBank. We include novel sequences from *A. aquatica, C. angustiarum* and *P. phalgngioides* and investigate whether the salt bridge structure in the C-terminus is conserved across all species and silk types sequenced to date. In the case of *A. aquatica* we also ask if changes in biophysical properties or the utilisation of silk in an "extreme" environment necessitates a change in how silk proteins are formed.

4.3. Materials and Methods

4.3.1. RNA extraction and transcriptome sequencing

Two *A. aquatica* transcriptomes were generated using total RNA extracted from the combined silk glands and whole abdomen, respectively, of two adult females.

A third transcriptome was generated from the whole abdomen of one *C. angustiarum* sub-adult (for details see section 2.2). Sequencing was performed using an Illumina NextSeq500 (DeepSeq, University of Nottingham).

4.3.2. Transcriptome assembly and identification of silk genes

Transcriptomes were *de novo* assembled using Trinity (Grabherr et al., 2011) then TransDecoder (Haas et al., 2013) was used to screen for open reading frames and generate predicted protein sequences. Custom BLAST databases of silk genes downloaded from GenBank and the *Nephila clavipes* [36] and *Stegodyphus mimosarum* genomes [37] were generated to screen transcriptomes for silk sequences, which were manually examined and BLASTed against GenBank for confirmation (e-value <0.0005).

Putative silk genes were identified through a series of BLAST searches and manually confirmed prior to being assigned as one of the seven spidroins or as a generic spidroin-like sequence, depending on the greatest similarity to the top results in the custom and GenBank databases (section 3.2.1).

4.3.3. Phylogenetic analysis

Putative silk genes and database silk sequences were manually examined for the presence of N- and C-terminal regions, which were separated for analysis. Silk sequences from GenBank, the *N. clavipes* and *S. mimosarum* genomes (see section 3, Table 3.1 and Appendix 9.2) and newly-identified spidroins (section 3 and Appendix Table 9.3) were filtered to only include N-terminal regions (manually confirmed visually in an alignment) or C-terminal regions (manually confirmed as containing a variant of the "QALLE" motif (Challis et al., 2006)). Duplicate sequences were discarded, leaving the longest sequence for analysis. Two sequences were considered identical if their nucleotide sequences had only a small number of variations which did not significantly alter the amino acid composition (e.g. variation causing a polar residue to change to acidic was considered significant, whereas polar to polar was not).

Selected sequences were manually trimmed leaving only the relevant terminal domain C-terminal regions were trimmed to include only the last 360 nucleotides/120 amino acids (Rising and Johansson, 2015, Hagn et al., 2010), whereas N-terminal-containing sequences were trimmed to a length of 501 nucleotides/167 amino acids counting from the first methionine (Motriuk-Smith et al., 2005, Rising et al., 2006).

Selected nucleotide sequences were aligned in Geneious v8.1.8 (Kearse et al., 2012) (Clustal W algorithm (Larkin et al., 2007)) then translated in order to allow refined alignment by eye. The refined untranslated nucleotide sequence alignments were imported into MEGA6.0 (Tamura et al., 2013) and used to construct a maximum-likelihood tree (MEGA6.0 default settings: Tamura Nei model, 1000 bootstrap replicates, complete deletion of gaps).

Secondary protein structures of both the N- and C-terminal regions were predicted using JPred4 (Drozdetskiy et al., 2015b), which utilises a secondary structure prediction algorithm and database of known structures to form a prediction (Cuff and Barton, 2000) (see section 4.4.2, Figures 4.6 and 4.7).

4.4. Results

4.4.1. Transcriptome analysis

Based on the presence of major ampullate, minor ampullate, piriform, aciniform and tubuliform glands and associated spigots, we expected to find spidroins with the same associations in the transcriptomes. Novel silk sequences were retrieved from the assembled transcriptomes using the silk database as described in section 3 and are listed in full in the Appendix, Table 9.3.

Table 4.1. Summary of unique novel N- and C-terminal silk spidroin sequences identified from the combined *A. aquatica* transcriptomes.
(N-terminal|C-terminal) Spidroins: AcSp = Aciniform, AgSp = Aggregate, Flag = Flagelliform, MaSp = Major ampullate, MiSp = Minor ampullate, PiSp = Piriform, TuSp = Tubuliform/Cylindrical, Uk = Unknown/Uncharacterised spidroin

Species	AcSp	AgSp	Flag	MaSp	MiSp	PiSp	TuSp	Uk
A. aquatica	1|2	-	-	3|2	-	0|3	0|1	-
C. angustiarum	-	-	-	1|4	-	0|2	-	-
P. phalangioides	3|0	-	-	-	1|0	-	0|1	1|4

Within the *A. aquatica* and *P. phalangioides* transcriptomes, a number of sequences were retrieved from both transcriptome sets, whilst some were unique to a single transcriptome. Some sequences showing similarity to aggregate glue 1 sequences (ASG1) (Choresh et al., 2009) were found in the *A. aquatica* combined silk glands transcriptome, but as these are not spidroins (Collin et al., 2016) were not analysed further. As recorded in Table 4.1 and discussed in section 3, partial

spidroin sequences associated with each of these glands were identified, with the single exception of minor ampullate spidroins, which appear to be absent from the data (this study).

Of the silk sequences retrieved from the two *A. aquatica* transcriptomes, sequences homologous to aciniform, major ampullate, piriform and tubuliform spidroins were retrieved. Only major ampullate and piriform spidroins were detected in the *C. angustiarum* transcriptome (Table 4.1). Within the *P. phalangioides* transcriptomes multiple uncharacterised spidroin sequences were identified, along with two major ampullate N-termini and four other N-terminal sequences showing similarity to major ampullate and aciniform spidroins (Table 4.1).

Table 4.2. Examples of translated repeat regions from *Argyroneta aquatica* and *Argiope trifasciata*.
Structural motifs highlighted: GGX, [GA]n, [A]n GPG and GPGX (bold) and proline residues (underlined).

Argiope trifasciata MaSp2	GSAPQGT**GGP**ASQGPSQQGPVSQPSYGPSATVAVTAV**GGR** PQGPSAPRQQGPSQQ**GPG**QQ**GPGGR**GPYGPSA
Argiope trifasciata MaSp1	AA**GGQGGQGGY**DGLGSQGAGQ**GGY**GQ**GGAAAAAAAAA**SG AGSAQR**GGL**GA**GGA**GQGYGAGS**GGQGGA**GQ**GGA**A
Argyroneta aquatica MaSp variant 1	GA**GGY**GQGQYGQG**AAAGAGAAAGAGAGAGA**SAGATGVG RR**GGY**GQ**GAGAGAGA**YNQIGYGQVV**GAGA**GIGAG
Argyroneta aquatica MaSp variant 2	**AAAAAGGTGGY**YYEA**GGA**G**AAAA**G**AAAA**SAAG**AAAGGA** RRA**GGY**GYGA**GGA**AG**AAAAA**GAA**GGA**GTRNTAYN

A comparison of peptide motifs found within major ampullate spidroin (MaSp) sequences showed that *A. aquatica* MaSps contain the GGX, poly-GA and poly-A peptides typically found in all MaSp sequences, but lack the GPG motif commonly found in MaSp2 (Hinman and Lewis, 1992, Ayoub et al., 2007). Hence, transcribed *A. aquatica* major ampullate spidroins are proposed to be exclusively MaSp1 (Table 4.2).

4.4.2. Analysis of silk terminal domains

Due to the structure of silk genes and the difficulty in assembling entire sequences, structural and phylogenetic analysis of complete spidroin sequences is not yet possible. However, the terminal domains may be analysed

independently in order to gain a rudimentary understanding of the evolution of silk genes.

Description of sequences retrieved from dataset containing terminal domains

61 new silk sequences isolated from transcriptomes of the spiders *A. aquatica*, *C. angustiarum* and *P. phalangioides* included a C-terminal domain (this study, Appendix Table 9.3). Once refined (see section 4.3), 21 unique sequences remained for analysis. Three hundred and forty-four silk sequences were retrieved from GenBank and the *N. clavipes* and *S. mimosarum* genomes, of which 150 sequences were both unique and contained a C-terminal region.

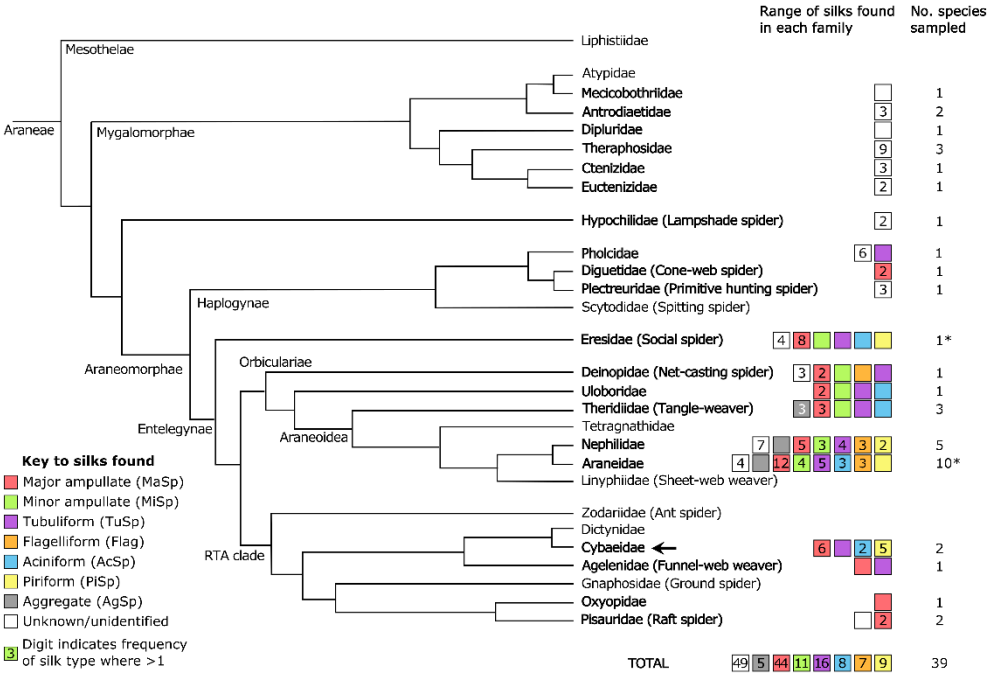


Figure 4.1. Familial origins of silk sequences used in this study.
Cladogram based on phylogeny presented by Nentwig (Nentwig et al., 2013), reduced to show positions of the different spider families and the silks previously described as being of a particular type. Location of *Argyroneta aquatica* and *Cybaeus angustiarum* (Cybaeidae) marked with an arrow.
Each coloured box represents a silk type; uncoloured boxes represent unclassified published spidroins. Numbers in each box indicate the total number of occurrences of each silk type. The total number of species sampled per family is detailed at the end of each row; * indicates where sequences have been retrieved from a genome.

The majority of the spidroin sequences remaining were identified in the Orbiculariae (n = 80). These, along with the spidroins obtained from the *S.*

mimosarum genome, also constituted the majority of the specialised spidroin sequences (Stark et al., 2007, Rising et al., 2007, Perez-Rigueiro et al., 2010, Garb et al., 2010, Babb et al., 2017, Sanggaard et al., 2014) (Figure 4.1). Five major ampullate spidroins were the only specialised silks found outside of this group prior to this study. Specialised spidroin sequences obtained from *A. aquatica, C. angustiarum* and *P. phalangioides* included major ampullate, tubuliform, aciniform, and piriform (Table 4.1) and marked the first evidence that the specialised piriform, tubuliform and aciniform spidroins exist outside of the Orbiculariae and Eresidae (Figure 4.1).

Within the five transcriptomes generated in this study, 16 sequences included a N-terminal domain, of which 8 were unique (Table 4.1). Within the data retrieved from GenBank and the *N. clavipes* and *S. mimosarum* genomes, 37 sequences were both unique and contained a N-terminal region. Similar to the C-terminal sequences, most of the N-terminal sequences (n = 31) were found in the Orbiculariae. One sequence represented the Filistatidae and two the Diguetidae, two families both members of the Haplogyne and closely related to *P. phalangioides*. Two sequences originated from the Agelenidae, which share an ancestor with the Cybaeidae and Dictynidae families, and another from the Pisauridae, another member of the RTA clade. No sequences found originated in the Mygalomorphae (see Appendix 9.2 and Figure 4.3 below).

Alignment of terminal domains

Evidence of sequence conservation is seen in both terminal domains in the BLAST results (see section 3.3, (Challis et al., 2006) and (Motriuk-Smith et al., 2005)) and can be visualised in an alignment of trimmed terminal domains. In the C-terminal domains the "QALLE" motif described by Challis and variants thereof are clearly conserved in all sequences; in particular, the final amino acid is maintained as an acidic amino acid in 100% of all sequences studied (Figure 4.2 A, page 45). A basic residue identified by (Hagn et al., 2010) as essential to the formation of the C-terminal secondary structure is also fully conserved in the data, although the precise position within the alignment varies (Figure 4.2 Region B, page 45).

A cysteine residue identified in major ampullate sequences by (Ittah et al., 2007), found immediately prior to the QALLE motif and described as crucial to the formation of the secondary structure, is conserved throughout the major ampullate group but rarely found outside of this group.

Within the limited data available for the N-terminal domain, there appears to be some conservation in the form of uncommon acidic and basic residues, not unlike what is seen in the C-terminal domain. The N-terminal domain is longer than the C-terminal domain and contains two partially conserved glutamic acids, one aspartic acid and one basic amino acid (arginine/histidine, Figure 4.3, page 46).

Maximum-Likelihood trees of terminal domains

Initial observations of the unrooted maximum-likelihood tree of C-terminal sequences (Figure 4.4, page 47) indicate that the data primarily cluster based on silk type, with groupings consisting of each specialised silk type (aciniform, aggregate, flagelliform etc.) and one grouping of uncharacterised silks belonging to the Mygalomorphae.

Not all groups are clearly defined; the major ampullate sequences form a paraphyletic group and several sequences are scattered across the tree. The Mygalomorphae also form a paraphyletic group. The flagelliform, piriform, minor ampullate, tubuliform and aggregate spidroins all form monophyletic groups. The aciniform group, whilst monophyletic, is poorly supported with no recorded bootstrap values – bootstrap values lower than 50% have been omitted.

The maximum-likelihood tree formed by the N-terminal data (Figure 4.5, page 48) contains monophyletic groups of tubuliform and aciniform sequences which are supported by the bootstrap values. Minor ampullate and major ampullate sequences group together, but some major ampullate sequences are distributed throughout the tree.

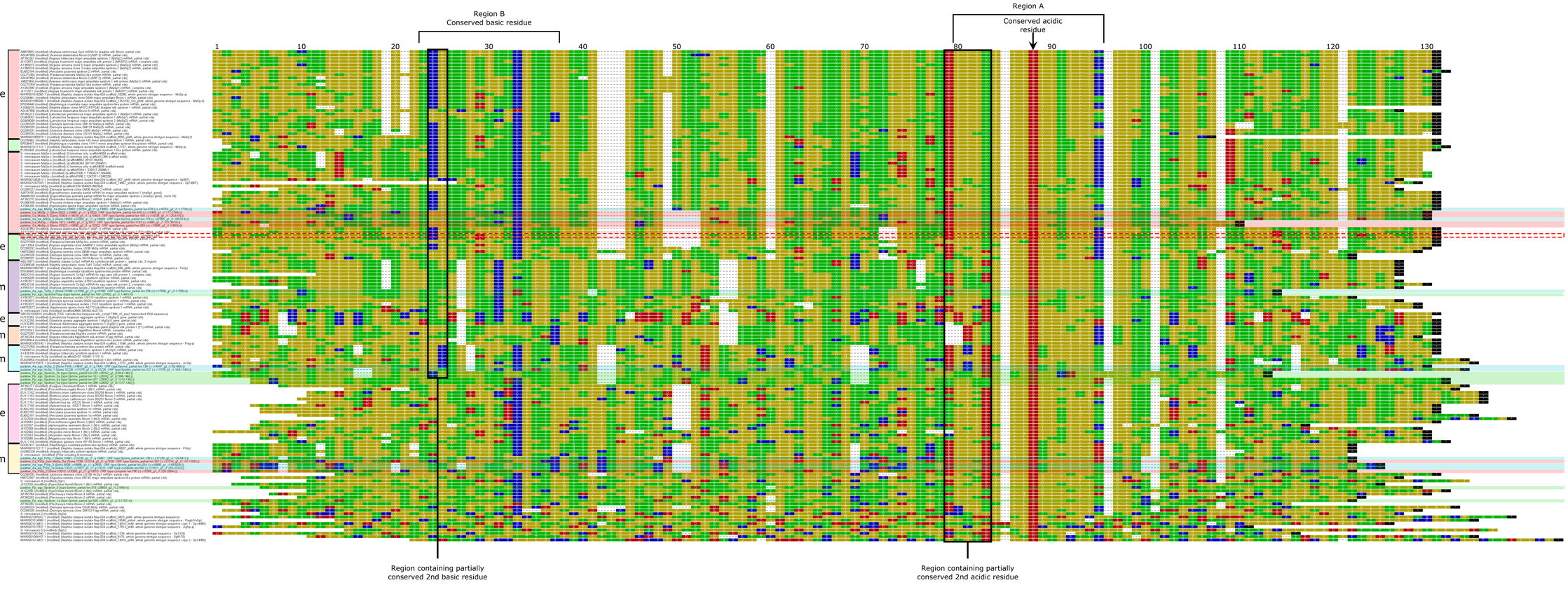


Figure 4.2. Translated alignment of all silk C-terminal nucleotide sequences used in this study.
Residues are coloured to show their charges: acidic (red), basic (blue), hydrophobic (yellow), hydrophilic (green), stop codon (black, marked *). Amino acids implicated in the formation of a salt bridge are marked (region A, arrow and indicated region B, basic residue only). Suspected secondary pair of conserved acidic and basic residues are highlighted with a black box. Novel sequences isolated in this study are highlighted: *Argyroneta aquatica* (blue), *Cybaeus angustiarum* (red), *Pholcus phalangioides* (green). The C-terminal sequence used to generate the 4RepCT synthertic silk sequence (Stark, 2007) is marked with a red dashed box.

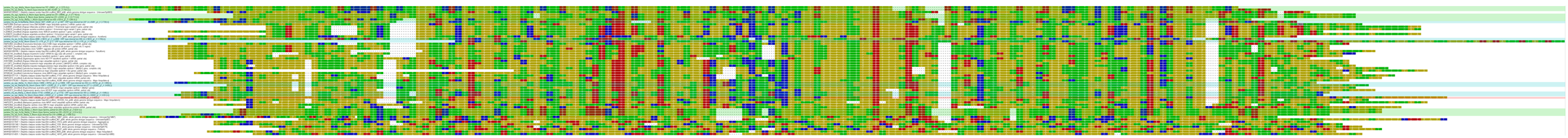


Figure 4.3. Translated alignment of all silk N-terminal nucleotide sequences used in this study.
Residues are coloured to show their charges: acidic (red), basic (blue), hydrophobic (yellow), hydrophilic (green), stop codon (black, marked *). Novel sequences isolated in this study are highlighted: *Argyroneta aquatica* (blue), *Cybaeus angustiarum* (red), *Pholcus phalangioides* (green).



Figure 4.4. Maximum-likelihood tree of all silk C-terminal nucleotide sequences used in this study. Maximum-likelihood tree (unrooted, 1000 bootstraps) showing relative relationships of silk C-terminal regions. Regions of characterised silks highlighted AcSp = aciniform spidroin (blue), AgSp = aggregate spidroin (grey), Flag = flagelliform spidroin (orange), MaSp = major ampullate spidroin (red), MiSp = minor ampullate spidroin (green), PiSp = piriform spidroin (yellow), TuSp/CySp = tubuliform spidroin (purple), Mygalomorphae (pink). Novel sequences isolated in this study are indicated by dashed boxes: *Argyroneta aquatica* (blue), *Cybaeus angustiarum* (red), *Pholcus phalangioides* (green).

Key:
- Aciniform
- Aggregate
- Flagelliform
- Major ampullate
- Minor ampullate
- Piriform
- Tubuliform
- Mygalomorphae

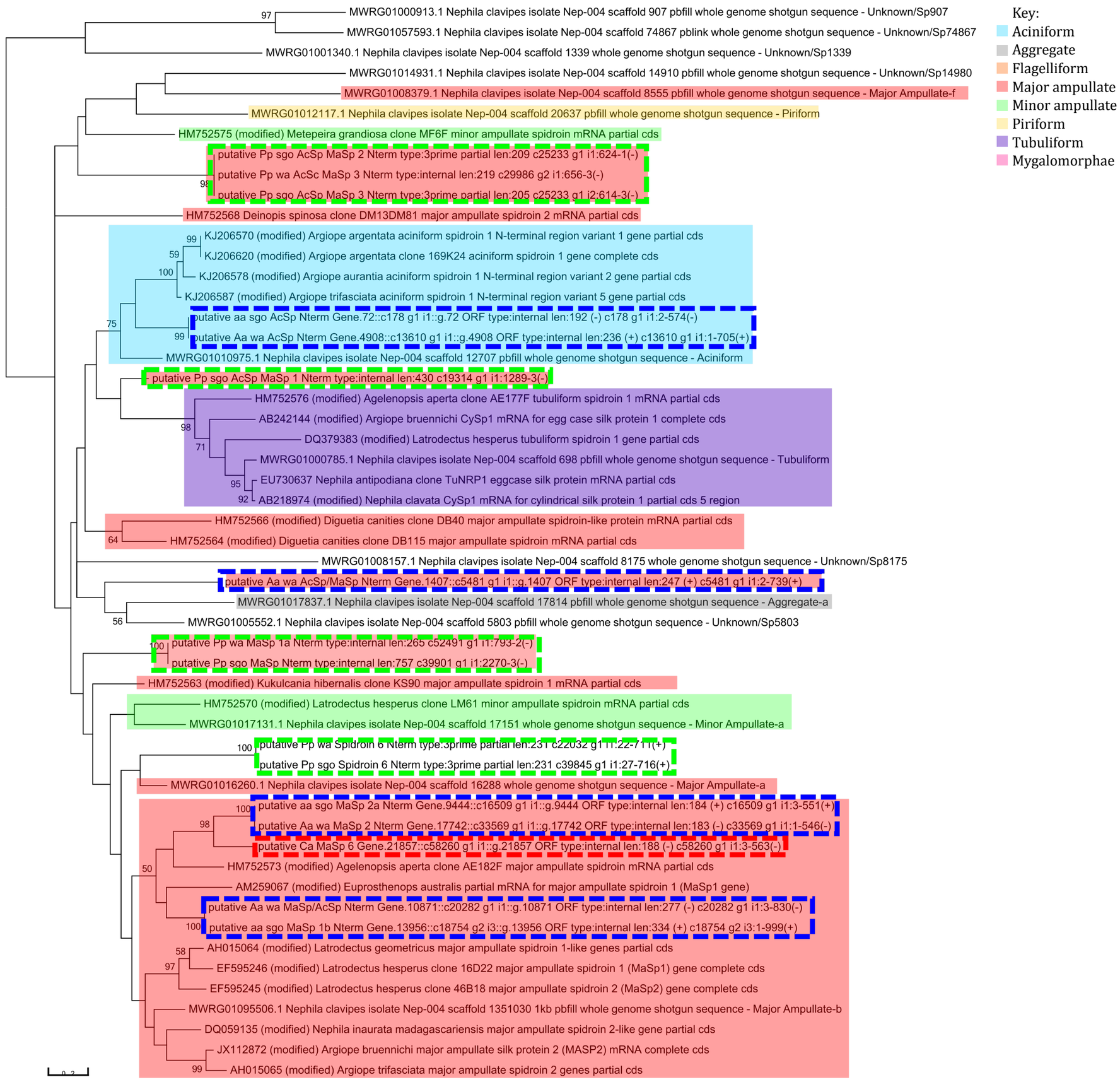


Figure 4.5. Maximum-likelihood tree of all silk N-terminal nucleotide sequences used in this study. Maximum-likelihood tree (unrooted, 1000 bootstraps) showing relative relationships of silk C-terminal regions. Regions of characterised silks highlighted AcSp = aciniform spidroin (blue), AgSp = aggregate spidroin (grey), Flag = flagelliform spidroin (orange), MaSp = major ampullate spidroin (red), MiSp = minor ampullate spidroin (green), PiSp = piriform spidroin (yellow), TuSp/CySp = tubuliform spidroin (purple), Mygalomorphae (pink). Novel sequences isolated in this study are indicated by dashed boxes: *Argyroneta aquatica* (blue), *Cybaeus angustiarum* (red), *Pholcus phalangioides* (green).

Analysis of secondary structure of terminal domains

Analysis of the predicted secondary structures of sequences representing each silk type with JPred4 (Drozdetskiy et al., 2015a) suggests the secondary structure of the C-terminus is conserved across silk types, containing up to five α-helices (Figure 4.6) and supporting previous studies (Ittah et al., 2007, Hagn et al., 2010). However, some predictions suggest Helix 1 may be absent, including in the *N. clavipes* MaSp-a and MaSp-b sequences analysed here (Figure 4.6). The conserved acidic and basic residues identified in Figure 4.6 are highlighted, with the acidic amino acid consistently nested within the fourth helix (Figure 4.6 A) and the basic amino acid within the second helix (Figure 4.6 Region B).

One of each specialised silk is represented in the alignment showing predicted secondary structures for the N-terminal domain (Figure 4.7), but due to the availability of spidroin sequences not all spidroin types can be shown. Six alpha helices are predicted in each sequence with the exception of piriform, which is shown with a short seventh helix at the 3' end of the region.

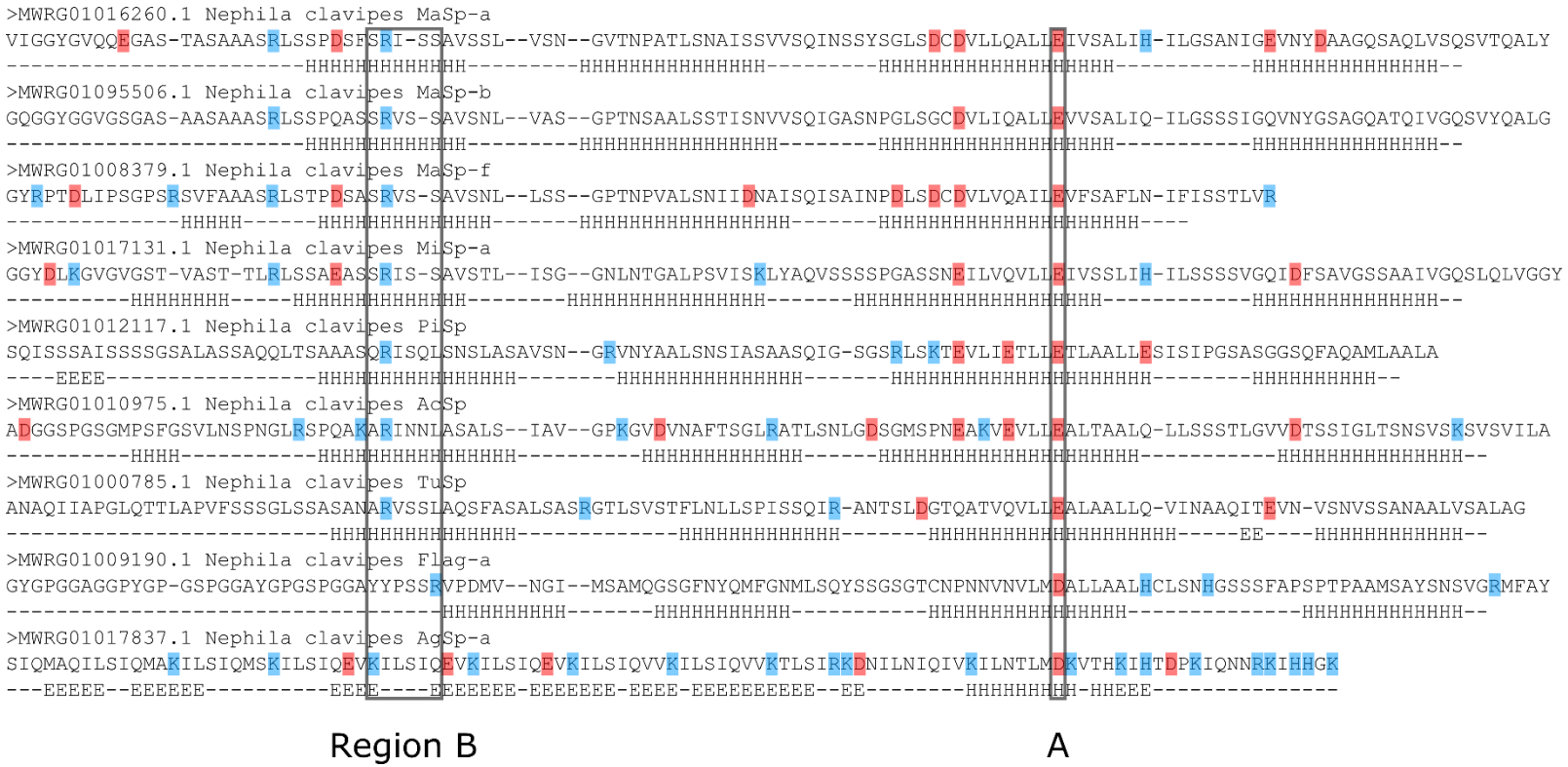


Figure 4.5. Alignment of *Nephila clavipes* C-terminal sequences showing individual JPred secondary structure predictions (Jnet) beneath.
Amino acid residues implicated in the formation of a salt bridge are marked A (acidic residue; glutamic or aspartic acid) and found within region B (basic residue; arginine or lysine). Acidic (red) and basic (blue) amino acids are highlighted.
Structures predicted by JPred are marked H (α-helix) and E (extended secondary structure (Degrève et al., 2014)).



Figure 4.6. Alignment of *Nephila clavipes* N-terminal sequences showing individual JPred secondary structure predictions (Jnet) beneath.

Amino acid residues that appear to be conserved in Figure 4.5 A1, A2, A3 and B are highlighted. Acidic (red) and basic (blue) amino acids are highlighted.

Structures predicted by JPred are marked H (α-helix) and E (extended secondary structure (Degrève et al., 2014)).

4.4.4. Conservation of a salt bridge structure across spider silk genes

In all of the C-terminal sequences analysed in this study, one acidic residue and one basic residue pair has been conserved (Figure 4.2). A conserved basic residue is consistently found in Helix 2 (Figures 4.2, 4.6 and 4.8), whilst an acidic base is always found towards the centre of Helix 4 (Figure 4.2, 4.6 and 4.8). Additionally, in major ampullate, minor ampullate, aciniform and piriform sequences a second acidic and basic residue appears to be partially conserved (Figure 4.2) as identified by Sponner (Sponner et al., 2004, Sponner et al., 2005). These pairs correspond with those shown to form a salt bridge in the secondary protein structure in previous studies (Hagn et al., 2010, Ittah et al., 2006).

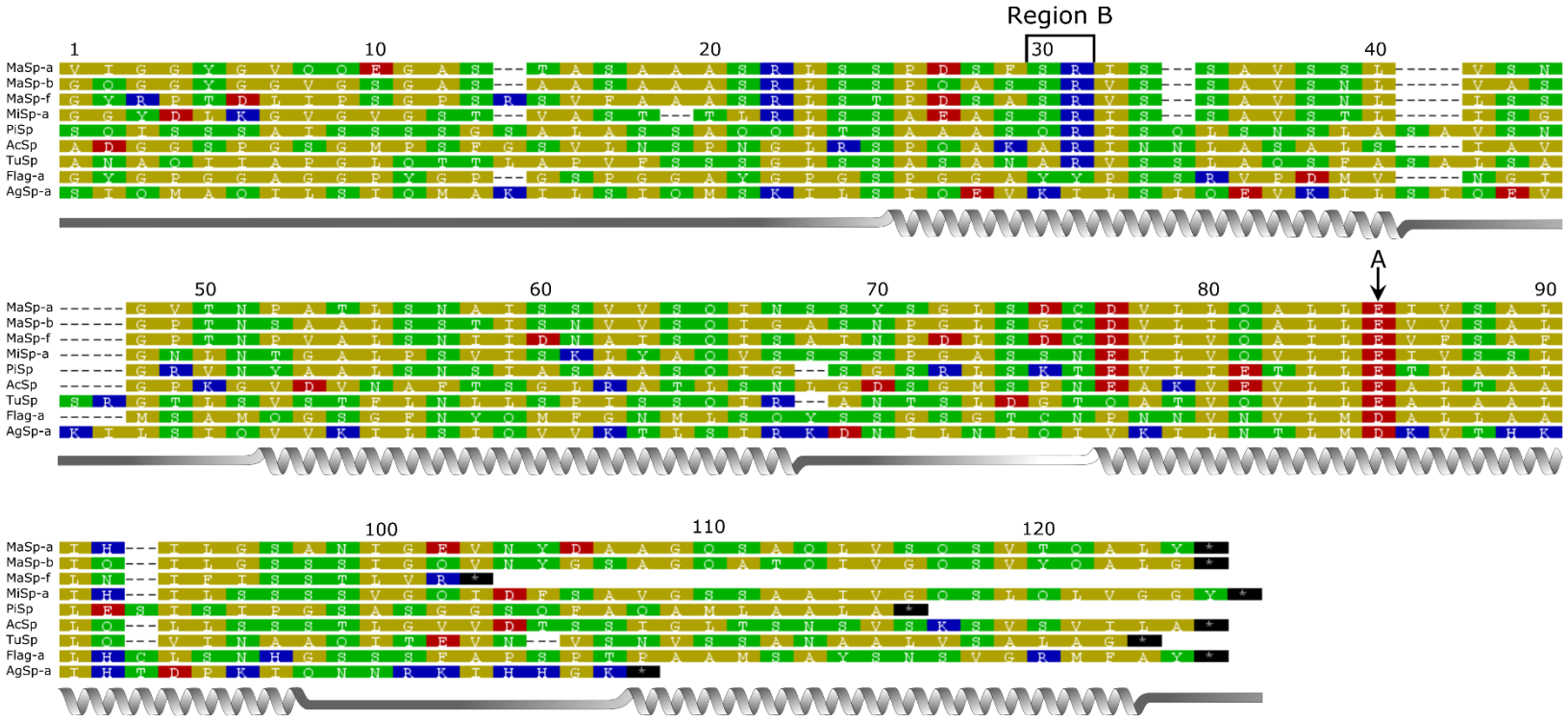


Figure 4.7. Colourised alignment of *Nephila clavipes* C-terminal sequences.
Representatives of all characterised silk types present, with predicted helical secondary structure of MaSp-a below. Amino acid residues implicated in the formation of a salt bridge are marked A (acidic residue; glutamic or aspartic acid) and found within region B (basic residue; arginine or lysine). Residues are highlighted to show whether acidic (red), basic (blue), hydrophilic (green) or hydrophobic (yellow).

Amino acid content of the C-terminal domain

The majority of the C-terminus is composed of hydrophobic and hydrophilic amino acids (average 53.5% and 37.0% respectively) including serine, alanine and leucine (average 21.2%, 13.2%, and 10.8% respectively) in an alternating arrangement. These residues promote the formation of alpha-helices in the protein structure, with the hydrophilic residues exposed, promoting solubility (Hagn et al., 2010).

Overall DNA sequence identity is highest in the acidic residue found in the "QALLE" motif (Challis et al., 2006), where the first two nucleotides (guanine, adenine) are completely conserved and the final nucleotide determines whether it is a glutamic or aspartic acid residue (89.3% and 10.7% respectively of all sequences in this study; Figure 4.2). The basic residues involved in the formation of the salt bridge are usually arginine, although some are lysine or histidine (94.7%, 3.3% and 2.0% respectively).

The charged residue content of the C-terminus is typically less than 10% (5.2% acidic, 4.1% basic; average across all sequences in this study) although this varies depending on the sequence (aggregate spidroins average 13% acidic, 8% basic). However, one acidic and basic residue pair (A and region B, respectively, Figure 4.2, 4.6 and 4.8) are present in each sequence examined, irrespective of species or silk type.

Additionally, we find a raised percentage of the amino acids surrounding the basic residue within spider silk are hydrophilic (43.0%, Figure 4.2 region B), reducing its pKa and hence promoting its protonation, whereas a majority of those surrounding the acidic residue are hydrophobic (68.8%, Figure 4.2 Region A), increasing its pKa and again promoting protonation.

4.5. Discussion

4.5.1. *A. aquatica* transcriptome

The successful identification of multiple spidroin sequences in the two *A. aquatica* transcriptomes using the spidroin BLAST databases demonstrated that these spidroins remain within the same gene family. What requires further investigation is the finding of multiple different copies of one spidroin and absence of a spidroin we expected to find. Whilst this may have been due to unrefined techniques, it cannot be eliminated that these absent and additional

sequences contain evidence of evolution in relation to the unique behaviour of this species.

Piriform spidroins

Three unique partial spidroin sequences, containing a section of repeat region and a C-terminal domain similar to that of known piriform spidroins, were identified in our data. Two of these sequences (PiSp variants 2 and 3) closely resemble each other – given their similarity and that one was identified from each transcriptome, it seems likely that these sequences represent the same piriform spidroin gene with intra-individual variation (in this case, the inclusion or deletion of three consecutive nucleotides). The remaining sequence, PiSp variant 1, was identified in both transcriptomes. As such, we propose that *Argyroneta aquatica* produce two PiSp-like proteins. To date, only one piriform spidroin has ever been identified in any species, including in the genomes of *Nephila clavipes* (Babb et al., 2017) and *Stegodyphus mimosarum* (Sanggaard et al., 2014).

Piriform silk is associated with the cement-like attachment discs of orb weaving spiders, used to attach major ampullate silk to a suitable frame during the web building process (Grawe et al., 2014). These attachment discs are composed of piriform fibres and glue, within which the major ampullate fibre is anchored (Grawe et al., 2014, Wolff et al., 2015). Notably, at the time of writing, no sequences have been proposed for the proteins that make up the glue, the composition of which is thought to contain unidentified salts, proteins (including nanofibrils) and hydrocarbons (Wolff et al., 2015, Kovoor and Zylberberg, 1980, Grawe et al., 2014).

The partial piriform spidroins we have identified in *A. aquatica* differ somewhat to those identified thus far in other species. In the repetitive region isolated, we note an increased acidic (PiSp1 6.7%, PiSp2 9.2%, PiSp3 8.4% vs *Nephila clavipes* PiSp 3.4%) and basic (PiSp1 8.4%, PiSp2 8.3%, PiSp3 7.5% vs *Nephila clavipes* PiSp 3.3%) amino acid content, and in phylogenetic analysis of the C-terminal region *A. aquatica* and *C. angustiarum* C-termini group on a separate branch to other piriform silks, although we note that they maintain a closer relationship with piriform than with other silk types. Given the difference between the two spiders, it is possible that the increase in acidic/basic residues in *A. aquatica* piriform spidroins is an adaptation specific to an underwater-dwelling spider.

Major ampullate spidroins

The presence of multiple MaSp1-like spidroins and apparent absence of MaSp2-like spidroins was unexpected, but may be an aquatic adaptation. Research by Boutry and Blackledge (2010) has shown that supercontraction in spidroins containing the GPGXX motif (typically MaSp2) is greater than in non-proline-containing spidroins (Boutry and Blackledge, 2010). Given the extreme humidity experienced by *A. aquatica* silks, MaSp2 spidroins would likely be of greater cost than benefit in the structure of a diving bell and anchor lines; they would likely contract more than MaSp1 on contact with water, resulting in a greater use of resources when producing a strand of MaSp2 equivalent in length to MaSp1.

Aciniform spidroins

We also identified three partial aciniform spidroin sequences (one N-terminal region and two C-terminal regions), a spidroin previously shown to form one of the finer and stronger silk fibres (Hayashi et al., 2004). AcSp are traditionally labelled the "prey-wrapping" silk in orb weavers (Blackledge and Hayashi, 2006) but they also form a significant component of sheet spiders' webs and are used as a lining in egg sacs. *A. aquatica* are not known to wrap their prey before consumption and the presence of multiple aciniform spigots on two pairs of spinnerets suggests it is a highly-used material. We therefore propose AcSp is likely to form a major component of the diving bell wall.

Absent minor ampullate spidroins

Comparing the spigots, glands and identified silk genes, we note that we were also unable to identify minor ampullate spidroin sequences. RNA degradation pre-sequencing, or removal of poor quality reads post-sequencing and prior to assembly cannot be completely excluded as possible explanations for the apparent absence. However even if this were the case, rather than a true loss of this particular silk in *A. aquatica*, the lack of sequences in our final assembly could also imply that they may be under-utilised and only present in small quantities. This reasoning also applies to the suspected absence of MaSp 2 mentioned above.

4.5.2. *C. angustiarum* transcriptome

The inclusion of *C. angustiarum* as a terrestrial cybaeid for comparison yielded interesting results in the form of multiple MaSp-like spidroins, but otherwise maintained expected evolutionary relationships, grouping closely with *A. aquatica* spidroins in the analyses (Figure 4.4).

The absence of aciniform spidroins and presence of only one partial piriform sequence, both of which we expected to find based on glands and spidroins found in adult specimens, may be due to a combination of low quantities of RNA and/or RNA degradation and use of a sub-adult in our analysis; further attempts to gather additional data through production of an additional transcriptome were not completed.

4.5.3. *P. phalangioides* transcriptome

A number of spidroin sequences were retrieved that did not associate with the characterised spidroins of the Orbiculariae. Within the maximum-likelihood tree of C-terminal sequences, one partial *P. phalangioides* C-terminal sequence did associate as a basal sequence to the tubuliform group although the bootstrap support for this was very weak (<50%, Figure 4.4). A second C-terminal sequence (Spidroin 1a) grouped together with Deinopidae and Uloboridae sequences which may be interpreted as demonstrating an evolutionary step towards the Orbiculariae, but this relationship is again very weakly supported (<50%) in the tree. Otherwise *P. phalangioides* C-terminal sequences are scattered outside of major groupings within Figure 4.4.

N-terminal sequences tell a similar story and are also scattered throughout Figure 4.5. However, within the BLAST results the *P. phalangioides* N-terminal domains showed similarity to aciniform and major ampullate sequences as opposed to uncharacterised sequences. This is unlikely to be significant as this association is not maintained in the maximum-likelihood tree of N-terminal sequences; one such spidroin is placed between tubuliform and aciniform spidroins although is poorly supported (<50% bootstrap support).

4.5.4. Analysis of N- and C-terminal domains

The data available for the terminal domains differs greatly in quantity. Whilst over 150 sequences can be analysed for the C-terminal domain, less than 40 are available for the N-terminal domain. Why this is the case is not known, but there is a possibility (visible in the variation shown in Figure 4.7) that despite the conservation suggested by the larger alignment there is sufficient variation between the N-terminal domains that identifying novel N-terminal domains is especially difficult. An additional consideration is that there are two possible isoforms available for the N-terminal region (Motriuk-Smith et al., 2005); only the longer, “full length” isoforms have been considered in this study.

An observation made by Ittah *et al.* (Ittah et al., 2006, Ittah et al., 2007) identified a polar amino acid, cysteine, as being conserved in the C-terminal domain of major ampullate spidroins. They demonstrated that this amino acid affects the correct folding of the secondary structure of the MaSp C-terminal domain and whilst substituting an amino acid with the most similar structure to cysteine – in this case, serine, also a polar amino acid – permitted the formation of the same structure. Conversely, a vastly different amino acid (arginine, a basic amino acid) prevented the correct structure forming. Whilst the alignment in Figure 4.2 here demonstrates the conservation of the lone cysteine does not extend beyond the major ampullate spidroins, the same position is often occupied by a serine, supporting the hypothesis that this position is important to the correct formation, and likely function, of the C-terminal domain.

Within the maximum-likelihood tree of C-terminal sequences, the aggregate group form a branch of the tubuliform spidroins suggesting these silks may have originated from within the tubuliform. The minor ampullate group is weakly supported by bootstrap values and falls within the major ampullate group, supporting previous research that proposes minor ampullate silk C-termini evolved from major ampullate (Garb et al., 2010). This is not reflected in the maximum-likelihood tree formed by the N-terminal sequences, which appears to be much less organised. The addition of further N-terminal data, spanning a broader range of material including both underrepresented and absent spidroin types and covering a greater range of species would lend weight to any emerging evolutionary relationships.

Predicted secondary structures

The secondary structure of the N-terminal (Kronqvist et al., 2014) and C-terminal (Hagn et al., 2010) domains have been determined in the major ampullate spidroins of *Euprosthenops australis*. The JPred alignments of *N. clavipes* spidroin C-termini in Figure 4.8 agree with the literature and demonstrates that the secondary structure of the major ampullate terminal domains is broadly conserved across the terminal domains of all the spidroin types analysed, suggesting that this structure is conserved across the silk gene family and therefore performs the same function in all spidroin sequences irrespective of the final properties or application of the silk protein.

The JPred predictions of the N-terminal domain only partially agrees with the X-ray structure described by Kronqvist *et al.*; in the prediction here an additional, sixth α-helix is predicted to form. However, the structure predictions show a similar structure for all the types of spidroin examined, agreeing with other research (Rising et al., 2006, Motriuk-Smith et al., 2005, Garb et al., 2010).

4.5.5. Presence and conservation of salt bridges within the terminal domains

One salt bridge is conserved in all C-terminal domains

The C-terminal domain has proven to be necessary in the production of synthetic spider silk proteins. Investigations into the secondary structure of major ampullate silk identified a number of alpha helical structures which, once folded, are held in place by a salt bridge (Hagn et al., 2010). As this structure appears to be conserved across all silk types (Figure 4.6), we investigated whether this salt bridge is also conserved and examine whether the amino acid content is likely to promote the formation of this feature.

The residues responsible for the formation of the pH-sensitive salt bridge(s) are both uncommon and conserved within the restricted number of sequences studied thus far (Hagn et al., 2010, Challis et al., 2006). Analysis of the expanded range of sequences used in this study shows this trait is maintained, with >90% of the C-terminal domain typically composed of hydrophobic and hydrophilic residues. Of the charged residues present, at least one acidic and one basic residue is conserved in all sequences and as such inferred to be crucial to the formation of a salt bridge and therefore the correct folding of silk proteins.

5. Morphology of the web and spinning apparatus of the water spider, *Argyroneta aquatica* and comparison with *Cybaeus angustiarum*

5.1. Introduction

Major evolutionary and ecological changes are uncommon and of these, examples of species adapting to life under water whilst retaining functioning lungs are rare. In the case of the diving bell spider *Argyroneta aquatica*, this transition is achieved through use of a silken 'gill' that takes the form of a diving bell, facilitating oxygen storage and gaseous exchange underwater.

As discussed in section 1, spider silks are a hugely diverse range of proteins which form specialised fibrous or glue-like materials (Swanson et al., 2006b, Swanson et al., 2007, Blackledge and Hayashi, 2006) (Wright and Goodacre, 2012, Allmeling et al., 2006) (Clarke et al., 2015) (Hinman et al., 2000, Rising and Johansson, 2015) (Blasingame et al., 2009). Whilst silks have historically been used in traditional medicine (Heimer, 1988), spider silks and their synthetic counterparts have a range of potential material and biomedical applications (Harvey et al., 2017) and a greater understanding of how *A. aquatica* make and use their silk will enable future biomaterial research.

