This is the peer reviewed version of the following article:

Expression of spidroin proteins in the silk glands of golden orb-weaver spiders

Inmaculada Jorge[a,b], Víctor Ruiz[c,d], Jesús Lavado-García[a,e], Jesús Vázquez[a,b], Cheryl Hayashi[f],

Francisco Javier Rojo[c,d], José Miguel Atienza[c,d], Manuel Elices[c,d,g], Gustavo Víctor Guinea[c,d,g],

José Pérez-Rigueiro[c,d,g]*

a. Cardiovascular Proteomics Laboratory. Centro Nacional de Investigaciones Cardiovasculares (CNIC). 28029, Madrid, Spain.

b. Centro de Investigación Biomédica en Red de Enfermedades Cardiovasculares (CIBERCV). 28029, Madrid, Spain.

c. Centro de Tecnología Biomédica. Universidad Politécnica de Madrid. 28223, Pozuelo de Alarcón (Madrid), Spain.

d. Departamento de Ciencia de Materiales. ETSI Caminos, Canales y Puertos. Universidad Politécnica de Madrid. 28040, Madrid, Spain.

e. Grup d'Enginyeria Cel·lular i de Bioprocessos (GECIB). Departament d'Enginyeria Química, Biològica i Ambiental. Universitat Autònoma de Barcelona. 08193, Barcelona, Spain.

f.   Division of Invertebrate Zoology and Sackler Institute for Comparative Genomics, American Museum of Natural History. New York, 10024 U.S.A.

g.   Biomedical Research Networking Center in Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN). Madrid, Spain.

*Corresponding Author

ABSTRACT

The expression of spidroins in the major ampullate, minor ampullate, flagelliform and tubuliform silk glands of *Trichonephila clavipes* spiders was analysed using proteomics analysis techniques. Spidroin peptides were identified and assigned to different gene products based on sequence concurrence when compared with the whole genome of the spider. It was found that only a relatively low proportion of the spidroin genes are expressed as proteins in any of the studied glands. In addition, the expression of spidroin genes in different glands presents a wide range of patterns, with some spidroins being found in a single gland exclusively, while others appear in the content of several glands. The combination of precise genomics, proteomics, microstructural and mechanical data provides new insights both on the design principles of these materials and how these principles might be translated for the production of high performance bioinspired artificial fibers.

INTRODUCTION

The biological success of a vast group of animal species is closely related to the exogenous materials that they may generate. The wax used by bees to build some structural elements of the hive and the different calcium-rich ceramics that constitute the shells of bivalve molluscs are just two of the wide number of examples that might be cited. However, it is difficult to find an example in which the evolutionary success of a lineage has been more dependent on the production of materials (Vollrath, 1999) than the one provided by the spinning of spider silk fibers. The very appearance of the order Araneae (spiders) is defined, in part, by the presence of specialized organs to produce the first silk fibers (Selden, Shear, & Sutton, 2008). From that initial event, the production of silk fibers has evolved along two complementary directions to reach its present level of diversification (Bond et al., 2014): the appearance of sets of specialized silk glands that produce either fibers or gluey solutions for different applications (Blackledge & Hayashi, 2006), and the addition of new motifs to the original silk protein (spidroin) sequences (Blackledge et al., 2009).

Up to seven different silk types are identified in the spiders of the Araneoidea group. This variety allows producing fibers and adhesive solutions that cover the wide range of distinct biological functions required by the spider. Thus, the outstanding combination of high tensile strength and strain at breaking exhibited by the major ampullate gland silk (MAS) (Heim, Keerl, & Scheibel, 2009) (Heidebrecht et al., 2015) has been a driving motivation for the development of bioinspired materials within the framework of Biomimetics. MAS is used for essential functions, such as constituting structural elements of the prey capture web and also as the lifeline spun by the spider while crawling, climbing, or descending (Swanson, Blackledge, Beltran, & Hayashi, 2006). Additionally, the requirements imposed by the building of efficient prey capture webs were probably a major driving force in the diversification of silk gland products since an additional fiber

4

and adhesive solution are found in the structure of the prototypical bidimensional orb-webs produced by Araneoidea: flagelliform silk (Adrianos et al., 2013)(Aparecido dos Santos-Pinto, Arcuri, Esteves, Palma, & Lubec, 2018) and viscous aggregate silk (Stellwaagen & Burns, 2021). Flagelliform silk is used to build the capture spiral threads of the web and is covered by the viscous fluid produced in the aggregate gland. Orb-weaving spiders also make minor ampullate silk, which is used as an auxiliary spiral during web construction, and pyriform silk that is utilized to attach the web to the surrounding elements, such as branches or leaves. In addition, spiders synthetize two additional silk fibers; tubuliform silk, from which the outer layer of the egg sac (Hu et al., 2005) is made, and aciniform silk, that is employed to swathe the prey (Ayoub, Garb, Kuelbs, & Hayashi, 2013).

Unveiling the principles of the spinning process requires a combined effort from different research fields, including genomics, proteomics and materials science. Thus, proteomics studies (Aparecido dos Santos-Pinto et al., 2019; Whaite, Wang, Macdnald, & Cummins, 2018) identified the presence of a numerous set of proteins in each of the spinning glands that can be grouped in six subsets whose functions range from housekeeping proteins to proteins related to the folding and modification of spidroins (Aparecido dos Santos-Pinto et al., 2015). Needless to say, the characterization of the subset that comprises the structural proteins with which the fibers are generated, the spidroin proteins, is especially relevant in order to understand the microstructure and properties of these materials.

Spidroins constitute a singular group of proteins, some of whose main features were apparent since the early studies of its primary structure (Xu & Lewis, 1990). In particular, the existence of a few characteristic motifs that appear frequently repeated along the sequence is considered as one of the

hallmarks of this group (Gatesy, Hayashi, Motriuk, Woods, & Lewis, 2001). In this regard, the presence and proportion of these motifs in the different proteins was used to establish an initial classification of the spidroins, that has been maintained thereafter (Hayashi, Shipley, & Lewis, 1999). Thus, major ampullate gland silk spidroin 1 (MaSp1) proteins were defined, in part, by the presence of the motif polyalanine $-A_n-$. MaSp1 proteins also present a few isolated isolated $-GA-$ motifs and a number of $-GGX-$ motifs, in which X is an amino acid from a small set that includes glutamine and leucine. Major ampullate gland spidroin 2 (MaSp2) proteins also show the presence of the $-A_n-$ motif, but additionally show a characteristic $-GPG-$ motif (Bittencourt, Dittmar, Lewis, & Rech, 2010). The $-GPG-$ motif also appears in the flagelliform gland silk spidroin (Flag), a spidroin that lacks the $-A_n-$ motif found in MaSp1 and MaSp2. In contrast, minor ampullate gland silk spidroin (MiSp) proteins are characterized by the presence of the polyalanine motif $-A_n-$ in combination with tandem repeats of the $-GA-$ motif and a few dispersed $-GGX-$ motifs. Other spidroins, such as tubuliform gland (TuSp) and aciniform gland spidroin (AcSp) contain very few of these simple motifs within their more complicated tandem repeats (Gatesy et al., 2001)·(Hayashi, Blackledge, & Lewis, 2004)·(Garb & Hayashi, 2005). Spidroin proteins also tend to show highly conserved N- and C-terminal domains (Askarieh et al., 2010; Hagn et al., 2010) that play a critical role in the self-assembly of the proteins in the gland and their subsequent conversion into a solid fiber (Jin & Kaplan, 2003; Vollrath & Knight, 2001).

In this context, it may seem apparent the necessity of performing a thorough characterization of the genes used to produce the silk fibers and gluey solutions. However, the peculiarities found in the sequences of these genes represent a major difficulty for completing this task. As a matter of fact, the first complete silk gene, the heavy side chain fibroin of *Bombyx mori* (domesticated silkworm), was fully sequenced in 2000 (Zhou et al., 2000), only a few years before the complete

sequence of the *Bombyx mori* genome (Xia et al., 2004). It was not until 2007 that the first complete spider silk genes were reported, MaSp1 and MaSp2 from the black widow spider, *Latrodectus hesperus* (Ayoub, Garb, Tinghitella, Collin, & Hayashi, 2007). The more challenging endeavour of characterizing the full set of genes through whole genome sequencing (Wang & Jin, 2017)˙(Posnien et al., 2014) required some technical refinements to reconstruct complete spidroins and succeeded in 2017 (Babb et al., 2017) with *Trichonephila clavipes* (golden orb-weaver). This initial work was soon complemented with the complete set of spidroin genes determined from the whole genome of a second orb-weaving spider: *Araneus ventricosus* (Kono et al., 2019).

Knowing the complete set of spidroin genes found in a spider species in combination with proteomics studies and a thorough characterization of the materials offers the possibility to establish reliable correlations among gene sequence, protein sequence, processing, microstructure and mechanical behaviour of the silk materials. Following this rationale, we focus our efforts in this study on the usage of proteomics tools to identify the expression of spidroins in four silk glands: major ampullate, minor ampullate, flagelliform and tubuliform. As presented below, the expression of spidroins in *Trichonephila clavipes* silk glands shows a complex pattern in the different glands. The usage of the proteomics data allows establishing correlations between these expression patterns and the behaviour described for the silk fibers. In this regard, new details are found that not only expand our present knowledge on the basic aspects of silk fibers, but also constitute a major source of inspiration in the search for new developments and findings at the frontier between materials science and molecular biology.

EXPERIMENTAL SECTION

Trichonephila clavipes (Linnaeus, 1767) spiders were maintained in captivity and reared by Oscar Campos (Reptilmadrid S.L. Spain). *Drosophila* flies were used to feed the spiders up to the 2nd-3rd moults and, subsequently, the diet was changed to *Musca* flies. Silk glands were retrieved from two adult females after being anesthetized with ethyl acetate before dissection. The following fiber-producing glands were included in this study: major ampullate, minor ampullate, flagelliform and tubuliform. The schematic position of these glands in the opisthosoma of the spider is shown in Figure 1. The initial anatomical identification of the glands was confirmed by the proteomics analysis and also by the formation of spider silk guts from silk glands retrieved from a different set of *T. clavipes* spiders as described elsewhere (Jiang et al., 2014; Perez-Rigueiro, Ruiz, Luis Cenis, Elices, & Victor Guinea, 2020; Ruiz et al., 2019). No attempt was made to retrieve aciniform glands due to their small size and, additionally, to the difficulty of separating them from the pyriform glands.

Glands were frozen at -80 ºC immediately after being retrieved and stored under this condition until initiating the proteomics analysis protocol. Glands were washed with cold PBS and immersed in 120 µl of homogenization buffer (50 mM Tris-HCl, pH 6.8, 10 mM DTT, 4% (w/v) SDS). Glands were boiled for 5 min, incubated overnight at 4 ºC and centrifuged at 4 ºC and 13000 rpm for 10 min. The whole gland proteome of each gland was concentrated as described elsewhere (Bonzon-Kulichenko, Garcia-Marques, Trevisan-Herraz, & Vazquez, 2015). Briefly, samples were loaded into an SDS-PAGE gel (0.5 mm-thick, 4% stacking, and 10% resolving) and the electrophoresis process was stopped when the front entered 2 mm into the resolving gel. The unseparated protein bands were then visualized by Coomassie staining, excised and incubated with 50mM iodoacetamide to block reduced cysteine residues. Proteins were subjected to digestion at 37 ºC overnight with 500 µl of 25 µg/µl chymotrypsin (Promega) in 100 mM Tris-HCl, pH 7.8

containing 10 mM CaCl$_2$. The cleaved peptides were extracted by incubation for 2 h in 100 mM Tris-HCl, pH 7.8 on shaking, acidized with 1% (v/v) trifluoroacetic acid, desalted onto C18 cartridges (Oasis, Waters Corporation, Milford, MA, USA), and dried down.

The analysis of the samples proceeded through liquid chromatography-tandem mass spectrometry (LC-MS/MS) using an Easy nLC 1000 nano-HPLC coupled to a Q-Exactive Orbitrap mass spectrometer (Themo Scientific). Peptides were injected onto a C18 reverse-phase precolumn (Acclaim PepMap100, 75-μm i.d., 3-μm particle size, 2-cm length, Thermo Scientific), and a reverse-phase analytical column (Acclaim PepMap 100, 75-μm, i.d., 3-μm particle size, 50-cm length, Thermo Scientific) in buffer A [0.1% formic acid (v/v)] and eluted with a 60 min linear gradient of buffer B [90% acetonitrile, 0.1% formic acid (v/v)], at 200 nL/min. Mass spectra (MS) were acquired using full ion-scan mode over the 390-1200 m/z range and 70000 FT resolution. MS/MS was acquired with the Top15 data-dependent MS/MS scan method (TopN method). For peptide fragmentation, normalized collision energy was set to 27% and the isolation window for the parent ion mass was established at 2 Da.