We have shown in an earlier chapter that *A. aquatica* spidroin sequences are not noticeably divergent from those of terrestrial spiders and include aciniform, piriform and major ampullate spidroins (Strickland et al., 2018). In order to better understand the silken structure that enabled the transition to underwater life, we use environmental scanning electron microscopy (eSEM) to examine the structure of a fully hydrated diving bell for the first time. eSEM varies from traditional SEM by allowing a gaseous atmosphere in the specimen chamber and permitting the imaging of wet or fully hydrated samples. This negates the need to dehydrate and coat the sample in a conductive material, meaning the "true" structure (unaffected by the sample preparation process) can be observed at high magnification, avoiding the inadvertent destruction of microstructures or creation of artefacts as a result of sample preparation (Uwins et al., 1993, Danilatos, 1981). These images are examined alongside those of (De Bakker et al., 2006) and (Neumann and Kureck, 2013) who both presented SEM images of dehydrated samples.

We also use traditional scanning electron microscopy (SEM) to examine *A. aquatica* spigots and spinnerets, and view the dissected silk glands using light microscopy. By comparing *A. aquatica* with the closely-related *Cybaeus angustiarum,* a small terrestrial cybaeid that lives exclusively in habitats with close to 100% humidity and a species closely related to *A. aquatica* (see section 1.1.2 and 1.1.3), we hope to better understand the process by which *A. aquatica* transitioned to an aquatic lifestyle.

5.3. Materials and methods

5.3.1. Study animals

Approximately 50 adult and sub-adult *Argyroneta aquatica* were collected from the Máchovo jezero Lake, Doksy, Czechia in October 2014 and June 2015. Approximately 20 *Cybaeus angustiarum* adults and sub-adults were collected from the Krkonoše region of Czechia in June 2015 and autumn 2016. (For full details of collection and husbandry see section 2.1.)

5.3.2. Preparation of diving bell and web specimens for environmental Scanning Electron Microscope (eSEM) imaging

Production of fresh *A. aquatica* diving bells and *C. angustiarum* webs was encouraged by first destroying existing structures, both to allow for investigation at different stages and to minimise the amount of debris caught in the mesh. Spiders were allowed to build and add to their diving bell/web for 7 days, with day 1 the day a diving bell/web was first visible. Samples were retrieved from the spider's enclosure, laid across a mounting stub and immediately imaged with a FEI Quanta 650 FEG eSEM in environmental mode, at low voltage (5kV) and a working distance of 7 – 20mm, with H_2O vapour atmosphere. No other preparatory methods were used (e.g. dehydration, sputter coating).

5.3.3. Preparation of spider specimens for Scanning Electron Microscope (SEM) imaging

Spider specimens stored in 70% ethanol were gradually transferred to 100% ethanol (30 min at each concentration, with one repeat at 100%) and critical-point dried (Leica EM CPD 3000). Specimens were then affixed to stubs with double sided tape and sputter-coated with gold *en vacuo* (Leica EM SCD005) before examination with a JEOL 6060LV SEM.

5.3.4. Imaging of silk glands

The spinning glands of fresh spiders were dissected in embryo dishes in physiological solution (0.9% aqueous solution of sodium chloride) and viewed under an Olympus SZX12 stereomicroscope. The glands were subsequently transferred in a drop of physiological solution onto a microscope glass slide with a small, prefabricated circular impression and photographed under a Nikon Eclipse 80i light microscope. Voucher specimens are deposited in the Crop Research Institute, Prague (Czechia).

5.4. Results

5.4.2. Morphology of the diving bell

Environmental scanning electron microscope (eSEM) images of the diving bell of an adult female show a variety of silk fibres. Fibres ranged in size from less than 0.25µm to greater than 16µm (see Figure 5.1 C, F & H), with smaller fibres apparently laid in parallel but in otherwise haphazardly placed layers throughout the structure. The few larger fibres found occurred rarely, in samples of the wall as opposed to the "roof" of the bell. The majority of the fibres found were between 0.2 - 5µm in diameter.

Regions of the diving bell appeared to be thickly coated in a hydrogel-like substance that obscured any silk threads it coated (originally described in De Bakker *et al.* 2006 and Neuman & Kurek 2014). This substance was typically associated with ribbon-like structures (Figure 5.1D) and flat, amorphous regions (Figure 5.1G). When the pressure and humidity inside the eSEM was lowered, these regions began to shrink and tear (Figure 5.1 G & H). In thicker areas of the bell, the surface coating of hydrogel cracked, exposing the fibrous structures underneath. Any fibres within the gel also appeared to shrink and break (Figure 5.1 H). Occasionally, individual or clusters of diatoms (likely *Sellaphora* sp.)(Wetzel et al., 2015), fragments of *Daphnia* sp. and other micro debris were found, adhered to or partially embedded in the structure (Figure 5.1 D).

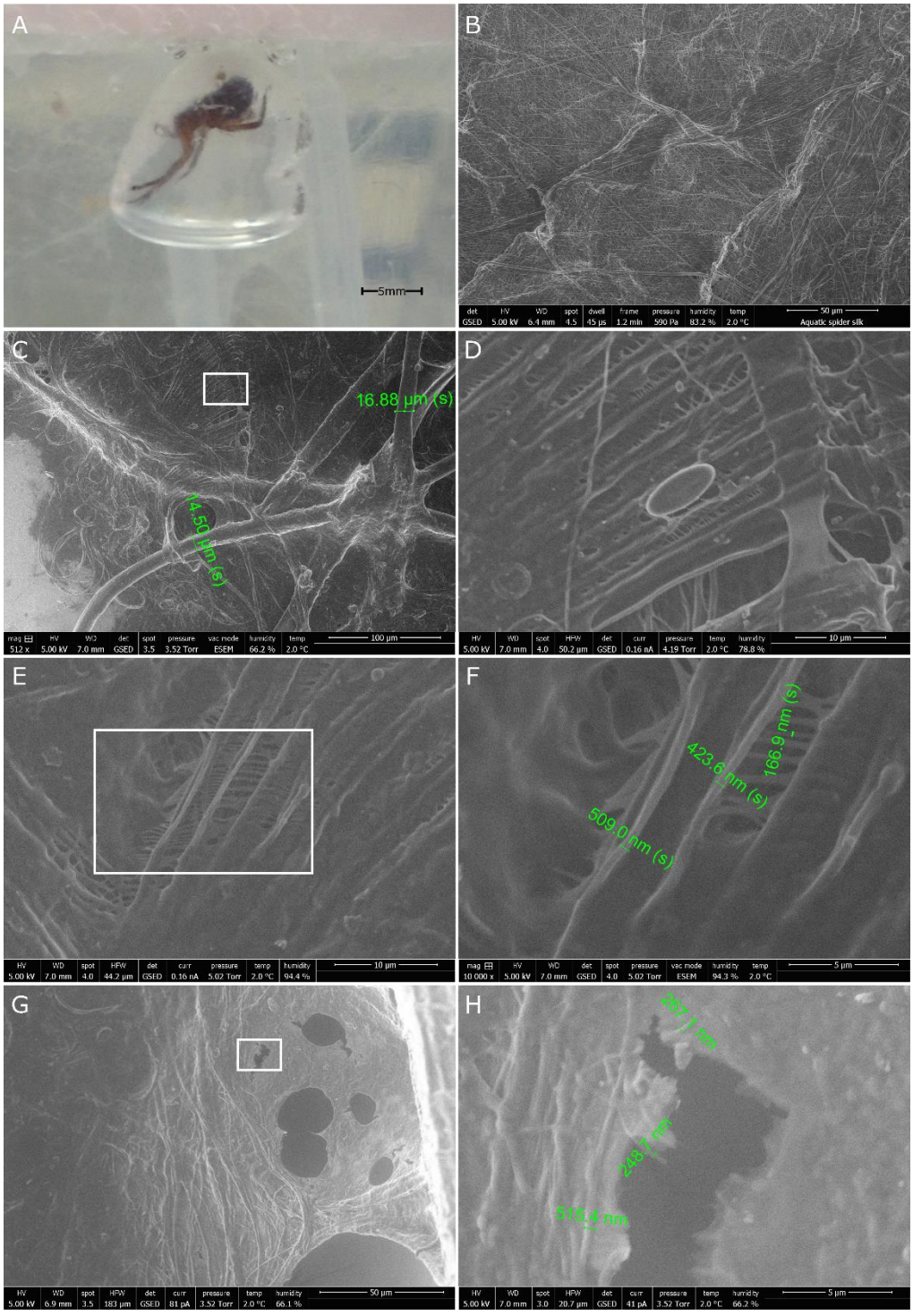


Figure 5.1. Photographic and eSEM images of the diving bell formed by one adult female *A. aquatica.*
Images B – H are eSEM images of one sample of the wall of the diving bell pictured in A. B: Overview of the main wall of the bell (exterior). C: Edge of the bell with anchor lines visible, seemingly attached together (right).
D: High magnification image showing marked area in C; fine fibres and ribbon-like structures are visible, with a diatom in the centre of the image. E: A second section of the exterior wall of the diving bell showing structures similar to image D. F:

Higher magnification of area marked in E showing fine fibres held together by an amorphous gel-like coating.
G: A third area of the diving bells showing a partially dehydrated amorphous region demonstrating perforations and tearing due to shrinkage. H: High magnification view of torn area highlighted in G, showing miniscule fibres embedded in the structure.

5.4.3. Structure of *C. angustiarum* web

C. angustiarum were reluctant web spinners, and we were unable to collect a dragline. Individuals occasionally produced small, flat sheet webs across the top of the moss substrate, but more commonly cocoon-like webs attached underneath bark strips, in which the spiders resided. When removed from the enclosure, these webs were inelastic and appeared to have a film-like coating, making patches of web appear amorphous to the naked eye.

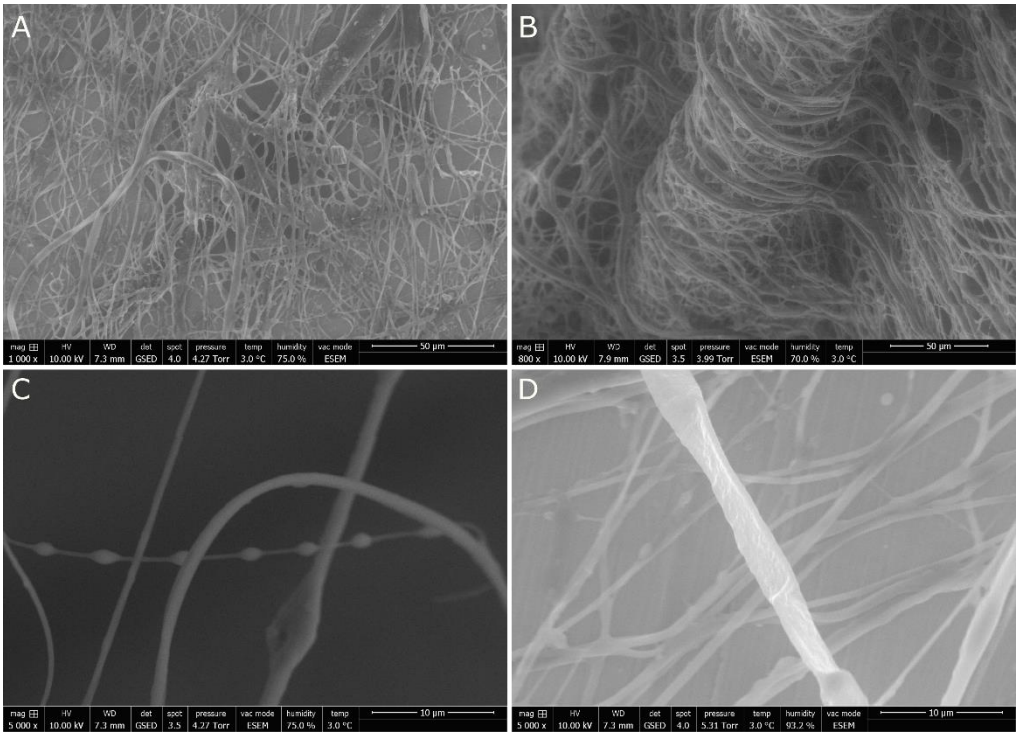


Figure 5.2. eSEM images of the web formed by one adult female *C. angustiarum.*
A & B: Overview of the web showing haphazardly laid strands of silk.
C: Single strand of "beads on a string" silk.
D: Ribbon-like structure formed of two strands of silk joined by glue. Another strand of "beads on a string" silk is visible in the background.

Viewed under the eSEM, the *C. angustiarum* web was composed of small threads haphazardly woven in a mesh, not unlike the *A. aquatica* web. However, there was no apparent layer of any hydrogel-like substance coating the threads; a small amount of amorphous material was viewed initially (see Figure 5.2 A) but disappeared shortly after the humidity in the specimen chamber was reduced, suggesting this material was condensed water. Reducing the humidity in the chamber below 70% induced no obvious change in the structure or arrangement of the web. Some instances of what appear to be silk coated in glue were seen ("beads on a string", typically seen as flagelliform silk coated with aggregate glue in orb webs) (Figure 5.2. C & D) and ribbon-like structures (Figure 5.2. D). A small number of structures and some debris were visible in the samples viewed, including diatoms, fragments of cuticle (likely from a previous moult), a single plumose setae and stacks of diatoms. A small ball formed of cube-like shapes measuring approximately 25µm was also observed.

5.4.4. Spinnerets

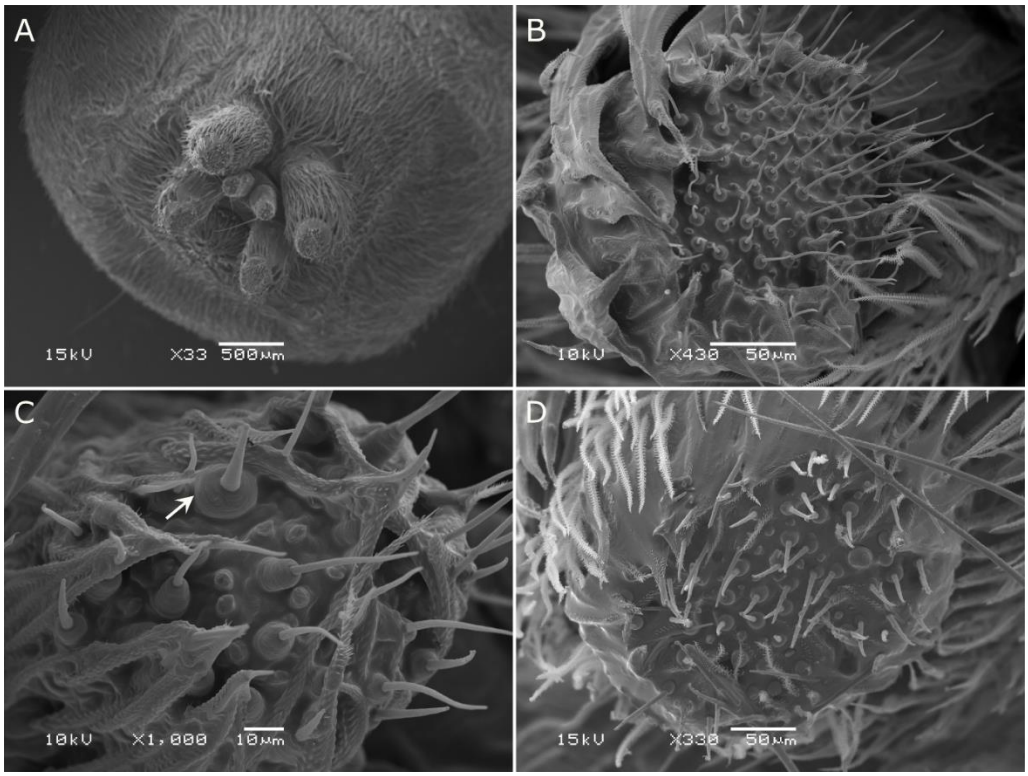


Figure 5.3. SEM images of the spinnerets of an adult female *A. aquatica.*
A: View of the spinnerets with anterior pair of spigots to top right. A layer of plumose setae (feathered hairs) is present, coating all spinnerets with the exception of the spinning fields, and across the abdomen.

B: Image of the spinning field of the posterior lateral spinneret (left of bottom pair in A) showing multiple aciniform spigots and some tartipores (vestigial spigots formed after a moult, now reduced in size, closed and missing extrusion nozzle). Tubuliform spigots are not visible. The hairs and spigots around the edge of the extrusion area appear to be coated in an unknown substance which obscures the structures beneath.
C: Spinning field of the posterior medial spinneret (left of centre pair in A) showing multiple aciniform spigots, a single minor ampullate spigot (arrow) and some tartipores. The hairs and spigots around the edge of the spinning field appear to be coated in an unknown substance although some parts of these structures are visible.
D: Spinning field of the anterior lateral spinneret (left of top pair) showing multiple piriform spigots and tartipores. Major ampullate spigot not visible. The hairs and spigots around the edge of the spinning field appear to be coated in an unknown substance which extends onto the surface of the spinneret; some hairs can be seen protruding from this coating and some spigots are partially obscured.

A. aquatica has three pairs of spinnerets: anterior lateral, posterior medial and posterior lateral. SEM imaging shows them to be mostly covered by the same hydrophobic hairs as found on the abdomen (Neumann and Woermann, 2013) but with the tip, where the spigots are mounted, exposed (Figure 5.3 B, C & D). Table 5.3 shows a summary of the spigots present on each pair. Where observed, spigots appeared similar to those of related species with no unusual or novel structures; the major ampullate spigot is not visible in these images, likely due to the gel-like coating.

Clear SEM images were difficult to obtain due to an unusual gel-like coating adhered to the plumose setae and coating a number of spigots on the end of each spinneret. This coating was not observed on any other surface of the spider. The spigots of adult *A. aquatica* specimens examined also appeared similar to related taxa.

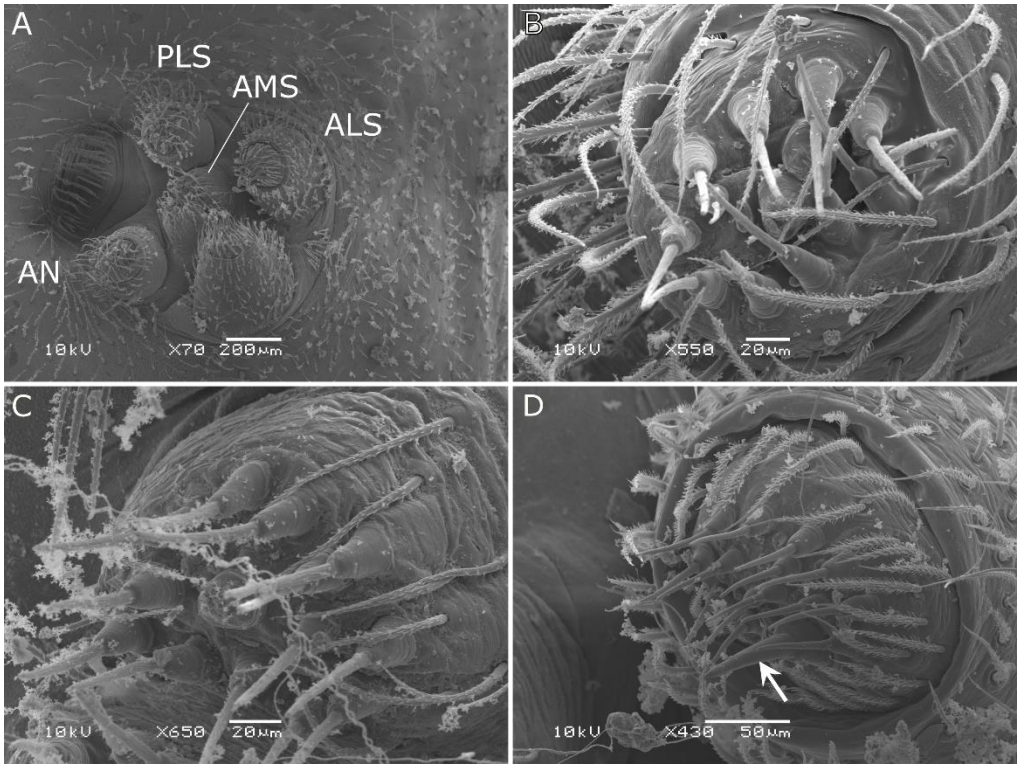


Figure 5.4. SEM images of the spinnerets of an adult female *C. angustiarum.*
A: View of the spinnerets with the anterior lateral pair of spigots to right (ALS) and anus (AN) on far left. A layer of plumose setae (feathered hairs) is present, coating all spinnerets and across the abdomen.
B: Image of the spinning field of the posterior lateral spinneret (left of bottom pair in A) showing multiple aciniform spigots
C: Spinning field of the posterior medial spinneret (PMS) showing multiple aciniform spigots and one tartipore to the centre. Some debris is present, likely fungal growth.
D: Spinning field of the anterior lateral spinneret showing multiple piriform spigots and a single major ampullate spigot (arrow).

Like *A. aquatica*, *C. angustiarum* has three pairs of spinnerets. They are covered by a finer coating of plumose setae, but these appear smaller and are fewer in number. A summary of the spinnerets and spigots found can be found below in Table 5.1. In Figure 5.4 B – D, debris and structures resembling fungal hyphae can be observed. The spinnerets appear partly deflated, with the spinning field partly retracted into the spinneret. Overall, the number of spigots present appears to be reduced compared with *A. aquatica*.

Table 5.1. Description of the spigots found in adult *A. aquatica* and *C. angustiarum.*

	Spinneret	Spigots	Number and location
A. aquatica	Anterior lateral	Major ampullate	Single, medial-frontal edge
		Piriform	Multiple
	Posterior medial	Minor ampullate	Single, on medial frontal edge
		Aciniform	Multiple
		Tubuliform (female only)	Multiple, along frontal edge
	Posterior lateral	Aciniform	Multiple
		Tubuliform (female only)	Multiple, along medial-frontal edge
C. angustiarum	Anterior lateral	Major ampullate	Single, medial-frontal edge
		Piriform	
	Posterior medial	Aciniform	Multiple
	Posterior lateral	Aciniform	Multiple

5.4.5. Silk gland morphology

A. aquatica silk glands

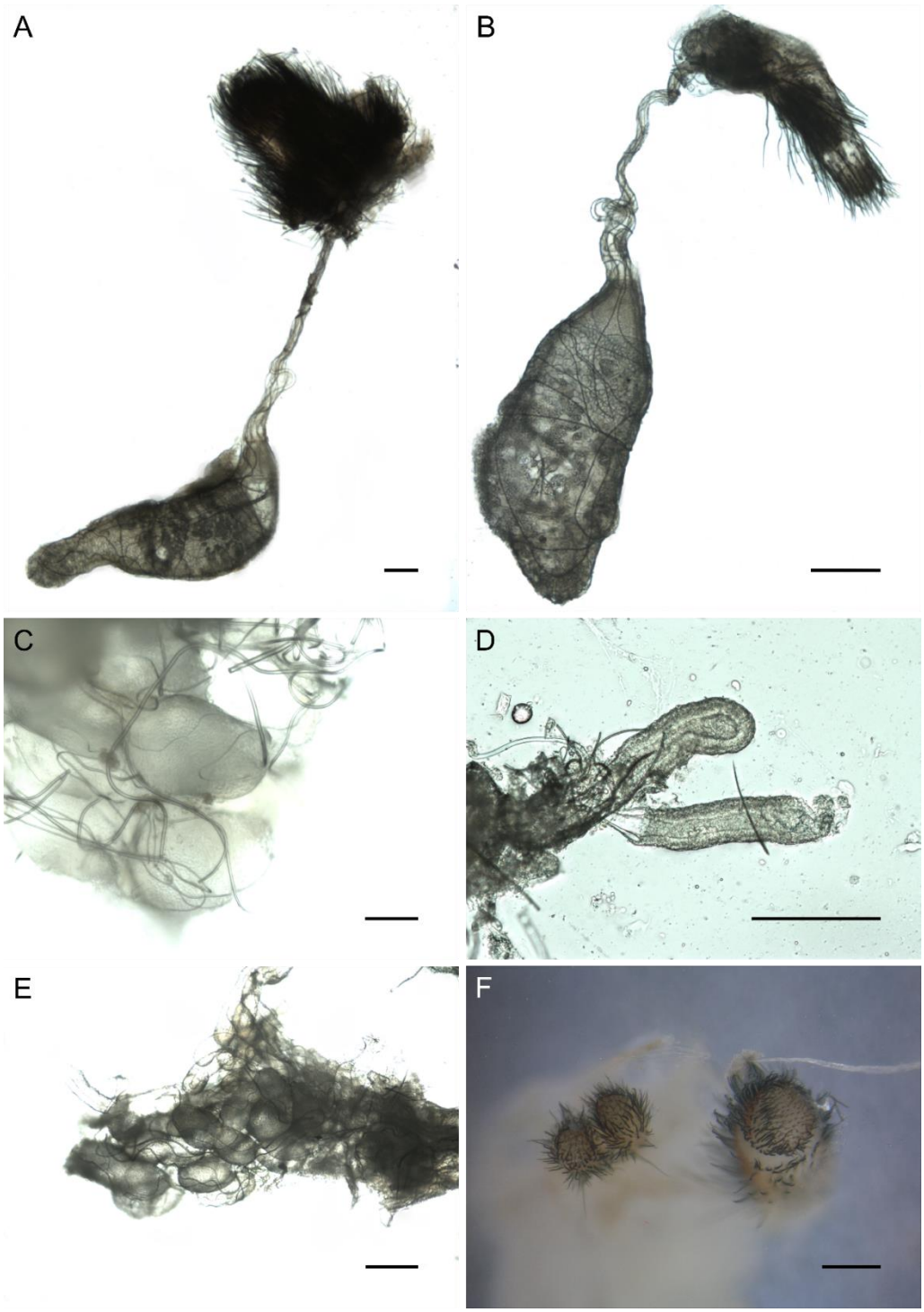


Figure 5.5. Optical microscope images of the dissected silk glands of *A. aquatica*.
A: Major ampullate gland with duct terminating in the dissected anterior lateral spinneret. B: Minor ampullate gland with duct terminating in posterior medial spinneret. C: Piriform glands. D: Tubuliform glands. E: Aciniform glands. F: Exterior edge of

the spinning field of spinnerets, showing plumose setae and spigots. Scale bar 200μm.

Anterior lateral spinnerets: Two types of gland, one large pair of major ampullate (Figure 5.5 A) and a number of small piriform glands open on the anterior lateral spinnerets. The secretory part of the major ampullate gland is relatively short and tubular; secretory zone B is shorter and more translucent compared with secretory zone A. The spinning duct is relatively short and carries a short loop on its most distal part, which lies close to the duct funnel (Figure 5.5 A). The piriform glands are small, ovoid, with short spinning ducts. Separate secretory zones are not apparent (Figure 5.5 C).

Posterior median spinnerets: Three types of glands including one large pair of minor ampullate glands, several tubuliform and multiple aciniform glands open on the posterior median spinnerets. The minor ampullate glands are smaller and have a relatively shorter secretory part when compared with the major ampullate glands, although they are similar in that secretory zone B is shorter and more translucent than secretory zone A (Figure 5.5 B). The tubuliform glands have white, tubular secretory parts composed of a single secretory zone (Figure 5.5D) and are small when compared with the relative size of tubuliform glands in related groups. The aciniform glands are small with ovoid secretory parts; although separate secretory zones are not apparent, the ducts of some of the aciniform glands are relatively long (Figure 5.5 E).

Posterior lateral spinnerets: A number of tubuliform and aciniform glands open on the posterior lateral spinnerets. These have the same morphology as seen in those that open on the posterior median spinnerets.

C. angustiarum silk glands

The silk glands of *C. angustiarum* are reduced both in number of types and in size when compared with those of *A. aquatica*.

Anterior lateral spinnerets: One large pair of major ampullate glands and many small piriform glands open on the anterior lateral spinnerets. The secretory part of the major ampullate glands is tubular, slightly more elongated than in *A. aquatica* and the border between secretory zones A and B is unclear. The spinning duct is relatively short and carries a short loop on its most distal part, which lies

close to the duct funnel. The piriform glands are elliptical and more elongated than in *A. aquatica*, with short, relatively wide spinning ducts.

Posterior median and lateral spinnerets: Small aciniform glands are the only glands to open on the posterior median and posterior lateral spinnerets.

5.5. Discussion

Argyroneta aquatica silks operate in a physical environment that could be considered "extreme" for a spider, as it remains to date the only species known to live a submerged lifestyle. Nevertheless, our study shows that the morphology of their spinning apparatus (silk glands, spinnerets and spigots) does not significantly differ from that of the related species *C. angustiarum* and does not appear to have atypical features, as it is unlikely the unusual material coating the plumose setae in the SEM images (Figure 5.3) may be concealing some small structures. This finding suggests that the hydrogel coating on the diving bell must be produced in the silk glands and extruded from the spigots.

5.5.2. Structure of the diving bell of *A. aquatica* and web of *C. angustiarum*

The eSEM images of an *A. aquatica* diving bell show a correlation between the location of the hydrogel and smaller-diameter silk fibres, with the thinner (<1µm) fibres seemingly merging into amorphous sheets and ribbons (Figure 5.1 D-H). This would coincide with the spider utilising multiple spigots to produce silk threads and hydrogel simultaneously. The presence of two piriform spidroins and two aciniform spidroins in our dataset (see section 4), and multiple spigots of each on the posterior and anterior lateral spinnerets (Figure 5.2 D), suggests these silks and glue make up a significant proportion of the diving bell. Piriform glue is known to be hydrophilic and will spread on contact with hydrophilic surfaces (Grawe et al., 2014) and it is possible that the fibrous hydrogel we see in the eSEM may comprise of piriform spidroins and glue, used to cement together layers of aciniform thread. The ability of piriform glue to spread on hydrophilic surfaces suggests it may behave in a similar manner in an aquatic environment. However, it is difficult to ascertain the exact origin of individual structures; the aciniform spigots on the posterior lateral and posterior medial spinnerets are of a similar size to the piriform spigots on the anterior lateral spinnerets, so an assumption made on fibre diameter is unfounded.

On the surface of the wall of the diving bell we identified a small number of individual fibres with a diameter of approximately 5µm. Based on the larger size of the minor ampullate spigot compared with the aciniform and piriform spigots (Figure 5.1B-D), we suggest the minor ampullate gland is the source of these fibres. These and the thicker fibres (>15µm diameter, Figure 5.1 C), likely to have been produced from the unseen major ampullate spigots, appeared to be consistently uncoated with hydrogel. In Figure 5.1C, the proposed MaSp fibres appear to be connected to the wall of the diving bell at a single point in the manner of an attachment disc, with the large major ampullate fibres protruding from an apparently solid mass (Grawe et al., 2014).

The web of *C. angustiarum* initially appeared to have a partial coating of an amorphous substance. However, as the humidity was decreased from 100% to around 75% this feature disappeared, indicating this was a layer of water. The structure of the web reflected that of *A. aquatica*'s diving bell in that multiple layers of silk fibre were laid haphazardly and a number of ribbon-like features. The samples examined did not contain any structures that appeared to form an attachment disc or anchor line. Unlike in the *A. aquatica* web sample, at least two occurrences of "strings of beads" structures were visible. As *C. angustiarum* is not known to produce aggregate glue or flagelliform silk (the silk types associated with this formation in orb webs (Elettro et al., 2016)), speculation as to the composition of this structure is merely that, as the only known glue produced by this spider is piriform glue, known to be used for attachment discs as it dries once produced. It is also worth noting that when viewed at 800x magnification the diameters of the strands of silk used in the *C. angustiarum* appear to be greater than those in the diving bell (Figures 5.1B and 5.2 B).

5.5.3. SEM of spinnerets

The presence of densely packed hydrophobic hairs around the edge of the spinning field of *A. aquatica*'s spinnerets, much greater than seen in *C. angustiarum*, may allow for a large enough air space to provide a "dry" environment in which extruded silk dope may harden. If so, this would demonstrate an adaptation by *A. aquatica* to underwater life that allows the retention of a spinning process otherwise better suited to terrestrial spiders.

The appearance of the material coating the hydrophobic hairs surrounding the spinning field in Figure 5.3 is not unlike the amorphous hydrogel coating the

diving bell in Figure 5.1. It is interesting to note that this material does not appear to have cracked or shrunk as part of the sample preparation process when preparing the spider specimen for SEM imaging; we would expect the hydrogel to shrink as it dehydrates, as seen in (Neumann and Kureck, 2013, De Bakker et al., 2006). If this is hydrogel, it may be possible that storage in ethanol has allowed for the material to partly dehydrate prior to critical point drying and preserving the original structure somewhat. This occurrence was not unique to the specimen shown here and was seen in a second specimen imaged.

The *C. angustiarum* spinnerets were coated in a large amount of fungal hyphae, seen in a number of specimens and completely obscuring the spigots in some (Figure 5.4 C). During the study a number of specimens died within a short period of time and it is possible this is related to the material seen. Otherwise, no unusual structures were observed during this study and all spigots anticipated could be viewed.

5.6. Conclusions

Combining these results with those of the previous chapters, we propose that the sheet-web component of the diving bell of *Argyroneta aquatica* is likely composed of aciniform and piriform silks, perhaps with minor ampullate silk used to lay a temporary structure to which these silks may be attached. The larger-diameter major ampullate silk fibres appear to be used exclusively for anchor lines. Together, these silks have been used in combination to form a structure that is strong and secure, but also able to hold a bubble of air underwater and allow gaseous diffusion.

In addition to furthering our understanding of the structure of the diving bell of *A. aquatica*, these results are significant to research into the utilisation of synthetic silks as *A. aquatica* show that silk fibres may be readily used in a structure with hydrophilic qualities without significantly altering the form of the silk or the apparatus designed to produce it. *A. aquatica* are unique in their adaptation to a fully aquatic lifestyle and their utilisation of silk as a physical gill adds yet another exciting characteristic to the list of functions that may be exploited in both material and medical research, as well as adding an interesting alternate function to piriform silk.

6. Modifying the physical properties of spider silk by combining with dispersed carbon nanomaterials

6.1. Introduction

Identifying possible composites and understanding how different compounds and polymers interact is important in the process of developing novel materials and advancing materials science. With recent advancements in the production of synthetic spider silk (Andersson et al., 2017, Stark et al., 2007, Xia et al., 2010), there is the potential to study how spider silk interacts with a range of compounds (Harvey et al., 2017, Lepore et al., 2017, Steven et al., 2013, Lee et al., 2009) and to know the physical properties of a silk composite, particularly if we are to modify the silk protein itself (Harvey et al., 2017). This not only allows a better understanding of the individual materials, but also allows for future development of technologically advanced products such as flexible, biodegradable or biocompatible electronics (Steven et al., 2013, Lepore et al., 2017) and biocompatible tissue scaffolds (Allmeling et al., 2008, Allmeling et al., 2006, Steins et al., 2015, Harvey et al., 2017). With the ready availability of carbon nanomaterials and the production of synthetic silk with native silk-like properties, more data and analysis of the combinations formed by carbon nanomaterials (CNM) and native spider silk is useful in predicting the material properties of silk-CNM composites, particularly where significant modifications to specific properties such as the tensile strength or toughness can be made (Lepore et al., 2017, Lee et al., 2009).

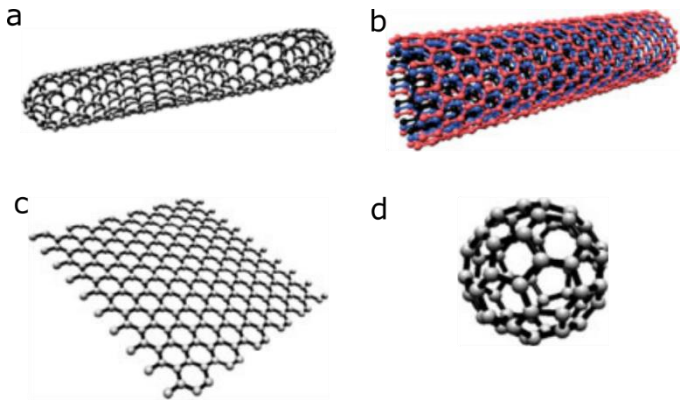


Figure 6.1. Schematic representation of some of the available carbon nanomaterials.
a. Single-walled carbon nanotube, b. multi-walled carbon nanotube, c. graphene, d. buckyball. Modified from Figure 1, Jiménez-Soto et al. (2012) (Jiménez Soto and Aranzana, 2012).

Carbon nanomaterials (CNM; also known as carbon nanostructures) come in a range of shapes and sizes based on a single-atom thick sheet of carbon, named graphene. These can include buckyballs, single-walled carbon nanotubes (SWNT), multi-walled carbon nanotubes (MWNT; formed either of a number of concentric single-walled nanotubes, or a single sheet or ribbon of carbon rolled into a multi-layered tube)(Lee et al., 2014) and graphite (small blocks formed of layers of graphene) (Figure 6.1) (Jiménez Soto and Aranzana, 2012). Each of these structures can be produced with a range of dimensions, formed only of a layer of carbon a single atom thick. CNM can be combined in multiple ways, forming fibres (Dalton et al., 2003), nanobuds (where a CNM is attached to the outside of a nanotube), peapods (a CNM such as a buckyball held inside a nanotube) (Lu and Chen, 2005) and pillared graphene (two sheets of graphene separated by perpendicular single-walled carbon nanotubes). CNM can also be modified with the addition of an amine group to allow covalent bonding with a range of compounds and polymers (Ramanathan et al., 2005).

6.1.1. Properties of carbon nanomaterials

Structure and electrical properties of CNM

Carbon nanotubes (CNT) have become a popular area of study in the search for technological advancement, as they demonstrate extraordinary thermal (Han and Fina, 2011), electrical (Scarselli et al., 2012) and mechanical (Yu et al., 2000a, Yu et al., 2000b) properties as a result of their structure and composition. Each nanotube is formed of a single-atom thick sheet of covalently-bonded carbon atoms with three of the available bonding sites bonded; the fourth electron is free to migrate across the surface of the tube. These C-C bonds can be orientated in three ways, resulting in formations known as armchair, zigzag and chiral. This directly affects the electrical properties of the nanotube, as a flow of electrons may only be directed along the axis of each respective carbon nanotube. The armchair variant (along with multi-walled carbon nanotubes), for example, demonstrates conductive properties similar to highly conductive metals, whereas zigzag and chiral orientated nanotubes act as semiconductors (Zsoldos et al., 2005, Scarselli et al., 2012, Lu and Chen, 2005, Treacy et al., 1996). In addition to electrical conductivity, the tubes of concentrically arranged multi-walled nanotubes can slide with minimal friction (Cumings and Zettl, 2000). Individual layers of graphite demonstrate similar low-friction properties; graphite layers are held

together by Van der Waals forces and hence are able to slide across each other with little force, making them useful as a lubricating agent (Berman et al., 2013).

Mechanical properties of CNM

Single-walled carbon nanotubes (SWNT), multi-walled carbon nanotubes (MWNT) and graphene have a high strength and stiffness (Table 6.1) due to the covalent bonds between individual carbon atoms. However, SWNT and MWNT are more readily deformed in the radial dimension than longitudinally; the force imposed by Van de Waal's interactions is sufficient to deform the tubes (Ruoff et al., 1993) but they demonstrate a high level of elasticity (radial Young's Modulus of SWNT and MWNT reaching a maximum value of 59GPa and 30GPa respectively), although this is relative to the diameter of both SWNT (Yang and Li, 2011) and MWNT, the latter of which is also affected by the number of nested tubes (Palaci et al., 2005).

Table 6.1. Average values of tensile properties of CNM, dragline silks, steel and Kevlar.

Material	Tensile Stress (MPa)	Strain	Young's Modulus (GPa)	Toughness (MJm^{-3})	Source
Single walled carbon nanotubes	13,000 – 52,000	0.05	260 – 1,040	345 – 1378	(Wang et al., 2010, Yu et al., 2000a)
Multi-walled carbon nanotubes	11,000 – 63,000	N/A	720 - 950	N/A	(Yu et al., 2000b)
Graphene	130,500	0.13	1,004	8548	(Lee et al., 2008)
Dragline silk (*Caerostris darwinii*)	1,652	0.52	11.5	354	(Agnarsson et al., 2010)
Dragline silk (*Nephila clavipes*)	1,215	0.17	7.14	111	(Swanson et al., 2006a)
Dragline silk (*Holocnemus pluchei*)	1,244	0.15	8.29	115	(Swanson et al., 2006a)
Dragline silk (*Araneus diadematus*)	1,100	0.27	4.07	160	(Gosline et al., 1999)
Steel (ASTM-A36)	400 – 550	0.20	2 – 2.75	40 – 55	(MatWeb)
Kevlar 29	3,600	0.04	90	72	(K. Seth et al., 2009)

6.1.2. Modification

The production of synthetic spider silk has been achieved by several commercial entities, all of which are exploring the potential of silk-based products (Arthur, 2017, Hornyak, 2013, Franklin-Wallis, 2017). This, combined with opportunities in material development and application, increases the possibility to test how native silks behave when coated with materials with properties of interest, such as carbon nanomaterials. With recent developments in synthetic silk research, rearing spiders for their silk will eventually become unnecessary and it will be possible to combine CNM with soluble or unformed synthetic spider silk proteins as well as pre-spun synthetic spider silk fibres. Knowledge of the application methods and subsequent effect of coating materials will indicate whether such studies should extend to synthetic silk.

A significant factor in silk modification is the silk itself; spider silk reacts to high levels of humidity (>70%) or wetting by supercontracting (Agnarsson et al., 2009a); a permanent molecular change in the structure of the silk proteins where the fibre absorbs water and shrinks by up to 50% of its original length (Blackledge et al., 2009). Dragline silk also demonstrates a cyclic response to relative humidity, in addition to supercontraction that can be fine-tuned by modifying environmental levels of humidity (Agnarsson et al., 2009b, Blackledge et al., 2009). Besides changes in the length of silk fibres, supercontraction affects both the tensile strength and toughness of silk fibres, resulting in an overall reduction in tensile strength and stiffness (Blackledge et al., 2009, Steven et al., 2013).

Although the structure and mechanical properties of raw silk have been studied extensively, the effect on silk after exposing silk-producing animals and raw silk to CNM has not been investigated thoroughly. Silk, originating both from spiders and the silkworm *Bombyx mori*, has been modified a number of ways. Both have been coated with metal ions (Lin et al., 2014, Lee et al., 2009) and the process of producing synthetic spider silk has been modified in such a way that functionalised compounds such as fluorophores and antibiotics can be bound to the individual silk spidroins (Harvey et al., 2017).

Perhaps the first publication demonstrating CNM-coated silk was presented by Steven *et al.* (Steven et al., 2013). Here, the authors coated the dragline silk of *Nephila clavipes* with amine-functionalised multi-walled nanotubes, mixing a bundle of silk fibres with powdered nanotubes before adding water, pressing and

shearing the bundles. This resulted in a uniformly-coated bundle, which could be separated into individual fibres that demonstrated a decrease in electrical resistance (native silk is highly resistive and acts as an electrically-insulating material). The fibres also showed a greater extensibility and toughness and equal tensile strength relative to that of supercontracted, uncoated silk fibres (Steven et al., 2013). However, there nevertheless remained a decrease in tensile properties compared with uncoated silk, likely due to the effect of supercontraction experienced due to the process by which the CNM were applied.

More recently, a publication by Lepore *et al.* (Lepore et al., 2017) presented data showing a significant increase in the tensile strength and toughness of silk produced by spiders that had been exposed to dispersions of either graphene or one of two types of single-walled carbon nanotube. The authors sprayed a dispersion of CNM into the enclosures of a range of spiders and measured the tensile properties of silk collected. According to the research, the tensile properties of silk combined with single-walled carbon nanotubes resulted in an increase in strength from 1.71GPa to 5.4GPa and toughness from 476.4MJm^{-3} to 2143.6MJm^{-3}. With these values, the silk-CNM composite vastly exceeds that of any biological material and the majority of man-made composites (Table 6.1) and highlights a material composite which may later be shown demonstrate other interesting properties. However, whist this data has been available since 2015 in the form of a preprint (Lepore et al., 2015), as yet the results are not yet reported to have been reproduced by others.

6.1.3. Overview

This study examines the tensile properties of silk collected from the spiders *Pholcus phalangioides* (as a direct comparison with the study by Lepore *et al.*) and *Araneus diadematus* (a more commonly studied species; the tensile properties of the silk are recorded in Gosline *et al.* (Gosline et al., 1999)). The use of multiple species allows for examination of the effect of CNM on dragline silks with different diameters and tensile strengths, as these are known to vary between species (Swanson et al., 2006a, Madsen et al., 1999).

Both *A. diadematus* and *P. phalangioides* use webs to catch prey, although the shape and use of these webs differ. As an orb-weaving spider *A. diadematus* builds an orb web composed of different types of silk, with dragline silk forming the radial framework which supports a glue-coated, elastic capture spiral. These webs

are used to catch airborne prey and as such are designed to absorb the energy of impact whilst the capture spiral adheres to the prey, preventing escape (Soler and Zaera, 2016). *P. phalangioides* build tangle webs in which they wait; when vibrations produced by another insect (or spider) in the web are felt, *P. phalangioides* rapidly move to capture and wrap the intruder in silk before it can escape (*BAS*).

Here, we measure native silk collected and spray-coated with CNM dispersions, and silk which has been collected directly from spiders after they have been exposed to the same dispersions. We utilise the same CNM as in the study by Lepore *et al.* (Lepore et al., 2017) and add multi-walled carbon nanotubes to expand the dataset. We collect dragline silk in a repeatable controlled manner which also allows for the testing of individual draglines, and care for the spiders in an attempt to reduce the mortality rate previously reported and negate any potential influence of starvation on the properties of the silk.

We utilise this data to explore the reproducibility of incorporation of carbon nanomaterials into spider silk and evaluate whether using CNM to increase the tensile properties of native silk is viable. Ultimately, this will inform future experimental research into combinations of synthetically-produced spidroin fibres and CNM.

We compare our results with those presented by Lepore and discuss our respective methodologies in the context of reproducibility and experimental design.

6.2. Methods

6.2.1. Carbon nanomaterial preparation and Raman spectroscopy

Note: When not referring to a specific carbon nanomaterial used in these experiments, "CNM" (carbon nanomaterials) is used as shorthand for any or multiples of the four carbon nanomaterials used in this study (CoMoCAT single-walled nanotubes, P2 single-walled nanotubes, multi-walled nanotubes and graphite).

All reagents were purchased from Sigma-Aldrich, UK and used without further purification. Carbon nanotubes were purchased from a range of suppliers, as detailed below (Table 6.2).

Table 6.2. Commercial suppliers of CNM and their dimensions.
d_{NT} and L_{NT} refer to the mean diameter and length of carbon nanotubes as determined by statistical analysis of transmission electron micrographs.