For peptide identification, the MS/MS spectra were analysed with the SEQUEST HT algorithm integrated in Proteome Discoverer 2.1 (Thermo Scientific) using a database containing *Bos taurus* and Araneae protein sequences (UniProt 2016, October Release). For database searching, parameters were selected as follows: chymotrypsin digestion with two maximum missed cleavage sites, 20 ppm precursor ion tolerance and 0.02 Da fragment ion tolerance. Cysteine carbamidomethylation (+57.021 Da) was chosen as fixed modification and methionine oxidation

(+15.995 Da) as dynamic modification. The same collections of the MS/MS spectra were searched against an inverted (decoy) database built from the target database. Peptide identification from MS/MS data was performed using the probability ratio method (Martinez-Bartolome et al., 2008) and the false discovery rate (FDR) of peptide identification was calculated after competition between the target and decoy results using the refined method (Navarro & Vazquez, 2009). The proportion of proteins identified in the gland was calculated through spectral counting, assuming as positive events those with 1% FDR (Supporting Information Table 1 includes the list of assigned peptide-spectrum-match (PSM) from spidroins at 1% FDR). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2019) partner repository with the dataset identifier PXD023141.

Peptides identified as originating from spidroins were subsequently aligned specifically with the proteins released from the genome sequencing of *Trichonephila clavipes* spiders as found in Babb et al. (Babb et al., 2017) (See Supporting Information Table 2). The alignment was performed with the Emboss Matcher for the Pairwise Sequence Alignment (www.ebi.ac.uk/Tools/psa/emboss_matcher/) using the BLOSUM 62 matrix, with a gap open of 1 and a gap extend of 4. An identity of 85% or higher was established as the condition for assigning a peptide to a given spidroin protein.

RESULTS

The complete set of peptides assigned to spidroin proteins found in the four silk glands included in this study (major ampullate, minor ampullate, flagelliform and tubuliform) is presented in the Supporting Information Table 3 and summarized in Table 1. In correspondence with the specificity of chymotrypsin for hydrolysis of peptide bonds formed by the carboxyl groups of Tyr, Phe, Trp and Leu, 93% of the identified amino acid sequences contained Tyr residues, and the remaining 7% Leu residues on the C-terminus. A total of 104 peptides belonging to five spidroin type proteins (MaSp1, MaSp2, MiSp, Flag and TuSp) were sequenced by liquid chromatography coupled to tandem mass spectrometry, although some peptides were found to be present in more than one gland. The assignment of each peptide to a spidroin type based on the presence of its characteristic motifs of the sequence is also indicated in Table 1. Numbers in parentheses in Table 1 refer to those peptides of the corresponding spidroin type based on the presence of their characteristic motifs, but with an identity score lower than 85% when aligned with the sequences of the complete set of spidroins as found in Babb et al. (Babb et al., 2017). The sequence match spectrum identification details of the peptides with an identity score between 95-100% are shown in Supporting Information Figure 2.

From a total number of 28 genes coding for spidroin proteins (Babb et al., 2017), peptides with an identity score of 85% or higher were identified in the protein sequences of six of them: MaSp-c (Major ampullate gland silk spidroin 1), MaSp-g and MaSp-h (Major ampullate gland silk spidroin 2), Flag-a (Flagelliform gland spidroin), MiSp-d (minor ampullate gland silk spidroin) and TuSp (Tubuliform gland spidroin). Some of these peptides were also identified in three additional proteins: MaSp-b, MiSp-b and MiSp-c, although with lower identity scores compared with the other six proteins. The covering of the identified peptides on the set of nine proteins is illustrated in Figures 2-6 and the protein with the highest identity score for each peptide is indicated in

Supporting Information Table 3. In cases where the identity score for a peptide was lower than 85%, a new alignment was performed by using the whole set of UniProt KB sequences and this information is also included in Supporting Information Table 3.

Figure 2 illustrates the protein coverage attained by the peptides that were identified from Major ampullate gland 1- type spidroins. All peptides were found in the major ampullate gland but, additionally, some of them were also identified in the flagelliform, minor ampullate and tubuliform glands. The gland in which the peptide was found is indicated using the following color code: Red- major ampullate gland, Green- flagelliform gland, Yellow- minor ampullate gland and Violet- tubuliform gland. Peptides found in more than one gland are labelled correspondingly with a combination of these colours. The highest coverage is found in the protein MaSp-c (Figure 2a), where the identity score reaches a value of 100% for the peptide GLGSQGAGRGGQGAGAAAAAGGAGQGGY, while that of most of the other peptides is typically around a value of 95%. Figure 2b shows the same analysis on the spidroin MaSp-b. In this case, lower coverage and identity scores are found. Besides, no peptide characteristic of MaSp-b was identified that was not also found in MaSp-c with a higher identity score. Consequently, and based on this analysis, it can be concluded that, although the data on the expression of the MaSp-b are not conclusive, the expression of this protein cannot be completely discarded either.

The same analysis was performed with Major ampullate gland silk spidroin-2 proteins and the results are shown in Figure 3. Four different peptides were identified that had an identity score of 88% or higher, when aligned with two of the proteins of this spidroin type: MaSp-g (three peptides) and MaSp-h (one peptide). Three additional peptides were identified that had an identity score

lower than 85%, when aligned with these two spidroins. However, these peptides were found to show an identity of 100 % with other UniProt KB entries of Major ampullate gland 2 spidroins of the species *Trichonephila inaurata madagascariensis* (see Supporting Information Table 3). All peptides were found in the major ampullate gland, and two of them were also found in the minor ampullate gland.

The coverage of the peptides with the sequence motifs characteristic of flagelliform spidroin on the protein expressed by the gene *Flag-a* is presented in Figure 4. It was found that these peptides are exclusively expressed in the flagelliform gland. Two peptides (GPGGAGPGGY and GPGGAGGPY) show an identity score of 100%, while the values of the identity score for the other peptides is always higher than 85% with just one exception (GPGGAGGPGGEGPGGAGGPY). However, this peptide shows an identity score of 100% when aligned with a Flag spidroin from *Trichonephila clavipes* as sequenced in an independent work to that of Babb et al. (Babb et al., 2017) . When the same analysis was performed on the other Flag spidroin (Flag-b, a spidroin shown to be expressed in venom glands rather than silk glands (Babb et al., 2017)) identity scores were always below the value of 85%.

Figure 5 shows the coverage attained by the peptides of Minor ampullate gland silk spidroins. The large coverage found in MiSp-d spidroin is apparent from Figure 5a, including three peptides with an identity score of 100% (GAGAGAAAAAGAGAGGAGGY, GAGAGAGAAAAAGAGAGGAGGY and GRGAGAGAGAAAGAGAGAGGY). Some of the peptides are also compatible with the sequences of the MiSp-b and MiSp-c spidroins. However, in most cases the identity scores found in these two latter proteins is lower than those yielded by the MiSp-d protein and always below a value 95%.

As previously discussed in relation to MaSp-b, these results do not entirely dismiss the possible expression of MiSp-b and MiSp-c spidroins in the minor ampullate gland. Finally, it was found that a few peptides with low identity scores for the preferred set of spidroins used in this work, yielded values of 100% when aligned with Minor ampullate gland silk spidroins of different spider species when the whole UniProt KB database was used for the alignment analysis.

Lastly, Figure 6 shows the large coverage of the peptides found on the Tubuliform spidroin. In addition, the identity score in this case is very high with 12 peptides with an identity score of 100%. No peptides found in Tubuliform spidroin were detected in a gland different from the tubuliform gland.

The identification of the peptides was complemented with a quantitative analysis of the proteomics data that allows estimating the relative proportion of proteins from a given spidroin group found in each gland from the number of recorded PSMs (Supporting Information Table 1). Only peptides with a number of PSMs greater than one were used for this quantitative analysis. The results are summarized in Table 2.

DISCUSSION

Spidroin genes in the genome of T. clavipes

The complete sequencing of the *T. clavipes* genome has revealed that the spidroin protein group in this species comprises 28 genes. These genes can be assigned, in most cases, to the previously defined spidroin families depending on a few characteristic motifs of the sequence (Babb et al., 2017). For instance, eight genes were grouped in the family of major ampullate gland silk spidroins and four in that of minor ampullate gland silk spidroins (see Supporting Information Table 2).

Ultimately, one of the most pressing questions with regard to this genetic organization is its significance on the properties of the different silk fibers spun by a spider. Addressing this question required obtaining information on the translation of these genes, including the compartmentalization of expression in the different glands. Consequently, the proteomics analysis of the spidroins expressed in different silk glands is a basic tool to complement the genetic analysis. Following this rationale, the proteomics analysis performed in this work allowed the characterization of the spidroins expressed in four of the five glands used by mature female *T. clavipes* spiders to spin silk fibers: major ampullate, minor ampullate, flagelliform and tubuliform. It was not possible to include the aciniform gland in this study due to its small size and difficult manipulation.

One of the most immediate results obtained from the proteomics analysis of the spidroins expressed in the different glands of adult females is the relatively large proportion of spidroin genes, for which no translated protein could be identified. Even when considering that there are three additional glands (aciniform, aggregate and pyriform glands) to those included in this study, only products from nine out of the 28 genes assigned to the spidroin family could be identified. In addition, there are doubts on the expression of three of these genes, since they could not be identified with a proteotypic peptide, that was neither found in another representative of the family with higher coverage and/or identity score. Thus, from the eight major ampullate gland silk genes, only four of them were found to be compatible with the identified peptides. A similar situation is found with minor ampullate gland spidroins (one protein, MiSp-a, not expressed out of four genes) and in the flagelliform gland spidroins (Flag-a is expressed, while no product of Flag-b was detected, as expected because Flag-b is associated with venom glands). Besides, there seems to be

a certain tendency to favour the expression of the shortest proteins within a type that is observed empirically from the previous covering analyses.

The number of reported proteins is broadly similar to that found by Aparecido dos Santos et al. (Aparecido dos Santos-Pinto et al., 2019; Whaite et al., 2018) in which three major ampullate spidroins, one minor ampullate spidroin and one tubuliform spidroin were identified in *T. clavipes*. Similarly, Whaite et al. (Aparecido dos Santos-Pinto et al., 2019; Whaite et al., 2018) identified spidroins homologous to the proteins MaSp-c, MaSp-d, MaSp-g and MaSp-h, and to the proteins MiSp-a and MiSp-d (following the labelling by Babb et al. (Babb et al., 2017)), in the species *Trichonephila plumipes* and *Nephila pilipes*.

Additionally, one of the major contributions of this work is probably the finding that several spidroin types are shared among glands. In this regard, most spidroins were found to be expressed in more than one gland, and MiSp-d and MaSp-c products were identified in all four glands analysed. In contrast, the Flag-a and TuSp genes were found to be expressed exclusively in the flagelliform and tubuliform glands, respectively.

This complex expression pattern immediately leads to the necessity of quantifying the proportion in which a given spidroin is found in each gland. In this regard, the quantitative expression profiles presented in Table 2 show the existence of distinct patterns for different spidroins and glands. The quantitative proteomics findings are again broadly similar to the proteomics data as found in Aparecido dos Santos et al. (Aparecido dos Santos-Pinto et al., 2019; Whaite et al., 2018) and to the qPCR results reported in Babb et al. (Babb et al., 2017). Thus, the previous proteomics analysis identified a preferential expression of major ampullate spidroins in the major ampullate gland silk, and a reduced presence in the flagelliform silk. Minor ampullate gland silk was also found in a

significant amount in the major ampullate gland. The complementary results obtained by qPCR indicated that not all of the 28 spidroin genes could be shown by qPCR to be expressed in any silk gland at an appreciable level (>10-fold expression change compared to a control tissue). From the major ampullate, minor ampullate, flagelliform, and tubuliform silk glands, Babb et al. detected a total of 20 different spidroin genes with such an expression level, but it is expected that qPCR, as a nucleic acid amplification-based method, could detect the expression of more genes than through peptide identification. Furthermore, within every silk gland that was examined, multiple types of spidroin transcripts were detected, as was seen with peptide identification (Table 1). Other notable qPCR results are the higher expression of MaSp1 over MaSp2 in the major ampullate glands, predominance of MiSp in minor ampullate glands, and high abundance of Flag and MiSp in flagelliform glands, all of which are consistent with the quantitative protein estimates from spectral counting (Table 2). Considering specific spidroin genes and their relative expression levels, however, the qPCR and proteomics results differ from each other. For example, MaSp-g (a MaSp2) expression was supported by three peptides (Figure 3) but was not detected at a substantial level in the qPCR study (Babb et al., 2017). These discrepancies may be due to the different detection limits and sensitivities of both techniques (assaying proteins versus transcripts), biological variation among individual spiders, or a combination of these factors.

Correlation between protein expression and different silk fibers

The information available on the microstructure and mechanical properties of the materials produced from these glands, either as naturally spun fibers (Anton et al., 2017; Blackledge & Hayashi, 2006; Vollrath, 2000) or as spider silk guts (Jiang et al., 2014; Perez-Rigueiro et al., 2020; Ruiz et al., 2019) allows exploring the correlations between protein expression and fiber properties mentioned above. Spider silk guts are fibers synthetised directly from a spider gland

through a two-step process consisting of submerging the gland into a mild acidic solution and a subsequent mechanical stretching step. Spider silk guts were produced from the four glands included in this study, and it was found that the microstructure and mechanical properties of naturally spun silk and spider silk gut fibers concur in the cases in which this comparison was possible. Consequently, and in the absence of direct information on the behaviour of the naturally spun material (such as the microstructure of naturally spun tubuliform silk), the results obtained from the analysis of spider silk guts are considered to be representative of the natural material. At this point it may be worth relating the qualitative and quantitative features of the proteomics analysis with the basic characteristics of the fibers spun from each gland.