Carbon nanomaterial	TEM d_{NT} / nm	TEM L_{NT} / µm	Commercial name	Commercial supplier
CoMoCAT SWNT	0.7±0.1	1.5±1.0	6,5-SWeNT	Sigma Aldrich, UK
AD SWNT	1.3±0.1	3.2±1.5	P2 SWNT	Carbon Solutions Inc., USA
CVD MWNT	14.4±6.0	3.2±1.6	PD30L520 MWNT	Nanolab, USA
Graphite	-	-	Graphite	Fisher Scientific, UK

Water was purified (> 18 MΩ cm) using a Barnstead NANOPure II system. UV-vis-NIR spectra were recorded in solution using 1 cm quartz cuvettes using a Perkin-Elmer Lambda 750 UV-vis-NIR spectrophotometer at a scan rate of 240 nm min^{-1} over the range 400-1400 nm. The position and intensity of the absorbance maximum was obtained using the "find peaks" function in UV WinLab ES software with the threshold set at 0.01. Dynamic and phase analysis light scattering was performed using a Malvern Instruments Nano-ZS Zetasizer at room temperature. Raman spectra were recorded using a Horiba LabRAM Raman microscope, with a laser wavelength of 532 nm and a 600 lines/mm grating and fitted with a Synapse CCD detector. The Raman shift was calibrated using the Raleigh peak and the 520.7 cm^{-1} silicon line from an Si(100) reference sample. Quoted values are the average of three measurements.

Raman spectroscopy is a technique which takes advantage of the interaction of a material with a beam of monochromatic light of a known wavelength (i.e. laser). Individual molecules reflect, or scatter (as opposed to absorbing or transmitting) the light in a manner unique to each molecule, resulting in an increase or decrease in the wavelength of some of the scattered light unique to each molecule. By measuring the frequencies of light scattered by the sample, visualised on a chart (e.g. Figure 6.2), the chemical composition can be determined (Vandenabeele, 2013). Measurements are taken looking at individual points across a sample: if readings indicating a molecule of interest is present, multiple measurements can be taken across the sample and "mapped" to a light-microscope image (typically taken when positioning the sample inside the spectroscope) to show the location(s) of the molecule within the sample (for an example, see Figure 6.6).

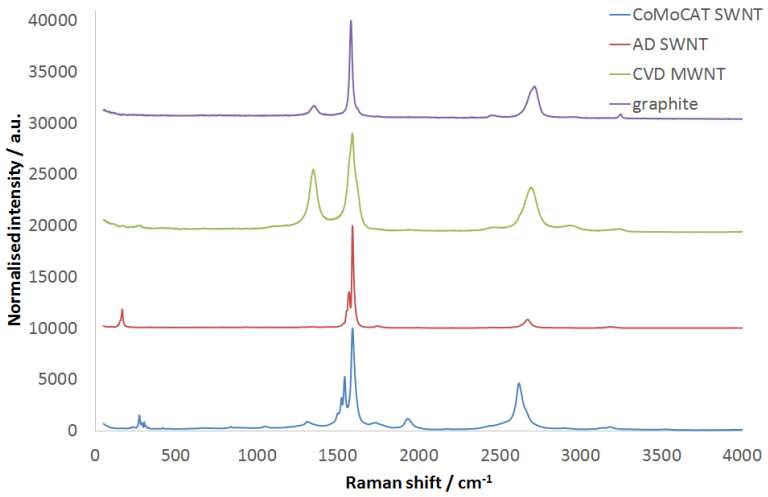


Figure 6.2. Chart showing the Raman spectra of the four CNM used in this study.
The G peak is visible at 1580nm, the D2 peak at 2700nm. D peaks are visible to the left of the G peak in the graphite (top) and MWNT (first from top) spectra. Graphite, MWNT and P2 SWNT (labelled above as AD SWNT, second from top) spectra are offset for clarity.

Graphene and CNT are characterised on Raman spectra by the presence of D (1350nm), G (1580nm) and D2 (2700nm, or twice the value of D) peaks. The G peak is induced by the vibration of carbon atoms within the CNM, whereas the D (or disorder) peak is caused by anomalies within the structure of the CNM, including structural damage, absent or substituted atoms such as nitrogen. This peak is minimal in completely perfect samples of CNM (such as the SWNT used in this experiment), but typically more prevalent in graphite and MWNT samples. As the disorder increases within samples the D peak will increase in size (Childres et al., 2013). The 2D peak is used in combination with the G peak as a diagnostic for CNM, as it situated away from the G peak (which may be surrounded by noise) and is easily distinguishable (Figure 6.2.).

The carbon nanotube suspensions were prepared according to the following procedure, modified from that as described by Lepore *et al.* (Lepore et al., 2017): Carbon nanotubes (4 mg) were added to a solution of sodium deoxycholate (40 mL, 2% w/v in deionised water) and the mixture homogenised by ultrasonication using a Sonics Vibracell CVX130 Ultrasonic Processor (20 kHz) for 2 hrs at room temperature. The dispersions were then subject to ultracentifugation using a JA-30.50 fixed angle rotor in a Beckman Coulter Avanti J-30I ultracentrifuge (100,000

x g) for 2 hours at 18°C. The pale-grey supernatants were decanted by pipette and the stored at 4°C.

Graphite (100 mg) was added to a solution of sodium deoxycholate (10 mL, 2% w/v in deionised water) and the mixture homogenised by ultrasonication using a Sonics Vibracell CVX130 Ultrasonic Processor (20 kHz) for 10 hrs at room temperature. The dispersions were then subject to ultracentrifugation using a Fixed Angle HIGHConic™ rotor in a Fisherbrand accuspin 400 centrifuge (4,300 x *g*) for 30 mins at room temperature. The grey supernatants were decanted by pipette and the stored at 4°C.

The suspensions were characterised by UV-vis-NIR spectroscopy and dynamic and phase analysis light scattering. The nanotubes were also characterised by Raman spectroscopy (Figure 6.2). All aggregates of CNM within the dispersions were negatively charged (zeta potential: CoMoCAT SWNT= -50.4mV (σ=10.3), P2 SWNT= -41.5mV (σ=8.3), MWNT= -52.9mV (σ=8.1), graphite= -47.4mV (σ=8.5)), reflecting the surface chemistry of the surfactant. The high negative value supports a high colloidal stability and therefore the dispersions were suitable for spraying onto spider silk fibres.

6.2.2. Application of carbon nanomaterials

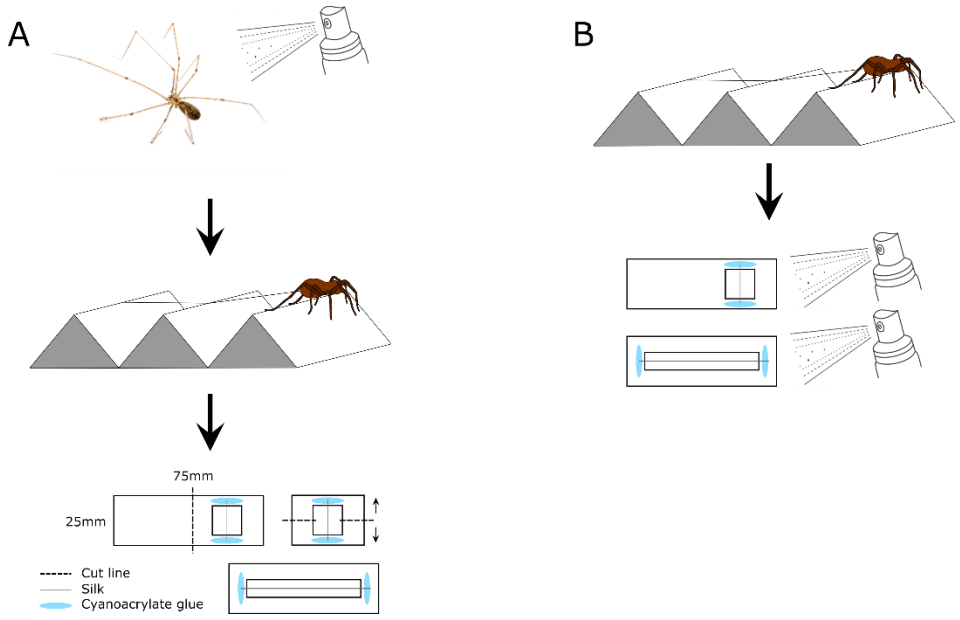


Figure 6.3. Representation of "spray on spider" (A) and "spray on silk" (B) application and collection process.
Dashed line and arrows in (A, bottom) show where card will be cut and the direction in which pieces will be separated during tensile testing.

"Spray on spider"

One day prior to exposure ("Day 0"), 10 dragline silk samples were obtained from 12 spiders for reference (for method see section 2.3). Leaving each spider in its own enclosure, spiders were sprayed twice directly onto the body with the aqueous dispersion of CNM or a surfactant-only control (100±5µL per spray) using a hand-held atomiser (Boots Travel Spray Bottle 50 mL capacity) on "Day 1". Silk samples were subsequently collected on days 3 and 10 and tensile tested (see 6.2.3 and Figure 6.3A).

"Spray on silk"

Pristine dragline silk samples from unsprayed spiders (collected as in section 2.3) were mounted on card frames and sprayed with the aqueous dispersion of CNM or a surfactant-only control (100±5µL per spray) using a hand-held atomiser (two pumps per sample; Boots Travel Spray Bottle 50 mL capacity) (see 6.2.3 and Figure 6.3B).

6.2.3. Tensile testing of silk samples

Initial testing of samples from a *P. phalangioides* and two *A. diadematus* was performed on a custom-modified piece of equipment based on a plotting machine. Treated and raw silk samples on card frames were placed in the holder and the supporting frame was cut, leaving the sample fibre intact and the sample was stretched to breaking point. Calibration of the equipment was performed at the beginning and end of each testing session. Further testing with new samples obtained from an additional *A. diadematus* at a later point was undertaken by transferring and securing the collected silk samples to plastic tabs with cyanoacrylate glue and removing the card frames. The samples were subsequently placed into a fibre tensile testing machine (Diastron Ltd. Linear Extensometer LEX 810) fitted with a 2000g load cell and stretched to breaking point. The temperature (19.6°C ±0.1°C) and humidity (42% ±1%) were recorded at the beginning and end of each testing session, and monitored throughout.

6.2.4. Measurement of sample diameter

Samples of dragline silk were secured to a metal stub with double-sided carbon tape then sputter-coated with gold *en vacuo* to a layer thickness of approximately 14nm using a Leica EM SCD005. The samples were imaged and the diameter of each of the two fibres measured using a JEOL 6060 Scanning Electron Microscope. As each sample of dragline silk includes two fibres, the cross-sectional area of the

sample was calculated by taking the mean diameter of both fibres at multiple positions and calculating the average cross-sectional area, which was then multiplied by 2.

6.2.5. Calculations of tensile strength, stress and Young's Modulus

Tensile strength is the maximum force that a material can sustain before fracturing, and this value, along with the known cross-sectional area (XS Area) of the material, can be used to calculate the maximum *stress* (Pa) the material can undergo.

$$Stress = \frac{Force\ (N)}{XS\ Area\ (m^2)}$$

The deformation of the sample material can be represented by *strain* (no units), calculated using the original and maximum length of extension.

$$Strain = \frac{Extention}{Original\ length}$$

Both *stress* and *strain* can then be analysed in respect to each other to determine the stiffness of a material, known as Young's Modulus (Pa).

$$Young's\ Modulus = \frac{Tensile\ Strength\ (\frac{kg}{m^2})}{Strain}$$

One further calculation that might be used in analysis is the *toughness* (Jm^{-3}) or deformation energy of a material; the amount of energy it can absorb before fracturing. This is relevant to the volume of the material involved, and is represented by the area under the stress-strain curve using the following integration equation (where MS = maximum strain, in this experiment also assumed to be strain at point of fracture):

$$\frac{Energy}{Volume} = \int_0^{MS} Stress * dStrain$$

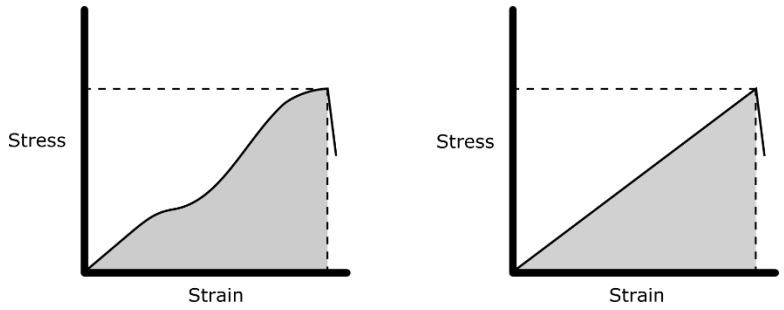


Figure 6.4. Area of toughness on a stress-strain chart.
Diagram showing representation of a tensile test (left) and simplified data using the same ultimate stress point (right). Toughness is calculated as the area under the line drawn from the data produced during the tensile test, with the point of ultimate tensile strength assumed to equal the point of fracture.

As this is an initial exploratory analysis, for speed and ease this equation has been simplified. The amended calculation has the disadvantage of assuming the boundary of the area under the stress-strain line is strictly linear (Figure 6.4), but the advantage of only requiring the input of the stress and strain at moment of fracture. This allows the calculation of the area under the stress-strain line to be substituted to that of the calculation of the area of a right-angled triangle;

$$Approximate\ Toughness = \frac{Stress\ (Pa) * Strain}{2}$$

6.3. Results

6.3.1. Measurements of silk fibres under SEM

Samples of untreated silk were measured under a SEM; all dragline silk samples examined were formed of two fibres of similar diameter (±0.1µm). Often only one fibre was fully visible during imaging as the samples became partially or fully twisted during or after spinning so all measurements taken were used to calculate the average diameter of both fibres (Table 6.3).

Table 6.3. Average measurements of diameter and subsequently-calculated cross-sectional area of silk fibre samples.
Sample marked * shows mean of multiple fibre measurements (n=23), standard error = 0.014.

Spider	Sample diameter (µm)	Cross-sectional area of individual fibres (µm²)	Cross-sectional area of sample (µm²)
P.p.	1.91	2.87	5.73
A.d. 1 & *A.d.*2	3.51	9.65	19.30
A.d. 3	1.64*	2.11	4.23

6.3.2. "Spray on spider" experimentation

Table 6.4. Record of spiders after application of CNM or control substances.
Spiders which went on to produce an egg sack after testing had completed are marked *

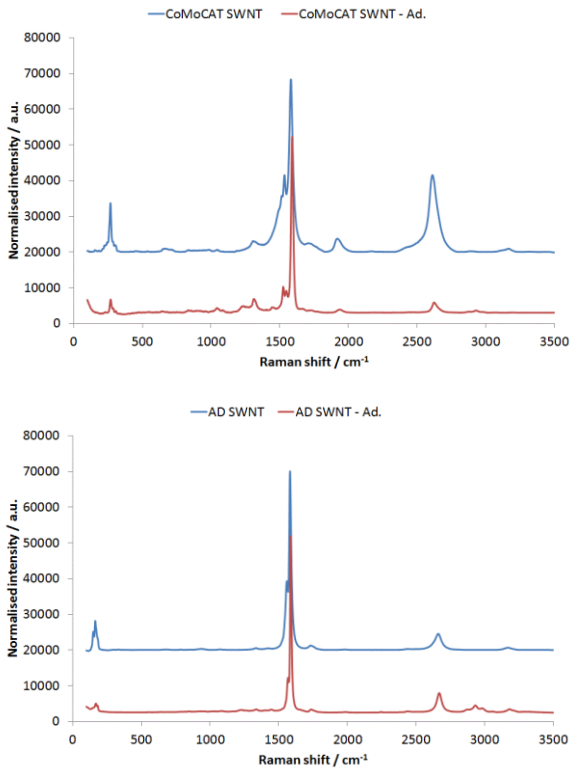
Treatment	Spider	Outcome
CoMoCAT single-walled nanotubes	*A. diadematus* (Adult ♀)	Died (Day 1)
	A. diadematus (Adult ♀)	Died (Day 1)
	P. phalangioides (Adult ♀)	Silk retrieved on Day 3, 10
	P. phalangioides (Adult ♀)	Silk retrieved on Day 3, 10
P2 single-walled nanotubes	*A. diadematus* (Adult ♀)	Silk retrieved on Day 3, 10 (produced egg sac on Day 8)
	P. phalangioides (Adult ♀)	Silk retrieved on Day 3, 10
Multi-walled nanotubes	*A. diadematus* (Adult ♀)	Silk retrieved on Day 3, 10*
	P. phalangioides (Adult ♀)	Silk retrieved on Day 3, 10
Graphite	*A. diadematus* (Adult ♀)	Silk retrieved on Day 3, 10*
	P. phalangioides (Adult ♀)	Silk retrieved on Day 3, 10
Surfactant in dH_20 (control)	*A. diadematus* (Adult ♀)	Silk retrieved on Day 3, 10*
	P. phalangioides (Adult ♀)	Silk retrieved on Day 3, 10

A total of 12 spiders were directly sprayed with either a CNM dispersion or a surfactant control (Table 6.4). Due to the results showing the greatest tensile strength of silks from CoMoCAT-treated spiders observed in Lepore *et al.* (Lepore et al., 2017), a larger group was used to test the CoMoCAT SWNT to increase the chance of replicating this finding. Both *A. diadematus* spiders sprayed with CoMoCAT single-walled nanotubes perished within 24 hours; all other spiders

survived experimentation and were eventually released. All surviving *A. diadematus* spiders produced egg sacks during or after testing. No *P. phalangioides* females produced egg sacks.

Raman spectroscopy analysis of silks from spiders sprayed with CNM dispersions

Silk samples collected from *A. diadematus* and *P. phalangioides* spiders at Day 0, and after spraying each spider with one of the stable dispersions of CNM on days 3 and 10, were examined with a Raman spectroscope using a 532nm laser. Multiple readings were taken along each sample to confirm homogeneity. No CNM were detected in any samples taken from spiders sprayed with CNM dispersions irrespective of species, CNM applied or time of collection (no 2D feature present in any of the evaluated silks; data not shown).



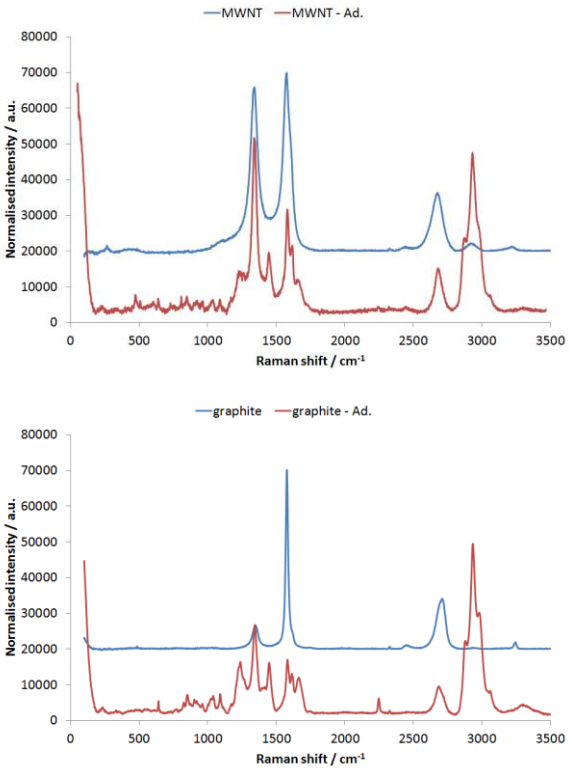


Figure 6.5. Raman spectra of original graphite and nanotube samples compared to those measured in *A. diadematus* silk after "spray on silk" experiments.
Graphite spectra offset for clarity.

The presence of CNM were, however, detected by Raman spectroscopy in *A. diadematus* silk samples where the silk itself had been directly sprayed with each of the CNM suspensions (Figure 6.5). There were no significant changes in the relative intensity and positioning for the CNM bands in spectra of CoMoCAT and P2 single-walled nanotubes that had been sprayed onto silk compared with controls, however there was some evidence that the spray application increases the number of, or extent of disorder, in both multi-walled nanotubes and graphite as seen by an increase in the D band (1350nm) relative to the G band (1580nm).

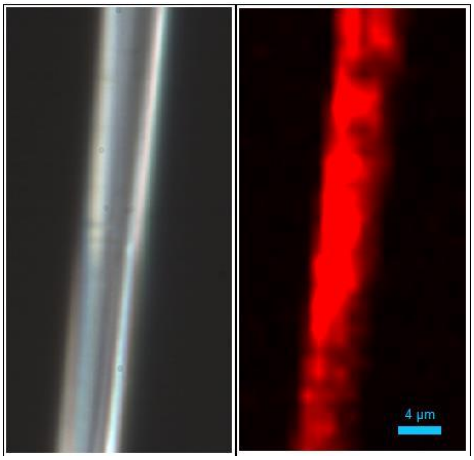


Figure 6.6. Images of *A. diadematus* silk after "spray on silk" experiments.
Optical microscope image (left) and Raman map of CoMoCAT SWNT incorporated into the same fibre (right).

The incorporation of CoMoCAT single-walled nanotubes was confirmed by 2D (x,y) Raman mapping. A point-spectra map allows for the visualisation of the coating of CNM which appears uneven (Figure 6.6).

6.3.3. Tensile test of silks incorporated with CNM by spraying

When examining charts of the tensile data it was observed that some samples had a second, smaller peak indicating the sample did not fully separate, which may occur if only one strand of the double-stranded fibre breaks initially. As this did not occur in all samples and the calculation of approximate toughness used in this study would exclude data obtained after the measurement of ultimate tensile strength (which was always the first break in all samples), the data presented only includes data up to and including the initial fracture point as a measure of tensile strength.

Initial collections of silk were taken from two *A. diadematus* individuals (adult female, gravid) to ensure adequate samples for a control group (Table 6.5). All *A. diadematus* 1, *A. diadematus* 2 and *P. phalangioides* samples were tested in 2016 using custom-built equipment. *A. diadematus* 3 samples were tested in 2018 using an LEX 810 linear extensometer (see section 6.2.4). As silk fibres are inherently fragile due to their size, samples were easily damaged before or during testing. These were discounted as no data was retrieved from them (Table 6.6).

Table 6.5. Number of silk samples collected and tested.
A hyphen indicates no samples were taken or tested for that category.

Spider	Untreated silk	CoMoCAT SWNT	P2 SWNT	MWNT	Graphite	Graphite (1:10 dilution)	Surfactant and water control
P. phalangioides (Adult ♀)	20	8	6	7	5	-	-
A. diadematus 1 (Adult ♀, gravid)	-	9	5	9	6	-	-
A. diadematus 2 (Adult ♀, gravid)	21	-	-	-	-	-	-
A. diadematus 3 (Adult ♀)	10	6	7	9	5	5	10

Table 6.6. Record of silk samples damaged after treatment and before tensile testing could be completed.
A hyphen indicates no samples were taken or tested for that category. Data lost to equipment failure is marked *, missing data marked $ and data discounted due to suspected multiple fibres in sample marked ^. Samples broken before testing are also shown as a percentage of total samples taken for each group (%).

Spider	Untreated silk	CoMoCAT SWNT	P2 SWNT	MWNT	Graphite	Graphite (1:10 dilution)	Surfactant control
P.p.	5 (20%)	1 (11%)	3 (33%)	3 (30%)	3 (11%) + 1*	-	-
A.d. 1	-	0	2 (22%) + 2*	1$	1*+1$ +1^	-	-
A.d. 2	2 (9%)	-	-	-	-	-	-
A.d. 3	0	4 (40%)	3 (30%)	1 (10%)	0	0	0

Silk samples from *P. phalangioides* which have been coated with CNM dispersions do not show large increases in tensile properties when compared with untreated silk. Conversely, the tensile properties generally decrease, making the coated fibres less strong, elastic or tough (Figure 6.7).

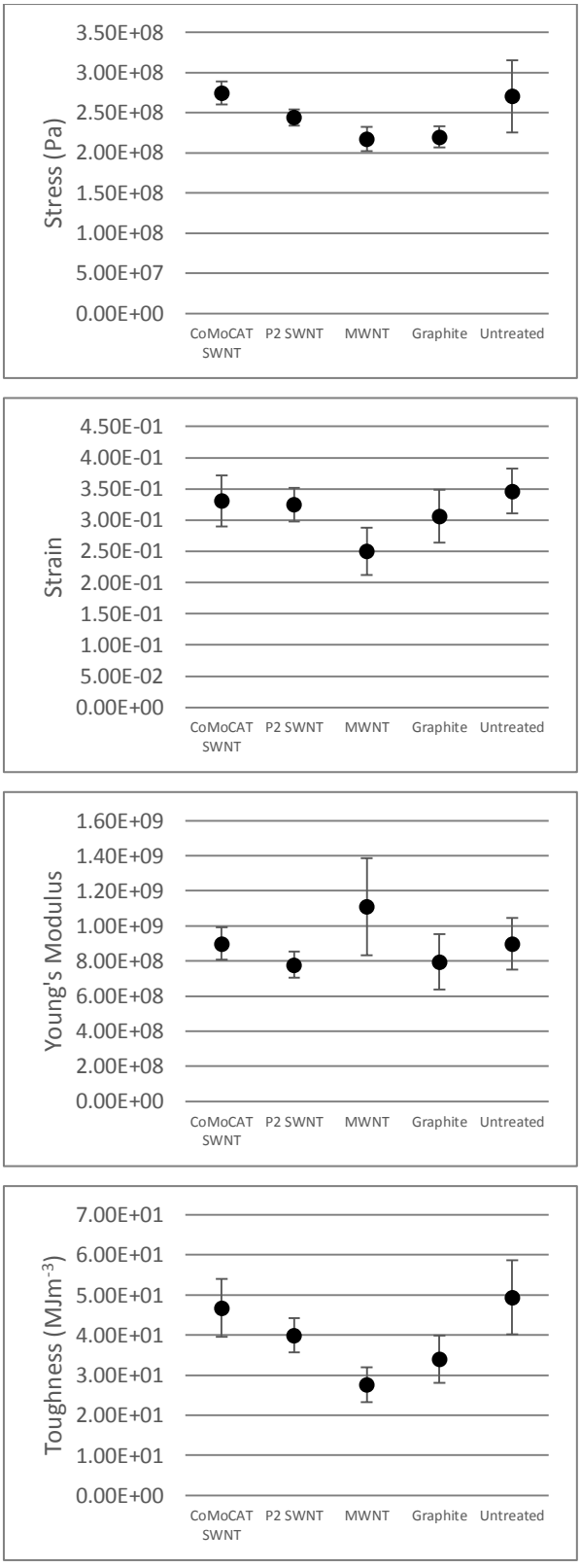


Figure 6.7. Results of tensile tests of *P. phalangioides* silk. Charts showing the average (dot) and standard error of each sample group from tensile test experiments on *P. phalangioides* silk (P.p.). Data can be found in Appendix 9.3.

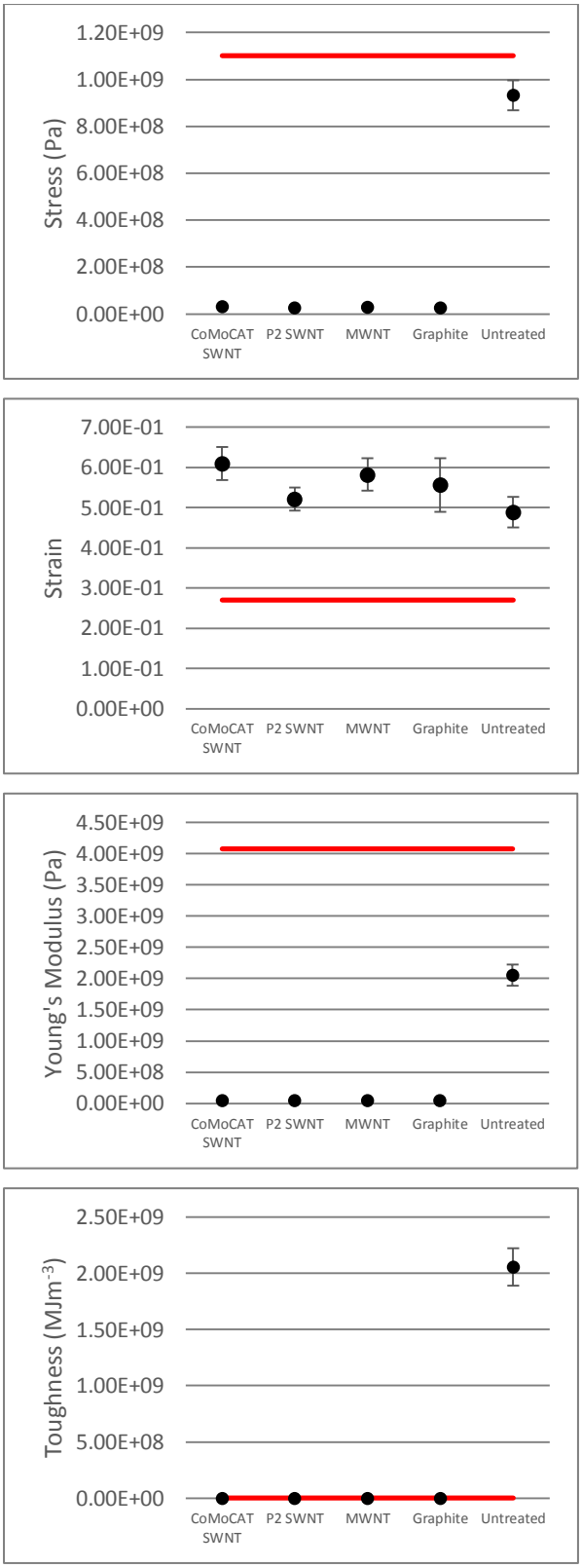


Figure 6.8. Results of tensile tests of *A. diadematus* 1 and 2 silk.
Charts showing the average (dot) and standard error of each sample group from the first set of tensile test experiments with *A. diadematus* silk (A.d. 1 and A.d. 2). Line is value of same

measurement as given in Gosline (1999) (Gosline et al., 1999). Data can be found in Appendix 9.3.

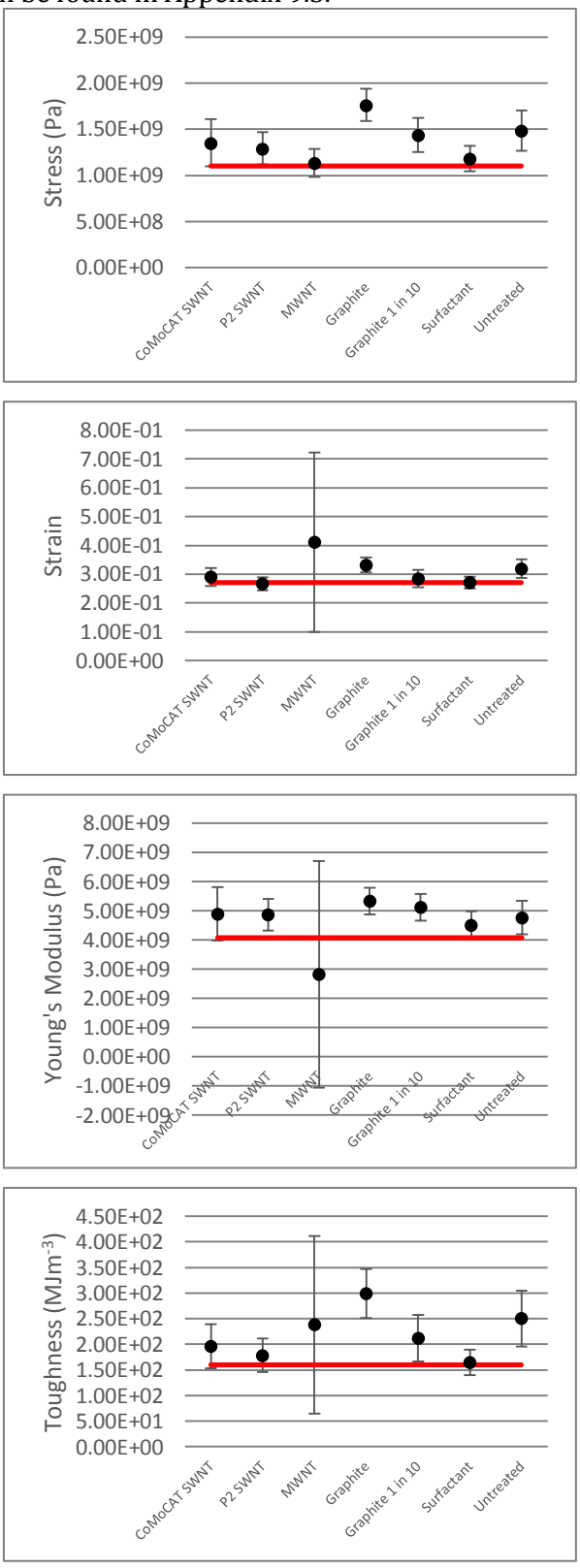


Figure 6.9. Results of tensile tests of *A. diadematus* 3 silk. Charts showing the average (dot) and standard error of each sample group from the second set of tensile test experiments on *A. diadematus* silk (A.d. 3). Line is value of same measurement

as given in Gosline (1999) (Gosline et al., 1999). Data can be found in Appendix 9.3.

The initial silk samples collected from *A. diadematus* show a decrease of one order of magnitude (9.33 x 10^{08}Pa for untreated silk vs 2.7 – 3.16 x 10^{07}Pa for coated silks) in the tensile strength and eight orders of magnitude in toughness (2.06 x 10^{9} for untreated silk vs 7.08 – 9.98 x 10) when coated in CNM dispersions and compared with the untreated samples. This does not appear to extend to the extensibility of the fibres, which remain consistent (allowing for minor variation between samples due to intra-individual variation) (Figure 6.8). Subsequent samples collected from *A. diadematus* 3 show a small decrease (from 1.49 x 10^{9}Pa in untreated silk vs 1.13 – 1.44 x 10^{9}Pa) in tensile strength and toughness compared with untreated silk. There is a small increase in the tensile properties of samples coated in graphite compared with untreated and surfactant-coated controls (Figure 6.9).

6.4. Discussion

In this study we compared the effects of CNM on the physical properties of silks of two spiders that come from distantly related parts of the overall spider phylogeny. Members of the Orbiculariae, to which *Araneus* belongs, are most commonly used as research subjects. For this reason we chose to study *A. diadematus*, which is readily found in the summer months in the UK and for which previously published measurements of the tensile properties of the silk can be compared with our own (Gosline et al., 1986). In Lepore *et al.* (Lepore et al., 2017) the authors use a mixture of spiders from the Pholcidae and Theredidae families, the former of which includes *P. phalangioides*. These are also commonly found in the UK, throughout the year. Using adult *P. phalangioides* spiders we attempted to replicate the data in (Lepore et al., 2017) whilst simultaneously applying the same treatment to *A. diadematus*.

There is variation in the diameters of silk fibres between species, and in the case of the *A. diadematus* used here, gravid and non-gravid females. However, this variation is of little significance in determining whether the combination of silk with CNM is mechanically improved over unmodified silk as each spider is treated as a separate group (*P. phalangioides,* gravid (1&2) and non-gravid (3) *A. diadematus*).

6.4.1. Integration of CNM into spider silks

We were unable to find any CNM in silk samples produced by spiders exposed to CNM by aerosol application using Raman spectroscopy, and preliminary tensile tests showed no change in tensile properties of these samples. Two *A. diadematus* spiders perished within 24 hours of application of CoMoCAT single-walled nanotubes. After initial testing it was decided not to reattempt to collect this dataset after results showed no presence of CNM in silks from *P. phalangioides* spiders coated with CoMoCAT single-walled nanotubes and *A. diadematus* spiders coated with other CNM did not produce silks containing CNM. This, in addition to our failure to reproduce the findings of the original study, led to us judge further experimentation unnecessary as it would likely have produced similar results or induced the death of further specimens.

As our assumption of the application method was based on a reduced amount of information given in a preprint (Lepore et al., 2015), our application of CNM direct to the spider is inconsistent with Lepore (Lepore et al., 2017). However, were we to continue to assume the conclusions drawn by the authors were correct, it is highly likely there will have been some transfer of CNM to the sides of the enclosure in our experiments and this, combined with inadvertent ingestion of CNM when the spider grooms, should be sufficient to allow enough CNM to be consumed to elicit a change in the tensile properties of subsequently produced silk; a result we have been unable to replicate in this instance.

A further consideration is that we have arbitrarily selected two time points in which to collect silk from the sprayed spiders. It is entirely possible that any CNM in the silks may have been passed before the first collection after exposure to CNM or later, after the first collection but before the second, and continuous or regular exposure may be required to guarantee a positive sample. It is noted that this does not appear to have been a consideration in the work by Lepore.

However, the inference that the silk tested in the study was taken from the same enclosure into which the spider was exposed suggests the results in Lepore may also have been due to inadvertent contamination of the silk during or after spinning. Therefore, in addition to the “spray on spider” samples, we also tested whether direct application of CNM to the silks could have produced the results previously reported.

Direct application of CNM to pre-spun spider silk

We hypothesised that the results seen in Lepore (2017) are the result of inadvertent contamination of the silk fibres either during the application by spraying or by inadvertent contamination as the spider spins silk in the same container into which the nanotubes were applied. To this end we intentionally applied CNM dispersions directly onto dragline silk to examine whether this induced the observed increase in tensile properties.

In total, 98 silk samples were collected and coated with one of the four CNM or a surfactant and water control then tensile tested (Figures 6.7 – 6.9 and Appendix 9.3). A number of samples were damaged before or during testing (see Table 6.6). The largest number of broken samples occurred in samples of *P. phalangioides* silk, seemingly irrespective of CNM coating. These accidental breakages are likely due to the silk being very fine (each fibre in the dragline measured ~940nm diameter), difficult to see and as such easily damaged during handling. However, the second greatest loss of samples occurred in *A. diadematus* fibres coated with CoMoCAT and P2 SWNT. Whilst the *P. phalangioides* samples may be discounted due to the fragility of the initial sample, these still account for a large number of breakages (40% in the CoMoCAT sample group), suggesting that the samples became more fragile as a result of being coated with SWNT. This correlates with the tensile data collected from A.d. 3.

The decrease in tensile strength and toughness, but consistency in strain (extensibility) of the samples from *A. diadematus* 1 and 2 (Figure 6.8) suggests that the coating does not affect the molecular structure of the silks as they are still able to stretch in the same way, to the same length, but the overall strength of the fibres is greatly affected. These results differ to those shown by the *P. phalangioides* data, which shows a much smaller decrease in the tensile properties of CNM-coated samples. Repetition of tests with *A. diadematus* silks from an additional spider more closely reflect the *P. phalangioides* data than the original *A. diadematus* data; the difference in the two datasets may therefore reflect an unobserved variation in experimental protocol or intra-species variation of silk properties, although the small variation in strain may suggest consistency between silk fibres produced by different individuals. It may also be possible that the application of CNM emphasises structural variation between silks taken from different individuals.

There is a small increase in the tensile properties of *A. diadematus* 3 silk samples coated in graphite (and to a lesser extent, an increase in the Young's Modulus of silk samples coated in a 1-in-10 dilution of graphite). This may indicate a positive interaction between silk and this CNM, but the small sample size does not allow for firm conclusions to be drawn. Samples from the same spider coated in MWNT appear to have stretched further under strain, but the small sample size and large error (see Figure 6.9) also do not allow for firm conclusions to be drawn.

Overall, the data we have collected shows a decrease in the range of tensile properties exhibited by samples sprayed with either the CNM dispersions or the control. This is likely due to supercontraction, which occurs when silk fibres are exposed to a humidity greater than 70% (Agnarsson et al., 2009a). The direct application of a water-based substance likely also induces greater supercontraction stress than a gradual increase in humidity would, and because these dispersions were applied whilst the silk fibres were restrained on a frame any molecular rearrangement that occurred as a result of interaction with the water molecules may not reach its maximum possible reconfiguration (Agnarsson et al., 2009a); ergo, if samples were unrestrained during application of CNM dispersions different results may be obtained.

We do not see any large increases in tensile properties of any of the silk samples directly coated with CNM dispersions. This finding is also inconsistent with data published in Lepore, assuming such results were a result of inadvertent contamination, and disagrees with our hypothesis that those results are due to direct or unintentional contamination. There was a slight increase in the calculated maximum strain observed in sprayed *A. diadematus* 1 samples compared with unsprayed control *A. diadematus* 2 samples, though it is small and may be accounted for by intra-individual variation, as the coated samples were obtained from one spider whilst the control samples were obtained from another of the same species and similar size. However, even with a small increase in range of results and average, these do not approach the large increases seen in Lepore and the ultimate tensile strength, Young's Modulus and toughness for the treated samples are reduced compared to the unsprayed control. Additionally, application of CNM by pump-action aerosol will introduce a random element to the application process, as fine control of droplet size and subsequent coating – both in coverage and thickness – is impossible to control.

As described here in section 6.1, the study by Steven uniformly coated dragline silk fibres reeled from *Nephila clavipes* by mixing fibres with dry SWNT, adding a few drops of water and applying shear (this method was not replicable with MWNT) (Steven et al., 2013). This appears to more closely represent our data than that of Lepore, although they nevertheless report a significant increase in toughness to 300% of a native silk fibre which is not observed in any of our results.

Integration of CNM into spider silk by ingestion – discussion of Lepore (2017)

The biological processes behind the transfer of carbon nanomaterials from the digestive organs into the lumen of the silk glands has been neither described nor proposed in any peer reviewed works. It is highly unlikely that an animal that would not naturally encounter any quantity of CNM during its evolution (especially those which are also coated in surfactant, although this is a naturally occurring bile salt that may be encountered elsewhere) would have a mechanism to not only transport it throughout its body but selectively, intentionally or accidentally introduce it into the silk gland and include this material in the subsequently produced silk. Therefore it is improbable that ingestion is the beginning of a process by which silk fibres may be reinforced.

Application of CNM

The introduction of CNM to the test spiders in Lepore describes the application by aerosol of carbon nanomaterials into one corner of the container in which the test specimens resided, avoiding direct application to the spiders. It is not mentioned whether the spiders were transferred to another, clean, container from which silk samples were later retrieved (see Supplementary Information, Figure S1c, d and "Collection from the box", Figure S1e in Lepore, leaving open the possibility that the samples collected were contaminated with residual CNM either as a result of the aerosol application or contamination by movement of the spider within its enclosure after application. No mention was made of the containers being cleaned between the experiments.

Silk collection

What is also not made clear in this research is precisely how the samples are collected; it is worth noting that a single sample of dragline silk collected from a spider contains just two fibres, whereas the field emission SEM image (Figure S1i, Lepore et al., 2017) shows a sample with a number of fibres greater than two, suggesting fresh silk was not collected directly from the spider as in this study;

the length of time between application of CNM and collection of silk also varies between 2 and 12 days and no note is made of how many collections were made from which spiders, making it difficult to gauge the speed of ingestion of CNM.

Additionally, samples of *P. phalangioides* dragline silk collected in this study measured an average of 1.91μm per fibre; in Lepore, samples of *P. phalangioides* fibres range from 0.69μm (±0.06μm) to 1.00μm (±0.12μm), although this disparity may be accounted for if the spiders in Lepore were not fully grown adults, this discrepancy may indicate that the silk fibres being examined may be a fibre other than dragline silk.

Collection of spun silk from a container without observing the spinning process means the authors cannot be certain that the silk being tested is exclusively dragline silk. Different silks have different properties (reviewed in Tokareva (Tokareva et al., 2014)), and the inability to identify or quantify the different fibre types within each sample leaves open the possibility that variation between silks accounts for the differences in tensile strength, although this does not solely explain the dramatic increase observed in the data. Whilst this would be unlikely to affect an investigation comparing the effect of CNM coatings on silk from the same study, it does make difficult the process of direct comparison with other studies.

Another explanation for an increase in the tensile properties of silks spun by spiders is starvation. Spiders such as the Theredіidae used in the experiment use their webs to ensnare wandering prey (Benjamin and Zschokke, 2002), whereas the Pholcidae rely on the vibration-transmission properties of the silk and as such, web-building spiders alter the properties of their webs to raise the chance of a successful capture of prey (Zevenbergen et al., 2008). This is a variable that could be explored in future experiments, although spiders in this study were fed and watered to minimise any effects from starvation or dehydration.

Independent corroboration

Independent studies looking into the incorporation of CNM into spider silk by forced feeding (administration of CNM suspension direct to the mouthparts of the spider) have so far failed to show any incorporation of CNM into subsequently produced silk dragline fibres. Tensile tests of dragline fibres produced after feeding showed no improvement in the tensile properties of the silk and there was

no indication of CNM in samples examined using Raman spectroscopy (Kelly, 2018, Vollrath, 2018).

6.4.2. Discussion of techniques and suggested improvements

Use of surfactant

It was necessary to use the surfactant in the dispersions, as intermolecular Van der Waal's forces between individual CNM lead to the agglomeration of CNM in dispersion, which prevents an even coating of CNM on the silk fibre and thus negatively affect the tensile properties (alternate methods of preventing agglutination include functionalisation, reduction by alkali metals or protonation with superacids (Bergin et al., 2008)). However, this surfactant may need to be removed if we need to perform chemical modification of CNT as seen in (Steven et al., 2013, O'Connor et al., 2009); higher concentrations of CNTs result in aggregations of CNM which, when coating fibres, do not infer mechanical strength and instead lead to a decrease in strength and modulus of the polymer fibre (Manchado et al., 2005, Nan et al., 2004). This was tested, albeit with a very small sample set, in the samples obtained from A.d. 3 and coated with a 1 in 10 dilution of graphite; the comparable results of these tests suggest that, in the case of the graphite dispersions at least, either an excess of CNM in the dispersion was not inducing an effect or a much greater dilution was required to negate such an effect.

Variation amongst *A. diadematus* samples

The average diameter of silk samples varies between species, which is not unexpected, however the diameter of samples taken from each group of *A. diadematus* spiders vary. This may partly be due to *A. diadematus* 3 being collected earlier in the year (June 2018) and despite being a mature female apparently not having mated with a male, hence not developing into a gravid (egg carrying) state – the sex of the spider was confirmed by microscopic examination of the epigyne after death (September 2018), and it did not moult whilst in captivity. *A. diadematus* spiders undergo a biyearly life cycle and fully mature at the end of their second summer, typically during the month of September (Bee et al., 2017, Lee and Thomas). The first group of these spiders were collected in September and were gravid, so likely produced larger, stronger dragline fibres to account for their increased weight. Additionally, spiders have a muscular valve which allows them to adjust the diameter of the silk fibres they produce (Wilson, 1962a, Wilson, 1962b) which may also allow for the production of different diameter fibres.

Methods of silk collection

In this study silk samples were collected by encouraging a spider to walk, where they automatically deposit a silk dragline which could be readily collected. This has the disadvantage of relying on the spider to deposit said dragline in such a way that it can be easily collected. In addition, the dragline samples collected are composed of two strands of silk, one from each of a pair of spinnerets, which are impossible to separate when following our method of silk collection.