Major and minor ampullate gland silks

As indicated above, the proteomics analysis suggests that the major and minor ampullate gland silk fibers originate from the same set of genes, although in varying ratios. Both silks share the polyalanine motif $-A_n-$, which is responsible for the transition from a liquid solution to a solid fiber and acts as reinforcement to improve the mechanical behaviour of the material (Dicko, Knight, Kenney, & Vollrath, 2005). Both fibers also show high values of tensile strength and strain at breaking, that yields a value of work to fracture of 240 MJ/m^3 for the minor ampullate gland silk (miS) fibers of *T. inaurata* spiders (Guinea et al., 2012), and a record value of 500 MJ/m^3 for *Argiope aurantia* MAS fibers (Madurga et al., 2016). The microstructural analysis of both fibers yields a semicrystalline microstructure (Riekel, Craig, Burghammer, & Muller, 2001) with a crystalline fraction of 12-13% in miS (Guinea et al., 2012), and of 15% in MAS fibers (Riekel et al., 1999).

Larger differences are found between major and minor ampullate gland fibers when the influence of water in their tensile behaviour is considered. miS fibers show a negligible contraction in water (Work, 1977) and, in addition, this contraction does not correspond to the existence of a ground state for this material (Guinea et al., 2012). In contrast, MAS exhibits the property of supercontraction (Work, 1977), that implies the existence of a ground state to which the fiber can return independently of its previous loading history upon wetting. The presence of the same set of proteins in both glands supports the hypothesis that the appearance of supercontraction does not depend on the presence of singular motifs in the sequence, but on the proportion in which these motifs appear and lead to the formation of given microstructural details. The quantitative origin of supercontraction was first proposed in relation with its emergence in regenerated silkworm (*Bombyx mori*) silk fibers under given processing conditions (Perez-Rigueiro et al., 2019), in spite of the absence of this phenomenon in the natural silkworm silk fibers.

In turn, the existence or not of supercontraction might influence the distinct behaviour of silks spun by different spider species. In this regard, the comparison of the tensile properties of fibers spun by different Entelegynae spiders (Blackledge & Hayashi, 2006; Guinea et al., 2012; Work, 1977) shows the remarkable concurrence of the stress-strain curves of miS fibers, even among spiders that diverged over 100 Mya ago. In contrast, the presence of supercontraction in MAS fibers seems to underlie the essential property of this material that allows the adaptation of its tensile behaviour to the immediate needs of the spider (Madsen, Shao, & Vollrath, 1999), and would explain the large variability found in the mechanical performance of MAS fibers across the Entelegynae group (Madurga et al., 2016). In particular, the presence of MaSp2 spidroins in the *T. clavipes* major and minor ampullate glands (although only traces are to be found in the latter)

acquires special significance, since the characteristic motif of this family of spidroins, –GPG–, is found in the Deinopidae, Uloboridae and Araneoidea (to which *T. clavipes* belongs) but is lacking in RTA- clade (Blackledge et al., 2009)·(Bond et al., 2014)·(Fernandez, Hormiga, & Giribet, 2014)·(Fernandez et al., 2018). In this regard, the relationship found between proline content and the mechanical properties of MAS fibers (Liu, Sponner, Porter, & Vollrath, 2008)(Craig, Piorkowski, Nakagawa, Kasumovic, & Blamires, 2020), suggests that the ratio between MaSp1 and MaSp2 might play a role in determining the variability found among the MAS fibers spun by different Entelegynae species (Madurga et al., 2016). However, the details of the molecular basis for this variability remains unknown at present. Lastly, it may be argued that the distinct properties exhibited by major and minor ampullate gland silk fibers are not a consequence of differences in the spinning process. On the contrary, the basic instructions for the self-assembly of both types of fibers seem to be encoded in their sequences, as proven by the possibility of producing minor and major ampullate gland silk guts from both glands under mild processing conditions (Perez-Rigueiro et al., 2020).

Flagelliform silk

Flagelliform silk (Aparecido dos Santos-Pinto et al., 2018) also exhibits supercontraction and a ground state, as well as high values of tensile strength, strain at breaking and work to fracture (Guinea, Cerdeira, Plaza, Elices, & Perez-Rigueiro, 2010). As occurs with miS and MAS fibers, flagelliform silk fibers show a semicrystalline microstructure with a nanocrystalline phase embedded in an amorphous matrix (Perea et al., 2013). In contrast to MAS and miS, however, a less clear correlation can be established between the motifs found in the sequence and the appearance of the nanocrystalline phase. In particular, Flag spidroin lacks the polyalanine motif –A_n– which constitutes the core of the nanocrystals in MAS and miS (in this case, with the additional

contribution of the –GA– tandem repeats) fibers. Thus, it was found that the nanocrystals in flagelliform silk correspond to a polyproline II unit cell (Perea et al., 2013), that is compatible with the presence of the –GPG– motifs. In this context, the complex combination of proteins revealed by the proteomics analysis and the high proportion of MaSp and, especially, MiSp spidroins might be understood by the requirement of seeds (or nuclei) for the initial formation of the less compact polyproline II nanocrystals. Under this hypothesis, the characteristic nanocrystal-forming motifs present in MAS and miS would fulfill this seeding function. Another aspect in which this combination of proteins might be relevant is related with the self-assembly process of the proteins to yield the solid fiber during spinning. It has been found that flagelliform silk guts can be produced, but only under much harsher processing conditions than those required to obtain silk glands from the other three glands included in this work (Perez-Rigueiro et al., 2020).

Tubuliform silk

Finally, tubuliform silk is singular when compared with the silk fibers discussed previously in a number of respects. The biological function of this fiber, which constitutes the outer layer of the egg sac, implies that it is fully operative during a limited period of a female spider's lifetime (Blackledge & Hayashi, 2006). During this period, however, it appears as three pairs of plump, elongated glands on each side of the opisthosoma whose combined sizes are even larger than that of the major ampullate glands. Tubuliform silk shows tensile properties characteristic of high-performance fibers with a work to fracture that may reach a value of 140 MJ/m^3 (Ruiz et al., 2019), although this figure is only half of the value reached by the other three fibers discussed in this work. In parallel with the behaviour of minor ampullate gland silk, tubuliform fibers do not supercontract and the stress-strain curves of fibers spun by different species of the Entelegynae group appear remarkably similar. Tubuliform silk also has a semicrystalline microstructure, and

the nanocrystals present a unit cell similar to that previously found in the minor ampullate gland silk. This latter result is explained by the proteomics analysis, since it was found that the spidroin content of the gland is composed mainly of Tubuliform spidroin (TuSp), a spidroin that is specifically expressed in the tubuliform gland, but with a significant contribution of Minor ampullate gland spidroin (MiSp) and a smaller contribution of Major ampullate gland spidroin (MaSp). As indicated above when discussing flagelliform silk, no motif in the sequence of TuSp spidroin offers clear evidence on the nature of the nanocrystals found in tubuliform silk. Consequently, the presence of the MiSp – again it may be argued that this protein is the product of the MiSp-d gene– plays an essential role in the microstructural organization of tubuliform silk fibers. Intriguingly, the proteomics analysis casts doubt on the role played by the specific TuSp spidroin on the tubuliform silk fibers. It might be argued that the higher proportion of amino acids other than glycine and alanine, for instance serine, in TuSp spidroin would prevent the depletion of the content of these amino acids during a period in which the metabolic costs for the spider are specially high (Craig, Hsu, Kaplan, & Pierce, 1999), due to the large amount of material that must be spun for creating the egg sac. The usage of these amino acids different from glycine and alanine would imply, however, a certain reduction on the performance of the material in terms of its mechanical behaviour compared with the other fibers.

CONCLUSIONS

The combination of data obtained from the genomics and proteomics analyses of spidroin genes and silk glands with the detailed microstructural and mechanical characterization of spider silk fibers allows gaining a deeper insight in the design principles of these materials. In some respects,

this combined analysis supports previously accepted ideas while, in others, some new and even unexpected findings were made.

Thus, it is shown that nine spidroins are expressed in the major ampullate, minor ampullate, flagelliform and tubuliform glands in mature female *T. clavipes* from a total number of 28 spidroin genes described in the genome of this species. The number of expressed proteins could be even smaller, since all the identified peptides can be accounted for by considering the translation of just six genes. The expression of the other three genes (one from the MaSp1 and two from the MiSp spidroin types), however, cannot be completely excluded.

In addition, a wide range of different patterns of gene translation is found among the analysed glands. Thus, it is observed that the MiSp-d gene (perhaps with some contributions of the MiSp-b and MiSp-c genes) is expressed ubiquitously and makes a significant contribution to the spidroin content of the four glands. MaSp genes are expressed mainly in the major ampullate gland, although they can be also expressed in the minor ampullate silk and, to a lesser extent, in the flagelliform and tubuliform glands. The expression of the MaSp-c gene accounts for the results of the proteomics analysis, although the expression of MaSp-b cannot be discarded. In contrast, this analysis supports the simultaneous expression of both MaSp2 genes (MaSp-g and MaSp-h), and traces of these proteins are even found in the minor ampullate gland.

The presence of the same set of proteins in the major and minor ampullate silk glands indicates that the differences in the mechanical behaviour between both materials must originate from the distinct proportions of each spidroin expressed in each gland. As suggested by previous studies, these quantitative differences and, in particular, the ratio between MaSp1 and MaSp2 spidroins in spiders of the Entelegynae group might contribute to the variation in the mechanical properties of major ampullate gland silk among different representatives of this lineage.

The expression of TuSp and Flag spidroins, in contrast to MaSp and MiSp spidroins, appears to be confined to the tubuliform and to the flagelliform glands, respectively. In both cases, a large proportion of the spidroin content in the gland corresponds to MiSp spidroins that, in the case of the tubuliform silk, play a fundamental role in the organization of the nanocrystalline phase. It may be argued that MiSp might play a similar role in flagelliform silk and, perhaps in combination with MaSp, provide the seeds for the formation of the characteristic polyproline II nanocrystals in this material.

Needless to say, the full description of the complex system of silk producing proteins and glands in spiders and its origin and evolution will require more extended analyses. However, it is apparent that the combination of data retrieved from different fields such as genomics and proteomics, as well as the thorough mechanical characterization of the silk fibers offers a significantly deeper insight in the production and properties of this material. This enhanced understanding, in turn, must be an invaluable guide in the search for bioinspired artificial fibers based on natural silks.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA SHARING

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Adrianos, S. L., Teule, F., Hinman, M. B., Jones, J. A., Weber, W. S., Yarger, J. L., & Lewis, R.
    V. (2013). Nephila clavipes flagelliform silk-like GGX motifs contribute to extensibility and
    spacer motifs contribute to strength in synthetic spider silk fibers. *Biomacromolecules,*
    14(6), 1751-1760. doi:10.1021/bm400125w

Anton, A. M., Heidebrecht, A., Mahmood, N., Beiner, M., Scheibel, T., & Kremer, F. (2017).
    Foundation of the outstanding toughness in biomimetic and natural spider silk.
    Biomacromolecules, 18(12), 3954-3962. doi:10.1021/acs.biomac.7b00990

Aparecido dos Santos-Pinto, J. R., Arcuri, H. A., Esteves, F. G., Palma, M. S., & Lubec, G.
    (2018). Spider silk proteome provides insight into the structural characterization of nephila
    clavipes flagelliform spidroin. *Scientific Reports, 8*, 14674. doi:10.1038/s41598-018-33068-
    9

Aparecido dos Santos-Pinto, J. R., Arcuri, H. A., Priewalder, H., Salles, H. C., Palma, M. S., &
    Lubec, G. (2015). Structural model for the spider silk protein spidroin-1. *Journal of*
    Proteome Research, 14(9), 3859-3870. doi:10.1021/acs.jproteome.5b00243

Aparecido dos Santos-Pinto, J. R., Esteves, F. G., Sialana, F. J., Ferro, M., Smidak, R., Rares, L.
    C., . . . Luebec, G. (2019). A proteotranscriptomic study of silk-producing glands from the
    orb-weaving spiders. *Molecular Omics, 15*(4), 256-270. doi:10.1039/c9mo00087a

Askarieh, G., Hedhammar, M., Nordling, K., Saenz, A., Casals, C., Rising, A., . . . Knight, S. D. (2010). Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. *Nature, 465*, 236-239.

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). *Molecular Biology and Evolution, 30*(3), 589-601. doi:10.1093/molbev/mss254

Ayoub, N. A., Garb, J. E., Tinghitella, R. M., Collin, M. A., & Hayashi, C. Y. (2007). Blueprint for a high-performance biomaterial: Full-length spider dragline silk genes. *PLoS ONE, 2*(6), e514.