An alternate method of collection is forced silking: the spider is restrained on its back with the spinnerets exposed and the dragline silk is captured using a small brush; the two fibres can then be separated and attached to a rotating frame, which mechanically and consistently reels the silk from the spinnerets. Forced silking allows for a long length of silk to be collected from the spider faster and in a more controlled way, and each strand of dragline silk can be reeled separately. However, this is easily undertaken on larger spiders such as *Nephila* sp. and whilst possible with care using the much smaller *A. diadematus* and *P. phalangioides*, the equipment was not available in this instance. Additionally, spiders are able to control the diameter of silk strands as they are extruded from the spigots through a small muscular valve [40, 41]. This becomes relevant if the spider recognises its silk is being reeled and attempts to restrict the event, thereby affecting the fibre size and subsequent tensile properties (spigots are not able to be used to cut through a silk fibre; instead spiders use their legs to break the fibre, therefore resulting in the need for restraint when silking). Whilst the presence of spigot valves is also relevant to the "chasing" method of silk collection, the spider is more likely to keep a consistent minimum diameter as the silk strands are actively being used as a safety line whilst maintaining a minimum expenditure of resources.

6.5. Conclusions

The effect of exposing spiders to CNM on subsequently produced silk was examined and found to not have an effect on the tensile properties of the silk and no CNM structures could be identified in the silk using Raman spectroscopy. Additionally, direct application of CNM dispersions to spiders may cause the death of the spider within 24 hours.

The adhesion of carbon nanomaterials to silk fibres when applied by spraying was examined using Raman spectroscopy and the mechanical properties of treated silk samples tested to destruction. Raman spectroscopy identified a coating of

CNM on samples; results show a greater range of tensile properties within uncoated controls and a small decrease in tensile properties of silks exposed to the CNM dispersions or surfactant carrier. Not all CNM dispersions appear to have had the same effect, although inter-individual variation may also cause minor variations. The small overall decrease in tensile properties may be attributed to supercontraction as water was used as the carrying agent for the CNM dispersions and surfactant, as seen in Steven *et al.* 2013 (Steven et al., 2013).

7. General discussion

7.1. Views, perspective and future directions

In sections 3 to 6 of this thesis I have discussed each area of research individually. However, as these discussions are specific to each section I would like to use this opportunity to discuss how they might be used to inform and direct future research of spider silk, *Argyroneta aquatica*, and synthetic silk as a biomaterial.

7.1.1. Further sequencing of *Argyroneta aquatica*, *Cybaeus angustiarum* and *Pholcus phalangioides* spidroins

Confirmation of expression of piriform spidroins in piriform glands and sequencing of full-length spidroins

There remains some uncertainty around the *A. aquatica* piriform spidroins identified in our study and the variation in their sequence from piriform spidroins identified in other species. The incomplete spidroin sequences identified here, in particular the absent N-terminal regions, necessitates further study. The same applies to the aciniform and tubuliform N-termini and the minor ampullate spidroins(s) that are also absent from the data presented here. Obtaining complete information for these, and the spidroins of the other two species studied, will likely inform future research into silk evolution, structure and function.

Experimental PCR of a cDNA library

One option is to generate an *A. aquatica* cDNA library and attempt to amplify complete spidroin sequences using custom-designed primers. Speculative primer design, using known spidroins from other species to generate primers is unlikely to be successful due to the evolutionary distance between *A. aquatica* and known complete spidroin sequences. Using the partial sequences described in this thesis, primers could be designed to anneal to regions within the terminal domains. Due to the sequences being obtained from a transcriptome (expressed regions only), it would not be possible to design primers in such a way that the entire terminal domain is sequenced, but this information would be retained in the transcriptome data.

Using PCR to amplify spidroins is complicated by an inability to determine the correct N-terminal/C-terminal pairings, however combining primer pairs from opposing terminal regions of known silk types may result in the successful amplified regions of DNA of approximately the length of spidroins (estimated based on the length of known sequences in other species), thereby indicating a

"correct" pairing. This both allows for pairs to be indicated in the transcriptome sequence and provides full length (or longer partial-length) sequences for sequencing in order to determine the sequence of the unknown repetitive region between the terminal domains.

However, this is complicated by the incomplete data obtained, which does not include all the terminal domains of all the spidroins produced by *A. aquatica;* for some spidroins (e.g. piriform) only the C-terminal end are known. This makes it especially difficult to obtain complete sequences through use of PCR.

Sequencing individual silk glands

Historically, it has been assumed that specialised silk glands produce single, specialised silk spidroins – thus, spidroins have been named after the gland in which they were identified. Following this line of thought, our putative piriform spidroins should therefore only be produced by the piriform glands. By examining the RNA of individual glands, we should be able to confirm that the piriform spidroins identified are indeed produced by the piriform gland in *A. aquatica*. However, some studies have identified that some specialised glands produce multiple spidroins e.g. MaSps in tubuliform glands [54, 172]. As such, I propose that further investigation into the precise origin of the piriform silks be undertaken. One method would be to individually examine RNA expression of the silk glands, using rtPCR of each gland with primers designed from the sequences we have obtained. However, this only permits identification of partial sequences already known and would only confirm which glands express which spidroins. Due to the evolutionary distance between *A. aquatica* and the nearest species with a known MiSp, speculative rtPCR of this spidroins are unlikely to be unsuccessful. Generating a transcriptome of each gland would not only allow us to confirm by BLAST which spidroins are expressed in each gland, but could also provide additional data on currently unidentified regions of the same spidroins. This would likely be the best way to identify the so-far unobserved N-termini and minor ampullate spidroin(s).

Protein sequencing of the diving bell

Some attempts by other researchers to protein sequence silk have been successful [173]. We could attempt to repeat and optimise the preparation procedure specific to the diving bell, as if successful, the sequencing data would confirm which spidroins (major ampullate, minor ampullate, aciniform and piriform) are

present in the diving bell. Additionally, this may also produce data permitting identification of the components of the hydrogel, allowing for further study of the properties and structure of this material.

Exploiting advances in sequencing technologies

As in all research, new data and improved tools become available with time. Sequencing technologies continue to advance, with ultra-long read lengths of over 2 million bases now possible with the MinIon (Oxford Nanopore) (Payne et al., 2018). This could be useful when sequencing repetitive DNA such as that found within spidroin sequences; nanopore technology allows for the sequencing of a single strand of DNA without requiring digestion or fragmentation that is seen in Illumina sequencing, negating the need for post-sequencing reassembly. However, as a small number of spidroin sequences have been observed to contain intron sequences (Hayashi and Lewis, 2000), DNA sequencing of spidroin genes may not be completely straightforward. Instead, nanopore sequencing of RNA may become the most effective method in terms of speed and coverage (Garalde et al., 2018). Alternatively, software dedicated to assembling gene sequences containing repetitive regions is under development but may ultimately be less effective than nanopore sequencing (Nowak, 2014).

Use of newly available genetic data

During this study, two highly annotated spider genomes and a transcriptome were published by other researchers (Babb et al., 2017, Schwager et al., 2017, Correa-Garhwal et al., 2018) and it is likely more will follow. Using the latest genetic data, our custom databases could be extended (particularly relevant given the transcriptome in (Correa-Garhwal et al., 2018) is of a semi-aquatic spider) and the transcriptomes reassembled using the *N. clavipes* genome as a scaffold. This was not undertaken during this study as the assemblies described earlier were exploratory and, whilst achieving our goal of retrieving spidroin sequences from previously unsequenced species, we determined that the quality of the data retrieved indicated that reassembly using the *N. clavipes* genome as a scaffold is unlikely to result in sufficient additional data. Attempts were made to include further transcriptome material that would assist with improving the amount of data retrieved but sequencing was not completed in time to allow for inclusion in the studies reported.

Further in silico analysis and annotation

Updated versions of the assembly and translation software have been released and reassembly of our transcriptomes may yield further data. There could also be some benefit in running the open reading frame sequences predicted by TransDecoder through databases such as Pfam (Finn et al., 2016) and Uniprot (The UniProt Consortium, 2017), using software such as HMMER (Eddy, 2011). This could allow identification and annotation of further silk sequences, possibly including sections of repetitive regions missed during earlier BLAST searches. Furthermore, tools such as BUSCO are available that can be used to assess the quality of a transcriptome assembly (Correa-Garhwal et al., 2018) and may allow greater insight into the effectiveness of dissection vs. whole abdomen sequencing.

Quantitative analysis of gene expression

Transcriptomes can be used for a range of lines of investigation other than sequencing specific genes of interest. By sequencing messenger RNA, each copy present in the sequenced tissue will be included in the transcriptome allowing us to quantify the levels of expression of genes of interest. However, to account for differences in expression levels due to external influences on the specimen, a comparison would be better made across a number of transcriptomes, likely requiring a minimum of three. This could also be extended to investigating different levels of expression of silk spidroins between spiders of different developmental stages, for example.

Identification of non-spidroin sequences

Other peptides of interest in the silk glands may also be identifiable in our transcriptomes. Expression of enzymes such as carbonic anhydrase (Chaw et al., 2015, Andersson et al., 2014, Pouchkina et al., 2003) and peptides like aggregate spider glue 1 (ASG1) (Clarke et al., 2014, Choresh et al., 2009, Vasanthavada et al., 2012) have been identified in multiple species. Other silk-associated peptides such as spider-coating peptides SCP1 & SCP2, egg-case silk proteins ECP1 & ECP2 and aggregate silk factors ASF1 & ASF2 have not been identified in existing studies, such as that by Correa-Garhwal but these may be species or family specific. Ultimately, if we know the sequence of a particular gene of interest, they can be used to probe transcriptomes for similar or conserved peptides.

7.1.2. *A. aquatica* silk - carbon nanomaterial composites

Perhaps the most obvious direction is to explore the application of carbon nanomaterial (CNM) to the diving bell web formed by *A. aquatica*. One observation made from the literature and research discussed in section 6 is that exposing spider silk to water induces supercontraction in silk dragline fibres. This is relevant to the synthetic production of spider silk, as synthetic silk utilises repetitive motifs from *Euprosthenops australis* (Stark et al., 2007, Harvey et al., 2017) or *Nephila clavipes* (Xia et al., 2010) major ampullate spidroins and will likely demonstrate similar contracting properties when exposed to water as the native fibre. However, *A. aquatica* silk is typically spun directly into water and as such application of CNM dispersions will not induce supercontraction of the silk. Additionally, the coating of hydrogel on the surface of the silk may improve adhesion of CNM to the structure.

One method of applying CNM to diving bells is by addition to, or gradually replacing, the water into which the diving bell has been spun. Introduction of CNM to the diving bell is recommended only after the spider has been removed, so as not to inadvertently cause death. There is a chance that the surfactant required to prevent accumulations of CNM will interfere with the hydrophobic hairs allowing the spider to retain a bubble of air whilst underwater and temporarily or permanently disrupting the bubble. Methods would need to be developed to collect, store and tensile test wet fibres and examine any effect on the tensile properties of the silk-CNM composite in order to test for any variation between wet and dry composites.

Potential applications of silk-CNM composites

Despite the likelihood that combining CNM with spider silk does not appear to result in a composite with greater tensile properties than native spider silk, there are still a multitude of potential modifications to and uses for these composites. CNM-coated silk is electrically conductive through the outer layer of CNM, and could still be utilised as strong, ultra-fine, electrically conductive fibres, suitable for use in biocompatible sensors (Steven et al., 2013). The carbon nanotubes (CNT) themselves may be modified before application to the silk; studies have examined the effect of Kevlar-coated CNTs on polymer structures (O'Connor et al., 2009, O'Connor et al., 2008, Chang et al., 2012) and the nanotubes used in (Steven et al., 2013) are amine-functionalised MWNT, capable of binding to a number of metal ions (CNT may be functionalised in a number of ways, as reviewed in

(Lamanna et al., 2012)). Another study explored modification of CNT by adding iron nanoparticles (Luo et al., 2016).

CNM-synthetic silk composites

Synthetic silk has yet to be combined with CNM of any kind and combining synthetic silk dope with CNM prior to spinning or formation into thin films or hydrogel structures may yield potentially interesting composites. The addition of CNTs to native and artificial silk fibroin solutions has been demonstrated to increase the tensile strength and Young's modulus of subsequently formed nanocomposite films (Pan et al., 2015, Blond et al., 2007) and also combined with solutions of *Bombyx mori* (silkworm) silk prior to electrospinning (Ayutsede et al., 2006) – however, there will be little control over how the nanotubes are incorporated into the fibre which may result in inconsistent results that are difficult to reproduce reliably.

Furthermore, the functionalisation of synthetic silk by substitution of a non-natural amino acid and subsequent copper catalyzed azide-alkyne cycloaddition "click chemistry" (Harvey et al., 2017) could allow for further functionalisation of a synthetic silk-CNM composite, with both silk and CNM bound to different compounds or proteins.

7.1.3. Further research into silk of semi-aquatic spiders

A. aquatica are not the only spiders adapted to a wet environment, although they remain the only species to actively spin silk whilst submersed. Other spiders are adapted to live around water, including members of the Desidae family and *Paratheuma*, *Dolomedes* and *Pirata* genera. Of interest to some silk researchers would perhaps be the intertidal spiders (including Desidae and *Paratheuma*). These spiders live between the high and low water mark and use their silk to seal themselves into small spaces whilst the tide is high. Once the water level drops low enough, these spiders emerge from their retreat to hunt on the newly-exposed surface. This behaviour indicates that the silken block applied by intertidal spiders likely functions as a physical gill in a similar way to *A. aquatica* silk, but little is known about the structure or content of the webs and no spidroin sequences currently exist in public databases. These spiders, and *A. aquatica*, are part of the poorly-supported "marronoid" group as described by Wheeler. (Wheeler et al., 2017). Greater study of this group may provide further insight into the evolution of the intertidal and diving bell spiders and their silk.

As a penultimate note, it is a pleasure to mention that the SpiderLab at the University of Nottingham were part of a successful team campaigning for the genome of *Dolomedes plantarius* to be sequenced as part of the 25 genomes project (Wellcome). It is hoped this data will aid in the future conservation of this species within the UK, but it is also relevant that the full spidroin sequences of a spider outside the Orbiculariae and Mygalomorphae will become known. This is particularly exciting as the majority of silk research has been conducted on Orbiculariae including *Nephila*, *Araneus*, and *Latrodectus* species (a summary of the majority of spidroin sequences available can be viewed in section 4.4.1 Figure 4.1). Additionally, with the recently published transcriptome of *Dolomedes triton*, the *D. plantarius* genome will enable investigation into the evolution of two closely related spiders which occupy the same niche but are geographically isolated from each other.

7.1.4. Synthetic *A. aquatica* silk

This thesis presents the first *A. aquatica* spidroin sequences, laying the groundwork for future research into synthetic silk derived from an aquatic spider. This is interesting for a number of reasons, not least because we do not yet know if synthetic spidroins based on *A. aquatica* silk will be suited to the "spinning" process as used in current synthetic silk production. The genetic analysis in section 4 showing conservation of the salt bridge in the C-terminal of *A. aquatica* spidroins suggests that the fibre-forming process is also conserved, however post-spinning conditions such as immersion in fresh water may be required as a final step to fully replicate native silk. Additionally, we have yet to identify the source of the hydrogel coating – if this is indeed a spidroin (one likely candidate would be piriform silk), synthetic production of a "natural" silk hydrogel could begin a new branch of synthetic silk research.

8. Bibliography

"PHOLCUS PHALANGIOIDES" (FUESSLIN. *Pholcus phalangioides (Fuesslin, 1775)* [Online]. NBN Atlas Partnership. Available: https://species.nbnatlas.org/species/NBNSYS0000008639 [Accessed 02/03/2019].

AGNARSSON, I., BOUTRY, C., WONG, S. C., BAJI, A., DHINOJWALA, A., SENSENIG, A. T. & BLACKLEDGE, T. A. 2009a. Supercontraction forces in spider dragline silk depend on hydration rate. *Zoology (Jena),* 112**,** 325-31.

AGNARSSON, I., DHINOJWALA, A., SAHNI, V. & BLACKLEDGE, T. A. 2009b. Spider silk as a novel high performance biomimetic muscle driven by humidity. *Journal of Experimental Biology,* 212**,** 1990.

AGNARSSON, I., KUNTNER, M. & BLACKLEDGE, T. A. 2010. Bioprospecting Finds the Toughest Biological Material: Extraordinary Silk from a Giant Riverine Orb Spider. *PLoS ONE,* 5**,** e11234.

ALLMELING, C., JOKUSZIES, A., REIMERS, K., KALL, S., CHOI, C. Y., BRANDES, G., KASPER, C., SCHEPER, T., GUGGENHEIM, M. & VOGT, P. M. 2008. Spider silk fibres in artificial nerve constructs promote peripheral nerve regeneration. *Cell Proliferation,* 41**,** 408-420.

ALLMELING, C., JOKUSZIES, A., REIMERS, K., KALL, S. & VOGT, P. M. 2006. Use of spider silk fibres as an innovative material in a biocompatible artificial nerve conduit. *Journal of Cellular and Molecular Medicine,* 10**,** 770-777.

ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. & LIPMAN, D. J. 1990. Basic local alignment search tool. *Journal of Molecular Biology,* 215**,** 403-410.

ANDERSSON, M., CHEN, G., OTIKOVS, M., LANDREH, M., NORDLING, K., KRONQVIST, N., WESTERMARK, P., JÖRNVALL, H., KNIGHT, S., RIDDERSTRÅLE, Y., HOLM, L., MENG, Q., JAUDZEMS, K., CHESLER, M., JOHANSSON, J. & RISING, A. 2014. Carbonic Anhydrase Generates CO(2) and H(+) That Drive Spider Silk Formation Via Opposite Effects on the Terminal Domains. *PLoS Biology,* 12**,** e1001921.

ANDERSSON, M., HOLM, L., RIDDERSTRÅLE, Y., JOHANSSON, J. & RISING, A. 2013. Morphology and Composition of the Spider Major Ampullate Gland and Dragline Silk. *Biomacromolecules,* 14**,** 2945-2952.

ANDERSSON, M., JIA, Q., ABELLA, A., LEE, X.-Y., LANDREH, M., PURHONEN, P., HEBERT, H., TENJE, M., ROBINSON, C. V., MENG, Q., PLAZA, G. R., JOHANSSON, J. & RISING, A. 2017. Biomimetic spinning of artificial spider silk from a chimeric minispidroin. *Nature Chemical Biology,* 13**,** 262.

ANDERSSON, M., JOHANSSON, J. & RISING, A. 2016. Silk Spinning in Silkworms and Spiders. *International Journal of Molecular Sciences,* 17**,** 1290.

ARTHUR, R. 2017. *Bolt threads Is Launching Its First Bioengineered Spider Silk Product At SXSW - A Necktie* [Online]. Jersey City, NJ, USA: Forbes Digital. Available: forbes.com/sites/rachelarthur/2017/03/10/bolt-threads-is-launching-its-first-bioengineered-spider-silk-product-at-sxsw-a-necktie/#78519f844d13 [Accessed 13/06/2018].

ASAKURA, T., UMEMURA, K., NAKAZAWA, Y., HIROSE, H., HIGHAM, J. & KNIGHT, D. 2007. Some Observations on the Structure and Function of the Spinning Apparatus in the Silkworm Bombyx mori. *Biomacromolecules,* 8**,** 175-181.

ASKARIEH, G., HEDHAMMAR, M., NORDLING, K., SAENZ, A., CASALS, C., RISING, A., JOHANSSON, J. & KNIGHT, S. D. 2010. Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. *Nature,* 465**,** 236-8.

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.

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.

AYUTSEDE, J., GANDHI, M., SUKIGARA, S., YE, H., HSU, C.-M., GOGOTSI, Y. & KO, F. 2006. Carbon Nanotube Reinforced Bombyx mori Silk Nanofibers by the Electrospinning Process. *Biomacromolecules,* 7**,** 208-214.

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. *Nature Genetics,* 49**,** 895.

BAEHR, B. C., RAVEN, R. & HARMS, D. 2017. "High Tide or Low Tide": Desis bobmarleyi sp. n., a new spider from coral reefs in Australia's Sunshine State and its relative from Sāmoa (Araneae, Desidae, Desis). *Evolutionary Systematics,* 1**,** 111-120.

BAS, P. P. Pholcus phalangioides [Online]. British Arachnological Society, Salticus Webhosting. Available: britishspiders.org.uk/wiki2015/index.php?title=Pholcus_phalangioides [Accessed 10/08/2018].

BECKWITT, R. & ARCIDIACONO, S. 1994. Sequence conservation in the C-terminal region of spider silk proteins (Spidroin) from Nephila clavipes (Tetragnathidae) and Araneus bicentenarius (Araneidae). *Journal of Biological Chemistry,* 269**,** 6661-6663.

BECKWITT, R., ARCIDIACONO, S. & STOTE, R. 1998. Evolution of repetitive proteins: spider silks from Nephila clavipes (Tetragnathidae) and Araneus bicentenarius (Araneidae)1. *Insect Biochemistry and Molecular Biology,* 28**,** 121-130.

BEDINI, S. A. 2005. Along Came A Spider - Spinning Silk for Cross-Hairs (The Search for Cross-Hairs for Scientific Instrumentation Part II). *The American Surveyor.* Frederick, MD 21701, USA: Spatial Media LLC.

BEE, L., OXFORD, G. & SMITH, H. 2017. *Britain's Spiders,* Princeton, New Jersey, USA, Princeton University Press.

BELL, A. L. & PEAKALL, D. B. 1969. Changes in fine structure during silk protein production in the ampullate gland of the spider *Araneus sericatus*. *The Journal of Cell Biology,* 42**,** 284-295.

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). *Journal of Insect Behavior,* 15**,** 791-809.

BERGIN, S. D., NICOLOSI, V., STREICH, P. V., GIORDANI, S., SUN, Z., WINDLE, A. H., RYAN, P., NIRAJ, N. P. P., WANG, Z.-T. T., CARPENTER, L., BLAU, W. J., BOLAND, J. J., HAMILTON, J. P. & COLEMAN, J. N. 2008. Towards Solutions of Single-Walled Carbon Nanotubes in Common Solvents. *Advanced Materials,* 20**,** 1876-1881.

BERMAN, D., ERDEMIR, A. & SUMANT, A. V. 2013. Reduced wear and friction enabled by graphene layers on sliding steel surfaces in dry nitrogen. *Carbon,* 59**,** 167-175.

BITTENCOURT, D., DITTMAR, K., LEWIS, R. V. & RECH, E. L. 2010. A MaSp2-like gene found in the Amazon mygalomorph spider Avicularia juruensis. *Comparative biochemistry and physiology. Part B, Biochemistry & molecular biology,* 155**,** 419-426.

BITTENCOURT, D., SOUTO, B. M., VERZA, N. C., VINECKY, F., DITTMAR, K., SILVA, P. I., JR., ANDRADE, A. C., DA SILVA, F. R., LEWIS, R. V. & RECH, E. L. 2007. Spidroins from the Brazilian spider Nephilengys cruentata (Araneae:

Nephilidae). *Comparative biochemistry and physiology. Part B, Biochemistry & molecular biology,* 147**,** 597-606.

BLACKLEDGE, T. A., BOUTRY, C., WONG, S. C., BAJI, A., DHINOJWALA, A., SAHNI, V. & AGNARSSON, I. 2009. How super is supercontraction? Persistent versus cyclic responses to humidity in spider dragline silk. *J Exp Biol,* 212**,** 1981-9.

BLACKLEDGE, T. A. & HAYASHI, C. Y. 2006. Silken toolkits: biomechanics of silk fibers spun by the orb web spider Argiope argentata (Fabricius 1775). *Journal of experimental biology,* 209 (Pt 13)**,** 2452-61.

BLAMIRES, S. J., BLACKLEDGE, T. A. & TSO, I.-M. 2017. Physicochemical Property Variation in Spider Silk: Ecology, Evolution, and Synthetic Production. *Annual Review of Entomology,* 62**,** 443-460.

BLAMIRES, S. J., KASUMOVIC, M. M., TSO, I. M., MARTENS, P. J., HOOK, J. M. & RAWAL, A. 2016. Evidence of Decoupling Protein Structure from Spidroin Expression in Spider Dragline Silks. *International Journal of Molecular Sciences,* 17**,** 1294.

BLASINGAME, E., TUTON-BLASINGAME, T., LARKIN, L., FALICK, A. M., ZHAO, L., FONG, J., VAIDYANATHAN, V., VISPERAS, A., GEURTS, P., HU, X., LA MATTINA, C. & VIERRA, C. 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**,** 29097-108.

BLOND, D., MCCARTHY, D. N., BLAU, W. J. & COLEMAN, J. N. 2007. Toughening of Artificial Silk by Incorporation of Carbon Nanotubes. *Biomacromolecules,* 8**,** 3973-3976.

BON, M. 1710. A Discourse upon the Usefulness of the Silk of Spiders. By Monsieur Bon, President of the Court of Accounts, Aydes and Finances, and President of the Royal Society of Sciences at Montpellier. Communicated by the Author. *Philosophical Transactions (1683-1775),* 27**,** 2-16.

BOUTRY, C. & BLACKLEDGE, T. A. 2010. Evolution of supercontraction in spider silk: structure-function relationship from tarantulas to orb-weavers. *J Exp Biol,* 213**,** 3505-14.

BRAUN, F. 1931. Beiträge zur Biologie und Atmungsphysiologie der Argyroneta aquatica Cl. *Zoologische Jahrbucher Abteilüng für Systematik,* 62**,** 175 - 262.

BUFFALO, V. 2014. Scythe - A Bayesian adapter trimmer (version 0.994 BETA) [Software].

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. *Journal of Arachnology,* 38**,** 99-103.

CATALOG, W. S. 2019. World Spider Catalog, version 20.0.

CHALLIS, R. J., GOODACRE, S. L. & HEWITT, G. M. 2006. Evolution of spider silks: conservation and diversification of the C-terminus. *Insect Mol Biol,* 15**,** 45-56.

CHANG, C.-M., HSU, K.-Y. & LIU, Y.-L. 2012. Matrix-polymer-functionalized multiwalled carbon nanotubes as a highly efficient toughening agent for matrix polymers. *Journal of Polymer Science Part B: Polymer Physics,* 50**,** 1151-1155.

CHANG, J. C., FLETCHER, M. J., GURR, G. M., KENT, D. S. & GILBERT, R. G. 2005. A new silk: Mechanical, compositional, and morphological characterization of leafhopper (Kahaono montana) silk. *Polymer,* 46**,** 7909-7917.

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. *Journal of Proteome Research,* 14**,** 4223-4231.

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 Evolutionary Biology,* 14**,** 31-31.

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-e52293.

CHILDRES, I., JAUREGUI, L. A., PARK, W., CAO, H. & CHEN, Y. P. 2013. Raman spectroscopy of graphene and related materials. *In:* JANG, J. I. (ed.) *New Developments in Photon and Materials Research.* Nova Science Publishers, Inc.

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., 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**,** 1856-70.

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**,** 365.

CLERCK, C. A. 1757. Svenska Spindlar uti sina hufvud-slågter indelte samt under några och sextio särskildte arter beskrefne och med illuminerade figurer uplyste / Aranei Svecici, descriptionibus et figuris æneis illustrati, ad genera subalterna redacti, speciebus ultra LX determinati.

CODDINGTON, J. & SOBREVILA, C. 1987. Web Manipulation and Two Stereotyped Attack Behaviors in the Ogre-Faced Spider Deinopis spinosus Marx (Araneae, Deinopidae). *The Journal of Arachnology,* 15**,** 213-225.

CODDINGTON, J. A. & LEVI, H. W. 1991. Systematics and evolution of spiders (Araneae). *Annual Review of Ecology and Systematics,* 22**,** 565-592.

CODDINGTON JONATHAN, A. 1986. Orb webs in "non-orb weaving" ogre-faced spiders (Araneae: Dinopidae): A question of genealogy. *Cladistics,* 2**,** 53-67.

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 Science : A Publication of the Protein Society,* 7**,** 667-672.

COLLIN, M. A., CLARKE, T. H., AYOUB, N. A. & HAYASHI, C. Y. 2016. Evidence from Multiple Species that Spider Silk Glue Component ASG2 is a Spidroin. *Sci Rep,* 6**,** 21589.

COLLIN, M. A., CLARKE, T. H., AYOUB, N. A. & HAYASHI, C. Y. 2018. Genomic perspectives of spider silk genes through target capture sequencing: Conservation of stabilization mechanisms and homology-based structural models of spidroin terminal regions. *International Journal of Biological Macromolecules,* 113**,** 829-840.

COLLIN, M. A., EDGERLY, J. S. & HAYASHI, C. Y. 2011. Comparison of fibroin cDNAs from webspinning insects: insight into silk formation and function. *Zoology (Jena),* 114**,** 239-46.

CORREA-GARHWAL, S. M., CHAW, R. C., DUGGER, T., CLARKE, T. H., CHEA, K. H., KISAILUS, D. & HAYASHI, C. Y. 2018. Semi-Aquatic Spider Silks: Transcripts, Proteins, and Silk fibers of the Fishing Spider, Dolomedes triton (Pisauridae). *Insect Molecular Biology,* 0.

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**,** 4598-605.

COYLE, F. A. 1986. The role of silk in prey capture by non-araneomorph spiders. *In:* SHEAR, W. (ed.) *Spiders: Webs, Behavior and Evolution.* Stanford: Stanford University Press.

CRAIG, C. L. 2003. *Spiderwebs and Silk: Tracing Evolution From Molecules to Genes to Phenotypes,* New York, Oxford University Press.

CROME, W. 1951. *Die Wasserspinne,* Leipzig, Die neue Brehm-Bücherei.

CUFF, J. A. & BARTON, G. J. 2000. Application of multiple sequence alignment profiles to improve protein secondary structure prediction. *Proteins: Structure, Function, and Bioinformatics,* 40**,** 502-511.

CUMINGS, J. & ZETTL, A. 2000. Low-Friction Nanoscale Linear Bearing Realized from Multiwall Carbon Nanotubes. *Science,* 289**,** 602.

DALTON, A. B., COLLINS, S., MUÑOZ, E., RAZAL, J. M., EBRON, V. H., FERRARIS, J. P., COLEMAN, J. N., KIM, B. G. & BAUGHMAN, R. H. 2003. Super-tough carbon-nanotube fibres. *Nature,* 423**,** 703.

DANILATOS, G. D. 1981. The examination of fresh or living plant material in an environmental scanning electron microscope. *Journal of Microscopy,* 121**,** 235-238.

DE BAKKER, D., BAETENS, K., VAN NIMMEN, E., GELLYNCK, K., MERTENS, J., VAN LANGENHOVE, L. & KIEKENS, P. 2006. Description of the structure of different silk threads produced by the water spider Argyroneta aquatica (Clerck, 1757) (Araneae: Cybaeidae). *Belgian Journal of Zoology,* 136**,** 137 - 143.

DEGRÈVE, L., FUZO, C. A. & CALIRI, A. 2014. Extended secondary structures in proteins. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics,* 1844**,** 384-388.

DEL-CLARO, K. & TIZO-PEDROSO, E. 2009. Ecological and evolutionary pathways of social behavior in Pseudoscorpions (Arachnida: Pseudoscorpiones). *acta ethologica,* 12**,** 13-22.

DEL RIO, T. G., CHERTKOV, O., YASAWONG, M., LUCAS, S., DESHPANDE, S., CHENG, J.-F., DETTER, C., TAPIA, R., HAN, C., GOODWIN, L., PITLUCK, S., LIOLIOS, K., IVANOVA, N., MAVROMATIS, K., PATI, A., CHEN, A., PALANIAPPAN, K., LAND, M., HAUSER, L., CHANG, Y.-J., JEFFRIES, C. D., ROHDE, M., PUKALL, R., SIKORSKI, J., GÖKER, M., WOYKE, T., BRISTOW, J., EISEN, J. A., MARKOWITZ, V., HUGENHOLTZ, P., KYRPIDES, N. C., KLENK, H.-P. & LAPIDUS, A. 2010. Complete genome sequence of Intrasporangium calvum type strain (7 KIP). *Standards in genomic sciences,* 3**,** 294-303.

DORAN, N. E., KIERNAN, K., SWAIN, R. & RICHARDSON, A. M. M. 1999. Hickmania Troglodytes, the Tasmanian Cave Spider, and its Potential Role in Cave Management. *Journal of Insect Conservation,* 3**,** 257-262.

DROZDETSKIY, A., COLE, C., PROCTER, J. & BARTON, G. J. 2015a. JPred4: a protein secondary structure prediction server. *Nucleic Acids Research,* 43**,** W389-W394.

DROZDETSKIY, A., COLE, C., PROCTER, J. & BARTON, G. J. 2015b. JPred4: a protein secondary structure prediction server. *Nucleic Acids Research*.

EBERHARD, W. G. 1977. Aggressive Chemical Mimicry by a Bolas Spider. *Science,* 198**,** 1173.

EDDY, S. R. 2011. Accelerated Profile HMM Searches. *PLOS Computational Biology,* 7**,** e1002195.

EISOLDT, L., SMITH, A. & SCHEIBEL, T. 2011. Decoding the secrets of spider silk. *Materials Today,* 14**,** 80-86.

ELETTRO, H., NEUKIRCH, S., VOLLRATH, F. & ANTKOWIAK, A. 2016. In-drop capillary spooling of spider capture thread inspires hybrid fibers with mixed solid–liquid mechanical properties. *Proceedings of the National Academy of Sciences,* 113**,** 6143.

FERNÁNDEZ, R., KALLAL, R. J., DIMITROV, D., BALLESTEROS, J. A., ARNEDO, M. A., GIRIBET, G. & HORMIGA, G. 2018. Phylogenomics, Diversification Dynamics, and Comparative Transcriptomics across the Spider Tree of Life. *Current Biology,* 28**,** 1489-1497.e5.

FINN, R. D., COGGILL, P., EBERHARDT, R. Y., EDDY, S. R., MISTRY, J., MITCHELL, A. L., POTTER, S. C., PUNTA, M., QURESHI, M., SANGRADOR-VEGAS, A., SALAZAR, G. A., TATE, J. & BATEMAN, A. 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research,* 44**,** D279-D285.

FOELIX, R. F. 2011. *Biology of spiders / Rainer F. Foelix,* Oxford, Oxford : Oxford University Press.

FRANKLIN-WALLIS, O. 2017. *These lightweight adidas shoes are made from spider silk grown in a lab* [Online]. Wired. Available: wired.co.uk/article/adidas-futurecraft-biofabric-shoes [Accessed 15/08/2018].

GAINES, W. A. T. & MARCOTTE, W. R., JR. 2008. Identification and characterization of multiple Spidroin 1 genes encoding major ampullate silk proteins in Nephila clavipes. *Insect molecular biology,* 17**,** 465-474.

GAO, Z., LIN, Z., HUANG, W., LAI, C. C., FAN, J.-S. & YANG, D. 2013. Structural Characterization of Minor Ampullate Spidroin Domains and Their Distinct Roles in Fibroin Solubility and Fiber Formation. *PLoS ONE,* 8**,** e56142.

GARALDE, D. R., SNELL, E. A., JACHIMOWICZ, D., SIPOS, B., LLOYD, J. H., BRUCE, M., PANTIC, N., ADMASSU, T., JAMES, P., WARLAND, A., JORDAN, M., CICCONE, J., SERRA, S., KEENAN, J., MARTIN, S., MCNEILL, L., WALLACE, E. J., JAYASINGHE, L., WRIGHT, C., BLASCO, J., YOUNG, S., BROCKLEBANK, D., JUUL, S., CLARKE, J., HERON, A. J. & TURNER, D. J. 2018. Highly parallel direct RNA sequencing on an array of nanopores. *Nature Methods,* 15**,** 201.

GARB, J. E. 2013. Spider silk: an ancient biomaterial for the 21st century. *In:* PENNEY, D. (ed.) *Spider research in the 21st century: trends and perspectives.* Manchester, UK: Siri Scientific 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. *Molecular biology and evolution,* 24**,** 2454-2464.

GARB, J. E., DIMAURO, T., VO, V. & HAYASHI, C. Y. 2006. Silk genes support the single origin of orb webs. *Science,* 312**,** 1762.

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**,** 11379-84.

GATESY, J., HAYASHI, C., MOTRIUK, D., WOODS, J. & LEWIS, R. 2001. Extreme diversity, conservation, and convergence of spider silk fibroin sequences. *Science (New York, N.Y.),* 291**,** 2603-2605.

GELLYNCK, K., VERDONK, P., ALMQVIST, F., VAN NIMMEN, E., DE BAKKER, D., VAN LANGENHOVE, L., MERTENS, J., VERBRUGGEN, G. & KIEKENS, P. 2006. *A spider silk supportive matrix used for cartilage regeneration,* Cambridge, Woodhead Publ Ltd.

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.

GLATZ, L. 1972. Der spinnapparat haplogyner spinnen (Arachnida, Araneae). *Zeitschrift für Morphologie der Tiere,* 72**,** 1-25.

GLATZ, L. 1973. Der spinnapparat der orthognatha (Arachnida, Araneae). *Zeitschrift für Morphologie der Tiere,* 75**,** 1-50.

GNESA, E., HSIA, Y., YARGER, J. L., WEBER, W., LIN-CEREGHINO, J., LIN-CEREGHINO, G., TANG, S., AGARI, K. & VIERRA, C. 2012. Conserved C-terminal domain of spider tubuliform spidroin 1 contributes to extensibility in synthetic fibers. *Biomacromolecules,* 13**,** 304-12.

GOSLINE, J. M., DEMONT, M. E. & DENNY, M. W. 1986. The structure and properties of spider silk. *Endeavour,* 10**,** 37-43.

GOSLINE, J. M., GUERETTE, P. A., ORTLEPP, C. S. & SAVAGE, K. N. 1999. The mechanical design of spider silks: from fibroin sequence to mechanical function. *J Exp Biol,* 202**,** 3295-303.

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 Biotech,* 29**,** 644-652.

GRAWE, I., WOLFF, J. O. & GORB, S. N. 2014. Composition and substrate-dependent strength of the silken attachment discs in spiders. *Journal of The Royal Society Interface,* 11.

GUINEA, G. V., ELICES, M., PLAZA, G. R., PEREA, G. B., DAZA, R., RIEKEL, C., AGULLÓ-RUEDA, F., HAYASHI, C., ZHAO, Y. & PÉREZ-RIGUEIRO, J. 2012. Minor ampullate silks from Nephila and Argiope spiders: tensile properties and microstructural characterization. *Biomacromolecules,* 13**,** 2087-2098.

HAAS, B. J., PAPANICOLAOU, A., YASSOUR, M., GRABHERR, M., BLOOD, P. D., BOWDEN, J., COUGER, M. B., ECCLES, D., LI, B., LIEBER, M., MACMANES, M. D., OTT, M., ORVIS, J., POCHET, N., STROZZI, F., WEEKS, N., WESTERMAN, R., WILLIAM, T., DEWEY, C. N., HENSCHEL, R., LEDUC, R. D., FRIEDMAN, N. & REGEV, A. 2013. De novo transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. *Nature protocols,* 8**,** 10.1038/nprot.2013.084.

HAGN, F., EISOLDT, L., HARDY, J. G., VENDRELY, C., COLES, M., SCHEIBEL, T. & KESSLER, H. 2010. A conserved spider silk domain acts as a molecular switch that controls fibre assembly. *Nature,* 465**,** 239-42.

HAN, Z. & FINA, A. 2011. Thermal conductivity of carbon nanotubes and their polymer nanocomposites: A review. *Progress in Polymer Science,* 36**,** 914-944.

HANSELL, M. H. 1996. The function of lichen flakes and white spider cocoons on the outer surface of birds' nests. *Journal of Natural History,* 30**,** 303-311.

HARVEY, D., BARDELANG, P., GOODACRE, S. L., COCKAYNE, A. & THOMAS, N. R. 2017. Antibiotic Spider Silk: Site-Specific Functionalization of Recombinant Spider Silk Using “Click” Chemistry. *Advanced Materials,* 29**,** 1604245-n/a.

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. *Molecular Biology and Evolution,* 21**,** 1950-1959.

HAYASHI, C. Y. & LEWIS, R. V. 1998. Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks1. *Journal of Molecular Biology,* 275**,** 773-784.

HAYASHI, C. Y. & LEWIS, R. V. 2000. Molecular Architecture and Evolution of a Modular Spider Silk Protein Gene. *Science,* 287**,** 1477.

HAYASHI, C. Y., SHIPLEY, N. H. & LEWIS, R. V. 1999. Hypotheses that correlate the sequence, structure, and mechanical properties of spider silk proteins. *International Journal of Biological Macromolecules,* 24**,** 271-275.

HAZAN, A., GERTLER, A., TAHORI, A. S. & GERSON, U. 1975. Spider mite webbing—III. Solubilization and amino acid composition of the silk protein. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry,* 51**,** 457-462.

HEIM, M., KEERL, D. & SCHEIBEL, T. 2009. Spider Silk: From Soluble Protein to Extraordinary Fiber. *Angewandte Chemie International Edition,* 48**,** 3584-3596.

HEIMER, S. 1988. *Wunderbare Welt der Spinnen*.

HINMAN, M. B., JONES, J. A. & LEWIS, R. V. 2000. Synthetic spider silk: a modular fiber. *Trends Biotechnol,* 18**,** 374-9.

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. *Journal of Biological Chemistry,* 267**,** 19320-4.

HORMIGA, G. & GRISWOLD, C. E. 2014. Systematics, Phylogeny, and Evolution of Orb-Weaving Spiders. *Annual Review of Entomology,* 59**,** 487-512.

HORNYAK, T. 2013. *Dress to kill in this synthetic spider silk outfit* [Online]. CNET Magazine. Available: cnet.com/news/dress-to-kill-in-this-synthetic-spider-silk-outfit/ [Accessed 15/08/2018].

HOUSELEY, J. & TOLLERVEY, D. 2009. The Many Pathways of RNA Degradation. *Cell,* 136**,** 763-776.

HU, X., KOHLER, K., FALICK, A. M., MOORE, A. M. F., JONES, P. R. & VIERRA, C. 2006. Spider Egg Case Core Fibers: Trimeric Complexes Assembled from TuSp1, ECP-1, and ECP-2†. *Biochemistry,* 45**,** 3506-3516.

HU, X., LAWRENCE, B., KOHLER, K., FALICK, A. M., MOORE, A. M. F., MCMULLEN, E., JONES, P. R. & VIERRA, C. 2005. Araneoid Egg Case Silk: A Fibroin with Novel Ensemble Repeat Units from the Black Widow Spider, Latrodectus hesperus. *Biochemistry,* 44**,** 10020-10027.

HUANG, W., LIN, Z., SIN, Y. M., LI, D., GONG, Z. & YANG, D. 2006. Characterization and expression of a cDNA encoding a tubuliform silk protein of the golden web spider Nephila antipodiana. *Biochimie,* 88**,** 849-58.

HUBER, B. A. 2011. Phylogeny and classification of Pholcidae (Araneae): an update. *The Journal of Arachnology,* 39**,** 211-222.

ITTAH, S., COHEN, S., GARTY, S., COHN, D. & GAT, U. 2006. An Essential Role for the C-Terminal Domain of A Dragline Spider Silk Protein in Directing Fiber Formation. *Biomacromolecules,* 7**,** 1790-1795.

ITTAH, S., MICHAELI, A., GOLDBLUM, A. & 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.

JACKSON, R. R. 1990. Predator-prey interactions between jumping spiders (Araneae, Salticidae) and Phokus phalangioides (Araneae, Pholcidae). *Journal of Zoology,* 220**,** 553-559.

JACKSON, R. R. & BRASSINGTON, R. J. 1987. The biology of Pholcus phalangioides (Araneae, Pholcidae): predatory versatility, araneophagy and aggressive mimicry. *Journal of Zoology,* 211**,** 227-238.

JIMÉNEZ SOTO, J. M. & ARANZANA, M. S. C. 2012. Conical Carbon Nanoparticles in Analytical Chemistry. *Encyclopedia of Analytical Chemistry*.

JOSHI, N. & FASS, J. 2011. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software].

K. SETH, R., QUINE, B. & ZHU, Z. H. 2009. *Feasibility of 20 km free-standing inflatable space tower*.

KANG, J.-H., LEE, Y.-J., OH, B.-K., LEE, S.-K., HYUN, B.-R., LEE, B.-W., CHOI, Y.-G., NAM, K.-S. & LIM, J.-D. 2014. Microstructure of the water spider (Argyroneta aquatica) using the scanning electron microscope. *Journal of Asia-Pacific Biodiversity,* 7**,** 484-488.

KARATZAS, C. N., TURNER, J. D. & LAZARIS, A. 2007. *Production of biofilaments in transgenic animals*. USA patent application.

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.

KELLY, S. 2018. Personal Communications. Tunghai University.

KOVOOR, J. 1987. Comparative Structure and Histochemistry of Silk-Producing Organs in Arachnids. *In:* NENTWIG, W. (ed.) *Ecophysiology of Spiders.* Berlin, Heidelberg: Springer Berlin Heidelberg.

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

KRÁL, J., MUSILOVÁ, J., ŠŤÁHLAVSKÝ, F., ŘEZÁČ, M., AKAN, Z., EDWARDS, R. L., COYLE, F. A. & ALMERJE, C. R. 2006. Evolution of the karyotype and sex chromosome systems in basal clades of araneomorph spiders (Araneae: Araneomorphae). *Chromosome Research,* 14**,** 859.

KRONENBERGER, K., DICKO, C. & VOLLRATH, F. 2012. A novel marine silk. *Naturwissenschaften,* 99**,** 3-10.

KRONQVIST, N., OTIKOVS, M., CHMYROV, V., CHEN, G., ANDERSSON, M., NORDLING, K., LANDREH, M., SARR, M., JÖRNVALL, H., WENNMALM, S., WIDENGREN, J., MENG, Q., RISING, A., OTZEN, D., KNIGHT, S. D., JAUDZEMS, K. & JOHANSSON, J. 2014. Sequential pH-driven dimerization and stabilization of the N-terminal domain enables rapid spider silk formation. *Nature Communications,* 5**,** 3254.

KURKA, A., REZAC, M., MACEK, R. & DOLANSKY, J. 2015. *Pavouci Ceske republiky,* Praha, Czechia, Academia, Praha.

LA MATTINA, 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,* 47**,** 4692-4700.

LAMANNA, G., BATTIGELLI, A., MÉNARD-MOYON, C. & BIANCO, A. 2012. Multifunctionalized carbon nanotubes as advanced multimodal nanomaterials for biomedical applications. *Nanotechnology Reviews,* 1**,** 17-29.

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.

LAZARIS, A., ARCIDIACONO, S., HUANG, Y., ZHOU, J.-F., DUGUAY, F., CHRETIEN, N., WELSH, E. A., SOARES, J. W. & KARATZAS, C. N. 2002. Spider Silk Fibers Spun from Soluble Recombinant Silk Produced in Mammalian Cells. *Science,* 295**,** 472-476.

LEE, C., WEI, X., KYSAR, J. W. & HONE, J. 2008. Measurement of the Elastic Properties and Intrinsic Strength of Monolayer Graphene. *Science,* 321**,** 385.

LEE, J.-K., LEE, S., KIM, J.-G., MIN, B.-K., KIM, Y.-I., LEE, K.-I., AN, K. H. & JOHN, P. 2014. Structure of Single-Wall Carbon Nanotubes: A Graphene Helix. *Small,* 10**,** 3283-3290.