Babb, P. L., Lahens, N. F., Correa-Garhwal, S. M., Nicholson, D. N., Kim, E. J., Hogenesch, J. B., . . . Voight, B. F. (2017). The nephila clavipes genome highlights the diversity of spider silk genes and their complex expression. *Nature Genetics, 49*(6), 895-903. doi:10.1038/ng.3852

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 B-Biochemistry & Molecular Biology, 155*(4), 419-426. doi:10.1016/j.cbpb.2010.01.005

Blackledge, T. A., & Hayashi, C. Y. (2006). Silken toolkits: Biomechanics of silk fibers spun by the orb web spider argiope argentata (fabricius 1775). *J.Exp.Biol., 209*(13), 2452-2461. doi:10.1242/jeb.02275 ER

Blackledge, T. A., Scharff, N., Coddington, J. A., Szuts, T., Wenzel, J. W., Hayashi, C. Y., & Agnarsson, I. (2009). Reconstructing web evolution and spider diversification in the molecular era. *Proceedings of the National Academy of Sciences of the United States of America, 106*(13), 5229-5234. doi:10.1073/pnas.0901377106 ER

Bond, J. E., Garrison, N. L., Hamilton, C. A., Godwin, R. L., Hedin, M., & Agnarsson, I. (2014). Phylogenomics resolves a spider backbone phylogeny and rejects a prevailing paradigm for orb web evolution. *Current Biology, 24*(15), 1765-1771. doi:10.1016/j.cub.2014.06.034

Bonzon-Kulichenko, E., Garcia-Marques, F., Trevisan-Herraz, M., & Vazquez, J. (2015). Revisiting peptide identification by high-accuracy mass spectrometry: Problems associated with the use of narrow mass precursor windows. *Journal of Proteome Research, 14*(2), 700-710. doi:10.1021/pr5007284

Craig, C. L., Hsu, M., Kaplan, D., & Pierce, N. E. (1999). A comparison of the composition of silk proteins produced by spiders and insects. *International Journal of Biological Macromolecules, 24*(2-3), 109-118. doi:10.1016/S0141-8130(99)00006-9

Craig, H. C., Piorkowski, D., Nakagawa, S., Kasumovic, M. M., & Blamires, S. J. (2020). Meta-analysis reveals materiomic relationships in major ampullate silk across the spider phylogeny. *Journal of the Royal Society Interface, 17*(170), 20200471. doi:10.1098/rsif.2020.0471

Dicko, C., Knight, D., Kenney, J. M., & Vollrath, F. (2005). Conformational polymorphism, stability and aggregation in spider dragline silks proteins. *International Journal of Biological Macromolecules, 36*(4), 215-224. doi:10.1016/j.ijbiomac.2005.06.004 ER

Fernandez, R., Hormiga, G., & Giribet, G. (2014). Phylogenomic analysis of spiders reveals

    nonmonophyly of orb weavers. *Current Biology, 24*(15), 1772-1777.

    doi:10.1016/j.cub.2014.06.035

Fernandez, R., Kallal, R. J., Dimitrov, D., Ballesteros, J. A., Arnedo, M. A., Giribet, G., &

    Hormiga, G. (2018). Phylogenomics, diversification dynamics, and comparative

    transcriptomics across the spider tree of life. *Current Biology, 28*(9), 2190-2193.

    doi:10.1016/j.cub.2018.03.064

Garb, J. E., & Hayashi, C. Y. (2005). Modular evolution of egg case silk genes across orb-

    weaving spider superfamilies. *Proceedings of the National Academy of Sciences of the*

    United States of America, 102(32), 11379-11384. doi:10.1073/pnas.0502473102 ER

Gatesy, J., Hayashi, C., Motriuk, D., Woods, J., & Lewis, R. (2001). Extreme diversity,

    conservation, and convergence of spider silk fibroin sequences. *Science, 291*(5513), 2603-

    2605.

Guinea, G. V., Cerdeira, M., Plaza, G. R., Elices, M., & Perez-Rigueiro, J. (2010). Recovery in

    viscid line fibers. *Biomacromolecules, 11*(5), 1174-1179. doi:10.1021/bm901285c ER

Guinea, G. V., Elices, M., Plaza, G. R., Perea, G. B., Daza, R., Riekel, C., . . . Perez-Rigueiro, J.

    (2012). Minor ampullate silks from nephila and argiope spiders: Tensile properties and

    microstructural characterization. *Biomacromolecules, 13*(7) doi:10.1021/bm3004644

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*, 239-242.

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*(10), 1950-1959. doi:10.1093/molbev/msh204 ER

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*(2-3), 271-275.

Heidebrecht, A., Eisoldt, L., Diehl, J., Schmidt, A., Geffers, M., Lang, G., & Scheibel, T. (2015). Biomimetic fibers made of recombinant spidroins with the same toughness as natural spider silk. *Advanced Materials, 27*(13), 2189-2194. doi:10.1002/adma.201404234

Heim, M., Keerl, D., & Scheibel, T. (2009). Spider silk: From soluble protein to extraordinary fiber. *Angew Chem Int Ed Engl, 48*(20), 3584-96.

Hu, X. Y., Lawrence, B., Kohler, K., Falick, A. M., Moore, A. M. F., McMullen, E., . . . Vierra, C. (2005). Araneoid egg case silk: A fibroin with novel ensemble repeat units from the black widow spider, latrodectus hesperus. *Biochemistry, 44*(30), 10020-10027. doi:10.1021/bi050494i

Jiang, P., Mari-Buye, N., Madurga, R., Arroyo-Hernandez, M., Solanas, C., Ganan, A., . . . Perez-Rigueiro, J. (2014). Spider silk gut: Development and characterization of a novel strong spider silk fiber. *Scientific Reports, 4*, 7326. doi:10.1038/srep07326

Jin, H. J., & Kaplan, D. L. (2003). Mechanism of silk processing in insects and spiders. *Nature, 424*(6952), 1057-1061. doi:10.1038/nature01809 ER

Kono, N., Nakamura, H., Ohtoshi, R., Moran, D. A. P. E., Shinohara, A., Yoshida, Y., . . . Arakawa, K. (2019). Orb-weaving spider araneus ventricosus genome elucidates the spidroin gene catalogue. *Scientific Reports, 9*, 8380. doi:10.1038/s41598-019-44775-2

Liu, Y., Sponner, A., Porter, D., & Vollrath, F. (2008). Proline and processing of spider silks. *Biomacromolecules, 9*(1), 116-121. doi:10.1021/bm700877g ER

Madsen, B., Shao, Z. Z., & Vollrath, F. (1999). Variability in the mechanical properties of spider silks on three levels: Interspecific, intraspecific and intraindividual. *Int.J.Biol.Macromol., 24*(2-3), 301-306.