LEE, K. S., KIM, B. Y., JE, Y. H., WOO, S. D., SOHN, H. D. & JIN, B. R. 2007. Molecular cloning and expression of the C-terminus of spider flagelliform silk protein from Araneus ventricosus. *Journal of biosciences,* 32**,** 705-712.

LEE, P. & THOMAS, T. *Summary for Araneus diadematus (Araneae)* [Online]. British Arachnological Society. Available: srs.britishspiders.org.uk/portal.php/p/Summary/s/Araneus+diadematus [Accessed 10/08/2018].

LEE, S.-M., PIPPEL, E., GÖSELE, U., DRESBACH, C., QIN, Y., CHANDRAN, C. V., BRÄUNIGER, T., HAUSE, G. & KNEZ, M. 2009. Greatly Increased Toughness of Infiltrated Spider Silk. *Science,* 324**,** 488.

LEPORE, E., BONACCORSO, F., BRUNA, M., BOSIA, F., TAIOLI, S., GARBEROGLIO, G., FERRARI, A. C. & PUGNO, N. M. 2015. Silk reinorced with graphene or carbon nanotubes spun by spiders.

LEPORE, E., BOSIA, F., BONACCORSO, F., BRUNA, M., TAIOLI, S., GARBEROGLIO, G., FERRARI, A. C. & PUGNO, N. M. 2017. Spider silk reinforced by graphene or carbon nanotubes. *2D Materials,* 4**,** 031013.

LIN, H., XU, S., ZHANG, Y.-Q. & WANG, X. 2014. Electron Transport and Bulk-like Behavior of Wiedemann–Franz Law for Sub-7 nm-Thin Iridium Films on Silkworm Silk. *ACS Applied Materials & Interfaces,* 6**,** 11341-11347.

LIN, Z., HUANG, W., ZHANG, J., FAN, J.-S. & YANG, D. 2009. Solution structure of eggcase silk protein and its implications for silk fiber formation. *Proceedings of the National Academy of Sciences of the United States of America,* 106**,** 8906-8911.

LISSNER, J. & SCHARFF, N. *Danish Spiders from A to Z: Ero furcata* [Online]. University of Copenhagen

Naturhistorik Museum Aarhus. Available: danmarks-edderkopper.dk/species-description?taxon=Ero%20furcata [Accessed 05/06/2018 2018].

LISSNER, J. & SCHARFF, N. *Danish Spiders from A to Z: Tetragnatha nigrita* [Online]. University of Copenhagen

Naturhistorisk Museum Aarhus. Available: danmarks-edderkopper.dk/species-description?taxon=Tetragnatha%20nigrita [Accessed 06/06/2018].

LU, X. & CHEN, Z. 2005. Curved Pi-Conjugation, Aromaticity, and the Related Chemistry of Small Fullerenes (<C60) and Single-Walled Carbon Nanotubes. *Chemical Reviews,* 105**,** 3643-3696.

LUO, C., LI, F., LI, D., FU, Q. & PAN, C. 2016. Bioinspired Single-Walled Carbon Nanotubes as a Spider Silk Structure for Ultrahigh Mechanical Property. *ACS Applied Materials & Interfaces,* 8**,** 31256-31263.

MADSEN, B., SHAO, Z. Z. & VOLLRATH, F. 1999. Variability in the mechanical properties of spider silks on three levels: interspecific, intraspecific and intraindividual. *International Journal of Biological Macromolecules,* 24**,** 301-306.

MANCHADO, M. A. L., VALENTINI, L., BIAGIOTTI, J. & KENNY, J. M. 2005. Thermal and mechanical properties of single-walled carbon nanotubes–polypropylene composites prepared by melt processing. *Carbon,* 43**,** 1499-1505.

MARTEL, A., BURGHAMMER, M., DAVIES, R. J., DI COLA, E., VENDRELY, C. & RIEKEL, C. 2008. Silk fiber assembly studied by synchrotron radiation SAXS/WAXS and Raman spectroscopy. *Journal of the American Chemical Society,* 130**,** 17070.

MASUMOTO, T., MASUMOTO, T., YOSHIDA, M. & MNISHIKAWA, Y. 1998a. Water conditions of the habitat of the water spider *Argyroneta aquatica* (Araneae: Argyronetidae) in Mizoro pond. *Acta arachnologica,* 47**,** 121 - 124.

MASUMOTO, T., MASUMOTO, T., YOSHIDA, M. & NISHIKAWA, Y. 1998b. Time budget of activity in the water spider *Argyroneta aquatica* (Araneae: Argyronetidae) under rearing condition. *Acta arachnologica,* 47**,** 125 - 131.

MATWEB, A. S. *ASTM A36 Steel* [Online]. MatWeb LLC. Available: matweb.com/search/datasheet.aspx?MatGUID=afc003f4fb40465fa3df05129f0e88e6&ckck=1 [Accessed 13/08/2018].

MESSNER, B. & ADIS, J. 1995. Es gibt nur fakultative Plastronatmer unter den tauchenden Webespinnen (Araneae). *Deutsche Entomologische Zeitschrift,* 42**,** 453-459.

MORGAN, E. 2016. *Gossamer Days: Spiders, Humans and Their Threads,* London, UK, Strange Attractor Press.

MOROZOVA, O. & MARRA, M. A. 2008. Applications of next-generation sequencing technologies in functional genomics. *Genomics,* 92**,** 255-264.

MORTON, S. & RENTZ, D. 1983. Ecology and Taxonomy of Fossorial, Granivorous Gryllacridids (Orthoptera: Gryllacrididae) from Arid Central Australia. *Australian Journal of Zoology,* 31**,** 557-579.

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**,** 3152-9.

MURAD, A. M. & RECH, E. L. 2011. Molecular dynamics simulations of the minor ampullate spidroin modular amino acid sequence from Parawixia bistriatra: insights into silk tertiary structure and fibre formation. *Journal of molecular modeling,* 17**,** 1183-1189.

NAN, C.-W., LIU, G., LIN, Y. & LI, M. 2004. Interface effect on thermal conductivity of carbon nanotube composites. *Applied Physics Letters,* 85**,** 3549-3551.

NENTWIG, W., ARNEDO, M. A., CODDINGTON, J. A., HORMIGA, G., JOCQUÉ, R., KROPF, C., MILLER, J. & SCHARFF, N. 2013. Appendix. Spider Phylogeny. *In:* NENTWIG, W. (ed.) *Spider Ecophysiology.* Berlin: Springer-Verlag.

NEUMANN, D. & 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.

NEUMANN, D. & WOERMANN, D. 2013. Stability of the volume of air trapped on the abdomen of the water spider Argyroneta aquatica. *SpringerPlus,* 2**,** 694.

NOWAK, R. M. 2014. Assembly of repetitive regions using next-generation sequencing data.

NYFFELER, M. & PUSEY, B. J. 2014. Fish Predation by Semi-Aquatic Spiders: A Global Pattern. *PLoS ONE,* 9**,** e99459.

O'CONNOR, I., HAYDEN, H., O'CONNOR, S., COLEMAN, J. N. & GUN'KO, Y. K. 2008. Kevlar coated carbon nanotubes for reinforcement of polyvinylchloride. *Journal of Materials Chemistry,* 18**,** 5585-5588.

O'CONNOR, I., HAYDEN, H., O'CONNOR, S., COLEMAN, J. N. & GUN'KO, Y. K. 2009. Polymer Reinforcement with Kevlar-Coated Carbon Nanotubes. *The Journal of Physical Chemistry C,* 113**,** 20184-20192.

ONO, H. 2002. New and Remarkable Spiders of the Families Liphistiidae, Argyronetidae, Pisauridae, Theridiidae and Araneidae (Arachnida) from Japan. *Bulletin of the National Science Museum. Series A, Zoology,* 28**,** 51 - 60.

PALACI, I., FEDRIGO, S., BRUNE, H., KLINKE, C., CHEN, M. & RIEDO, E. 2005. Radial Elasticity of Multiwalled Carbon Nanotubes. *Physical Review Letters,* 94**,** 175502.

PALMER JACQUELINE, M. 1985. The silk and silk production system of the funnel-web mygalomorph spider Euagrus (Araneae, Dipluridae). *Journal of Morphology,* 186**,** 195-207.

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. *Journal of Morphology,* 174**,** 269-274.

PAN, C., XIE, Q., HU, Z., YANG, M. & ZHU, L. 2015. Mechanical and biological properties of silk fibroin/carbon nanotube nanocomposite films. *Fibers and Polymers,* 16**,** 1781-1787.

PAN, H.-C., SONG, D.-X., ZHOU, K.-Y. & ZHU, G.-P. 2007. Cloning and prokaryotic expression of major ampullate spidroin gene of spider. *Sheng wu gong cheng xue bao = Chinese journal of biotechnology,* 23**,** 446-451.

PAYNE, A., HOLMES, N., RAKYAN, V. & LOOSE, M. 2018. Whale watching with BulkVis: A graphical viewer for Oxford Nanopore bulk fast5 files. *bioRxiv*.

PEERS, S. & GODLEY, N. 2011. *Golden Spider Silk* [Online]. Victoria and Albert Museum, London, UK. Available: vam.ac.uk/content/articles/g/golden-spider-silk/ [Accessed 13/06/2018].

PEREZ-RIGUEIRO, J., PLAZA, G. R., TORRES, F. G., HIJAR, A., HAYASHI, C., PEREA, G. B., ELICES, M. & GUINEA, G. V. 2010. Supercontraction of dragline silk spun by lynx spiders (Oxyopidae). *Int J Biol Macromol,* 46**,** 555-7.

PLAZAOLA, A. & C. CANDELAS, G. 1991. Stimulation of fibroin synthesis elicits ultrastructural modifications in spider silk secretory cells. *Tissue and Cell,* 23**,** 277-284.

POUCHKINA-STANTCHEVA, N. N. & MCQUEEN-MASON, S. J. 2004. Molecular studies of a novel dragline silk from a nursery web spider, Euprosthenops sp. (Pisauridae). *Comp Biochem Physiol B Biochem Mol Biol,* 138**,** 371-6.

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 Biochemistry and Molecular Biology,* 33**,** 229-238.

RAMANATHAN, T., FISHER, F. T., RUOFF, R. S. & BRINSON, L. C. 2005. Amino-Functionalized Carbon Nanotubes for Binding to Polymers and Biological Systems. *Chemistry of Materials,* 17**,** 1290-1295.

ŘEZÁČ, M. 2015. Personal communications. Crop Research Institute, Prague.

RISING, A., HJALM, G., ENGSTROM, W. & JOHANSSON, J. 2006. N-terminal nonrepetitive domain common to dragline, flagelliform, and cylindriform spider silk proteins. *Biomacromolecules,* 7**,** 3120-4.

RISING, A. & JOHANSSON, J. 2015. Toward spinning artificial spider silk. *Nat Chem Biol,* 11**,** 309-315.

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 Molecular Biology,* 16**,** 551-561.

ROBINSON, M. H. & ROBINSON, B. 1975. Evolution beyond the orb web: the web of the araneid spider Pasilobus sp., its structure, operation and construction. *Zoological Journal of the Linnean Society,* 56**,** 301-313.

RUOFF, R. S., TERSOFF, J., LORENTS, D. C., SUBRAMONEY, S. & CHAN, B. 1993. Radial deformation of carbon nanotubes by van der Waals forces. *Nature,* 364**,** 514.

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. *Nature Communications,* 5**,** 3765.

SCARSELLI, M., CASTRUCCI, P. & CRESCENZI, M. D. 2012. Electronic and optoelectronic nano-devices based on carbon nanotubes. *Journal of Physics: Condensed Matter,* 24**,** 313202.

SCHELLER, J., GUEHRS, K.-H., GROSSE, F. & CONRAD, U. 2001. *Production of spider silk proteins in tobacco and potato*.

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. *Evolutionary Ecology Research,* 5**,** 105 - 117.

SCHÜTZ, D. & TABORSKY, M. 2005. Mate choice and sexual conflict in the size dimorphic water spider *Argyroneta aquatica* (Araneae, Argyronetidae). *Journal of Arachnology,* 33**,** 767-775.

SCHÜTZ, D. & TABORSKY, M. 2011. Sexual Selection in the Water Spider: Female Choice and Male-Male Competition. *Ethology,* 117**,** 1101-1110.

SCHUTZ, D., TABORSKY, M. & DRAPELA, T. 2007. Air bells of water spiders are an extended phenotype modified in response to gas composition. *J Exp Zool A Ecol Genet Physiol,* 307**,** 549-55.

SCHWAGER, E. E., SHARMA, P. P., CLARKE, T., LEITE, D. J., WIERSCHIN, T., PECHMANN, M., AKIYAMA-ODA, Y., ESPOSITO, L., BECHSGAARD, J., BILDE, T., BUFFRY, A. D., CHAO, H., DINH, H., DODDAPANENI, H., DUGAN, S., EIBNER, C., EXTAVOUR, C. G., FUNCH, P., GARB, J., GONZALEZ, L. B., GONZALEZ, V. L., GRIFFITHS-JONES, S., HAN, Y., HAYASHI, C., HILBRANT, M., HUGHES, D. S. T., JANSSEN, R., LEE, S. L., MAESO, I., MURALI, S. C., MUZNY, D. M., NUNES DA FONSECA, R., PAESE, C. L. B., QU, J., RONSHAUGEN, M., SCHOMBURG, C., SCHÖNAUER, A., STOLLEWERK, A., TORRES-OLIVA, M., TURETZEK, N., VANTHOURNOUT, B., WERREN, J. H., WOLFF, C., WORLEY, K. C., BUCHER, G., GIBBS, R. A., CODDINGTON, J., ODA, H., STANKE, M., AYOUB, N. A., PRPIC, N.-M., FLOT, J.-F., POSNIEN, N., RICHARDS, S. & MCGREGOR, A. P. 2017. The house spider genome reveals an ancient whole-genome duplication during arachnid evolution. *BMC Biology,* 15**,** 62.

SELDEN, P. A., SHEAR, W. A. & SUTTON, M. D. 2008. Fossil evidence for the origin of spider spinnerets, and a proposed arachnid order. *Proceedings of the National Academy of Sciences,* 105**,** 20781-5.

SEPPO KOPONEN, T. P., NICLAS FRITZÉN, IKA ÖSTERBLAD 2015. Pholcus phalangioides in Finland (Araneae, Pholcidae). *Memoranda Societatis pro Fauna et Flora Fennica,* 91.

SEYMOUR, R. S. & HETZ, S. K. 2011. The diving bell and the spider: the physical gill of Argyroneta aquatica. *The Journal of Experimental Biology,* 214**,** 2175-2181.

SEYYAR, O. & DEMIR, H. 2009. Distribution and habitats of the water spider Argyroneta aquatica (Clerck, 1757) (Araneae, Cybaeidae) in Turkey. *Archives of Biological Sciences,* 61**,** 773-776.

SINGH, A., HEDE, S. & SASTRY, M. 2007. Spider Silk as an Active Scaffold in the Assembly of Gold Nanoparticles and Application of the Gold–Silk Bioconjugate in Vapor Sensing. *Small,* 3**,** 466-473.

SOLER, A. & ZAERA, R. 2016. The secondary frame in spider orb webs: the detail that makes the difference. *Scientific Reports,* 6**,** 31265.

SPONNER, A., UNGER, E., GROSSE, F. & 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., VATER, W., ROMMERSKIRCH, W., VOLLRATH, F., UNGER, E., GROSSE, F. & WEISSHART, K. 2005. The conserved C-termini contribute to the properties of spider silk fibroins. *Biochemical and Biophysical Research Communications,* 338**,** 897-902.

STARK, M., GRIP, S., RISING, A., HEDHAMMAR, M., ENGSTRÖM, W., HJÄLM, G. & JOHANSSON, J. 2007. Macroscopic Fibers Self-Assembled from Recombinant Miniature Spider Silk Proteins. *Biomacromolecules,* 8**,** 1695-1701.

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.

STEINS, A., DIK, P., MÜLLER, W. H., VERVOORT, S. J., REIMERS, K., KUHBIER, J. W., VOGT, P. M., VAN APELDOORN, A. A., COFFER, P. J. & SCHEPERS, K. 2015. In Vitro Evaluation of Spider Silk Meshes as a Potential Biomaterial for Bladder Reconstruction. *PLoS ONE,* 10**,** e0145240.

STEVEN, E., SALEH, W. R., LEBEDEV, V., ACQUAH, S. F. A., LAUKHIN, V., ALAMO, R. G. & BROOKS, J. S. 2013. Carbon nanotubes on a spider silk scaffold. *Nature Communications,* 4**,** 2435.

STRICKLAND, M., TUDORICA, V., ŘEZÁČ, M., THOMAS, N. R. & GOODACRE, S. L. 2018. Conservation of a pH-sensitive structure in the C-terminal region of spider silk extends across the entire silk gene family. *Heredity,* 120**,** 574-580.

SUTHERLAND, T. D., WEISMAN, S., TRUEMAN, H. E., SRISKANTHA, A., TRUEMAN, J. W. H. & HARITOS, V. S. 2007. Conservation of Essential Design Features in Coiled Coil Silks. *Molecular Biology and Evolution,* 24**,** 2424-2432.

SUTHERLAND, T. D., YOUNG, J. H., WEISMAN, S., HAYASHI, C. Y. & MERRITT, D. J. 2010. Insect Silk: One Name, Many Materials. *Annual Review of Entomology,* 55**,** 171-188.

SWANSON, B. O., BLACKLEDGE, T. A., BELTRÁN, J. & HAYASHI, C. Y. 2006a. Variation in the material properties of spider dragline silk across species. *Applied Physics A,* 82**,** 213-218.

SWANSON, B. O., BLACKLEDGE, T. A. & HAYASHI, C. Y. 2007. Spider capture silk: performance implications of variation in an exceptional biomaterial. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology,* 307A**,** 654-666.

SWANSON, B. O., BLACKLEDGE, T. A., SUMMERS, A. P. & HAYASHI, C. Y. 2006b. Spider dragline silk: Correlated and mosaic evolution in high-performance biological materials. *Evolution,* 60**,** 2539-2551.

TAI, P.-L., HWANG, G.-Y. & TSO, I. M. 2004. Inter-specific sequence conservation and intra-individual sequence variation in a spider silk gene. *International Journal of Biological Macromolecules,* 34**,** 237-243.

TAMURA, K., STECHER, G., PETERSON, D., FILIPSKI, A. & KUMAR, S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular biology and evolution,* 30**,** 2725-2729.

TEULÉ, F., MIAO, Y.-G., SOHN, B.-H., KIM, Y.-S., HULL, J. J., FRASER, M. J., LEWIS, R. V. & JARVIS, D. L. 2012. Silkworms transformed with chimeric silkworm/spider silk genes spin composite silk fibers with improved mechanical properties. *Proceedings of the National Academy of Sciences,* 109**,** 923-928.

THE UNIPROT CONSORTIUM 2017. UniProt: the universal protein knowledgebase. *Nucleic Acids Research,* 45**,** D158-D169.

TIAN, M. & LEWIS, R. V. 2005. Molecular Characterization and Evolutionary Study of Spider Tubuliform (Eggcase) Silk Protein†. *Biochemistry,* 44**,** 8006-8012.

TIAN, M., LIU, C. & LEWIS, R. 2004. Analysis of major ampullate silk cDNAs from two non-orb-weaving spiders. *Biomacromolecules,* 5**,** 657-60.

TILLINGHAST, E. K., CHASE, S. F. & TOWNLEY, M. A. 1984. Water extraction by the major ampullate duct during silk formation in the spider, Argiope aurantia Lucas. *Journal of Insect Physiology,* 30**,** 591-596.

TOKAREVA, O., JACOBSEN, M., BUEHLER, M., WONG, J. & KAPLAN, D. L. 2014. Structure-Function-Property-Design Interplay in Biopolymers: Spider Silk. *Acta biomaterialia,* 10**,** 1612-1626.

TREACY, M. M. J., EBBESEN, T. W. & GIBSON, J. M. 1996. Exceptionally high Young's modulus observed for individual carbon nanotubes. *Nature,* 381**,** 678.

TURNER, S. 1992. Spiders in the Crosshairs: Cobwebs,Instrument Makers, and the Search for the Perfect Line. *Journal of the Antique Telescope Society,* 1**,** 10-12.

Inside Porton Down: Britain's Secret Weapons Research Facility, 2016. Directed by USBORNE, T.: British Broadcasting Corporation.

UWINS, P. J. R., MURRAY, M. & GOULD, R. J. 1993. Effects of four different processing techniques on the microstructure of potatoes: Comparison with fresh samples in the ESEM. *Microscopy Research and Technique,* 25**,** 412-418.

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. *Proceedings of the National Academy of Sciences,* 99**,** 10266-10271.

VANDENABEELE, P. 2013. *Practical Raman Spectroscopy: An Introduction,* Chichester, West Sussex, U.K., John Wiley & Sons, Incorporated.

VASANTHAVADA, K., HU, X., FALICK, A. M., LA MATTINA, C., MOORE, A. M. F., 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. *The Journal of biological chemistry,* 282**,** 35088-35097.

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, C. 2012. Spider Glue Proteins Have Distinct Architectures Compared with Traditional Spidroin Family Members. *The Journal of Biological Chemistry,* 287**,** 35986-35999.

VOLLRATH, F. 2018. Personal Communications. University of Oxford.

VOLLRATH, F., BARTH, P., BASEDOW, A., ENGSTROM, W. & LIST, H. 2002. Local tolerance to spider silks and protein polymers in vivo. *In Vivo,* 16**,** 229-34.

VOLLRATH, F. & EDMONDS, D. T. 1989. Modulation of the mechanical properties of spider silk by coating with water. *Nature,* 340**,** 305.

VOLLRATH, F., FAIRBROTHER, W. J., WILLIAMS, R. J. P., TILLINGHAST, E. K., BERNSTEIN, D. T., GALLAGHER, K. S. & TOWNLEY, M. A. 1990. Compounds in the droplets of the orb spider's viscid spiral. *Nature,* 345**,** 526-528.

VOLLRATH, F. & KNIGHT, D. P. 1999. Structure and function of the silk production pathway in the Spider Nephila edulis. *International Journal of Biological Macromolecules,* 24**,** 243-249.

VOLLRATH, F. & KNIGHT, D. P. 2001. Liquid crystalline spinning of spider silk. *Nature,* 410**,** 541-548.

VOLLRATH, F., KNIGHT, D. P. & HU, X. W. 1998. Silk production in a spider involves acid bath treatment. *Proceedings of the Royal Society B: Biological Sciences,* 265**,** 817-820.

WANG, M.-S., GOLBERG, D. & BANDO, Y. 2010. Tensile Tests on Individual Single-Walled Carbon Nanotubes: Linking Nanotube Strength with Its Defects. *Advanced Materials,* 22**,** 4071-4075.

WELLCOME, G. F. Y. *25 Genomes for 25 Years* [Online]. Wellcome Sanger Institute, UK: Sanger Institute. Available: sanger.ac.uk/science/collaboration/25-genomes-25-years [Accessed 16/10/2018].

WENDT, H., HILLMER, A., REIMERS, K., KUHBIER, J. W., SCHÄFER-NOLTE, F., ALLMELING, C., KASPER, C. & VOGT, P. M. 2011. Artificial Skin – Culturing of Different Skin Cell Lines for Generating an Artificial Skin Substitute on Cross-Weaved Spider Silk Fibres. *PLoS ONE,* 6**,** e21833.

WETZEL, E. C., ECTOR, L., VAN DE VIJVER, B., COMPERE, P. & MANN, G. D. 2015. Morphology, typification and critical analysis of some ecologically important small naviculoid species (Bacillariophyta). *Fottea,* 15**,** 203-234.

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., ALMEIDA-SILVA, L., ALVAREZ-PADILLA, F., ARNEDO, M. A., BENAVIDES SILVA, L. R., BENJAMIN, S. P., BOND, J. E., GRISMADO, C. J., HASAN, E., HEDIN, M., IZQUIERDO, M. A., LABARQUE, F. M., LEDFORD, J., LOPARDO, L., MADDISON, W. P., MILLER, J. A., PIACENTINI, L. N., PLATNICK, N. I., POLOTOW, D., SILVA-DÁVILA, D., SCHARFF, N., SZŰTS, T., UBICK, D., VINK, C. J., WOOD, H. M. & ZHANG, J. 2017. The spider tree of life: phylogeny of Araneae based on target-gene analyses from an extensive taxon sampling. *Cladistics,* 33**,** 574-616.

WILSON, R. S. 1962a. The Control of Dragline Spinning in the Garden Spider. *Quarterly Journal of Microscopical Science,* s3-103**,** 557.

WILSON, R. S. 1962b. The Structure of the Dragline Control Valves in the Garden Spider. *Quarterly Journal of Microscopical Science,* s3-103**,** 549.

WOLFF, J. O., GRAWE, I., WIRTH, M., KARSTEDT, A. & GORB, S. N. 2015. Spider's super-glue: thread anchors are composite adhesives with synergistic hierarchical organization. *Soft Matter,* 11**,** 2394-2403.

WRIGHT, S. & GOODACRE, S. L. 2012. Evidence for antimicrobial activity associated with common house spider silk. *BMC Research Notes,* 5**,** 326.

XIA, X.-X., QIAN, Z.-G., KI, C. S., PARK, Y. H., KAPLAN, D. L. & LEE, S. Y. 2010. Native-sized recombinant spider silk protein produced in metabolically engineered *Escherichia coli* results in a strong fiber. *Proceedings of the National Academy of Sciences,* 107**,** 14059.

XU, D., SHI, X., THOMPSON, F., WEBER, W. S., MOU, Q. & YARGER, J. L. 2015. Protein secondary structure of Green Lynx spider dragline silk investigated by solid-state NMR and X-ray diffraction. *International journal of biological macromolecules,* 81**,** 171-179.

XU, M. & LEWIS, R. V. 1990. Structure of a protein superfiber: spider dragline silk. *Proceedings of the National Academy of Sciences of the United States of America,* 87**,** 7120-7124.

YANG, Y. H. & LI, W. Z. 2011. Radial elasticity of single-walled carbon nanotube measured by atomic force microscopy. *Applied Physics Letters,* 98**,** 041901.

YARGER, J. L., CHERRY, B. R. & VAN DER VAART, A. 2018. Uncovering the structure–function relationship in spider silk. *Nature Reviews Materials,* 3**,** 18008.

YEARGAN, K. V. 1994. Biology of Bolas Spiders. *Annual Review of Entomology,* 39**,** 81-99.

YONEMURA, N., SEHNAL, F., MITA, K. & TAMURA, T. 2006. Protein composition of silk filaments spun under water by caddisfly larvae. *Biomacromolecules,* 7**,** 3370.

YU, M.-F., FILES, B. S., AREPALLI, S. & RUOFF, R. S. 2000a. Tensile Loading of Ropes of Single Wall Carbon Nanotubes and their Mechanical Properties. *Physical Review Letters,* 84**,** 5552-5555.

YU, M.-F., LOURIE, O., DYER, M. J., MOLONI, K., KELLY, T. F. & RUOFF, R. S. 2000b. Strength and Breaking Mechanism of Multiwalled Carbon Nanotubes Under Tensile Load. *Science,* 287**,** 637.

ZEVENBERGEN, J. M., SCHNEIDER, N. K. & BLACKLEDGE, T. A. 2008. Fine dining or fortress? Functional shifts in spider web architecture by the western black widow Latrodectus hesperus. *Animal Behaviour,* 76**,** 823-829.

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. *Comparative biochemistry and physiology. Part B, Biochemistry & molecular biology,* 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,* 45**,** 3348-3356.

ZHAO, A., ZHAO, T., SIMA, Y., ZHANG, Y., NAKAGAKI, K., MIAO, Y., SHIOMI, K., KAJIURA, Z., NAGATA, Y. & NAKAGAKI, M. 2005. Unique molecular architecture of egg case silk protein in a spider, Nephila clavata. *Journal of biochemistry,* 138**,** 593-604.

ZHOU, C.-Z., CONFALONIERI, F., MEDINA, N., ZIVANOVIC, Y., ESNAULT, C., YANG, T., JACQUET, M., JANIN, J., DUGUET, M., PERASSO, R. & LI, Z.-G. 2000. Fine organization of Bombyx mori fibroin heavy chain gene. *Nucleic Acids Research,* 28**,** 2413-2419.

ZSOLDOS, I., KAKUK, G., JANIK, J. & PÉK, L. 2005. Set of carbon nanotube junctions. *Diamond and Related Materials,* 14**,** 763-765.

8. Bibliography

9. Appendix

9.1. Local BLAST database spidroin sequence list

Table 9.1. Full list of all spidroin sequences included in the custom local BLAST database.
Direct submission = sequence submitted directly to GenBank and was not associated with a publication.

Family	Species	Silk type	Accession	Source
Agelenidae	Agelenopsis aperta	Tubuliform	AY566305	(Tian et al., 2004)
	Agelenopsis aperta	Tubuliform	HM752572	(Garb et al., 2010)
	Agelenopsis aperta	Major ampullate	HM752573	
	Agelenopsis aperta	Tubuliform	HM752576	
Antrodiaetidae	Aliatypus gulosus	Unknown spidroin	EU117159	(Garb et al., 2007)
	Atypoides riversi	Unknown spidroin	JX102562	(Starrett et al., 2012)
	Atypoides riversi	Unknown spidroin	JX102563	
Araneidae	Araneus diadematus	Unknown spidroin	ADU47853	(Del Rio et al., 2010)
	Araneus diadematus	Unknown spidroin	ADU47854	
	Araneus diadematus	Unknown spidroin	ADU47855	
	Araneus diadematus	Unknown spidroin	ADU47856	
	Araneus ventricosus	Aciniform	AY174110	Direct submission
	Araneus ventricosus	Minor ampullate	AY177203	Direct submission
	Araneus ventricosus	Minor ampullate	AY587193	Direct submission
	Araneus gemmoides	Tubuliform	AY855100	(Tian and Lewis, 2005)
	Araneus gemmoides	Tubuliform	AY855101	
	Araneus ventricosus	Major ampullate	AY940091	Direct submission
	Araneus ventricosus	Major ampullate	AY945306	Direct submission
	Araneus ventricosus	Flagelliform	EF025541	(Lee et al., 2007)
	Araneus ventricosus	Aciniform	HQ008714	Direct submission
	Araneus gemmoides	Piriform	JF827034	Direct submission
	Araneus ventricosus	Minor ampullate	JX513954	(Chen et al., 2012)
	Araneus ventricosus	Minor ampullate	JX513955	
	Araneus ventricosus	Minor ampullate	JX513956.1	
	Argiope bruennichi	Tubuliform	AB242144	(Zhao et al., 2006)
	Argiope bruennichi	Tubuliform	AB242145	
	Argiope trifasciata	Flagelliform	AF350264	(Gatesy et al., 2001)
	Argiope trifasciata	Flagelliform	AF350265	
	Argiope trifasciata	Major ampullate	AF350266	
	Argiope trifasciata	Major ampullate	AH015065	(Motriuk-Smith et al., 2005)
	Argiope amoena	Aciniform	AY263390	Direct submission
	Argiope amoena	Major ampullate	AY365015	Direct submission
	Argiope amoena	Major ampullate	AY365016	Direct submission
	Argiope amoena	Major ampullate	AY365017	(Pan et al., 2007)
	Argiope amoena	Major ampullate	AY365018	Direct submission
	Argiope amoena	Major ampullate	AY365019	Direct submission
	Argiope amoena	Major ampullate	AY365020	Direct submission
	Argiope amoena	Major ampullate	AY365021	Direct submission
	Argiope trifasciata	Aciniform	AY426339	(Hayashi et al., 2004)
	Argiope amoena	Major ampullate	AY453691	Direct submission
	Argiope aurantia	Tubuliform	AY855098	(Tian and Lewis, 2005)
	Argiope aurantia	Tubuliform	AY855099	
	Argiope argentata	Tubuliform	AY953071	(Garb and Hayashi, 2005)
	Argiope aurantia	Tubuliform	AY953081	
	Argiope argentata	Tubuliform	AY953084	
	Argiope argentata	Tubuliform	AY953091	
	Argiope aurantia	Tubuliform	AY953092	
	Argiope trifasciata	Piriform	GQ980328	Direct submission
	Argiope argentata	Tubuliform	HM752577	(Garb et al., 2010)
	Argiope amoena	Aciniform	HQ008715	Direct submission
	Argiope amoena	Tubuliform	JQ291306	Direct submission
	Argiope argentata	Minor ampullate	JQ713003	(Guinea et al., 2012)
	Argiope argentata	Minor ampullate	JQ713004	
	Argiope bruennichi	Major ampullate	JX112872	(Zhang et al., 2013)
	Argiope bruennichi	Major ampullate	JX202781	
	Argiope argentata	Aciniform	KJ206570	(Chaw et al., 2014)

Argiope argentata	Aciniform	KJ206571	
Argiope argentata	Aciniform	KJ206572	
Argiope argentata	Aciniform	KJ206573	
Argiope argentata	Aciniform	KJ206574	
Argiope argentata	Aciniform	KJ206575	
Argiope argentata	Aciniform	KJ206576	
Argiope aurantia	Aciniform	KJ206577	
Argiope aurantia	Aciniform	KJ206578	
Argiope aurantia	Aciniform	KJ206579	
Argiope aurantia	Aciniform	KJ206580	
Argiope aurantia	Aciniform	KJ206581	
Argiope aurantia	Aciniform	KJ206582	
Argiope trifasciata	Aciniform	KJ206583	
Argiope trifasciata	Aciniform	KJ206584	
Argiope trifasciata	Aciniform	KJ206585	
Argiope trifasciata	Aciniform	KJ206586	
Argiope trifasciata	Aciniform	KJ206587	
Argiope trifasciata	Aciniform	KJ206588	
Argiope trifasciata	Aciniform	KJ206589	
Argiope trifasciata	Aciniform	KJ206590	
Argiope trifasciata	Aciniform	KJ206591	
Argiope trifasciata	Aciniform	KJ206592	
Argiope trifasciata	Aciniform	KJ206593	
Argiope trifasciata	Aciniform	KJ206594	
Argiope trifasciata	Aciniform	KJ206595	
Argiope trifasciata	Aciniform	KJ206596	
Argiope argentata	Aciniform	KJ206597	
Argiope aurantia	Aciniform	KJ206598	
Argiope trifasciata	Aciniform	KJ206599	
Argiope argentata	Aciniform	KJ206600	
Argiope argentata	Aciniform	KJ206601	
Argiope argentata	Aciniform	KJ206602	
Argiope argentata	Aciniform	KJ206603	
Argiope argentata	Aciniform	KJ206604	
Argiope argentata	Aciniform	KJ206605	
Argiope argentata	Aciniform	KJ206606	
Argiope aurantia	Aciniform	KJ206607	
Argiope aurantia	Aciniform	KJ206608	
Argiope trifasciata	Aciniform	KJ206609	
Argiope trifasciata	Aciniform	KJ206610	
Argiope trifasciata	Aciniform	KJ206611	
Argiope trifasciata	Aciniform	KJ206612	
Argiope trifasciata	Aciniform	KJ206613	
Argiope trifasciata	Aciniform	KJ206614	
Argiope trifasciata	Aciniform	KJ206615	
Argiope trifasciata	Aciniform	KJ206616	
Argiope trifasciata	Aciniform	KJ206617	
Argiope trifasciata	Aciniform	KJ206618	
Argiope trifasciata	Aciniform	KJ206619	
Argiope argentata	Aciniform	KJ206620	
Cyrtophora moluccensis	Major ampullate	AY666060	(Tai et al., 2004)
Cyrtophora moluccensis	Major ampullate	AY666061	
Cyrtophora moluccensis	Major ampullate	AY666062	
Cyrtophora moluccensis	Major ampullate	AY666063	
Cyrtophora moluccensis	Major ampullate	AY666078	
Cyrtophora moluccensis	Major ampullate	AY666079	
Cyrtophora moluccensis	Major ampullate	AY666080	
Cyrtophora moluccensis	Major ampullate	AY666081	
Cyrtophora moluccensis	Tubuliform	AY953083	(Garb and Hayashi, 2005)
Gasteracantha mammosa	Major ampullate	AF350272	(Gatesy et al., 2001)
Gea heptagon	Tubuliform	AY953082	(Garb and Hayashi, 2005)
Gea heptagon	Tubuliform	AY953093	
Metepeira grandiosa	Minor ampullate	HM752569	(Garb et al., 2010)
Metepeira grandiosa	Minor ampullate	HM752575	
Nephila clavata	Tubuliform	AB218973	(Zhao et al., 2005)
Nephila clavata	Tubuliform	AB218974	
Nephila clavipes	Minor ampullate	AF027735	(Colgin and Lewis, 1998)
Nephila clavipes	Minor ampullate	AF027736	
Nephila clavipes	Minor ampullate	AF027737	

Nephila clavipes	Flagelliform	AF027972	(Hayashi and Lewis, 1998)
Nephila clavipes	Flagelliform	AF027973	
Nephila clavata	Major ampullate	AF441245	Direct submission
Nephila inaurata	Flagelliform	AH009147	(Hayashi and Lewis, 2000)
Nephila pilipes	Major ampullate	AY666048	(Tai et al., 2004)
Nephila pilipes	Major ampullate	AY666049	
Nephila pilipes	Major ampullate	AY666050	
Nephila pilipes	Major ampullate	AY666051	
Nephila pilipes	Major ampullate	AY666052	
Nephila pilipes	Major ampullate	AY666053	
Nephila pilipes	Major ampullate	AY666054	
Nephila pilipes	Major ampullate	AY666055	
Nephila pilipes	Major ampullate	AY666056	
Nephila pilipes	Major ampullate	AY666069	
Nephila pilipes	Major ampullate	AY666070	
Nephila pilipes	Major ampullate	AY666071	
Nephila pilipes	Major ampullate	AY666072	
Nephila pilipes	Major ampullate	AY666073	
Nephila pilipes	Major ampullate	AY666074	
Nephila pilipes	Major ampullate	AY666075	
Nephila pilipes	Major ampullate	AY666076	
Nephila pilipes	Major ampullate	AY666077	
Nephila clavipes	Tubuliform	AY855102	(Tian and Lewis, 2005)
Nephila inaurata	Major ampullate	DQ059135	(Motriuk-Smith et al., 2005)
Nephila antipodiana	Tubuliform	DQ089048	(Huang et al., 2006)
Nephila antipodiana	Minor ampullate	DQ338462	
Nephila antipodiana	Unknown spidroin	DQ338463	
Nephila clavipes	Major ampullate	EU599241	(Gaines and Marcotte, 2008)
Nephila clavipes	Major ampullate	EU599242	
Nephila clavipes	Major ampullate	EU599243	
Nephila clavipes	Major ampullate	EU617338	Direct submission
Nephila antipodiana	Tubuliform	EU730637	(Lin et al., 2009)
Nephila clavipes	Aggregate	EU780014	Direct submission
Nephila clavipes	Aggregate	EU780015	Direct submission
Nephila clavipes	Piriform	GQ980329	Direct submission
Nephila clavipes	Piriform	GQ980330	Direct submission
Nephila clavipes	Piriform	HM020705	(Geurts et al., 2010)
Nephila clavipes	Piriform	HM020705.1	
Nephila clavipes	Tubuliform	MWRG01000785.1	(Babb et al., 2017)
Nephila clavipes	Unknown spidroin	MWRG01000913.1	
Nephila clavipes	Unknown spidroin	MWRG01001340.1	
Nephila clavipes	Minor ampullate	MWRG01004753.1	
Nephila clavipes	Unknown spidroin	MWRG01005552.1	
Nephila clavipes	Unknown spidroin	MWRG01008157.1	
Nephila clavipes	Major ampullate	MWRG01008379.1	
Nephila clavipes	Flagelliform	MWRG01009190.1	
Nephila clavipes	Aggregate	MWRG01010266.1	
Nephila clavipes	Aciniform	MWRG01010975.1	
Nephila clavipes	Piriform	MWRG01012117.1	
Nephila clavipes	Flagelliform	MWRG01014089.1	
Nephila clavipes	Minor ampullate	MWRG01014802.1	
Nephila clavipes	Unknown spidroin	MWRG01014931.1	
Nephila clavipes	Major ampullate	MWRG01015344.1	
Nephila clavipes	Unknown spidroin	MWRG01016148.1	
Nephila clavipes	Major ampullate	MWRG01016260.1	
Nephila clavipes	Unknown spidroin	MWRG01016846.1	
Nephila clavipes	Minor ampullate	MWRG01017131.1	
Nephila clavipes	Aggregate	MWRG01017837.1	
Nephila clavipes	Aggregate	MWRG01023134.1	
Nephila clavipes	Major ampullate	MWRG01038219.1	
Nephila clavipes	Unknown spidroin	MWRG01057593.1	
Nephila clavipes	Major ampullate	MWRG01085391.1	
Nephila clavipes	Minor ampullate	MWRG01085987.1	
Nephila clavipes	Major ampullate	MWRG01095506.1	
Nephila clavipes	Unknown spidroin	NCU20329	(Xu and Lewis, 1990)
Nephilengys cruentata	Flagelliform	EF638444	(Bittencourt et al., 2007)

	Nephilengys cruentata	Tubuliform	EF638445	
	Nephilengys cruentata	Major ampullate	EF638446	
	Nephilengys cruentata	Minor ampullate	EF638447	
	Nephilengys cruentata	Minor ampullate	EF638448	
	Nephilengys cruentata	Minor ampullate	EF638449	
	Nephilengys cruentata	Piriform	GU062417	Direct submission
	Parawixia bistriata	Aciniform	GQ275356	Direct submission
	Parawixia bistriata	Flagelliform	GQ275357	Direct submission
	Parawixia bistriata	Minor ampullate	GQ275358	(Murad and Rech, 2011)
Atypidae	Sphodros rufipes	Unknown spidroin	JX102564	(Starrett et al., 2012)
Deinopidae	Deinopis spinosa	Tubuliform	AY953073	(Garb and Hayashi, 2005)
	Deinopis spinosa	Unknown spidroin	DQ399323	(Garb et al., 2006)
	Deinopis spinosa	Minor ampullate	DQ399324	
	Deinopis spinosa	Flagelliform	DQ399325	
	Deinopis spinosa	Unknown spidroin	DQ399326	
	Deinopis spinosa	Unknown spidroin	DQ399327	
	Deinopis spinosa	Major ampullate	DQ399328	
	Deinopis spinosa	Major ampullate	DQ399329	
	Deinopis spinosa	Unknown spidroin	DQ399330	
	Deinopis spinosa	Major ampullate	HM752568	(Garb et al., 2010)
Diguetiddae	Diguetia canities	Major ampullate	HM752564	
	Diguetia canities	Major ampullate	HM752565	
	Diguetia canities	Major ampullate	HM752566	
	Diguetia canities	Major ampullate	HM752567	
Dipluridae	Euagrus chisoseus	Unknown spidroin	AF350271	(Gatesy et al., 2001)
	Euagrus chisoseus	Unknown spidroin	EU117165	(Garb et al., 2007)
Eresidae	Stegodyphus mimosarum	Piriform	S.m._PiSp	(Sanggaard et al., 2014)
	Stegodyphus mimosarum	Aciniform	Stegodyphus_mimosarum_AcSp	
	Stegodyphus mimosarum	Major ampullate	Stegodyphus_mimosarum_MaSp-a	
	Stegodyphus mimosarum	Major ampullate	Stegodyphus_mimosarum_MaSp-b	
	Stegodyphus mimosarum	Major ampullate	Stegodyphus_mimosarum_MaSp-c	
	Stegodyphus mimosarum	Major ampullate	Stegodyphus_mimosarum_MaSp-d	
	Stegodyphus mimosarum	Major ampullate	Stegodyphus_mimosarum_MaSp-e	
	Stegodyphus mimosarum	Major ampullate	Stegodyphus_mimosarum_MaSp-f	
	Stegodyphus mimosarum	Major ampullate	Stegodyphus_mimosarum_MaSp-g	
	Stegodyphus mimosarum	Major ampullate	Stegodyphus_mimosarum_MaSp-h	
	Stegodyphus mimosarum	Major ampullate	Stegodyphus_mimosarum_MaSp-i	
	Stegodyphus mimosarum	Major ampullate	Stegodyphus_mimosarum_MaSp-j	
	Stegodyphus mimosarum	Major ampullate	Stegodyphus_mimosarum_MaSp-k	
	Stegodyphus mimosarum	Minor ampullate	Stegodyphus_mimosarum_MiSp	
	Stegodyphus mimosarum	Unknown spidroin	Stegodyphus_mimosarum_S.m._2	
	Stegodyphus mimosarum	Unknown spidroin	Stegodyphus_mimosarum_S.m._3	
	Stegodyphus mimosarum	Unknown spidroin	Stegodyphus_mimosarum_S.m._4	
	Stegodyphus mimosarum	Unknown spidroin	Stegodyphus_mimosarum_S.m._5	
	Stegodyphus mimosarum	Tubuliform	Stegodyphus_mimosarum_TuSp	
Euctenizidae	Aptostichus sp.	Unknown spidroin	EU117160	(Garb et al., 2007)
	Aptostichus sp.	Unknown spidroin	EU117161	
Filistatidae	Kukulcania hibernalis	Major ampullate	AY571309	(Tian et al., 2004)
	Kukulcania hibernalis	Major ampullate	AY571310	
	Kukulcania hibernalis	Major ampullate	HM752563	(Garb et al., 2010)
Halonoproctidae	Bothriocyrtum californicum	Unknown spidroin	EU117162	(Garb et al., 2007)
	Bothriocyrtum californicum	Unknown spidroin	EU117163	
	Bothriocyrtum californicum	Unknown spidroin	EU117164	

Hypochilidae	Hypochilus thorelli	Unknown spidroin	JX102555	(Starrett et al., 2012)
	Hypochilus thorelli	Unknown spidroin	JX102556	
	Kukulcania hibernalis	Major ampullate	AY571307	(Tian et al., 2004)
	Kukulcania hibernalis	Major ampullate	AY571308	
Macrothelidae	Macrothele holsti	Major ampullate	AY666068	(Tai et al., 2004)
Mecicobothriidae	Hexura picea	Unknown spidroin	JX102565	(Starrett et al., 2012)
	Megahexura fulva	Unknown spidroin	JX102566	
Pisauridae	Dolomedes tenebrosus	Unknown spidroin	AF350270	(Gatesy et al., 2001)
	Euprosthenops australis	Major ampullate	AM259067	(Rising et al., 2006)
	Euprosthenops australis	Major ampullate	AM490170	(Rising et al., 2007)
	Euprosthenops australis	Major ampullate	AM490171	
	Euprosthenops australis	Major ampullate	AM490172	
	Euprosthenops australis	Major ampullate	AM490173	
	Euprosthenops australis	Major ampullate	AM490174	
	Euprosthenops australis	Major ampullate	AM490175	
	Euprosthenops australis	Major ampullate	AM490176	
	Euprosthenops australis	Major ampullate	AM490177	
	Euprosthenops australis	Major ampullate	AM490178	
	Euprosthenops australis	Major ampullate	AM490179	
	Euprosthenops australis	Major ampullate	AM490180	
	Euprosthenops australis	Major ampullate	AM490181	
	Euprosthenops australis	Major ampullate	AM490182	
	Euprosthenops australis	Major ampullate	AM490183	
	Euprosthenops australis	Major ampullate	AM490184	
	Euprosthenops australis	Major ampullate	AM490185	
	Euprosthenops australis	Major ampullate	AM490186	
	Euprosthenops australis	Major ampullate	AM490187	
	Euprosthenops australis	Major ampullate	AM490188	
	Euprosthenops australis	Major ampullate	AM490189	
	Euprosthenops australis	Major ampullate	AM490190	
	Euprosthenops australis	Major ampullate	AM490191	
	Euprosthenops australis	Major ampullate	AM490192	
	Euprosthenops sp.	Major ampullate	AY338244	(Pouchkina-Stantcheva and McQueen-Mason, 2004)
Plectreuridae	Plectreurys tristis	Unknown spidroin	AF350282	(Gatesy et al., 2001)
	Plectreurys tristis	Unknown spidroin	AF350283	
	Plectreurys tristis	Unknown spidroin	AF350284	
Psechridae	Psechrus sinensis	Major ampullate	AY666064	(Tai et al., 2004)
	Psechrus sinensis	Major ampullate	AY666065	
	Psechrus sinensis	Major ampullate	AY666066	
	Psechrus sinensis	Major ampullate	AY666067	
Scytodidae	Scytodes thoracica	Unknown spidroin	KM987233	(Correa-Garhwal and Garb, 2014)
	Scytodes thoracica	Unknown spidroin	KM987234	
Theraphosidae	Aphonopelma seemanni	Unknown spidroin	JX102557	(Starrett et al., 2012)
	Aphonopelma seemanni	Unknown spidroin	JX102558	
	Aphonopelma seemanni	Unknown spidroin	JX102559	
	Avicularia juruensis	Unknown spidroin	EU652181	(Bittencourt et al., 2010)
	Avicularia juruensis	Unknown spidroin	EU652182	Direct submission
	Avicularia juruensis	Unknown spidroin	EU652183	Direct submission
	Avicularia juruensis	Unknown spidroin	EU652184	(Bittencourt et al., 2010)
	Poecilotheria regalis	Unknown spidroin	JX102560	(Starrett et al., 2012)
	Poecilotheria regalis	Unknown spidroin	JX102561	
Theridiidae	Latrodectus geometricus	Major ampullate	AF350273	(Gatesy et al., 2001)
	Latrodectus geometricus	Major ampullate	AF350274	
	Latrodectus geometricus	Major ampullate	AH015064	(Motriuk-Smith et al., 2005)
	Latrodectus geometricus	Major ampullate	AY685201	Direct submission
	Latrodectus hesperus	Tubuliform	AY953070	(Garb and Hayashi, 2005)
	Latrodectus hesperus	Major ampullate	AY953074	
	Latrodectus hesperus	Major ampullate	AY953075	
	Latrodectus hesperus	Tubuliform	AY953076	
	Latrodectus mactans	Tubuliform	AY953077	
	Latrodectus tredecimguttatus	Tubuliform	AY953078	
	Latrodectus geometricus	Tubuliform	AY953079	
	Latrodectus hasselti	Tubuliform	AY953080	
	Latrodectus mactans	Tubuliform	AY953085	
	Latrodectus hesperus	Tubuliform	AY953086	
	Latrodectus tredecimguttatus	Tubuliform	AY953087	

	Latrodectus hasselti	Tubuliform	AY953088	
	Latrodectus geometricus	Tubuliform	AY953089	
	Latrodectus hesperus	Tubuliform	DQ109035	(Hu et al., 2005)
	Latrodectus hesperus	Tubuliform	DQ341220	(Hu et al., 2006)
	Latrodectus hesperus	Tubuliform	DQ379383	Direct submission
	Latrodectus hesperus	Major ampullate	EF595247	(Ayoub et al., 2007)
	Latrodectus hesperus	Major ampullate	EF595248	
	Latrodectus hesperus	Aciniform	EU025854	(Vasanthavada et al., 2007)
	Latrodectus hesperus	Unknown spidroin	EU394445	(La Mattina et al., 2008)
	Latrodectus hesperus	Piriform	FJ973621.1	(Blasingame et al., 2009)
	Latrodectus hesperus	Minor ampullate	HM752570	(Garb et al., 2010)
	Latrodectus hesperus	Minor ampullate	HM752571	
	Latrodectus hesperus	Piriform	HQ005791.1	Direct submission
	Latrodectus hesperus	Major ampullate	HQ005804	Direct submission
	Latrodectus hesperus	Tubuliform	HQ005885	Direct submission
	Latrodectus hesperus	Tubuliform	HQ005889	Direct submission
	Latrodectus hesperus	Aggregate	JX262189	(Vasanthavada et al., 2012)
	Latrodectus hesperus	Aggregate	JX262192	
	Latrodectus hesperus	Aggregate	JX262195	
	Steatoda grossa	Tubuliform	AY953090	(Garb and Hayashi, 2005)
Uloboridae	Octonoba varians	Major ampullate	AY666057	(Tai et al., 2004)
	Octonoba varians	Major ampullate	AY666058	
	Octonoba varians	Major ampullate	AY666059	
	Uloborus diversus	Tubuliform	AY953072	(Garb and Hayashi, 2005)
	Uloborus diversus	Major ampullate	DQ399331	(Garb et al., 2006)
	Uloborus diversus	Minor ampullate	DQ399332	
	Uloborus diversus	Aciniform	DQ399333	
	Uloborus diversus	Major ampullate	DQ399334	
	Uloborus diversus	Major ampullate	DQ399335	
	Uloborus diversus	Minor ampullate	HM752574	(Garb et al., 2010)

9.2. Sequences of novel spidroins identified in *Argyroneta aquatica*, *Cybaeus angustiarum* and *Pholcus phalangioides* transcriptomes.

61 novel partial spidroin sequences were retrieved from the 5 transcriptomes. 21 of these sequences were uploaded to GenBank as part of Strickland et al. 2018 and given the accession numbers MG744694–MG744714. The remaining 40 sequences are provided below in FASTA format. Aa =*Argyroneta aquatica*, Ca = *Cybaeus angustiarum*, Pp = *Pholcus phalangioides,* wa = whole abdomen transcriptome, sgo = combined silk glands transcriptome, MaSp = major ampullate, MiSp = minor ampullate, AcSp = aciniform, PiSp = piriform, TuSp = tubuliform.

>putative_aa_sgo_AcSp_Nterm Gene.72::c178_g1_i1::g.72 ORF type:internal len:192 (-) c178_g1_i1:2-574(-)
AGAGCAGAAAAGTTCATCGATTGTTTAACGGATGGAATCGCAAAATCGCCAATTTTCCCTCAGCAAGAAAAAGATGATTTGCAGTATATTGTCGAGACAATGACA
TCAGCCATAACTAGCACAAGCTCTCAAAGGGGTAATAGAGCACAACTTCAAGCTTTAAATATGGCATTTGCATCTTCGCTGGCAGAGCTTGTAATAGCAGAAGAC
ATAGGAAACACGCAAGCAATCTCGGTCAAAACACAGGCTCTAACACAGTCTCTGTCGTCGTGCTTCAGATCCGTTATTGGAACGGTTAACAAACGTTTCATCAAC
GAGATAAAAGATTTAATAAAGGTTTTTGCACAAGAAGCAAATTCAGAGCTAAATGAAGTAGACGTAGAAGATTTAGGATCTGATGTGGCTGGAAATATATACGGA
ACAGGATTTTACGGGGCAGCGTCGTCTGGTATATTCACAGCGGAAGGTGAAGCCACTAGTGACTCGGCCTCCGAAAGTAGTTCGACTAGTTCGTCGTCATCCACT
TCAACCTCTGCTTCCTCCGCCTCTGAAGCGGCTGCAGCATCCGCTGCT

>putative_Aa_sgo_MaSp_1b Gene.13954::c18754_g2_i2::g.13954 ORF type:5prime_partial len:305 (+)
c18754_g2_i2:2-916(+)
GGTCAAGGCGGCTACGGACAAGCAGCTGGTTCTGGAGCAGGAGCAGCAGCTGGTGCTGGCGCAGGAGCTGGCGCAAGCGCTGGTGCTACTGGTGTAGGAAGACGA
GGGGGATACGGACAAGGAGCTGGTGCAGGAGCCGGTGCATATAATCAAATTGGTTATGGACAAGTTGTGGGTGCTGGCGCAGGCATTGGTGCAGGAGCAGCAGCA
GCGGGTGCAGGAGTAGCAGCAACAGGTGCAGGAGCAGCAGCAGCAGCAGGTGCAGGAGCAGCAGCAGGTGCAGGAGCTGATGCAGGAAGACGAGGTGGTTATGGG
CAAGGATTTGGTGTCGGAGCAGGTGCTGCAGCTGGCGTAGATTCATATGGTCAAGGCCAATACGGACAGGGTGTAGGCGCTGGAGCTGGTGCAACAACAGCAGCA
GGAGGTGCATATAGTCAAGGTCAATACGGTCAGGGTGCTGGAGCTGGAGCAGGCGCAGGAAGACAAGGTGGATATGGACTAGGATTTGGTTCTGGCGCAGGGGTA
GGCGCCGCATATGTTTCTGGTCAAAGTTCCGCTGGTGTTTCTTCATCGGCATCAACAAAAGATTCAGTAACCAATACGGTTGTGAACACCGTATCCAGACTTACA
TCCTCTTCTGCTTCTCAGCGAGTCTCATCAGCTGTTTCCCTCCTTGTATCTGGCGGTAGAGTCAATATGGACTCTTTGTCTAATGTAATCTCGAATCTTTCTTCA
GCCATCAGTGCCTCTAGTCCCGGTCTTTCTGGATGTGAGATTTTGGTTCAAGTTCTGATGGAAGTAGCCTCGGCTCTTGTGCACGTCCTCAGTTCATCTAATCTC
GGACCAGTCAACGTCAATGCTTCATCTAATTCTGCTAGCCTGGTTGGTCAATCAATTTCGGCAGTTTATGCTTGA

>putative_aa_sgo_MaSp_1b_Nterm Gene.13956::c18754_g2_i3::g.13956 ORF type:internal len:334 (+)
c18754_g2_i3:1-999(+)
AGGCGATACACCAGCAATATGGGTTTCTTTGCACAGTTTGCTATTCCCTTGGTCTTCTCTGTCCTTCTCGGATCAGGATGCTTGGGACAAAGCAACAATCCATGG
TCCAATTCGGCTTTGGCTGAGCAATTCATCTCCGGTTTCATTGGGCAAATGTCCTACCAGACATGCTTCACAGTCGACCAACTCGAGGACATGGCGTCGATCGGA
GAGAGCATCATGGCTTCCATCGACAAAATGGAGAAAACCGGCAGAAGTACTCCCAAAAAACTGCAGGCCTTGAATATGGCTTTCGCTTCTGGCATCGCTGAAATC
GCGGTTTCTGAAGATGGTGGAGCGAGTTTAGATGTCAAGACAAACGCAATAGCCGACAGCTTGTCGTCAGCTTTCATGCAGACAACCGGAGCCGTCAACCAACAA
TTCGTCAATGAAATCAGATCTCTGGTCGGCATGTTCGCTCAAGCCAACGCTAACGAGATAGCAGCAAGTAGTGCATCTGCAGCTGGAGCGGCATCTGCAGCTGGA
GCGGCATCTGCCTCTTCGGGAGCAGGTTATGGTCAGGCGTCCAGAGGAGGAGGTTATGGACAGGGAGTCGGAGCAAGTGCCGCTGGAGGAGCTGCTGCAGGTGGA
TATGCACAAAGTATGGGATCTGGAGCAACAAGACCAAGTACATACACACAGACTGTAAGTACCTCAACATCAGGAGCTAGTGCAGCAGCTGGAGCAGGAGATGGT
AGATACGGTCAAGATCAATACGGTCAACGTAGCTACGGACAAGGTATGGGTGCTGGTGCTGCAACTGCAGGAGCTGGTGGATATGGTCAAGGTCAATACGGACAA
GGAGCTGCTGCCGGAGCAGGAGCAGCAGCTGGTGCGGGCGCAGGAACAGCTGCATACGGTCAAGGTGGTTACCGACAAGGAGCTGGTTACGGAGCAGGTGCAACA
GCTGGTGCAAGCGCAGGAGCGGACGCAAGCGCGGGTGCTACTGGTGCAGGAAGA

>putative_aa_sgo_MaSp_2a_Nterm Gene.9444::c16509_g1_i1::g.9444 ORF type:internal len:184 (+)
c16509_g1_i1:3-551(+)
AGTATCCCAACAAGAGTAATGGGTTGGTCTACTCAACTTGCACTTCCTGTCGTTCTTCTGGTGAGCTCCCTTTTGGGGAGCACACTAGCAGCGAGTTCTATGTTT
ACCGACGCTTCGAAGACGGAAGACTTCATGCGATCAATGGTTTCAAGTCTGTCCAGGAACCCATGTTTCACGGCCGACCAGATTTCGGACATGGAATCCTTGTCT
GATACGATGACAATGTCTGTAGACAGAATGGTACAAACCAATAAAGTATCTGAACATAGAATGAAAGCTATGTCCACAGCAATGGCGTCAGGTATTGCCGAAATT
GCTGCAAGCTCTGGCAGCTGTGGTCTGCAACAAAAATCCAGTGCCATAGCCAACGCAATGTCAGATGCCTTCTTACAAACAAATGGAAGAGTTGATAGCATTTTC
GTACAAGAGATGGTTTCAATGATAAATATGTTTGCTCAAGCTTCCTCTAACGAAATATCCTACAGTACTGGATCTACTAGAGGATCCGCCGGAGCAGCAGCAGCA
GCAGCTGCCGCCGCCGCCGCAGCA

>putative_Aa_sgo_PiSp_x Gene.9936::c16886_g1_i2::g.9936 ORF type:5prime_partial len:191 (-)
c16886_g1_i2:68-640(-)
GTGGAAGCATCGCCATTTGCTGCCTTTGAAACTGCCCGTGTTGAAACAAATCCTATTATGGTGAATTCTGATTCTCCTACTGCAGTCTCAAGTACAACATCGCCA
CCAAATGATGATTCAAATGTGTTACTAGTGAAGGGACTTCCATCTTTAAGCTCAGTTATGATGGAAATATCACCAACTGCTGATTTGGAAAGCTTCCATCATGCA
CCAGCTTCAGGTTATGCCAATATGGAATCGCCTGACAGACATGCAAATTCAGTTGATAGATCTCCAGTAGCATTTGCAGTTCTGACATCAGGTGTTCTTACATCT
CCTTCTGCAACAAACAGGATTACAGAGATCAAATCTGTTATATCTTCTAGAGGTGTAAGCGTTGATTCGCTTGCTAATGCTATTTTCACAGCAGTTAATGGAGTT
CAAAAATATATGAGAATGACACAAATGGAAGTCTTGCCGGAAGCCTTGTTGGAGACTTTTGCGGCGTTTCTGCAATGTGTATCCCCAGATTTTACCGATTCTGCT
ACTTTGGTATTCAATTCCTTACAAAATGCCCTGACTTTATCAAAATAA

>putative_Aa_wa__PiSp_1 Gene.10713::c20156_g2_i1::g.10713 ORF type:5prime_partial len:139 (+)
c20156_g2_i1:2-418(+)
TCCACATATACGAGTTCGGAAAGTTCTGCTGCTGCCTCTGCCGCCGCAGCATCACAAGCCACTTCTAGTGTGTCTTCGGCTGCGACTCAATCTATGATTTCGTCT
ACTTCTGCTAGTGTGTCTTCCGCTGCGACTCAAGCTATATCTAGTGCTTCTGCAAAATTATTATCACCAATAGCGTCTCAAAGAATTTCAAAAATCAAATCCATG
TTTGCCTCTGATGTTAGTCCTTCTTCAGTTGCAAATGCTATTTTAGAGGTAATTAATTCTCTGAAGAGAGACTCAAGCCTGAAAAAATATGAGATTATTATTGAA
GCTATGCTAGAAACAATTGTAGCTCTTTTGGAATGCCTCTCTCCGTCATATGGTAATACTGCCAGTGAATTTTCTAGATCGTTGGTAAAAGCTTTCAACTAG

>putative_Aa_wa_AcSp_1 Gene.9278::c18865_g1_i1::g.9278 ORF type:5prime_partial len:327 (-)
c18865_g1_i1:1097-2077(-)
ACCTCTGCTTCCTCCGCCTCTGCAGCGGCTGCAGCATCCGCTGCTGGAGCAAGTTATGCAGCTTCCAGTGCCTTCGGAGCTAGCGGTTTCTCGACTCTTATCGGA
AACGTAGCATCAGGTCTGTTGGGCAGTGGTGCCTTCCAGTCTTTGTTTGGAGCAGGCATAGCCGTTAACGTTCGACTCAGTATTATCCAGTCGAGCCTTCAAGCC
GCATTCAGGACACTGAACGTCAGTTACGCCGTATCGTCAAGCATCTCGAGGGCTGTCAGCAGCGCCATTTCTTCCACTTCATCGTCGTCTGCCTGTGCAGAGGCT
ATTGCAAGTGTTGTCGTCAATGGTCTCGCAAATGCCGGAATCTTGAACACCAGCAACAGTTACGCATTCACTTCCAACTTCGTAAGTTCGTTCGTCCAGACCGCG
TCATCCGCCGCAGCCTCTGCCGCAGCTTCCTCTAGTGCCTCGTCCTCCCAAACCAGCTCGAGCACTTCTTCATCCACTTCGACCTCAGCTTCCTCCGCCTCTGGA
GCAGGAGCAGCATCTGCTGCGGGCAACGGTGCTGGGGCCTCAAGTGCTTTATCTTCGGGTCTATTGTATGATAGAGACTTGTTTGATGGTGTCGGAATCGTCGAT
AACGTACCCTATGGAAGTGGGTCTTTGCAGTGGGCATCGCCTGATGGTTTAACATCTCCCCAAGCTAGGGCTAGAATAGAGCAACTAACGTCTGAACTAGCTCAA

9. Appendix

TCGTTCGGATCCGGGAAGGCAGACCTCAGTGCTGTTTCACGTGCCATTGTGTCTTCGTTTTCGCAGGCGAAATCTTATGGCCTCTCCGAAGTTGACGCGAAGGTG
GAAGCTCTGTTCGAAACTATTACCGCATTGATTCAGATTTTATCTTCTTCTCAAGTCGGTCAGGTTGATACTTCACTGACGGAATCGATTTCAAGTTTTCTGGTT
GATTCATTGGAAGACAAATTTTGTCTTGGCATATGA

>putative_Aa_wa_AcSp_2 Gene.12143::c21247_g2_i1::g.12143 ORF type:5prime_partial len:195 (-) c21247_g2_i1:152-736(-)
TTGGTCAGTGGATTCTTGCAGGCTGCTGCGGGAGTTGCCGCTCAATTCGGAATCCGAATCTCCCAAAGTGAGGTTGCTGCGGACATTAGTAAGATAAGGAATACC
CTTTCAACTTCGACATCTACCTCCATCTCCGCTGGTGCTCAGTCTTCTTTCGATTTAGGTTTATCGGCTGGCATTGATGTAACAGACGTCGGCTTTGAAGCGGGA
TTGGACTTAGGTGTTAAGGGTTATGGAGCTGAGTCTAACGCCGGGGATATTGGTATTGGAGCTGTGTTGGTTTCTTCGACTGGTCTAAGATCACCTCAAGCTATT
TCAAGGATTCAAGGTATGACGTCAGATATTGTGAACGCTCTTCGGCCTGATGGCTTAGATTTCAATGTCCTTGCCCAGATGATGAAGAATATGAGCGGCCAGGCT
GTGAAATCTGGTTTGTCACCATCTGATGCCAAGATTGAAATTTTGTTGGAAACAATAACAGCCTTACTCCAAATATTATCATCAGCTAGAGTTGGTCCTATCAAC
AATAATAACATAAGCAATATTGTTCTTGGAATATCTAGTTCTCTTTCCAATGCTCTTTAA

>putative_Aa_wa_AcSp_Nterm Gene.4908::c13610_g1_i1::g.4908 ORF type:internal len:236 (+) c13610_g1_i1:1-705(+)
GTTCACCCACTCAAGCTTTCCAAGATGTACTGGTTAACAGGCTTCACAGTATTAGTGTTGCTACTACAGCAATACACCAGCGTTGAGGCAGATTCCAGGTATGAA
ACACTAAGTCCTTGGACTAACCCAACTAGAGCAGAAAAGTTCATCGATTGTTTAACGGATGGAATCGCAAAATCGCCAATTTTCCCTCAGCAAGAAAAAGATGAT
TTGCAGTATATTGTCGAGACAATGACATCAGCCATAACTAGCACAAGCTCTCAAAGGGGTAATAGAGCACAACTTCAAGCTTTAAATATGGCATTTGCATCTTCG
CTGGCAGAGCTTGTAATAGCAGAAGACATAGGAAACACGCAAGCAATCTCGGTCAAAACACAGGCTCTAACACAGTCTCTGTCGTCGTGCTTCAGATCCGTTATT
GGAACGGTTAACAAACGTTTCATCAACGAGATAAAAGATTTAATAAAGGTTTTTGCACAAGAAGCAAATTCAGAGCTAAATGAAGTAGACGTAGAAGATTTAGGA
TCTGATGTGGCTGGAAATATATACGGAACAGGATTTTACGGGGCAGCGTCGTCTGGTATATTCACAGCGGAAGGTGAAGCCACTAGTGACTCGGCCTCCGAAAGT
AGTTCGACTAGTTCGTCGTCATCCACTTCAACCTCTGCTTCCTACGCCTCTGAAGCGGCTGCAGCATCCGCTGCT

>putative_Aa_wa_AcSp/MaSp_Nterm Gene.1407::c5481_g1_i1::g.1407 ORF type:internal len:247 (+) c5481_g1_i1:2-739(+)
AAGCTTTTAGCTGACAAAATGACTTCCAACAGACATTTTCATCGGTTTTATGTGGCCTTTGCTTTCCTCATTGCATTATGCGGCTACACTGCTGGACTGGGAGCA
CCAGCAAATCCCTTGCTAGACCCGCACAGAGCGAACCAGTTCATCTCTAACTTCATGAGGAACCTGGACCAGTCGGGAATATTTCCCGCCCAGCAGGAAGAAGAT
ATGGCTGAAGTTATGAGTACGTTAACTCAAGCTATGGCCGGTAATAAACATTCAAATTCAGCACAAAATAAAGCTATGTCGTATGCCTTCGCATCTGTAATTGCT
GAATTGGTCGTGGCGAGTAAAGCTCCAGACATTTCAATGGACGAAAGAAGATGGGTCATCAAACACTGCCTCGAAAAGACGTTCAGAGACTGTGGAACAAAAGTC
AACACGGCATTTGTCCAAGAAATAGAGCATCTCATACAATTATTCGATGAAGAACAGAAAGAGAGCGGCTACGACAACGAAATACTCGTAAACCTTGATGCAGGT
GTGGGAGTAGAGGCAGATGTAGGAGTAGATGTAGGGGTAGCTGCGGACGTTGGAGTGGGGGTAGGCGTAAATGTTGATGCCGATGTCGGAGTAGGTGTCGATGCT
GATGTCGGTGTGGAAGTAGGTGTTGGTGTTGATGCAGACGTTGGCGTAGACGTAGATGCCGACGTAGATATCGATGCTGATGTCGATGTGGGAGTAGGTGTAGAT
GTC

>putative_Aa_wa_MaSp_1a Gene.10293::c19802_g1_i2::g.10293 ORF type:5prime_partial len:292 (+) c19802_g1_i2:2-877(+)
GGTGCCGGAGCTGGTGCCGGAGCAGGTGCAGCAGCTGGTGCAAGCGCAGGAGCTGAGGCAAGCGCTGGTGCTACTGGTGCAGGAAGACGAGGGGGATACGGACAA
GGAGCTGGTGCAGGAGCCGGTGCATATAATCAAATTGGTTATGGACAAGGTGTGGGTGCTGGCGCAGGCATTGGTGCAGGAGCAGCAGCAGGTGCAGGAGCA
GCAGCAGCAGGCGCAGGAGCAGCAGCAGGTGCAGGAGCTGATGCAGGAAGACGAGGTGGTTATGGGCAAGGATTTGGTGTCGGAGCAGGTGCTGCAGGTGGCGTA
GATTCATACGGTCAAGGCCAATACGGACAGGGTGTAGGTGCTGGAGCTGGTGCAACAACAGCAGCAGGAGGTGCATATAGTCAAGGTCAATACGGTCAGGGTGCT
GGAGCTGGAGCAGGCGCAGGAAGACAAGGTGGATATGGACTAGGATTTGGTTCTGGTGCAGGGGTAGGCGCCGCATATGTTTCTGGTCAAAGTTCCGCTGGTGTT
TCTTCATCGGCATCAACAAAAGATTCAGTAACCAATACGGTTGTGAACACCGTATCCAGACTTACATCCTCTTCTGCTTCTCAGCGAGTCTCATCAGCTGTTTCC
CTCCTTGTATCTGGCGGTAGAGTCAATATGGACTCTTTGTCTAATGTAATCTCGAATCTTTCTTCAGCCATCAGTGCCTCTAGTCCCGGTCTTTCTGGATGTGAG
ATTTTGGTTCAAGTTCTGATGGAAGTAGCCTCGGCTCTTGTGCACGTCCTCAGTTCATCTAATCTCGGACCAGTCAACGTCAATGCTTCATCTAATTCTGCTAGC
CTGGTTGGTCAATCAATTTCGGCAGTTTATGCTTGA

>putative_Aa_wa_MaSp_1b Gene.10296::c19802_g1_i3::g.10296 ORF type:5prime_partial len:312 (+) c19802_g1_i3:1-936(+)
GGAGCAGGTGCAGCAGTTGGTGCAAGCGCAGGAACAGGTGCATACGGTCAAGGCGGCTACGGGCAAGGAGCTGGTTCTGGAGCAGGTGCAGCAGCTGGTGCAAGC
GCAGGAGCTGAGGCAAGCGCTGGTGCTACTGGTGCAGGAAGACGAGGGGGATACGGACAAGGAGCTGGTGCAGGAGCCGGTGCATATAATCAAATTGGTTATGGA
CAAGGTGTGGGTGCTGGCGCAGGCATTGGTGCAGGAGCAGCAGCAGCAGGTGCAGGAGCAGCAGCAGCAGGCGCAGGAGCAGCAGCAGGTGCAGGAGCTGATGCA
GGAAGACGAGGTGGTTATGGGCAAGGATTTGGTGTCGGAGCAGGTGCTGCAGGTGGCGTAGATTCATACGGTCAAGGCCAATACGGACAGGGTGTAGGTGCTGGA
GCTGGTGCAACAACAGCAGCAGGAGGTGCATATAGTCAAGGTCAATACGGTCAGGGTGCTGGAGCTGGAGCAGGCGCAGGAAGACAAGGTGGATATGGACTAGGA
TTTGGTTCTGGTGCAGGGGTAGGCGCCGCATATGTTTCTGGTCAAAGTTCCGCTGGTGTTTCTTCATCGGCATCAACAAAAGATTCAGTAACCAATACGGTTGTG
AACACCGTATCCAGACTTACATCCTCTTCTGCTTCTCAGCGAGTCTCATCAGCTGTTTCCCTCCTTGTATCTGGCGGTAGAGTCAATATGGACTCTTTGTCTAAT
GTAATCTCGAATCTTTCTTCAGCCATCAGTGCCTCTAGTCCCGGTCTTTCTGGATGTGAGATTTTGGTTCAAGTTCTGATGGAAGTAGCCTCGGCTCTTGTGCAC
GTCCTCAGTTCATCTAATCTCGGACCAGTCAACGTCAATGCTTCATCTAATTCTGCTAGCCTGGTTGGTCAATCAATTTCGGCAGTTTATGCTTGA

>putative_Aa_wa_MaSp_2_Nterm Gene.17742::c33569_g1_i1::g.17742 ORF type:internal len:183 (-) c33569_g1_i1:1-546(-)
ATCCCAACAAGAGTAATGGGTTGGTCTACTCAACTTGCACTTCCTGTCGTTCTTCTGGTGAGCTCCCTTTTGGGGAGCACACTAGCAGCGAGTTCTATGTTTACC
GACGCTTCGAAGACGGAAGACTTCATGCGATCAATGGTTTCAAGTCTGTCCAGGAACCCATGTTTCACGGCCGACCAGATTTCGGACATGGAATCCTTGTCTGAT
ACGATGACAATGTCTGTAGACAGAATGGTACAAACCAATAAAGTATCTGAACATAGAATGAAAGCTATGTCCACAGCAATGGCGTCAGGTATTGCCGAAATTGCT
GCAAGCTCTGGCAGCTGTGGTCTGCAACAAAAATCCAGTGCCATAGCCAACGCAATGTCAGATGCCTTCTTACAAACAAATGGAAGAGTTGATAGCATTTTCGTA
CAAGAGATGGTTTCAATGATAAATATGTTTGCTCAAGCTTCCTCTAACGAAATATCCTACAGTACTGGATCTACTAGAGGATCCGCCGGAGCAGCAGCAGCAGCA
GCTGCCGCCGCCGCCGCAGCA

>putative_Aa_wa_MaSp/AcSp_Nterm Gene.10871::c20282_g1_i1::g.10871 ORF type:internal len:277 (-) c20282_g1_i1:3-830(-)
TTCTTTGCACAGTTTGCTATTCCCTTGGTCTTCTCTGTCCTTCTCGGATCAGGATGCTTGGGACAAAGCAACAATCCATGGTCCAATTCGGCTTTGGCTGAGCAA
TTCATCTCCGGTTTCATTGGGCAAATGTCCTACCAGACATGCTTCACAGTCGACCAACTCGAGGACATGGCGTCGATCGGAGAGAGCATCATGGCTTCCATCGAC
AAAATGGAGAAAACCGGCAGAAGCACTCCCAAAAAACTGCAGGCCTTGAACATGGCTTTCGCTTCTGGCATCGCTGAAATCGCCGTTTCTGAAGATGGTGGAGCG
AGTTTAGATGTCAAGACAAACGCAATAGCCGACAGCTTGTCGTCAGCTTTCATGCAGACAACCGGAGCCGTCAACCAACAATTCGTCAATGAAATCAGATCTCTG
GTCGGCATGTTCGCTCAAGCCAACGCTAACGAGATAGCAGCAAGTAGTGCATCTGCAGCTGGAGCAGCATCTGCCTCTTCGGGAGCAGGTTATGGTCAGGCGTCC
AGAGGAGGAGGTTATGGACAGGGAGTCGGAGCAAGTGCCGCTGGAGGAGCTGCTGCAGGTGGATATGCACAAAGTATGGGATCTGGAGCAACAAGACCAAGTACA
TACACACAGACTGTAAGTACCTCAACATCAGGAGCTAGTGCAGCAGCTGGAGCAGGAGATGGTAGATACGGTCAAGATCAATACGGTCAACGTAGCTACGGACAA
GGTATGGGTGCTGGAGCTGCAACTGCAGGAGCTGGTGGATATGGTCAAGGTCAATACGGACAAGGAGCTGCTGCCGGAGCAGGAGCAGCAGCT

>putative_Aa_wa_PiSp_2b Gene.13024::c21821_g1_i2::g.13024 ORF type:complete len:602 (-) c21821_g1_i2:263-2068(-)
ATGTCAGTGTTACGGGTCCTTCTTTGCTCTGCTGCAATAACCTCTTTTCTGCAGGTATCTAGTGGAAAAGCAAGGTTGCCAAAGACGTTCAAAGACCAAAGATCC
TTTAATAGTTTTGTAGAGAAACTCGTCTATTATTTCAAGGATTTTAACGTATTGCATGCTGGAGATGAGGGTGAATTCCTAGAAGCATTGAATTACATATTTACG
ACTAACGAAGATAAGGAGACGACGGTATTAAAGCACGCTTTAGCAAGCGAAATTGCTGACGTTATTGGGAAGGCAATTGAAGCAGGTGTGAACTGGAAATATTGC
CTTGCTAATGCAGTCCAAACCACTGCTTTTGCTTTCGAGGATGTCTATGAAACGGTCGATGACTCGTTTCTGGAACTGATAATTCGCATGGTCACGGATAAATAT
GACGACGACTGCAAAATTCATGAATGTGAAGAAGTTTCAAACGATGGATCATTTATATCGACTTTTCTAGATACTGGTGTAGATTCAGAATCTGCAGTATCTCTT
ACGACTACAGATATCAAACAAAAATCTTTCGAAAAACTACAGATTATTCCTGAATCTATATCTAATTCTGAATCTTCAACAATGGTGGACGTCAGCTTGAACGAT
CTTGGCATAAAAACTATTGAACATAAATCCGAATCATTGGAGCAAGATGTGACTCCACGTACCACAATAGCGCCAGAAACAGTGCTGCAGCCCTCGGTTATTGCT
TCTAAGAACAACAGTTCAACAACTGAATTTTCAGAACCTTTAAGCGCACCTGTCACTATTCCGGCCTCTGAAGCATTTTTAACCACGGGCACAGAAGTCGAAAAA
ACTGTTTCTGTTATTTATAGAACGATACAGAAGAACTTAAAGAACAATTTTATGACTTCTAATATATTTTGTGATGAAGCGGGGACAGTCTTTGCAGCTAATAAC
CTCAAAAGAGCAATCGTTAGTACACTGGCTTCTTATGGAGTTCCTGAGTCGACATCTCAAAATATTCTATCCAACTATATTTTTGATTTTACTGTCGAATCTTTT
CCAGCTGTGACGGATGCAGTTTCGCTGGCCGTATCGAATTGCCTTGCTGTGAGTAATGTAGTTAATGCTGGAGACGAATCAGCCGTTTCAAAGAGTATATCATCC
GCTATTATCGAATCACTGAAGACCTTGAAATCCCAAAGCGAGCTGGATTTAAGTTTACACCAATCTCCTTCTGTTACTTCTATCTCACCACAGCAGGATACACGA
ACAACTCTGTCCGAAACAGATCCTGGTGTCATGACAAATTCTGATTCTCCTACTACAGTCTCAAGTACAACTTCGCCACCAAATGATGGTTCAAATGTGTTACTA
GTGAAGAGACTTCCATCTTTAAGCTCAGTTACGATGGAAATATCACCAACTGCTGATTTGGAAAGCTTCCATCATGCACCAGCTTCAGGTTATGCCAATATGGAA
TCGCCTGACAGACATGCAAATTCAGTTGATAGATCTCCAGTAGCATTTGCAGTTCTGACATCAGGTGTTCTTACATCTCCTTCTGCAACAAACAGGATTACAGAG
ATCAAATCTGTTATATCTTCTAGAGGTGTAAACATTGATTCGCTTGCTAATGCTATTTTCACAGCAGTTAATGGAGTTCAAAAAGATAAGAGAATGACACAAATG
GAAGTCTTGCCGGAAGCCTTGTTGGAGACTTTTGCGGCGTTTCTGCAATGTGTATCCCCAGATTTTACCGATTCTGCTACTTTGGTATTCAATTCCTTACAAAAT
GCCCTGACTTTATCAAAATAA

>putative_Aa_wa_TuSp_1 Gene.12686::c21586_g1_i1::g.12686 ORF type:5prime_partial len:311 (-) c21586_g1_i1:120-1052(-)

ACATCCGGATCCCAGGCAGCCAGTCAAGCAGCTGCACGTGCAGCAGCACAGGCGGCCAGCCAGGCGTCTTCCAGCAGTTACGCATCAGCCAGTGCCTCTGCCTTC
TCCCAGGCCTCTGCTGCCAGCCTCGCATCCTCCAGCGCCTTCTCCAGTGCCTTCGCTTCTGCTAGTTCAGCAGCAGCAGCCGGAAGCATCGCGTACAACCTTGCT
CTCCAGACAGCCAACACTCTCGGAATCTCAAACGCAGTCGGCATTGCAAGTGCCGTGGCTCAGGCCGTCAGCGCAGTTGGTGCAGGAGCTAGCTCATTCGCCTAC
GCATCTGCCGTCTCTAACGCCGCAGCACAATTCTTAGCAGCACAGGGAATACTGTCTCAGGCTAACGCTTCGGCTTTGGCCTCGTCTTTCGCAAGTGCATTCGCC
TCGGCAGCAGCATCGGCCTCTGCTTCCGCAGCAAGCTCAGCTTCTGCACAGTCCGCCGCAGCCGCTAGGAGTCAAGCCGCAGCCTCTGCCTTCTCCCAGGCAGCC
TCACAAGCCTCTTCACGATCTGCAAGTCAGGCGGCAGGCCAAGCCGGTGCCGCTGCTGGTGCTACCGCTGTTTCTAGTCTTCTAAGACCAGCTGCTGTTAATTTA
CCCAGCTTAGTAAGGAGTCTCTTTCCAGGTTTCTCTTCAAGTTCTGGGTTGTCATCAGCTGCAGCCTCTGCACGTGTTGATAGCTTGGCCTCGTCTTTAGGGACA
CTACTATCCTCTGGCCAGCAACTTAACGTTCAGAACTTTGCATACACTCTTTTGTCCCAAGCTTCTCAAGGTCAAAATAGCAATGTAGGAATCAGCTCGAATGAG
GTTATTGTTCAGGCACTGTTGGAAGGTCTTTCTGCGCTTATTCAGGTGATAAACAGCTCCCAGATTTCTTCAGTTAGCTCCCGGAGTGGCTAG

>putative_Ca_MaSp_6 Gene.21857::c58260_g1_i1::g.21857 ORF type:internal len:188 (-) c58260_g1_i1:3-563(-
)
ATCCCAACAAGCGAAATGGGTTGGTCTACCGGACTTGTGCTTCCCGTTGTTCTTCTCCTGGCGAGCTCTCTTTTTCAGAGTGCCCTAGGAGCTAGCAGCCAGGTA
CTACAATCGCCTGAAATGGCGGATGCTTTCATGCAAACCGCGATTCAAAGTTTATCTCAGAATGAAAATTTCACACCGGATCAGATATCTGATATGAATTCGATG
GCTGACACCTTGATAGCTTCTGTTGAGAGCATGGCAAAATCAGGACAAACTTCAACACATAGAATGAATGCGATCCGTATGGCAATGGCTTCGTCGATCGCTGAA
ATTGCGGCAGACGACGGAGGAGGCCCGTTGGCCGTAAAAGCAAGTGCAATAGCGAATGCATTGGCTGACGGTTTTTTGCAAACAACTGGAACAGTGGATGGAGTT
TTTATAGGAGAAATAATGTCAATGATCAAAATGTTTGCTCAAAGTTCTGCAAATGCTATTTCTTACAGTTCCTCAACATCAGGTGGAAGAGGATATGGTAGTGGA
GGTTCCGCAGCCGCAGCTGCTGCAGCAGCAGCAGCA

>putative_Pp_sgo_AcSp_MaSp_1_Nterm type:internal len:430 c19314_g1_i1:1289-3(-)
TTACAGTCTTCTTCACGGTCTACCCGACGTGCAGTCATGGACCGTTGGGGACATTTTACGTTTCTGCTAGTGATCGCAACCACTCTTTTTGCAATCAGTGATGCA
CGAGGTAAATCAGCAACATCACCTTGGTCGGACCCGGCAAAGGCCAAAGTATTCATAACCAAATTCATCGATGTGATCGGACAATCCAGTGCTTTCCCCGACCAG
CAAAAGGAAGATATGGAGTCCATTGCCGAAACAATGATGTCTGCTGTTGATGGGATGGCGAGAAACAAAAGAGGATATGCAGCAAAGATGCAAGCCATGAACATG
GCATTTGCTTCCTCGATGGCTGATCTTATTGTTTTGGAAAATGACATGAGTGGTAGTGTTAACGAAAAGGAAGAAGCACTTCTAAGTGCTCTCGAAGCAAGCATG
TACGCAGCAACTGGAACAGTTAATAAGGCATATATATCAGAAATTCATAATCTAATACAGATGTTTTCGCAAGCTAATATGAATGAAGTTGAGCCGGAGCCATCA
CCTCCCGCTATTCCTAAACCACGAAGAATACCGCAATCACCTATTCAATCCCAACCACCACCACAAGTTCCCCCGGCAATAGGACTAACATTTCCGCAGCAACCA
ATGAGATCTGGTGACATAAAATCTCCTCCGATGGCTAAAACGCTTGTGAGTTCAACAAGTGCATCCACATCTTCTGCGTCTTCCTCTTCGTCTGTTTCTACATCA
TCTACTACTTCCACTTACACACAAGCCCAGCAAGCGGTACAACTTCCAATCCAGCCAGCGATACAACTTCCAACCCAGCCAGCGATACAACTTCCAACTCAGCAA
AACCTGGCAACGCAGTACCAGTTTGCTTCGCAAGACAGGTCGGAAAAGTTTTCGTCACAGCTCACGAACATGCTCATTTCAAAAACTGAATTCGGAACATTTGTA
TCAACATTTCCCATATCATCCTTGATTAATGTGATACAAGGTGCATTGACAGAAATTTATGGATCTCAAGTCGCGATAAGCTGTGCAAATGCAGTAAATGTATAC
ATGTCAAACTTACAACAAACAACAAACGTTTACGCTACCGCACAAAGCATATCAAAAGCTTTAAGCACGACACTATCATCACTAGGCAAAATAGATATGGCGGAT
GCGCAGAGAATCGGAAATGAAATAGTAATGGCTATTCTACGAATAACAAATCAACAACAAACACAACGCAGACAGCCAGCAATAGCACCGATAGCCACCGTACCT
GCTATTAGCAGTCAAACAACTATGTCG

>putative_Pp_sgo_AcSp_MaSp_2_Nterm type:3prime_partial len:209 c25233_g1_i1:624-1(-)
ATGGAGCTGTTCTTACCGTTCTTACTCCTGCTATCCACAGCCCAACTACAAACGGTATGTGGAAGCCACAGTTCACCGGGCTCCCCATGGACTGATAAGCAGACT
GCAATACTATTCATGAGGAATTTCGTTCAAGAGGTCGAGAATAGCGAAGCTATGAGTGATGAAGAAAAGGCGGACGTTGGCAATGTGGCATCTCTGATGGAGACG
TCGGTGATGAGCAGCCGTTCCAAAACAGCCTTGCAAGCCATGAACGCGGGGTTTTGACTGCCGTGACCGAGGTATTGGTACTGGATAATGACAACGGGATCAAC
CAAATACCGATTGTCATCAATGCCTACCAGAATGCGCTGATAAAGACGGTTGGTTACGCTGACAACAACTACATGAGAGAAACCATGAGTATCGTGCAATCAGTC
GCTCAGACCAGCGGTGACAATGACATACCGGGATCGAACCAGCCGTCGGGAAATCAAGGACCTTATGGACCTGGACAAGGTGGGCCCGGGGGACCAAATTCTCAA
GCTTACGCTTTTGCTCAAAGCGGAACTCAAGGCCCTAATGGTCCAAGGGGACCATCAGGTCCGTATGGGCCAGGAGGACCACAAGGACCAGGTGGACCACAAGGA

>putative_Pp_sgo_AcSp_MaSp_3_Nterm type:3prime_partial len:205 c25233_g1_i2:614-3(-)
ATGGAGCTGTTCTTACCGTTCTTACTCCTGCTATCCACAGCCCAACTACAAACGGTATGTGGAAGCCACAGTTCACCGGGCTCCCCATGGACTGATAAGCAGACT
GCAATACTATTCATGAGGAATTTCGTTCAAGAGGTCGAGAATAGCGAAGCTATGAGTGATGAAGAAAAGGCGGACGTTGGCAATGTGGCATCTCTGATGGAGACG
TCGGTGATGAGCAGCCGTTCCAAAACAGCCTTGCAAGCCATGAACGCGGGGTTTTTGACTGCCGTGACCGAGGTATTGGTACTGGATAATGACAACGGGATCAAC
CAAATACCGATTGTCATCAATGCCTACCAGAATGCGCTGATAAAGACGGTTGGTTACGCTGACAACAACTACATGAGAGAAACCATGAGTATCGTGCAATCAGTC
GCTCAGACCAGCGGTGACAATGACATACCGGGATCGAACCAGCCGTCGGGAAATCAAGGACCTTATGGACCTGGACAAGGTGGGCCCGGGGGACCAAATTCTCAA
GCTTACGCTTTTGCTCAAAGCGGAACTCAACGCCCTAATGGTCCAAGGGGGCCAGGAGGACCACAAGGACCAGGTGGACCACAAGGA

>putative_Pp_sgo_MaSp_Nterm type:internal len:757 c39901_g1_i1:2270-3(-)
AAGATCAGTCAGACCTTAAGCGCAAACGAAGGTCCCACCACGATGCGTTGGTTTGCAGTGTCGCTGGTCGTCACGACCTTCCTGGTGCTCTACCAATCGGCAGAA
TGTAAGAAAAAGGAGAAAAACTACAAAAACAACCCTTGGGCAAGCAAAGAGAAAGCTTCGATTTTCGCTAATACTTTCATAAACTCGGCCAAGTCTTCAGATGCC
TTCAGCCAGGATCAAAAAGACAGTATCGAAGCCATCGGGAATACCATAACGTACGCTATGGGGAAAGTGGACACGTCTAAGAAGAACGCGAAAGCTACGTACCAA
AGTCTTGACATGGCTTTAGCTTCCGCAATCGCTGAGATCATAGCGAACGGGGACAATATTAAAGTACAGACAACAGCAGCGATGGAAGCTATGGCGGAAGCATTC
AGAAGAACAACTGGTAAAGTGAACAAGGTATACATAGCTGACATAAAAGAACTGATCGAAATGTTTGCTAATAGCAAGCTTGAAAGTGGAGAAGATGACTATGAT
TACGATGACGACGACGATGATGATTCATTGAATTCGATCCTTGGCTCCAGTTCACCTTCATCATCTGTCATGTTACCACCACCACCACCTCTTCCGCAGTTCTTC
CAACCTCCACCAATGCCAGTGAGCCCAGCTCCACCTAGCACAATTATAACCTCTTCCAGCTCGTCTTCGTCTAGCAGTAGTTCGTCAAGCTCCTTATCTTCCAGT
TCTCAGACGACTTTTTCCCAACAATCACCATCGGCACCTCCACCGTCACCACCGCCTCCCCCTCCTGCACCAGTACCAGTGTTCAGCCCGCCACCACCTCCTCCT
CCCCCTCCTGCACCAGTACCGGTGTTCAGCCCACCACCACCAGCTCCAGCACCCAGTCCCTCTCCACCTAGCCCGGTGACATCAACGACATCGTCTTCTACAAGT
TCAACATCAGCTAGCAGTAGTTCTTCAACCTCGTCTTCTTCCAGTTCTCAAACATCCTTCCAACAAGCACCGGCTCCACCAGCTTTTAGTCCACCCCCTCCTCCT
GTACAGTTACCAGGTTTCAGTTCTGAAGACATCACACGATTTGCATCAGCAGTATCAACGGCTTTCGTGTCATCAGATCTATTTTCTTCCCTCTTTAATACGGAC
ACTACTCTCACAGATGCTCAGATTTACACAGAAGACATGACCCAAACAGCAGGGTACTTTCGTGGATTCTATGGTTCTTCACTTTTGAATATGCAGCAGGCGGTG
AATCAGGCCGTCATCGAAGTTGGCAGCGGAGCAGGCACATACGCCTATGCTAAAAGCATCGCTTACGCTTTGGCTACATATATGACCAAGTCGCAGATATCACTT
GCATGGTTTGAGCCGAACTCGTTTGCGAGTAGCGTGCTTTCTGCACTGGAAATATACAGGCCCGTGCAGGTCGACCAAACCTCTCAGGGTTTTGATATTTCAGCC
GAAAGTTTTCTTTCTAATGAAATTTCAGAAACTCCGGCTACGGATACAGCCTTCGCCGAAGCACTTTCTTCTGTTTTCTTGGCTCATGATTTGTTTAGCACATTA
TTCTATACAGGGACAAATATTAATGACGCTCAAACAATGGCATCTGCTATGGCCAAAGTTGCCGCTATTTTAGTGAATTCCGATCAGACTACAACCTCAGGAATC
ACCGACATAGTACAGCAATCCGTAGCGGAAGTGGGAACCGGTGGAGCTACATACGATTATGCTAATGCTATAGCTTACTCCGTTGTCTCGTATTTGAACACGCAA
CAGGTATCTTTTGATAATTATGAACCGAACGCATTCGGAGCAACCATGCTCTCAGCGCTGTCTTTATTCAAACCTCAAGAAACACAGCCTGAAGCTACAGTTCAA
ATGTCATCATCGAGCATGTCCTCCTCCTCCTCCTCTTCTTCTTCCTCTTCAAGTTCAAGTATTACCACTACTCAACAAATTAGTGAGGCACCTCAAGCCCCTGCT
GCTCCTGATTATAACAAAATCAACCAAGACTTCACGAATACGATGACTTCAGTTTTCATATCTTCAGAGCTGTTCTCAACTCTTTTCTCTGGCACAACCGATGTC
ACCAACGCTCAAAAGTACGCATCGGCAATGGCCCAGTCAGCTGGAATGTTATTAGGATCTGACCAATCAACTGTAACGGATTTAACAAATAATGTGGCGCAAGGT
ACGACAGGCATCGGCAGTAATGCTGGAACATATGCCTATGCCAGTTCCATTGCGTCCACTACG

>putative_Pp_sgo_PiSp type:5prime_partial len:289 c27315_g1_i1:986-120(-)
TCGCAAGCCACTTCTGTCAGTTATCAAAGCGTATACAGCGCTCTGACTTCAGCGCTACAGCTTGACCGATTTGTCGGGTATTTGAATAGCGGAGACCTGTCAGTT
CAAGACTTACTGGAAGTGACTGCATCTGCGTGCGCAAGAGCGGGAGTTCTGATTTCCAGAGATATTCTAACTACGTTTCGTCAGTCTCATCTAGCGCCTCGGCA
TCAGCTGTGGCAAGCGCAGTAGCTCGCTCGTCTGTTGCTGCTCTAAGAGGAAGCGGTCTCGTACGTGCCGGACAGGAAGCAGACTATGGAAGACGATTAGCCGGT
TACCTGAACGCGGGCTTCAACGGACTCGTTTCAAGAGCATACTCTACCGCACAAGCAGCAGCGTATGCTGAAAGCAGCAGTTCCGCGGCCGCCTCATCTGCAGCA
GCTGCCACATCGGCAGCTTCAAGCACTACATCCACATATACGAGTTCGGAAAGTTCTGCTGCTGCCTCTGCCGCCGCAGCATCACAAGCCACTTCTAGTGTGTCT
TCGGCTGCGACTCAATCTATGATTTCGTCTACTTCTGCTAGTGTGTCTTCCGCTGCGACTCAAGCTATATCTAGTGCTTCTGCAAAATTATTATCACCAATAGCG
TCTCAAAGAATTTCAAAAATCAAATCCATGTTTGCCTCTGATGTTAGTCCTTCTTCAGTTGCAAATGCTATTTTAGAGGTAATTAATTCTCTGAAGAGAGACTCA
AGCCTGAAAAAATATGAGATTATTATTGAAGCTATGCTAGAAACAATTGTAGCTCTTTTGGAATGCCTCTCTCCGTCATATGGTAATACTGCCAGTGAATTTTCT
AGATCGTTGGTAAAAGCTTTCAACTAG

>putative_Pp_sgo_Spidroin_1b type:5prime_partial len:603 c28051_g1_i3:3-1811(+)
GGATCGGACCAGTCTATCACTTCCAGTCTCACCAGTACTGTAACGCAAGGTGTAACGGGTGTCGGAAATAATGCTGGAACCTATGCCTATGCAAGTTCCATTGCG
TCCACTACGGTATCATCCCTTTCATCGTTGCAGATTCCTCTTTCAAATTACGATGCCAGCGCATTTGCTGCGACGATGCTCTCAGCGCTTTCACGATTCAAGCCT
CAAGAAGCCCAGCCTGTCGCTCCAGTGCAAACAACGTCATCGACCTCGGTTGCGTCTTCTTCCTCCTCTTCTGCTTCTTCGTCGGCTACCACTTCGTGACAACC
TCAGTGGCACCTCAAGCTCCCGCAGCACCTGATTATAGCAAAATGAACCAAGAGTTCACAAATACCTTAACCTCTATTTTCGTTTCCTCACAGCTGTTCCAAACG
CTGTTCTCCGGAACAACTGATGTCAACAATGCTCAAACTTACGCATCAGCAATGGCCCAATCTGCTGGAACTCTATTGGGATCTGACCAGTCAACTACCAGTAGT
TTTACCAATGCGGTATCGCAAGGTGTAGGCAGTAATGCTGGAACTTACTCTTACGCCAGTACTATTGCATCTACTACGACTTCATTCTTGTCAAAC
ATTCAGGTAATTTCCAATTACGATGCTAACACATTTGCAGCGACAATGCTCTCCGCGCTTTCCAGATTCAAGCCTCAAGAAGCTCAGCCAACAGCACCAGTTCAA
GTGTCTTCGACCAGTACGTCTTCTTCTTCCTCCGCTAGCGTGTCTTCTTCTTCCTCGCTAGCGTGTCTTCTTCTTCCTCGCTAGCGTGTCTTCTTCTGTTGCC
ACAACGCAAGTCAGTGTAGCTCCTCAGCCACCACAAGCCCCGGCTGCCCCTGACTACTCTGAATTCACTAACTCTCTGAGCTCTTCTCTTGTGGCTTCAGAACTG
TTCACAACCCTATTCTCAAGCATCACTGATGTTTCCACAGCTCAAAGTTACTCATCGTCGATGGCTCAAGGTGCAGGGATGTTTGTAGGTTTAGATTCTACCACA
TCCGGGTTGACAAACTCTGTCGTGCAGAGCATCACTGCTGTTGGTGCTGGCAGTGACGTATCAACTTATGCCACATCAATTGCTTCCTCAATTGTAAGCTCTATT
TCCAGTTCTCAGACAATTCTTACTAACTATGATGTTACTTCCTTCAGTGCAGCAATGCTTTCGGCTCTCGAGAAATTCAGACCACAAGCTCCTAGTGCTCCTGCT
GTGTCAATATCATCCAGTATTGCTTCTTCCTCCTCTTCAACTACATCATCTAGCAGTACTTCCTCCTCCTCCTCTGCCTCATCCAGTAGCGCTTCCTCTACTACT
TCAACTGGTGATGCTTCCCCTTCGGCTCCTGCGTCACCTGGTGCAGATGTGGCAACTGTTCCTTCTGGTGGTTGTGCAAATACTGGGCCAAGCTCAGGCAGCGAT
GCTTACCCTCCTATTTCTATGGACACTGCTTTAGCCTTTTTCACTCCCTCGGTTCCAATGTCGACTAGTACTTCCCTGCGAATGTGTAACTTTGCAACAACGATC
CTGTCTTCGATAGCATCGGAGTCTGGGTTCAACTATATGGACTTTTCACGTGCTTTGAATGTCCTGAGCTCTTCAATACGGTCCAGTGGCTCTGGTTCTTCCTCG

9. Appendix

ACAAGTCTTATCATTCAGTTCCTCTTCGAAATCGTTGTATCACTGGTCCAAGTACTGCTCCGTATGAATGCATCTACTGATGATACAGCGAATGCTCTTTACACT
GCTATATCGCAGTCATTTGCTTGA

>putative_Pp_sgo_Spidroin_4a type:5prime_partial len:427 c29082_g1_i1:1410-130(-)
GGACCGCAAGGACCTGGACAAGGCGGACCAAGCAGTGCAACAGCTATGGCGATTGCTTCGTCCGGACCGGGAAGTGGCCAAGGCGGTCCTTCTGGTGGAGATGGT
GCGGCGGTTTCAAGTTTCGCCTCCGCTTTTGCTTCTGCTTTGGTGAGCAGCAATGCATACAACCAGAGGATTGGTGGTGTAGGGCCATCAGGAGCAGGATCAGCG
ATGGACATGTCAATGGGTCAAGCTTATGGACAGTCTATTGCGCAGGCAATTGCACAAGCCACAGCGCAGGCATATTCTAACTCTGGAGGAAATAGTCAAGCCACA
GCCTATGCTATCTCAAACGTTGTGGCACAGTATCTGAACAATTATGGAGTCATACAATCTGCCGACGCCACATCACTCGCTCTAGCGTTCGCATCAGCTCTGTCT
TCTGCAGGGGGTCAAGGACCATATGGACCGGGATCTGGACAAGGAGGTCCACAAGGAGGTTCCAGTTCGATGGCTTTTGCTACAGCTTCCAGTGGACCCAATGGC
GCTGGAGGGCCGCAAGGACCAGGAGGCCCACAAGGAGGTGGCCCATTCGGACCTGGAGGCCCACAAGGAAATGGCCCATTCGGCCCAGGAGGACCACAGAGTGGT
GGCCCTTACGGACCAGGTGGACCACAAGGACCGCAGGGTGGTGGCCCTTACGGACCAGGGGGATCACAAGCTTTCGCTCAAGCCTCCAGTGGGCCTGGACAAGGT
GGACCACGAGGACCAGGCAATCAAGGACCTGGACGAGGCCGGACCACAAGGACCTGGACAAGGTGGGCCACAAGGACCTGGACCTTACGGACCAGGAGGACCCGGC
TCTCAAGCTTTCGCTTTTGCTTCGGGTGTACCTGGTGGTCCACAAGGAGGACCACAAGGGCCAGGTAACCAGGGACCCGGAGGGCCTCGAGGACCTGGACAAGGG
GGACCACAAGGTTCGTCTTCAACAGCCTATGCGTTTTCTTCTTCATCTGGCTCTCAAGAAAGGATGAATTCTTTGGCGTCTGCTCTTTCATCGATGGGATCATCT
GCAGATTACCCAGCCCTGAATAAACTTCTGTCTGGACTCATTTCTAGCATCAGGCAGAGCAATCCTGGTATTTCGCCACAGAATGCGATGACTCAAGCTCTGCTA
GAAGCGCTTGCAGTTATGTTCGAACTGCTCGTTGCAGTGCTGTCAGGAGGTGGAGGCGTCAATTCAGTCTCCTCTTACGGCGTAAGTGCTCTCGTTCAAGCAATG
ATTCAAGCACAGTCTTATTAA

>putative_Pp_sgo_Spidroin_4c type:5prime_partial len:416 c29082_g1_i11:1377-130(-)
GGACCGCAAGGACCTGGACAAGGCGGACCAAGCAGTGCAACAGCTATGGCGATTGCTTCGTCCGGACCGGGAAGTGGCCAAGGCGGTCCTTCTGGTGGAGATGGT
GCGGCGGTTTCAAGTTTCGCCTCCGCTTTTGCTTCTGCTTTGGTGAGCAGCAATGCATACAACCAGAGGATTGGTGGTGTAGGGCCATCAGGAGCAGGATCAGCG
ATGGACATGTCAATGGGTCAAGCTTATGGACAGTCTATTGCGCAGGCAATTGCACAAGCCACAGCGCAGGCATATTCTAACTCTGGAGGAAATAGTCAAGCCACA
GCCTATGCTATCTCAAACGTTGTGGCACAGTATCTGAACAATTATGGAGTCATACAATCTGCCGACGCCACATCACTCGCTCTAGCGTTCGCATCAGCTCTGTCT
TCTGCAGGGGGTCAAGGACCATATGGACCGGGATCTGGACAAGGAGGTCCACAAGGAGGTTCCAGTTCGATGGCTTTTGCTACAGCTTCCAGTGGACCCAATGGC
GCTGGAGGGCCGCAAGGACCAGGAGGCCCACAAGGAGGTGGCCCATTCGGCCCAGGAGGACCACAGAGTGGTGGCCCTTACGGACCAGGTGGACCACAAGGACCG
CAGGGTGGTGGCCCTTACGGACCAGGGGGATCACAAGCTTTCGCTCAAGCCTCCAGTGGGCCTGGACAAGGTGGACCACGAGGACCAGGCAATCAAGGACCTGGA
CGAGGCGGACCACAAGGACCTGGACAAGGTGGGCCACAAGGACCTGGACCTTACGGACCAGGAGGACCCGGCTCTCAAGCTTTCGCTTTTGCTTCGGGTGTACCT
GGTGGTCCACAAGGAGGACCACAAGGGCCAGGTAACCAGGGACCCGGAGGGCCTCGAGGACCTGGACAAGGGGGACCACAAGGTTCGTCTTCAACAGCCTATGCG
TTTTCTTCTTCATCTGGCTCTCAAGAAAGGATGAATTCTTTGGCGTCTGCTCTTTCATCGATGGGATCATCTGCAGATTACCCAGCCCTGAATAAACTTCTGTCT
GGACTCATTTCTAGCATCAGGCAGAGCAATCCTGGTATTTCGCCACAGAATGCGATGACTCAAGCTCTGCTAGAAGCGCTTGCAGTTATGTTCGAACTGCTCGTT
GCAGTGCTGTCAGGAGGTGGAGGCGTCAATTCAGTCTCCTCTTACGGCGTAAGTGCTCTCGTTCAAGCAATGATTCAAGCACAGTCTTATTAA

>putative_Pp_sgo_Spidroin_5b type:5prime_partial len:422 c29082_g1_i4:1395-130(-)
GGACCGCAAGGACCTGGACAAGGCGGACCAAGCAGTGCAACAGCTATGGCGATTGCTTCGTCCGGACCGGGAAGTGGCCAAGGCGGTCCTTCTGGTGGAGATGGT
GCGGCGGTTTCAAGTTTCGCCTCCGCTTTTGCTTCTGCTTTGGTGAGCAGCAATGCATACAACCAGAGGATTGGTGGTGTAGGGCCATCAGGAGCAGGATCAGCG
ATGGACATGTCAATGGGTCAAGCTTATGGACAGTCTATTGCGCAGGCAATTGCACAAGCCACAGCGCAGGCATATTCTAACTCTGGAGGAAATAGTCAAGCCACA
GCCTATGCTATCTCAAACGTTGTGGCACAGTATCTGAACAATTATGGAGTCATACAATCTGCCGACGCCACATCACTCGCTCTAGCGTTCGCATCAGCTCTGTCT
TCTGCAGGGGGTCAAGGACCATATGGACCGGGATCTGGACAAGGAGGTCCACAAGGAGGTTCCAGTTCGATGGCTTTTGCTACAGCTTCCAGTGGACCCAATGGC
GCTGGAGGGCCGCAAGGACCAGGAGGCCCACAAGGAGGTGGCCCATTTGGCCCAGGAGGCCCACAAAGAGGTGGCCCATTCGGACCTGGAGGCCCACAAGGAAAT
GGCCCATTCGGCCCAGGAGGACCACAGAGTGGTGGCCCTTACGGACCGGGTGGACCACAAGGACCACAGGGTGGCGGCCCTTACGGACCAGGGGGACCACAAGGA
CCCGGACAAGGCGGACCACAAGGACCTGGACAAGGTGGGCCACAAGGACCTGGACCTTACGGACCAATGGGACCCGGCTCTCAATCTTTCGCTTTTGCTTCGGGT
GGACCTGGTGGTCCACAAGGAGGACCACAAGGGCCAGGTAACCAGGGACCCGGAGGGCCTCGAGGACCTGGACAAGGGGGACCACAAGGTTCGTCTTCAACAGCC
TATGCGTTTTCTTCTTCATCTGGCTCTCAAGAAAGGATGAATTCTTTGGCGTCTGCTCTTTCATCGATGGGATCATCTGCAGATTACCCAGCCCTGAATAAACTT
CTGTCTGGACTCATTTCTAGCATCAGGCAGAGCAATCCTGGTATTTCGCCACAGAATGCGATGACTCAAGCTCTGCTAGAAGCGCTTGCAGTTATGTTCGAACTG
CTCGTTGCAGTGCTGTCAGGAGGTGGAGGCGTCAATTCAGTCTCCTCTTACGGCGTAAGTGCTCTCGTTCAAGCAATGATTCAAGCACAGTCTTATTAA

>putative_Pp_sgo_Spidroin_5c type:5prime_partial len:418 c29082_g1_i5:1383-130(-)
GGACCGCAAGGACCTGGACAAGGCGGACCAAGCAGTGCAACAGCTATGGCGATTGCTTCGTCCGGACCGGGAAGTGGCCAAGGCGGTCCTTCTGGTGGAGATGGT
GCGGCGGTTTCAAGTTTCGCCTCCGCTTTTGCTTCTGCTTTGGTGAGCAGCAATGCATACAACCAGAGGATTGGTGGTGTAGGGCCATCAGGAGCAGGATCAGCG
ATGGACATGTCAATGGGTCAAGCTTATGGACAGTCTATTGCGCAGGCAATTGCACAAGCCACAGCGCAGGCATATTCTAACTCTGGAGGAAATAGTCAAGCCACA
GCCTATGCTATCTCAAACGTTGTGGCACAGTATCTGAACAATTATGGAGTCATACAATCTGCCGACGCCACATCACTCGCTCTAGCGTTCGCATCAGCTCTGTCT
TCTGCAGGGGGTCAAGGACCATATGGACCGGGATCTGGACAAGGAGGTCCACAAGGAGGTTCCAGTTCGATGGCTTTTGCTACAGCTTCCAGTGGACCCAATGGC
GCTGGAGGGCCGCAAGGACCAGGAGGCCCACAAGGAGGTGGCCCATTTGGCCCAGGAGGCCCACAAAGAGGTGGCCCATTCGGACCTGGAGGCCCACAAGGAAAT
GGCCCATTCGGCCCAGGAGGACCACAGAGTGGTGGCCCTTACGGACCGGGTGGACCACAAGGACCACAGGGTGGCGGCCCTTACGGACCAGGGGGACCACAAGGA
CCCGGACAAGGCGGACCACAAGGACCTGGACAAGGTGGGCCACAAGGACCTGGACCTTACGGACCAATGGGACCCGGCTCTCAATCTTTCGCTTTTGCTTCGGGT
GGACCTGGTGGTCCACAAGGAGGACCACAAGGGCCAGGTAACCAGGGACCCGGAGGGCCTCGAGGACCTGGACAAGGGGGACCACAAGGTTCGTCTTCAACAGCC
TATGCGTTTTCTTCTTCATCTGGCTCTCAAGAAAGGATGAATTCTTTGGCGTCTGCTCTTTCATCGATGGGATCATCTGCAGATTACCCAGCCCTGAATAAACTT
CTGTCTGGACTCATTTCTAGCATCAGGCAGAGCAATCCTGGTATTTCGCCACAGAATGCGATGACTCAAGCTCTGCTAGAAGCGCTTGCAGTTATGTTCGAACTG
CTCGTTGCAGTGCTGTCAGGAGGTGGAGGCGTCAATTCAGTCTCCTCTTACGGCGTAAGTGCTCTCGTTCAAGCAATGATTCAAGCACAGTCTTATTAA

>putative_Pp_sgo_Spidroin_5d type:5prime_partial len:407 c29082_g1_i9:1350-130(-)
GGACCGCAAGGACCTGGACAAGGCGGACCAAGCAGTGCAACAGCTATGGCGATTGCTTCGTCCGGACCGGGAAGTGGCCAAGGCGGTCCTTCTGGTGGAGATGGT
GCGGCGGTTTCAAGTTTCGCCTCCGCTTTTGCTTCTGCTTTGGTGAGCAGCAATGCATACAACCAGAGGATTGGTGGTGTAGGGCCATCAGGAGCAGGATCAGCG
ATGGACATGTCAATGGGTCAAGCTTATGGACAGTCTATTGCGCAGGCAATTGCACAAGCCACAGCGCAGGCATATTCTAACTCTGGAGGAAATAGTCAAGCCACA
GCCTATGCTATCTCAAACGTTGTGGCACAGTATCTGAACAATTATGGAGTCATACAATCTGCCGACGCCACATCACTCGCTCTAGCGTTCGCATCAGCTCTGTCT
TCTGCAGGGGGTCAAGGACCATATGGACCGGGATCTGGACAAGGAGGTCCACAAGGAGGTTCCAGTTCGATGGCTTTTGCTACAGCTTCCAGTGGACCCAATGGC
GCTGGAGGGCCGCAAGGACCAGGAGGCCCACAAGGAGGTGGCCCATTTGGCCCAGGAGGCCCACAAAGAGGTGGCCCATTCGGACCTGGAGGCCCACAAGGAAAT
GGCCCATTCGGCCCAGGAGGACCACAGAGTGGTGGCCCTTACGGACCGGGTGGACCACAAGGACCACAGGGTGGCGGCCCTTACGGACCAGGGGGACCACAAGGA
CCCGGACAAGGCGGACCACAAGGACCTGGACAAGGTGGGCCACAAGGACCTGGACCTTACGGACCAATGGGACCCGGCTCTCAATCTTTCGCTTTTGCTTCGGGT
GGACCTGGTGGTCCACAAGGAGGACCACAAGGGCCAGGTAACCAGGGACCCGGAGGGCCTCGAGGACCTGGACAAGGGGGACCACAAGGTTCGTCTTCAACAGCC
TATGCGTTTTCTTCTTCATCTGGCTCTCAAGAAAGGATGAATTCTTTGGCGTCTGCTCTTTCATCGATGGGATCATCTGCAGATTACCCAGCCCTGAATAAACTT
CTGTCTGGACTCATTTCTAGCATCAGGCAGAGCAATCCTGGTATTTCGCCACAGAATGCGATGACTCAAGCTCTGCTAGAAGCGCTTGCAGTTATGTTCGAACTG
CTCGTTGCAGTGCTGTCAGGAGGTGGAGGCGTCAATTCAGTCTCCTCTTACGGCGTAAGTGCTCTCGTTCAAGCAATGATTCAAGCACAGTCTTATTAA

>putative_Pp_sgo_Spidroin_5e type:5prime_partial len:222 c29082_g1_i10:795-130(-)
CAAGGAGGCGGCCCATTCGGACCAGGAGGCCCACAAGGAGGTGGCCCATTCGGCCCAGGAGGACCACAGAGTGGTGGCCCTTACGGACCGGGTGGACCACAAGGA
CCACAGGGTGGCGGCCCTTACGGACCAGGGGGACCACAAGGACCCGGACAAGGCGGACCACAAGGACCTGGACAAGGTGGGCCACAAGGACCTGGACCTTACGGA
CCAATGGGACCCGGCTCTCAATCTTTCGCTTTTGCTTCGGGTGGACCTGGTGGTCCACAAGGAGGACCACAAGGGCCAGGTAACCAGGGACCCGGAGGGCCTCGA
GGACCTGGACAAGGGGGACCACAAGGTTCGTCTTCAACAGCCTATGCGTTTTCTTCTTCATCTGGCTCTCAAGAAAGGATGAATTCTTTGGCGTCTGCTCTTTCA
TCGATGGGATCATCTGCAGATTACCCAGCCCTGAATAAACTTCTGTCTGGACTCATTTCTAGCATCAGGCAGAGCAATCCTGGTATTTCGCCACAGAATGCGATG
ACTCAAGCTCTGCTAGAAGCGCTTGCAGTTATGTTCGAACTGCTCGTTGCAGTGCTGTCAGGAGGTGGAGGCGTCAATTCAGTCTCCTCTTACGGCGTAAGTGCT
CTCGTTCAAGCAATGATTCAAGCACAGTCTTATTAA

>putative_Pp_sgo_Spidroin_6_Nterm type:3prime_partial len:231 c39845_g1_i1:27-716(+)
ATGGGTTTTCTCAACGGAATCGTTCTACTCACCGTCCTTTGGACACAAGGCGTTCTCGTCCCGGTGCAAGCTAAAGCCGCTTCACCATGGGAAAACATGAACAAT
GGCATGGGTTTCATCCAAACGATGTTGAACAACTTACCGCCAAGTCAGTTCACTGGTGGACAGTTGAACGACATCCGACAGATTGGATCGTCATTGATGGACGTC
ATGTCTTCGATGCAAAGTTCTGGTAAGGGGTCAGCGTCTATCAGCCATGCTATGGATGTCGCTATGGCATCTGCTATAGCTGAAGTCATAGCCTCCGAAGGTGGT
GCCAGTGCCCAAGCGCAAGCAAGTGCTGTGGCAAATGCCATGAGCCAGGCCTTCATGCAAATATCCGGTAGAGTCAACAAGCCGTATATTAGAGAATTCACTGAC
CTTCTCGGTATTATGATCGCTGATGCCTCCTCAAGTAATGAAGTTGCTACATCCGCATCTTCCGCAGCTGCTGCATCAAGTGCTGCAGGTGGTGGATATGGTGGA
GGTTACGGTGGATACGGCGGTGCTGGCGCAGGAGCAGCAGCAGCCTCAGCCGCATCTGCTGGAGCAGCTGGTGGATATGGAGGACAAGGTGGATACAGTCAAGGT
GCTGGAGCCGGCGCAGGTGCAGCAGCAGCCGCAGCTGCAGGAGCTGGAGCAGGACAAGGA

>putative_Pp_wa_AcSc_MaSp_3_Nterm type:internal len:219 c29986_g2_i1:656-3(-)
TTGTTCGTCAAAAAAAGCGCAACCATGGAGCTGTTCTTACCGTTCTTACTCCTGCTATCCACAGCCCAACTACAAACGGTATGTGGAAGCCACAGTTCACCGGGC
TCCCCATGGACTGATAAGCAGACTGCAATACTATTCATGAGGAATTTCGTTCAAGAGGTCGAGAATAGCGAAGCTATGAGTGATGAAGAAAAGGCGGACGTTGGC
AATGTGGCATCTCTGATGGAGACGTCGGTGATGAGCAGCCGTTCCAAAACAGCCTTGCAAGCCATGAACGCGGGGTTTTTGACTGCCGTGACCGAGGTATTGGTA
CTGGATAATGACAACGGGATCAACCAAATACCGATTGTCATCAATGCCTACCAGAATGCGCTGATAAAGACGGTTGGTTACGCTGACAACAACTACATGAGAGAA
ACCATGAGTATCGTGCAATCAGTCGCTCAGACCAGCGGTGACAATGACATACCGGGATCGAACCAGCCGTCGGGAAATCAAGGACCTTATGGACCTGGACAAGGT
GGGCCCGGGGGACCAAATTCTCAAGCTTACGCTTTTGCTCAAAGCGGAACTCAAGGCCCTAATGGTCCAAGGGGACCATCAGGTCCGTATGGGCCAGGAGGACCA
CAAGGACCAGGTGGACCACAAGGA

>putative_Pp_wa_AcSp_MaSp_1_Nterm type:internal len:110 c20461_g1_i1:2-328(+)
TTTACGTTTCTGCTAGTGATCGCAACCACTCTTTTTGCAATCAGTGATGCACGAGGTAAATCAGCAACATCACCTTGGTCGGACCCGGCAAAGGCCAAAGTATTC
ATAACCAAATTCATCGATGTGATCGGACAATCCAGTGCTTTCCCCGACCAGCAAAAGGAAGATATGGAGTCCATTGCCGAAACAATTATGTCTGCTGTTGATGGG
ATGGCGAGAAACAAAAGAGGATATGCAGCAAAGATGCAAGCCATGAACATGGCATTTGCTTCCTCGATGGCTGATCTTATTGTTTTGGAAAATGACATGAGTGGT
AGTGTTAACGAA

>putative_Pp_wa_MaSp_1a_Nterm type:internal len:265 c52491_g1_i1:793-2(-)
TTAAGCGCAAACGAAGGTCCCACCACGATGCGTTGGTTTGCAGTGTCGCTGGTCGTCACGACCTTCCTGGTGCTCTACCAATCGGCAGAATGTAAGAAAAAGGAG
AAAAACTACAAAAACAACCCTTGGGCAAGCAAAGAGAAAGCTTCGATTTTCGCTAATACTTTCATAAACTCGGCCAAGTCTTCAGATGCCTTCAGCCAGGATCAA
AAAGACAGTATCGAAGCCATCGGGAATACCATAACGTACGCTATGGTGAAAGTGGACACGTCTAAGAAGAACGCGAAAGCTACGTACCAAAGTCTTGACATGGCT
TTAGCTTCCGCAATCGCTGAGATCATAGCGAACGGGGACAATATTAAAGTACAGACAACAGCAGCGATGGAAGCTATGGCGGAAGCATTCAGAAGAACAACTGGT
AAAGTGAACAAGGTATACATAGCTGACATAAAAGAACTGATCGAAATGTTTGCTAATAGCAAGCTTGAAAGTGGAGAAGATGACTATGATTACGATGACGACGAC
GATGATGATTCATTGAATTCGATCCTTGGCTCCAGTTCACCTTCATCATCTGTCATGTTACCACCACCACCACCTCTTCCGCAGTTCTTCCAACCTCCACCAATG
CCAGTGAGCCCAGCTCCACCTAGCACAATTATAACCTCTTCCAGCTCGTCTTCGTCTAGCAGTAGTTCGTCAAGCTCCTTATCTTCCAGTTCTCAGACGACTTTT
TCCCAACAATCACCATCGGCACCTCCACCGTCACCACCGCCTCCCCCTCCTGCACCA

>putative_Pp_wa_Spidroin_1c type:5prime_partial len:362 c36429_g1_i3:1212-127(-)
ACAGCACCAGTTCAAGTGTCTTCGACCAGTACGTCTTCTTCTTCCGCTAGCGTGTCTTCTTCTTCTTCCGCTAGCGTGTCTTCTTCTTCTTCCGCTAGCGTGTCT
TCTTCTGTTGCCACAACGCAAGTCAGTGTAGCTCCTCAGCCACCACAAGCCCCGGCTGCCCCTGACTACTCTGAATTCACTAACTCTCTAAGCTCTTCTCTTGTG
GCTTCAGAACTGTTCACAACCCTATTCTCAAGCATCACTGATGTTTCCACAGCTCAAAGTTACTCATCGTCGATGGCTCAAGGTGCAGGGATGTTTGTAGGTTTA
GATTCTACCACATCCGGGTTGACAAACTCTGTCGTGCAGAGCATCACTGCTGTTGGTGCTGGCAGTGACGTATCAACTTATGCCACATCAATTGCTTCCTCAATT
GTAAGCTCTATTTCCAGTTCTCAGACAATTCTTACTAACTATGATGTTACTTCCTTCAGTGCAGCAATGCTTTCGGCTCTCGAGAAATTCAGACCACAAGCTCCT
AGTGCTCCTGCTGTGTCAATATCATCCAGTATTGCTTCTTCCTCCTCTTCAACTACATCATCTAGCAGTACTTCCTCCTCCTCCTCTGCCTCATCCAGTAGCGCT
TCCTCTACTACTTCAACTGGTGATGCTTCCCCTTCGGCTCCTGCGTCACCTGGTGCAGATGTGGCAACTGTTCCTTCTGGTGGTTGTGCAAATACTGGGCCAAGC
TCAGGCAGCGATGCTTACCCTCCTATTTCTATGGACACTGCTTTAGCCTTTTTCACTCCCTCGGTTCCAATGTCGACTAGTACTTCCCTGCGAATGTGTAACTTT
GCAACAACGATCCTGTCTTCGATAGCATCGGAGTCTGGGTTCAACTATATGGACTTTTCACGTGCTTTGAATGTCCTGAGCTCTTCAATACGGTCCAGTGGCTCT
GGTTCTTCCTCGACAAGTCTTATCATTCAGTTCCTCTTTGAAATCGTTGTATCACTGGTCCAAGTACTGCTCCGTATGAATGCATCTACTGATGATACAGCGAAT
GCTCTTTACACTGCTATATCGCAGTCATTTGCTTGA

>putative_Pp_wa_Spidroin_1d type:5prime_partial len:445 c36429_g1_i1:1461-127(-)
AATGCTCAAACTTACGCATCAGCAATGGCCCAATCTGCTGGAACTCTATTGGGATCTGACCAGTCAACTACCAGTAGTTTTACCAATGCGGTATCGCAAGGTGTA
ACGGGTGTAGGCAGTAATGCTGGAACTTACTCTTACGCCAGTACTATTGCATCTACTACGACTTCATTCTTGTCAAACATTCAGGTAATTTCCAATTACGATGCT
AACACATTTGCAGCGACAATGCTCTCCGCGCTTTCCAGATTCAAGCCTCAAGAAGCTCAGCCAACAGCACCAGTTCAAGTGTCTTCGACCAGTACGTCTTCTTCT
TCCGCTAGCGTGTCTTCTTCTTCTTCCGCTAGCGTGTCTTCTTCTGTTGCCACAACGCAAGTCAGTGTAGCTCCTCAGCCACCACAAGCCCCGGCTGCCCCTGAC
TACTCTGAATTCACTAACTCTCTAAGCTCTTCTCTTGTGGCTTCAGAACTGTTCACAACCCTATTCTCAAGCATCACTGATGTTTCCACAGCTCAAAGTTACTCA
TCGTCGATGGCTCAAGGTGCAGGGATGTTTGTAGGTTTAGATTCTACCACATCCGGGTTGACAAACTCTGTCGTGCAGAGCATCACTGCTGTTGGTGCTGGCAGT
GACGTATCAACTTATGCCACATCAATTGCTTCCTCAATTGTAAGCTCTATTTCCAGTTCTCAGACAATTCTTACTAACTATGATGTTACTTCCTTCAGTGCAGCA
ATGCTTTCGGCTCTCGAGAAATTCAGACCACAAGCTCCTAGTGCTCCTGCTGTGTCAATATCATCCAGTATTGCTTCTTCCTCCTCTTCAACTACATCATCTAGC
AGTACTTCCTCCTCCTCCTCTGCCTCATCCAGTAGCGCTTCCTCTACTACTTCAACTGGTGATGCTTCCCCTTCGGCTCCTGCGTCACCTGGTGCAGATGTGGCA
ACTGTTCCTTCTGGTGGTTGTGCAAATACTGGGCCAAGCTCAGGCAGCGATGCTTACCCTCCTATTTCTATGGACACTGCTTTAGCCTTTTTCACTCCCTCGGTT
CCAATGTCGACTAGTACTTCCCTGCGAATGTGTAACTTTGCAACAACGATCCTGTCTTCGATAGCATCGGAGTCTGGGTTCAACTATATGGACTTTTCACGTGCT
TTGAATGTCCTGAGCTCTTCAATACGGTCCAGTGGCTCTGGTTCTTCCTCGACAAGTCTTATCATTCAGTTCCTCTTTGAAATCGTTGTATCACTGGTCCAAGTA
CTGCTCCGTATGAATGCATCTACTGATGATACAGCGAATGCTCTTTACACTGCTATATCGCAGTCATTTGCTTGA

>putative_Pp_wa_Spidroin_2a type:5prime_partial len:167 c31056_g1_i1:647-147(-)
GGAGCAGGACAAGGTGGATATGGTGGCCAAGGATACGGTCAAGGAGCCGGAGCAGGTGCAGCCGCAGCTGCCGCCGCAGGAGCTGGACAAGGTGGATACGGAGGT
CAAGGATACGGTCAAGGTGCAGGAGCAGGAGCAGCAGCAGCAGGAGCAGCTGGTGGATATGGACCACAATATGGAGCAAGCACCGCATCTAGTGCAGCAGCA
GCAAGCGCAGCAGCAGCAGCAGCTTCACCAATGCTAAGACCATCCGGTGGAATCAACCAAGCAGTATCATCACAAGGAATCAGCGCAGTTGCCAACGCAGTCGTT
TCCATGGGCTCAAGCATCAGCTACCAAGCTCTCGGCGGACAGATGCAAGCACTCGCATCAAGCATCGCATCTGCCAATGCCGGTCTATCTCAAGCCGAAGTCACG
ATCGAAGTTTTGCGCGAAGTTATCATCGCTCTGATGCAAGTCATTGCATCAGCCTCAGCTTCCGTATCCGTCTACGGTTGA

>putative_Pp_wa_Spidroin_2b type:5prime_partial len:225 c31056_g1_i2:821-147(-)
GGAGCAGGACAAGGTGGATATGGTGGCCAAGGATACGGTCAAGGAGCCGGAGCAGGTGCAGCCGCAGCAGCAGCCGCAGGTGGTGCAGGAGGTTATGGCGCAGGA
CTCGCAGGAGCAGCAGCCGCAGCCGCTCAAGCCGAAGCAGCAGCAGCCGCCCAAGCTCAAGCACAAGCCGCCGCCGAAGCACAAGCTTCCGCCGCAGCCCAAGCA
GAAGCTCAAGCAGCAGCAGCAGCCGCAGGAGCTGAATCTGCATTAGCCGCTGGTCTAGGAGGCTACGGTGGATATGGTGCAGGAGCAGCATCAGCAGCATCAGCA
GCAGCATCAGGAGCAGCTGGTGGATATGGACCACAATATGGAGCAAGCACCGCATCTAGTGCAGCAGCAGCAAGCGCAGCAGCAGCAGCAGCTTCACCAATGCTA
AGACCATCCGGTGGAATCAACCAAGCAGTATCATCACAAGGAATCAGCGCAGTTGCCAACGCAGTCGTTTCCATGGGCTCAAGCATCAGCTACCAAGCTCTCGGC
GGACAGATGCAAGCACTCGCATCAAGCATCGCATCTGCCAATGCCGGTCTATCTCAAGCCGAAGTCACGATCGAAGTTTTGCGCGAAGTTATCATCGCTCTGATG
CAAGTCATTGCATCAGCCTCAGCTTCCGTATCCGTCTACGGTTGA

>putative_Pp_wa_Spidroin_3 type:5prime_partial len:269 c34483_g1_i1:1-807(+)
GCAGTAAACACTGCAATCGGTCCAATGTATGGTCCGCAAGTTTCGTCGTCGGCAGCGAGCTCAACCGTAAGCGCGGTATCTAGGATCGGACCCGGAGCAAGCCCA
ACCAGTACTGCTACAGCGATATCATCGTCTGTTGCAAGCACACTGTCCTCGTACGGTTACCTGAACACAGCTAGCGCGACACAAGTGGGCAGTTCCGTTGCATCT
GCTTTAATGCAAGCACCAAGACAGCCGCAGCAAGCACCGACAATTGCACCCAGCAGTCAAACAACTATGTCGACGGCGTCATCTGCGTCGTCAGCATCTTCTGCT
TCTTCAGCGACTACTACAGTGACCACTGCTCAAGCCCCATTTCTTCTCCCACAAGCAGCACCAACTACTGTTGGTTCCGCTTTTCCTCCTACTAGTCTTGGACTT
CCAAGTTTCCCAGCCTCTTATGGCTCCCAGCCTTCTCTTGGTAATGCAGTTCTTCCGTACGCTGGAATCCCTCCTTCGTTATCAGCTTCATCGCCTTCTGGTCAA
GGACGAATTAATTCAATTATGTCTTCTATGAGATCACTCCTTTCTCCTGGATCTCAGTTTGACTATTCGAGCTTGGCAAATACTCTTACAGGCCTTGGCAGATTT
TTGCGTCAAAGTAATTCAGGTTTGTCTTCTTCGCAAATTTTGATTGAAGTTCTGATTGAGGCTCTAATTGCCATCATAGAAATGCTGAAAACCGCAAGTATCACG
AATGCTACGCCGATGATTTCGTCCAGCACAGGAGCAGAGTTGACGTCTTTTATATCGAGTATGATGGTTTAA

>putative_Pp_wa_Spidroin_5f type:5prime_partial len:167 c35144_g1_i1:1-501(+)
CAAGGACCTGGACAAGGTGGGCCACAAGGACCTGGACCTTACGGACCAATGGGACCCGGCTCTCAATCTTTCGCCTTTGCTTCGGGTGGACCTGGTGGTCCACAA
GGAGGACCACAAGGGCCAGGTAACCAGGGACCCGGAGGGCCTCGAGGACCTGGACAAGGGGACCACAAGGTTCGTCTTCAACAGCCTATGCGTTTTCTTCTTCA
TCTGGCTCTCAAGAAAGGATGAATTCTTTGGCGTCTGCTCTTTCATCGATGGGATCATCTGCAGATTACCCAGCCCTGAATAAACTTCTGTCTGGACTCATTTCT
AGCATCAGGCAGAGCAATCCTGGTATTTCGCCACAGAATGCGATGACTCAAGCTATGCTAGAAGCGCTTGCAGTTATGTTCGAACTGCTCGTTGCAGTGCTGTCA
GGAGGTGGAGGCGTCAATTCAGTCTCCTCTTACGGCGTAAGTGCTCTCGTTCAAGCAATGATTCAAGCACAGTCTTATTAA

>putative_Pp_wa_Spidroin_6_Nterm type:3prime_partial len:231 c22032_g1_i1:22-711(+)
ATGGGTTTTCTCAACGGAATCGTTCTACTCACCGTCCTTTGGACACAAGGCGTTCTCGTCCCGGTGCAAGCTAAAGCCGCTTCACCATGGGAAAACATGAACAAT
GGCATGGGTTTCATCCAAACGATGTTGAACAACTTACCGCCAAGTCAGTTCACTGGTGGACAGTTGAACGACATCCGACAGATTGGATCGTCATTGATGGACGTC
ATGTCTTCGATGCAAAGTTCTGGTAAGGGGTCAGCGTCTATCAGCCATGCTATGGATGTCGCTATGGCATCTGCTATAGCTGAAGTCATAGCCTCCGAAGGTGGT
GCCAGTGCCCAAGCGCAAGCAAGTGCTGTGGCAAATGCCATGAGCCAGGCCTTCATGCAAATATCCGGTAGAGTCAACAAGCCGTATATTAGAGAATTCACTGAC
CTTCTCGGTATTATGATCGCTGATGCCTCCTCAAGTAATGAAGTTGCTACATCCGCATCTTCCGCAGCTGCTGCATCAAGTGCTGCAGGTGGTGGATATGGTGGA
GGTTACGGTGGATACGGCGGTGCTGGCGCAGGAGCAGCAGCAGCCTCAGCCGCATCTGCTGGAGCAGCTGGTGGATATGGAGGACAAGGTGGATACAGTCAAGGT
GCTGGAGCCGGCGCAGGTGCAGCAGCAGCCGCAGCTGCAGGAGCTGGAGCAGGACAAGGA

9.3. Tensile test data

Table 9.2 Transformed tensile data from the three groups of tensile-tested silk analysed in this study.
Samples which were broken before being fitted to equipment for testing, or were broken during fitting resulting in a failure to obtain data, have been excluded (see section 6.3.3 Table 6.6). The data below has been visualised in section 6.3.3 Figures 6.7 – 6.9.

Spider	Treatment	Sample number	Stress (Pa)	Strain	Young's Modulus (Pa)	Toughness (Pa or Jm^{-3})	Toughness (MPa or MJm^{-3})
A.d. 2	Untreated silk	4	5.13E+08	2.00E-01	2.57E+09	2.57E+09	51.30
		5	4.04E+08	1.30E-01	3.11E+09	3.11E+09	26.27
		7	1.18E+09	3.70E-01	3.18E+09	3.18E+09	217.62
		8	9.74E+08	5.10E-01	1.91E+09	1.91E+09	248.43
		9	1.07E+09	4.60E-01	2.32E+09	2.32E+09	245.53
		10	3.01E+08	2.50E-01	1.20E+09	1.20E+09	37.57
		11	8.08E+08	1.80E-01	4.49E+09	4.49E+09	72.76
		12	1.15E+09	6.60E-01	1.74E+09	1.74E+09	377.93
		13	8.14E+08	5.10E-01	1.60E+09	1.60E+09	207.46
		14	1.09E+09	5.90E-01	1.84E+09	1.84E+09	321.03
		15	1.11E+09	7.30E-01	1.52E+09	1.52E+09	404.77
		16	1.11E+09	6.40E-01	1.74E+09	1.74E+09	356.53
		17	3.73E+08	3.80E-01	9.82E+08	9.82E+08	70.89
		18	8.45E+08	5.90E-01	1.43E+09	1.43E+09	249.18
		19	1.12E+09	5.70E-01	1.96E+09	1.96E+09	319.01
		20	1.13E+09	6.30E-01	1.80E+09	1.80E+09	357.49
		21	1.13E+09	6.10E-01	1.86E+09	1.86E+09	346.14
		22	1.14E+09	6.50E-01	1.75E+09	1.75E+09	370.52
		23	1.07E+09	5.20E-01	2.05E+09	2.05E+09	277.55
		24	1.17E+09	5.90E-01	1.98E+09	1.98E+09	343.96
		25	1.09E+09	5.10E-01	2.13E+09	2.13E+09	277.50
P.p.	Untreated silk	1	2.79E+08	1.60E-01	1.75E+09	1.75E+09	22.34
		2	2.44E+08	6.60E-01	3.70E+08	3.70E+08	80.62
		4	2.27E+08	4.20E-01	5.40E+08	5.40E+08	47.64
		5	3.66E+08	7.10E-01	5.16E+08	5.16E+08	130.10
		6	2.79E+08	3.90E-01	7.16E+08	7.16E+08	54.45
		7	1.75E+08	2.10E-01	8.31E+08	8.31E+08	18.32
		8	1.92E+08	2.60E-01	7.38E+08	7.38E+08	24.95
		9	2.27E+08	3.40E-01	6.67E+08	6.67E+08	38.57
		10	1.22E+08	8.00E-02	1.53E+09	1.53E+09	4.89
		11	2.62E+08	5.50E-01	4.76E+08	4.76E+08	71.98
		12	2.27E+08	1.80E-01	1.26E+09	1.26E+09	20.42
		13	2.27E+08	3.40E-01	6.67E+08	6.67E+08	38.57
		14	1.05E+08	1.60E-01	6.54E+08	6.54E+08	8.38
		16	2.44E+08	3.40E-01	7.19E+08	7.19E+08	41.53

		18	1.75E+08	3.00E-01	5.82E+08	5.82E+08	26.18
		19	1.92E+08	3.60E-01	5.33E+08	5.33E+08	34.55
		20	2.09E+08	4.40E-01	4.76E+08	4.76E+08	46.07
		21	1.08E+09	3.30E-01	3.28E+09	3.28E+09	178.52
		23	2.79E+08	3.90E-01	7.16E+08	7.16E+08	54.45
		24	2.97E+08	3.10E-01	9.57E+08	9.57E+08	45.98
A.d. 1	CoMoCAT	1	2.90E+07	5.60E-01	5.18E+07	5.18E+07	8.13
		2	3.37E+07	6.50E-01	5.18E+07	5.18E+07	10.95
		3	3.94E+07	7.60E-01	5.18E+07	5.18E+07	14.97
		4	2.69E+07	5.20E-01	5.18E+07	5.18E+07	7.01
		5	2.38E+07	4.60E-01	5.18E+07	5.18E+07	5.48
		6	2.23E+07	4.30E-01	5.18E+07	5.18E+07	4.79
		7	3.83E+07	7.40E-01	5.18E+07	5.18E+07	14.19
		8	3.68E+07	7.10E-01	5.18E+07	5.18E+07	13.06
		9	3.42E+07	6.60E-01	5.18E+07	5.18E+07	11.29
	P2	11	2.44E+07	4.70E-01	5.18E+07	5.18E+07	5.72
		12	2.44E+07	4.70E-01	5.18E+07	5.18E+07	5.72
		16	2.54E+07	4.90E-01	5.18E+07	5.18E+07	6.22
		18	2.95E+07	5.70E-01	5.18E+07	5.18E+07	8.42
		19	3.16E+07	6.10E-01	5.18E+07	5.18E+07	9.64
	MWNT	21	3.11E+07	6.00E-01	5.18E+07	5.18E+07	9.33
		22	2.69E+07	5.20E-01	5.18E+07	5.18E+07	7.01
		23	4.04E+07	7.80E-01	5.18E+07	5.18E+07	15.76
		24	2.95E+07	5.70E-01	5.18E+07	5.18E+07	8.42
		25	3.32E+07	6.40E-01	5.18E+07	5.18E+07	10.61
		26	2.33E+07	4.50E-01	5.18E+07	5.18E+07	5.25
		28	3.21E+07	6.20E-01	5.18E+07	5.18E+07	9.96
		29	3.52E+07	6.80E-01	5.18E+07	5.18E+07	11.98
		30	1.97E+07	3.80E-01	5.18E+07	5.18E+07	3.74
	Graphite	32	2.38E+07	6.90E-01	3.45E+07	3.45E+07	8.22
		33	3.58E+07	3.00E-01	1.19E+08	1.19E+08	5.36
		34	1.55E+07	7.10E-01	2.19E+07	2.19E+07	5.52
		38	3.37E+07	4.20E-01	8.02E+07	8.02E+07	7.07
		39	2.18E+07	6.10E-01	3.57E+07	3.57E+07	6.64
		40	3.16E+07	6.10E-01	5.18E+07	5.18E+07	9.64
P.p.	CoMoCAT	26	2.97E+08	4.20E-01	7.06E+08	7.06E+08	62.30
		27	1.92E+08	2.10E-01	9.14E+08	9.14E+08	20.16
		28	3.14E+08	3.60E-01	8.73E+08	8.73E+08	56.54
		30	2.44E+08	2.70E-01	9.05E+08	9.05E+08	32.98
		31	2.62E+08	2.20E-01	1.19E+09	1.19E+09	28.79
		32	2.97E+08	4.60E-01	6.45E+08	6.45E+08	68.23
		33	2.97E+08	2.20E-01	1.35E+09	1.35E+09	32.63
		34	2.97E+08	4.90E-01	6.05E+08	6.05E+08	72.68
	P2	37	2.79E+08	3.80E-01	7.35E+08	7.35E+08	53.05
		38	2.62E+08	3.90E-01	6.71E+08	6.71E+08	51.04
		40	2.44E+08	2.20E-01	1.11E+09	1.11E+09	26.87

		41	2.44E+08	3.20E-01	7.63E+08	7.63E+08	39.09
		42	2.27E+08	2.80E-01	8.10E+08	8.10E+08	31.76
		43	2.09E+08	3.60E-01	5.82E+08	5.82E+08	37.69
	MWNT	47	1.92E+08	2.70E-01	7.11E+08	7.11E+08	25.91
		48	1.92E+08	3.10E-01	6.19E+08	6.19E+08	29.75
		50	2.44E+08	2.50E-01	9.77E+08	9.77E+08	30.54
		52	2.79E+08	1.90E-01	1.47E+09	1.47E+09	26.53
		53	2.27E+08	3.70E-01	6.13E+08	6.13E+08	41.97
		54	2.27E+08	3.00E-01	7.56E+08	7.56E+08	34.03
		55	1.57E+08	6.00E-02	2.62E+09	2.62E+09	4.71
	Graphite	57	2.27E+08	1.60E-01	1.42E+09	1.42E+09	18.15
		59	2.62E+08	3.80E-01	6.89E+08	6.89E+08	49.73
		63	2.27E+08	4.00E-01	5.67E+08	5.67E+08	45.37
		64	1.92E+08	2.90E-01	6.62E+08	6.62E+08	27.83
		65	1.92E+08	3.00E-01	6.40E+08	6.40E+08	28.79
A.d. 3	CoMoCAT	1	9.28E+08	4.16E-01	2.23E+09	2.23E+09	193.09
		2	9.28E+08	2.15E-01	4.32E+09	4.32E+09	99.79
		5	1.39E+09	2.93E-01	4.75E+09	4.75E+09	204.00
		6	9.28E+08	3.01E-01	3.08E+09	3.08E+09	139.71
		7	1.39E+09	2.05E-01	6.79E+09	6.79E+09	142.73
		8	2.55E+09	3.12E-01	8.18E+09	8.18E+09	398.24
	P2	1	6.96E+08	1.92E-01	3.63E+09	3.63E+09	66.84
		3	1.62E+09	2.49E-01	6.52E+09	6.52E+09	202.25
		4	1.16E+09	3.52E-01	3.30E+09	3.30E+09	204.23
		6	2.09E+09	3.26E-01	6.41E+09	6.41E+09	340.46
		8	1.16E+09	3.02E-01	3.84E+09	3.84E+09	175.22
		9	1.39E+09	2.26E-01	6.16E+09	6.16E+09	157.35
		10	9.28E+08	2.23E-01	4.16E+09	4.16E+09	103.51
	MWNT	1	1.16E+09	3.36E-01	3.45E+09	3.45E+09	194.94
		2	1.16E+09	3.12E-01	3.72E+09	3.72E+09	181.02
		3	1.62E+09	2.91E-01	5.58E+09	5.58E+09	236.37
		4	4.64E+08	2.29E-01	2.03E+09	2.03E+09	53.15
		5	9.28E+08	3.43E-01	2.71E+09	2.71E+09	159.20
		6	1.62E+09	2.02E-01	8.04E+09	8.04E+09	164.08
		7	4.64E+08	3.63E-01	1.28E+09	1.28E+09	84.24
		9	1.16E+09	4.11E-01	2.82E+09	2.82E+09	238.46
		10	1.62E+09	3.08E-01	5.27E+09	5.27E+09	250.18
	Graphite	1	2.09E+09	4.13E-01	5.06E+09	5.06E+09	431.31
		2	1.86E+09	2.70E-01	6.88E+09	6.88E+09	250.64
		3	2.09E+09	3.61E-01	5.79E+09	5.79E+09	377.01
		4	1.16E+09	2.77E-01	4.19E+09	4.19E+09	160.71
		5	1.62E+09	3.40E-01	4.78E+09	4.78E+09	276.17
	Graphite 1 in 10	1	1.62E+09	3.28E-01	4.95E+09	4.95E+09	266.42
		2	2.09E+09	3.44E-01	6.07E+09	6.07E+09	359.25
		3	1.16E+09	1.93E-01	6.01E+09	6.01E+09	111.98
		4	1.16E+09	3.27E-01	3.55E+09	3.55E+09	189.72

	5	1.16E+09	2.31E-01	5.02E+09	5.02E+09	134.02
Surfactant	1	2.09E+09	2.88E-01	7.25E+09	7.25E+09	300.77
	2	6.96E+08	2.99E-01	2.33E+09	2.33E+09	104.09
	3	1.39E+09	2.62E-01	5.31E+09	5.31E+09	182.41
	4	1.16E+09	2.69E-01	4.31E+09	4.31E+09	156.07
	5	1.62E+09	3.48E-01	4.67E+09	4.67E+09	282.67
	6	6.96E+08	1.20E-01	5.80E+09	5.80E+09	41.77
	7	9.28E+08	3.04E-01	3.05E+09	3.05E+09	141.10
	8	1.16E+09	2.17E-01	5.35E+09	5.35E+09	125.90
	9	1.16E+09	2.94E-01	3.95E+09	3.95E+09	170.58
	10	9.28E+08	3.08E-01	3.01E+09	3.01E+09	142.96
Untreated silk	1	2.78E+09	4.62E-01	6.03E+09	6.03E+09	643.32
	2	9.28E+08	4.67E-01	1.99E+09	1.99E+09	216.76
	3	2.09E+09	2.94E-01	7.10E+09	7.10E+09	307.04
	4	9.28E+08	3.16E-01	2.94E+09	2.94E+09	146.67
	5	2.09E+09	2.83E-01	7.38E+09	7.38E+09	295.55
	6	6.96E+08	2.38E-01	2.93E+09	2.93E+09	82.85
	7	1.86E+09	4.42E-01	4.20E+09	4.20E+09	410.31
	8	9.28E+08	2.16E-01	4.30E+09	4.30E+09	100.26
	9	1.16E+09	2.27E-01	5.11E+09	5.11E+09	131.70
	10	1.39E+09	2.44E-01	5.71E+09	5.71E+09	169.88

9. Appendix

10. Acknowledgements

International borders posed no barrier and I have the deepest appreciation for my invaluable poster-of-spiders and mountain-exploring collaborator Milan Řezáč, who collected, dissected and imaged spider silk glands on request.

Early investigations were spearheaded by Victor Tudorica, who performed all the RNA extractions and initial transcriptome assemblies and showed only patience when instructing. Sequencing undertaken by Sunir Malla of the DeepSeq facility, University of Nottingham. Thanks also to Angus Davison for the loan of a server for transcriptome assembly.

I've grown to appreciate the skill involved when operating electron microscopes, acknowledging the expertise of Nikki Weston (eSEM), Christine Grainger-Boultby and Denise McLean (Nanoscale and Microscale research centre, UoN) who assisted with the collection of high-quality silk images and let me loose on a microscope to collect images of my own.

I received assistance with tensile testing from Sassan Emadi-Coffin, Patrick Pope, Dominic Charles and Finn O'Conor (School of Physics & Astronomy) and thank Andrew Parsons (School of Engineering) for assistance with and access to equipment.

Raman spectroscopy, methodology, and acquirement and preparation of carbon nanomaterials were performed with the utmost care by Graham Rance, who was ever patient in helping me understand the basics.

Technical assistance by Sheila Keeble, Julie Rodgers and Liaque Latif is also gratefully acknowledged.

Thanks also go to the Biological and Biosciences Research Council and the School of Life Sciences at the University of Nottingham for awarding me a studentship and funding to supporting this work.

This work was supported by Biotechnology and Biological Sciences Research Council doctoral training studentships to Michelle Strickland and Dr Victor Tudorica [Grant Number BB/M008770/1], the Royal Society [Grant Number IE141413 to Dr Sara L Goodacre and Dr Milan Řezáč], and the Ministry of Agriculture of the Czech Republic [Grant Number MZe Ro0415 to Dr Milan Řezáč].

10. Acknowledgements

11. PIP Reflective Statement

I completed my 3-month Professional Internship for PhD Students (PIPS) with Net4FS, based in Pr. Dabing Zhang's developmental plant biology group at Shanghai Jiao Tong University (SJTU). This placement satisfies the compulsory training placement required as part of my BBSRC DTP-funded PhD, and was funded by the Net4FS mobility scheme.

The three aims of the placement were; to collect data on the effect of heat stress on rice plant fertility and flower/seed development, work within the areas of student recruitment and communication with the University of Nottingham (UoN), and present my own research at international conference in Chongqing (October 2017).

- Working on a project in collaboration between Pr. Zoe Wilson and Pr. Dabing Zhang's groups in the Life Sciences department at SJTU allowed me to gain a range of experience in plant research methods, including;
- field collection
- staining and imaging of anthers in preparation for assessing pollen count and viability
- recording and preparation of mature rice panicle specimens for imaging
- development of a reproducible imaging process
- generating image data to enable the development of seed-count software at UoN
- supervising a Masters student and auxiliary staff in preparation of panicles for imaging
- training a student in the imaging process to allow continuation/reproduction of images after my departure from SJTU

I was also able to join a bi-weekly language class held for international students, where I practiced beginner's level Mandarin as both an aural and written language and was introduced to Chinese culture.

In addition to my research efforts at SJTU, I also assisted the University of Nottingham's International office at the Chinese Education Expo (Shanghai), a student recruitment fair, where I assisted Ryan Parker and Lily Ma of the University of Nottingham's International Office with the promotion of the University as an international destination for aspiring undergraduate and postgraduate students. Quickly gaining an understanding of the Chinese

education system and student recruitment process, I spoke with many students and their parents about my experiences in Nottingham as both resident and student.

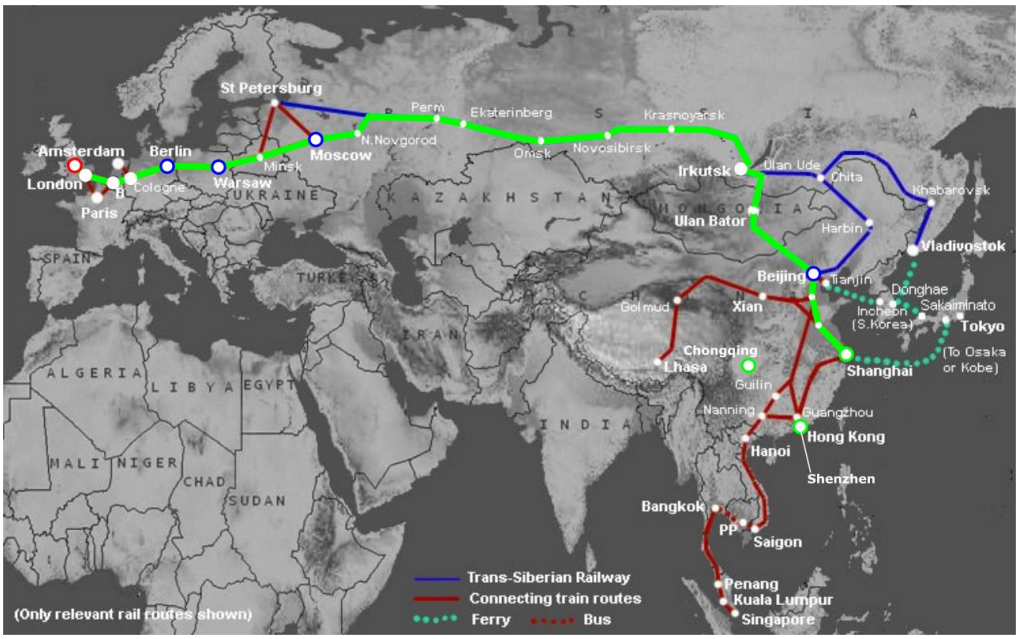
Finally, I was able to attend an international research conference at South West University, Chongqing, in October 2017. Not only was I able to present a talk on my own research, but I was able to meet and discuss this research with respected scientists from around the globe and within China, who suggested alternate research techniques and offered advice on improving my research to a high standard. Invitations to visit research groups were also extended, both at the conference and through Pr. Wilson, and I was able to visit Pr. Li-Jia Qu at Peking University and Pr. Shuqiang Li at the Chinese Academy of Sciences in Beijing. These visits not only allowed me to experience a city very different to Shanghai, but also allowed for further development of my research skills and knowledge.

I initially found the placement to be quite difficult – whilst I settled into my accommodation and quickly became familiar with my surroundings, I didn't have the best time starting the project. Initially I had very little to do and spent a lot of time waiting for rice to grow – in hindsight I could perhaps have used this downtime to better explore Shanghai, but without any local support (everyone had their own work to do) it was stressful going to the lab each day with no tasks to complete. However, once I had made some contacts and the plants matured, there was plenty to do and I was free to develop my own working strategies. Interestingly, I didn't feel particularly uncomfortable spending time in a country where I was obviously not a local!

Overall, not only have I had an immersive experience of Chinese culture, society and research, but I have also made invaluable contacts, expanded my knowledge and experience of academia in educational, research and administrative areas and laid the groundwork for pursuing an international career after the completion of my PhD. I am greatly indebted to Net4FS, Pr. Wilson, Dr Miriam Columbi, Pr. Zhang and my SJTU supervisor Pr. Wanqi Liang for this opportunity.

PS. The details of my travel both to and from the placement were unspecified. As such, I took the opportunity to make the return leg of my journey by rail (as opposed to the more typical method of long-distance transport, the aeroplane). I highly recommend this method of travel!

Shanghai – Nottingham by rail, December 2017 (green).



Map: modified from seat61.com/Trans-Siberian









11. PIPS Reflective Statement

12. Publications

Some of the data from Chapter 4 has been published as a research article in the *Nature* journal *Heredity*. A copy of this publication as it appears in print is included in the following pages.

Heredity
https://doi.org/10.1038/s41437-018-0050-9

the geneticssociety

ARTICLE

Conservation of a pH-sensitive structure in the C-terminal region of spider silk extends across the entire silk gene family

Michelle Strickland[1] · Victor Tudorica[2] · Milan Řezáč[3] · Neil R. Thomas[2] · Sara L. Goodacre[1]

Received: 15 June 2017 / Revised: 1 December 2017 / Accepted: 2 December 2017
© The Author(s) 2018. This article is published with open access

Abstract

Spiders produce multiple silks with different physical properties that allow them to occupy a diverse range of ecological niches, including the underwater environment. Despite this functional diversity, past molecular analyses show a high degree of amino acid sequence similarity between C-terminal regions of silk genes that appear to be independent of the physical properties of the resulting silks; instead, this domain is crucial to the formation of silk fibers. Here, we present an analysis of the C-terminal domain of all known types of spider silk and include silk sequences from the spider *Argyroneta aquatica*, which spins the majority of its silk underwater. Our work indicates that spiders have retained a highly conserved mechanism of silk assembly, despite the extraordinary diversification of species, silk types and applications of silk over 350 million years. Sequence analysis of the silk C-terminal domain across the entire gene family shows the conservation of two uncommon amino acids that are implicated in the formation of a salt bridge, a functional bond essential to protein assembly. This conservation extends to the novel sequences isolated from *A. aquatica*. This finding is relevant to research regarding the artificial synthesis of spider silk, suggesting that synthesis of all silk types will be possible using a single process.

Introduction

Spider silk proteins form multiple types of materials, including fibers and glues, each with specific mechanical properties and functions (Blamires et al. 2017; Eisoldt et al. 2011; Garb 2013; Hormiga and Griswold 2014). Gene duplication, recombination and diversification are thought to have generated this huge diversity of different silk types, each having a different ecological function (Clarke et al. 2015). These functions include safety lines, the structural frameworks of webs, external layers of protection around egg sacs, and securing prey (Hinman et al. 2000; Rising and Johansson 2015).

In the case of *Argyroneta aquatica* (Clerck 1757) (Araneae: Cybaeidae (Catalog 2016)), silk is used in a typical fashion to construct a web in which the spider resides and performs a number of actions, from feeding to mating and storing eggs (Schütz and Taborsky 2003, 2005, 2011; Schutz et al. 2007; Seymour and Hetz 2011). What is unusual about this spider is that it is the only known species to spin silk whilst submersed in water. The subsequent sheet web is then inflated with air drawn down from the surface and is utilised as an air reservoir. The silken "diving bell" allows for oxygen diffusion to occur, allowing the spider to avoid surfacing for extended periods of time (Seymour and Hetz 2011). *Cybaeus angustiarum*, a fellow cybaeid, is a terrestrial species found in dense forests (often on north facing, scree slopes) in areas under stones or in decaying wood with humidity close to 100%.

In all spiders, liquid silk dope is passed through specialised, elongated glands within the body of the spider and extruded through spinnerets (Askarieh et al. 2010; Rising and Johansson 2015; Vollrath and Knight 2001).

Electronic supplementary material The online version of this article (https://doi.org/10.1038/s41437-018-0050-9) contains supplementary material, which is available to authorized users.

✉ Michelle Strickland
michelle.strickland@nottingham.ac.uk

[1] School of Life Sciences, University of Nottingham, Nottingham NG7 2RD, UK

[2] Centre for Biomolecular Sciences, School of Chemistry, University of Nottingham, Nottingham NG7 2RD, UK

[3] Biodiversity Lab, Crop Research Institute, Drnovská 507, 16106 Prague 6 - Ruzyně Czechia

Published online: 15 February 2018

All silk genes have three components; a type-specific repetitive region flanked by conserved N-terminal (Motriuk-Smith et al. 2005; Rising et al. 2006) and C-terminal (Challis et al. 2006; Collin et al. 2016; Gnesa et al. 2012; Hagn et al. 2010) domains. Whilst the repetitive region determines the mechanical properties of each silk product (Hayashi and Lewis 1998; Hayashi et al. 1999), the terminal domains work together ensuring that individual proteins assemble correctly and that the fibre forms at the correct stage of the spinning process (Andersson et al. 2014; Andersson et al. 2017; Rising and Johansson 2015). The N-terminal domain restricts the formation of silk fibers to a precise point in the silk duct, preventing silk proteins stored in the silk gland from agglutinating (Askarieh et al. 2010). The C-terminal domain drives spontaneous fibre formation, likely through use of a pH-sensitive "salt bridge" (Ittah et al. 2006; Stark et al. 2007), where noncovalent interactions between one basic and one acidic residue are disrupted at low pH because the acidic residue becomes protonated and is no longer charged.

Salt bridges, either individual or paired, have been proposed to explain the dimeric bundling of 4 or 5 alpha helices in the C-termini of one particular silk, the major ampullate (Hagn et al. 2010; Sponner et al. 2004a, 2005). Whilst the miniature spidroin 4RepCT, formed of four copies of a MaSp repetitive region and one C-terminus, has been sufficient to produce self-assembling silk fibers (Stark et al. 2007), it has recently been shown in the minispidroin NT2RepCT that including the N-terminal region achieves greater efficiency in the production of synthetic fibers (Andersson et al. 2017). However, studies have shown that the level of solubility demonstrated by the terminal regions differs between spider species, leading to further questions around their function and the suitability of individual terminal regions for use in silk protein synthesis (Andersson et al. 2014, 2017; Askarieh et al. 2010).

A degree of amino acid sequence conservation has been observed from studies of small numbers of different silks (Beckwitt and Arcidiacono 1994; Challis et al. 2006; Collin et al. 2016; Gnesa et al. 2012; Hagn et al. 2010; Sponner et al. 2004b, 2005). What is not clear is the extent to which this conservation is maintained, particularly where silk proteins exhibit vastly different properties (e.g., the glue-like aggregate and piriform silk versus the superior strength of major ampullate and aciniform silks). Additionally, the diving bell spider *Argyroneta aquatica* spins silk whilst completely submersed in fresh water; given how silk proteins are dehydrated as part of the spinning process, does this "extreme" environment necessitate variation within the silk protein and spinning process?

Here, we analyse the genetic sequences silks of all known types from spiders of as many groups as are available in GenBank. We investigate whether the salt bridge structure is conserved across all species and silk types and ask if changes in biophysical properties or the utilisation of silk in an "extreme" environment necessitates a change in how silk proteins are formed.

Materials and methods

Transcriptome assembly

RNA was extracted from the silk glands and whole abdomen of adult females and sequenced on an Illumina NextSeq500 (DeepSeq, University of Nottingham). Transcriptomes were trimmed using Scythe (Buffalo 2014) and Sickle (Joshi and Fass 2011) and assembled using Trinity (Grabherr et al. 2011). Protein sequences were predicted using TransDecoder (Haas et al. 2013).

Identification of silk genes

Custom blast databases of silk genes downloaded from GenBank and the *Nephila clavipes* (Babb et al. 2017) and *Stegodyphus mimosarum* genomes (Sanggaard et al. 2014) were generated to screen transcriptomes for silk sequences, which were manually examined and blasted against GenBank for confirmation (e-value < 0.0005).

Phylogenetic analysis

Silk sequences from GenBank, the *N. clavipes* and *S. mimosarum* genomes and newly-identified spidroins were filtered to only include C-terminal regions, identified by the "QALLE" motif (Challis et al. 2006). Duplicate sequences were discarded, leaving the longest sequence for analysis. Two sequences were considered identical if their nucleotide sequences had only a small number of variations which did not significantly alter the amino acid composition (e.g., variation causing a polar residue to change to acidic was considered significant, whereas polar to polar was not). Selected sequences were manually trimmed leaving only the C-terminal region (maximum length: 360 bp (Hagn et al. 2010; Rising and Johansson 2015)).

Selected nucleotide sequences were aligned in Geneious v8.1.8 (Kearse et al. 2012) (Clustal W algorithm (Larkin et al. 2007)) and subsequently translated in order to allow refined alignment by eye. The refined untranslated nucleotide sequence alignment was imported into MEGA6.0 (Tamura et al. 2013) and used to construct a maximum-likelihood tree (1000 bootstrap replicates; see supplementary file 2).

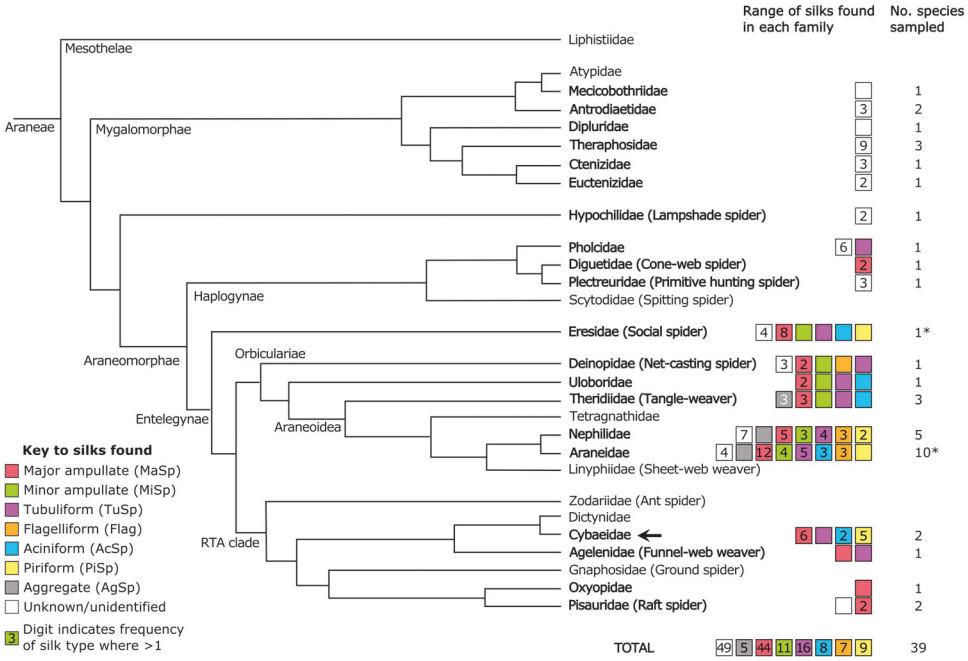


Fig. 1 Familial origins of silk sequences used in this study. Phylogenetic tree from Nentwig (Nentwig et al. 2013), modified to show positions of the different spider families and the silks previously described as being of a particular type. Location of *Argyroneta aquatica* and *Cybaeus angustiarum* (Cybaeidae) marked with an arrow. Each colored box represents a silk type; uncoloured boxes represent unclassified published spidroins. Numbers in each box indicate the total number of occurrences of each silk type. The total number of species sampled per family is detailed at the end of each row; * indicates where sequences have been retrieved from a genome

Results

Diversity of silk C-termini

Forty-four new silk sequences isolated from transcriptomes of the spiders *A. aquatica*, *C. angustiarum*, and *Pholcus phalangioides* included a C-terminal domain (this study; GenBank Accession Numbers MG744694–MG744714). Once refined (see methods), 21 unique sequences remained for analysis. Three hundred and thirty-four silk sequences were retrieved from GenBank and the *N. clavipes* and *S. mimosarum* genomes, of which 150 sequences were both unique and contained a C-terminal region. Specialised silks were mostly found in the Orbiculariae and *S. mimosarum* genome, with the exception of one tubuliform and a small number of major ampullate sequences (Babb et al. 2017; Garb et al. 2010; Perez-Rigueiro et al. 2010; Rising et al. 2007; Sanggaard et al. 2014; Stark et al. 2007) (Fig. 1). An unrooted maximum-likelihood tree of nucleotide sequences (1000 bootstraps) shows the clustering together of sequences that are identified as belonging to the same silk type by BLAST (see supplementary figure 2).

Conservation of a salt bridge structure across spider silk genes

Analysis of the predicted secondary structures of sequences representing each silk type with JPred4 (Drozdetskiy et al. 2015) suggests the secondary structure of the C-terminus is conserved (data not shown), containing up to five α-helices, supporting previous studies (Hagn et al. 2010; Ittah et al. 2007). However, some predictions suggest Helix 1 may be absent, including in the *N. clavipes* MaSp-a sequence analysed here (Fig. 2).

In all sequences analysed in this study, one acidic residue and one basic residue pair has been conserved. A conserved basic residue is consistently found in Helix 2 (position 30–31 Fig. 2, 23–37 supplementary figure 1), whilst an acidic base is always found towards the center of Helix 4 (position 85 in Fig. 2; 88 in supplementary figure 1). Additionally, in major ampullate, minor ampullate, aciniform, and piriform sequences a second acidic and basic residue appears to be partially conserved (supplementary figure 1) as identified by Sponner et al. (2004b, 2005). These pairs correspond with those already shown to form a salt bridge in the secondary protein structure in previous studies (Hagn et al. 2010; Ittah et al. 2006).

Content of the C-terminal domain

The majority of the C-terminus is composed of hydrophobic and hydrophilic amino acids (average 53.5 and 37.0% respectively) including serine, alanine, and leucine (average 21.2, 13.2, and 10.8% respectively) in an alternating arrangement. These residues promote the formation of alpha-helices in the protein structure, with the hydrophilic residues exposed, promoting solubility (Hagn et al. 2010).

Overall DNA sequence identity is highest in the acidic residue found in the "QALLE" motif (Challis et al. 2006), where the first two nucleotides (guanine, adenine) are completely conserved and the final nucleotide determines whether it is a glutamic or aspartic acid residue (89.3 and

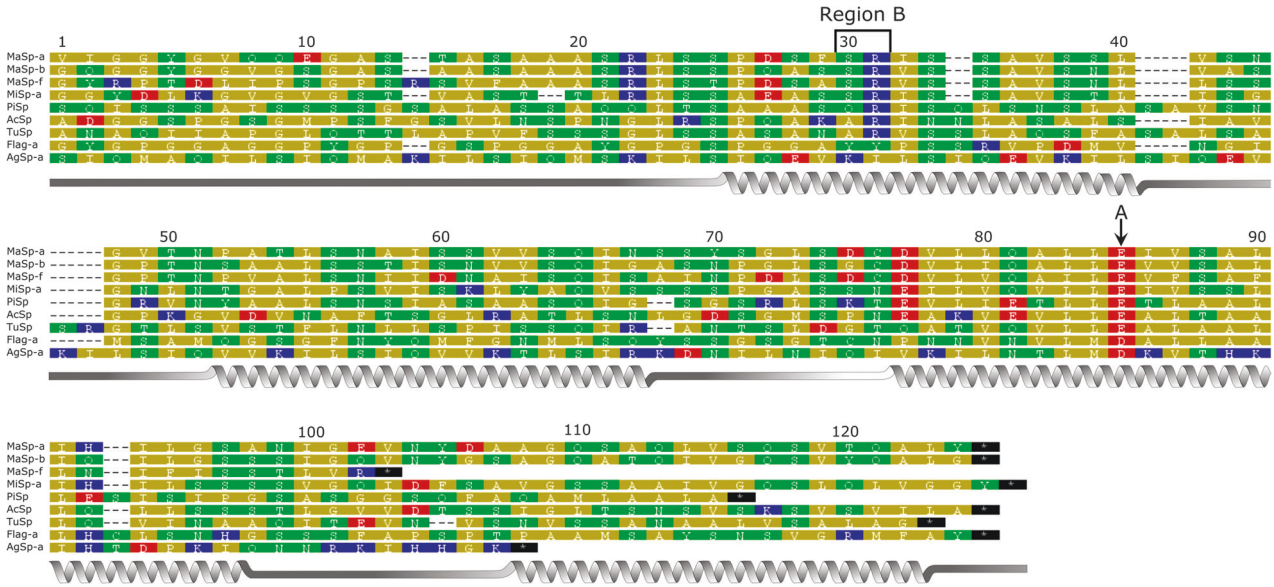


Fig. 2 Alignment of *Nephila clavipes* C-terminal sequences representing characterized silk types, with predicted helical secondary structure of MaSp-a below. Amino acid residues implicated in the formation of a salt bridge are marked A (acidic residue; glutamic or aspartic acid) and found within region B (basic residue; arginine or lysine). Residues are highlighted to show whether acidic (red), basic (blue), hydrophilic (green) or hydrophobic (yellow)

10.7%, respectively of all sequences in this study; see supplementary figure 1). The basic residues involved in the formation of the salt bridge are usually arginine, although some are lysine or histidine (94.7, 3.3, and 2.0%, respectively).

Overall, the charged residue content of the C-terminus is typically less than 10% (5.2% acidic, 4.1% basic; average across all sequences in this study) although this varies depending on the sequence (aggregate spidroins average 13% acidic, 8% basic). However, one acidic and basic residue pair (A and region B, respectively, Fig. 2 and supplementary figure 1) are present in each sequence examined, irrespective of species or silk type.

Additionally, we find a raised percentage of the amino acids surrounding the basic residue within spider silk are hydrophilic (43.0%, region B, supplementary figure 1), reducing its pKa and hence promoting its protonation, whereas a majority of those surrounding the acidic residue are hydrophobic (68.8%, region A, supplementary figure 1), increasing its pKa and again promoting protonation.

Discussion

This analysis encompasses 19 families from within the Araneae and shows conservation of a pH-sensitive salt bridge, typically composed of an arginine-glutamic acid pairing, within all the silk types and species examined. Our finding of this degree of conservation confirms an essential role for this feature in the correct assembly of silk fibers. The extended coverage in terms of species and phylogenetic diversity suggests this feature has been conserved for the entire evolutionary history of the group—some 360 million years. Moreover, this conservation has persisted despite the diversification of other spider traits. For example, terrestrial spiders such as the mygalomorphs produce two to three silk proteins from a large, undifferentiated silk gland whereas the Orbiculariae have evolved several specialised glands from which different types of silk are extruded (Blackledge and Hayashi 2006; Blamires et al. 2017; Clarke et al. 2017; Garb 2013; Garb et al. 2010; Perez-Rigueiro et al. 2010; Rising et al. 2007; Stark et al. 2007).

Conservation of C-terminus in all *A. aquatica* silk genes

Argyroneta aquatica predominantly spins its silk in water, although will occasionally leave a dragline when walking in a terrestrial environment. Comparatively, individual *Cybaeus* did not appear to produce a dragline but given time would spin silken webs in which they resided, which appear visually similar to the dehydrated diving bells of *A. aquatica*. In sequences obtained from both these species, and also from the first fully-annotated spider genome of *Nephila clavipes* (Babb et al. 2017) in which a number of new silk sequences were identified, the residue pairing within the C-terminal domain has been conserved. This suggests that the formation of silk fibers by *A. aquatica* occurs using the same method as in terrestrial spiders. This leads to questions around how this spider is adapted to spinning silk

underwater and the subsequent mechanical properties of *A. aquatica* silk in such an "extreme" environment. This is particularly relevant as studies of silk in humid conditions suggest the structure may be temporarily affected by the level of humidity; the degree to which major ampullate silk supercontracts depends on the species, whereas minor ampullate silk does not supercontract (Agnarsson et al. 2009; Blackledge et al. 2009; Boutry and Blackledge 2010).

Conservation of a physical structure

The residues responsible for the formation of the pH-sensitive salt bridge(s) are both uncommon and conserved within the restricted number of sequences studied thus far (Challis et al. 2006; Hagn et al. 2010). Analysis of the expanded range of sequences used in this study shows this trait is maintained, with >90% of the C-terminal domain typically composed of hydrophobic and hydrophilic residues. Of the charged residues present, at least one acidic and one basic residue is conserved in all sequences and as such inferred to be crucial to the formation of a salt bridge and therefore the correct folding of silk proteins into a fibre. The conservation of an arginine-glutamic acid pair of residues suggests this may be the optimum pairing for a salt bridge in silk, but the presence of lysine and aspartic acid in some sequences implies this is not always essential, although the effect of these substitutions on the final protein structure and its physical properties is currently unknown.

The higher level of hydrophilic residues around the basic residue and hydrophobic residues around the acidic residue may be necessary to ensure the correct formation of the salt bridge during the protein folding stage and allowing its regulation by smaller changes in pH due to the local environment created by the protein. Where two pairs of residues are conserved in major ampullate, minor ampullate, piriform, and some aciniform sequences it is likely that there are two salt bridges present in the C-terminus, although further structural analysis would confirm that the second is not merely sequence duplication or an evolutionary artefact suggesting the presence of a second bridge.

Conclusion

In silk research, the C-terminus is a key component of the minimal sequence used for recombinant spider silk production, as without this region fibers will not form (Stark et al. 2007). What we find is an overall model for all silk types irrespective of physical traits that illustrates the range of environments in which this single protein family may be utilised.

Our results allow us to identify conserved amino acid residues essential to the correct formation of silk proteins, thereby enabling the identification of less essential residues that may be chemically functionalised in artificially synthesised silk using techniques such as click-chemistry (Harvey et al. 2016). This study will aid researchers in selecting suitable sequences without repetitive testing in vivo whilst predicting sites which may be suitable for modification, as seen in Harvey et al. (2016), and build a repository of spidroin parts that could be combined to achieve novel and custom characteristics.

Summary

Here, we compare C-terminal sequences representing all known silk types, and from all the major clades within the Araneae. We illustrate how the salt bridge structure is conserved throughout the gene family as a whole, and show by inclusion of new data that this conservation extends to silks found in all the major divisions within the Araneae. This degree of conservation includes groups that are highly diverged in terms of their basic morphology, such as their silk spinning apparatus (Coddington 2005). We include new data on silk of the Cybaeid *Argyroneta aquatica*, the only extant species adapted to spin silk in an underwater environment, and a terrestrial Cybaeid for comparison, *Cybaeus angustiarum*. The insight gained allows us to characterize amino acid residues essential to correct protein assembly and aids the identification of residues that might be substituted to chemically functionalised alternatives for use in artificially synthesised silk proteins (Harvey et al. 2016).

Declarations

Availability of data and materials

The novel sequences generated in this study are deposited in the GenBank repository, Accession Numbers MG744694–MG744714 (https://www.ncbi.nlm.nih.gov/genbank/).

Acknowledgements Spider collection and dissection by MR. RNA extracted by VT and MS, transcriptomics and analysis by MS, with genomics assistance by James Fowkes. All living authors read and approved the final manuscript. Sequencing and data quality control by Sunir Malla (DeepSeq facility, University of Nottingham). In memory of Victor Tudorica (1990–2016).

Funding This work was supported by the Royal Society [Grant Number IE141413 to SLG and MR], Biotechnology and Biological Sciences Research Council doctoral training studentships to MS and VT [Grant Number BB/M008770/1] and the Ministry of Agriculture of the Czech Republic [Grant Number MZe RO0415 to MR].

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. If you remix, transform, or build upon this article or a part thereof, you must distribute your contributions under the same license as the original. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/4.0/.

References

Agnarsson I, Boutry C, Wong SC, Baji A, Dhinojwala A, Sensenig AT et al. (2009) Supercontraction forces in spider dragline silk depend on hydration rate. Zoology 112:325–331

Andersson M, Chen G, Otikovs M, Landreh M, Nordling K, Kronqvist N et al. (2014) Carbonic anhydrase generates CO(2) and H(+) that drive Spider silk formation via opposite effects on the terminal domains. PLoS Biol 12(8):e1001921

Andersson M, Jia Q, Abella A, Lee X-Y, Landreh M, Purhonen P et al. (2017) Biomimetic spinning of artificial spider silk from a chimeric minispidroin. Nat Chem Biol. 13:262–264

Askarieh G, Hedhammar M, Nordling K, Saenz A, Casals C, Rising A et al. (2010) Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. Nature 465:236–238

Babb PL, Lahens NF, Correa-Garhwal SM, Nicholson DN, Kim EJ, Hogenesch JB 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, 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

Blackledge TA, Boutry C, Wong SC, Baji A, Dhinojwala A, Sahni V (2009) How super is supercontraction? Persistent versus cyclic responses to humidity in spider dragline silk. J Exp Biol 212:1981–1989

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

Blamires SJ, Blackledge TA, Tso I-M (2017) Physicochemical property variation in spider silk: ecology, evolution, and synthetic production. Annu Rev Entomol 62:443–460

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

Buffalo V (2014) Scythe—A Bayesian adapter trimmer (version 0.994 BETA) [Software]

Catalog WS (2016) World Spider Catalog, version 17

Challis RJ, Goodacre SL, Hewitt GM (2006) Evolution of spider silks: conservation and diversification of the C-terminus. Insect Mol Biol 15:45–56

Clarke TH, Garb JE, Haney RA, Chaw RC, Hayashi CY, Ayoub NA (2017) Evolutionary shifts in gene expression decoupled from gene duplication across functionally distinct spider silk glands. Sci Rep 7:8393

Clarke TH, Garb JE, Hayashi CY, Arensburger P, Ayoub NA (2015) Spider transcriptomes identify ancient large-scale gene duplication event potentially important in silk gland evolution. Genome Biol Evol 7:1856–1870

Clerck CA (1757). Svenska Spindlar uti sina hufvud-slågter indelte samt under några och sextio särskildte arter beskrefne och med illuminerade figurer uplyste/Aranei Svecici, descriptionibus et figuris æneis illustrati, ad genera subalterna redacti, speciebus ultra LX determinati

Coddington JA (2005) Phylogeny and classification of spiders. In: Ubick D, Paquin P, Cushing PE, & Roth V (eds) Spiders of North America: an identification manual. American Arachnological Society, Poughkeepsie, New York, USA

Collin MA, Clarke TH, Ayoub NA, Hayashi CY (2016) Evidence from multiple species that spider silk glue component ASG2 is a spidroin. Sci Rep 6:21589

Drozdetskiy A, Cole C, Procter J, Barton GJ (2015) JPred4: a protein secondary structure prediction server. Nucleic Acids Res 43: W389–W394

Eisoldt L, Smith A, Scheibel T (2011) Decoding the secrets of spider silk. Mater Today 14:80–86

Garb JE (2013) Spider silk: an ancient biomaterial for the 21st century. In: Penney D (ed.) Spider research in the 21st century: trends and perspectives. Siri Scientific Press, Manchester, pp 252–281

Garb JE, Ayoub NA, Hayashi CY (2010) Untangling spider silk evolution with spidroin terminal domains. BMC Evolut Biol 10:243

Gnesa E, Hsia Y, Yarger JL, Weber W, Lin-Cereghino J, Lin-Cereghino G (2012) Conserved C-terminal domain of spider tubuliform spidroin 1 contributes to extensibility in synthetic fibers. Biomacromolecules 13:304–312

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29:644–652

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J et al. (2013) De novo transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. Nat Protocols 8(8): https://doi.org/10.1038/nprot.2013.084

Hagn F, Eisoldt L, Hardy JG, Vendrely C, Coles M, Scheibel T (2010) A conserved spider silk domain acts as a molecular switch that controls fibre assembly. Nature 465:239–242

Harvey D, Bardelang P, Goodacre SL, Cockayne A, Thomas NR (2017). Antibiotic spider silk: site-specific functionalization of recombinant spider silk using "Click" chemistry. Adv Mater 29:1604245-n/a.

Hayashi CY, Lewis RV (1998) Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks1. J Mol Biol 275:773–784

Hayashi CY, Shipley NH, Lewis RV (1999) Hypotheses that correlate the sequence, structure, and mechanical properties of spider silk proteins. Int J Biol Macromol 24:271–275

Hinman MB, Jones JA, Lewis RV (2000) Synthetic spider silk: a modular fiber. Trends Biotechnol 18:374–379

Hormiga G, Griswold CE (2014) Systematics, phylogeny, and evolution of Orb-weaving spiders. Annu Rev Entomol 59:487–512

Ittah S, Cohen S, Garty S, Cohn D, Gat U (2006) An essential role for the C-terminal domain of A dragline spider silk protein in directing fiber formation. Biomacromolecules 7:1790–1795

Ittah S, Michaeli A, Goldblum A, 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

Joshi N, Fass J (2011). Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files (Version1.33) [Software]

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H (2007) Clustal W and Clustal X version 2A0. Bioinformatics 23:2947–2948

Motriuk-Smith D, Smith A, Hayashi CY, Lewis RV (2005) Analysis of the conserved N-terminal domains in major ampullate spider silk proteins. Biomacromolecules 6:3152–3159

Nentwig W, Arnedo MA, Coddington JA, Hormiga G, Jocqué R, Kropf C et al. (2013) Appendix. Spider Phylogeny. In: Nentwig W (ed) Spider ecophysiology. Springer, Berlin

Perez-Rigueiro J, Plaza GR, Torres FG, Hijar A, Hayashi C, Perea GB (2010) Supercontraction of dragline silk spun by lynx spiders aOxyopidaea. Int J Biol Macromol 46:555–557

Rising A, Hjalm G, Engstrom W, Johansson J (2006) N-terminal nonrepetitive domain common to dragline, flagelliform, and cylindriform spider silk proteins. Biomacromolecules 7:3120–3124

Rising A, Johansson J (2015) Toward spinning artificial spider silk. Nat Chem Biol 11:309–315

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

Sanggaard KW, Bechsgaard JS, Fang X, Duan J, Dyrlund TF, Gupta V (2014) Spider genomes provide insight into composition and evolution of venom and silk. Nat Commun 5:3765

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. Evolut 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

Schütz D, Taborsky M (2011) Sexual selection in the water spider: female choice and male-male competition. Ethology 117:1101–1110

Schutz D, Taborsky M, Drapela T (2007) Air bells of water spiders are an extended phenotype modified in response to gas composition. J Exp Zool Part A 307:549–555

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

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

Sponner A, Vater W, Rommerskirch W, Vollrath F, Unger E, Grosse F et al (2005) The conserved C-termini contribute to the properties of spider silk fibroins. Biochem Biophys Res Commun 338:897–902

Stark M, Grip S, Rising A, Hedhammar M, Engström W, Hjälm G et al. (2007) Macroscopic fibers self-assembled from recombinant miniature spider silk proteins. Biomacromolecules 8:1695–1701

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729

Vollrath F, Knight DP (2001) Liquid crystalline spinning of spider silk. Nature 410:541–548