Madurga, R., Plaza, G. R., Blackledge, T. A., Guinea, G. V., Elices, M., & Perez-Rigueiro, J. (2016). Material properties of evolutionary diverse spider silks described by variation in a single structural parameter. *Scientific Reports, 6*, 18991. doi:10.1038/srep18991

Martinez-Bartolome, S., Navarro, P., Martin-Maroto, F., Lopez-Ferrer, D., Ramos-Fernandez, A., Villar, M., . . . Vazquez, J. (2008). Properties of average score distributions of SEQUEST. *Molecular & Cellular Proteomics, 7*(6), 1135-1145. doi:10.1074/mcp.M700239-MCP200

Navarro, P., & Vazquez, J. (2009). A refined method to calculate false discovery rates for peptide identification using decoy databases. *Journal of Proteome Research, 8*(4), 1792-1796. doi:10.1021/pr800362h

Perea, G. B., Riekel, C., Guinea, G. V., Madurga, R., Daza, R., Burghammer, M., . . . Perez-Rigueiro, J. (2013). Identification and dynamics of polyglycine II nanocrystals in argiope trifasciata flagelliform silk. *Scientific Reports, ,* 1-6. doi:10.1038/srep03061

Perez-Rigueiro, J., Madurga, R., Ganan-Calvo, A. M., Elices, M., Guinea, G., Tasei, Y., . . . Asakura, T. (2019). Emergence of supercontraction in regenerated silkworm (bombyx mori) silk fibers. *Scientific Reports, 9*, 2398. doi:10.1038/s41598-019-38712-6

Perez-Rigueiro, J., Ruiz, V., Luis Cenis, J., Elices, M., & Victor Guinea, G. (2020). Lessons from spider and silkworm silk guts. *Frontiers in Materials, 7*, 46. doi:10.3389/fmats.2020.00046

Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D. J., . . . Vizcaino, J. A. (2019). The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Research, 47*(D1), D442-D450. doi:10.1093/nar/gky1106

Posnien, N., Zeng, V., Schwager, E. E., Pechmann, M., Hilbrant, M., Keefe, J. D., . . . Extavour, C. G. (2014). A comprehensive reference transcriptome resource for the common house spider parasteatoda tepidariorum. *Plos One, 9*(8), e104885. doi:10.1371/journal.pone.0104885

Riekel, C., Branden, C., Craig, C., Ferrero, C., Heidelbach, F., & Muller, M. (1999). Aspects of X-ray diffraction on single spider fibers. *International Journal of Biological Macromolecules, 24*(2-3), 179-186.

Riekel, C., Craig, C. L., Burghammer, M., & Muller, M. (2001). Microstructural homogeneity of support silk spun by eriophora fuliginea (C.L. koch) determined by scanning X-ray microdiffraction. *Naturwissenschaften, 88*(2), 67-72.

Ruiz, V., Jiang, P., Mueller, C., Jorge, I., Vazquez, J., Ridruejo, A., . . . Perez-Rigueiro, J. (2019). Preparation and characterization of nephila clavipes tubuliform silk gut. *Soft Matter, 15*(14), 2960-2970. doi:10.1039/c9sm00212j

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 of the United States of America, 105*(52), 20781-20785. doi:10.1073/pnas.0809174106

Stellwagen, S. D., & Burns, M. (2021). Repeat variation resolves a complete aggregate silk sequence of bolas spider mastophora phrynosoma. *Integrative & Comparative Biology, ,* icab048. doi:https://doi.org/10.1093/icb/icab048

Swanson, B. O., Blackledge, T. A., Beltran, J., & Hayashi, C. Y. (2006). Variation in the material properties of spider dragline silk across species. *Applied Physics A-Materials Science & Processing, 82*(2), 213-218. doi:10.1007/s00339-005-3427-6 ER

Vollrath, F. (1999). Biology of spider silk. *International Journal of Biological Macromolecules, 24*(2-3), 81-88.

Vollrath, F. (2000). Strength and structure of spiders' silks. *Journal of Biotechnology, 74*(2), 67-83.

Vollrath, F., & Knight, D. P. (2001). Liquid crystalline spinning of spider silk. *Nature, 410*(6828), 541-548.

Wang, X., & Jin, Y. (2017). Predicted networks of protein-protein interactions in stegodyphus mimosarum by cross-species comparisons. *Bmc Genomics, 18*, 716. doi:10.1186/s12864-017-4085-8

Whaite, A. D., Wang, T., Macdnald, J., & Cummins, S. F. (2018). Major ampullate silk gland transcriptomes and fibre proteomes of the golden orb-weavers, nephila plumipes and nephila pilipes (araneae: Nephilidae). *Plos One, 13*(10), e0204243. doi:10.1371/journal.pone.0204243

Work, R. W. (1977). Dimensions, birefringences, and force-elongation behavior of major and minor ampullate silk fibers from orb-web-spinning spiders - effects of wetting on these properties. *Text.Res.J., 47*(10), 650-662.

Xia, Q. Y., Zhou, Z. Y., Lu, C., Cheng, D. J., Dai, F. Y., Li, B., . . . Yang, H. M. (2004). A draft sequence for the genome of the domesticated silkworm (bombyx mori). *Science, 306*(5703), 1937-1940. doi:10.1126/science.1102210 ER

Xu, M., & Lewis, R. V. (1990). Structure of a protein superfiber - spider dragline silk. *Proc.Natl.Acad.Sci.U.S.A., 87*(18), 7120-7124.

Zhou, C. Z., Confalonieri, F., Medina, N., Zivanovic, Y., Esnault, C., Yang, T., . . . Li, Z. G. (2000). Fine organization of bombyx mori fibroin heavy chain gene. *Nucleic Acids Research, 28*(12), 2413-2419. doi:10.1093/nar/28.12.2413

Tables

Gland / Assigned spidroin	major ampullate	minor ampullate	flagelliform	tubuliform
MaSp1	8	5	2	3
MaSp2	4 + (3)	2	0	0
MiSp	4	19 + (2)	10 + (2)	8 + (2)
Flag	0	0	14 + (1)	0
TuSp	0	0	0	15
Total number of peptides	**19**	**28**	**29**	**28**

Table 1. Summary of the peptides identified from the major ampullate, minor ampullate, flagelliform and tubuliform glands of *T. clavipes* spiders, and assignment to different spidroin types following the work by Babb et al. (Babb et al., 2017).

Gland / Spidroin	Major Ampullate	Minor Ampullate	Flagelliform	Tubuliform
MaSp1	84%	19%	7%	7%
MaSp2	6%	2%	0%	0%
MiSp	10%	79%	32%	18%
Flag	0%	0%	61%	0%
TuSp	0%	0%	0%	76%

Table 2. Quantitative analysis of the proteins expressed in each silk gland as determined from the proteomics analysis based on spectral counting.

Figure Legends

Figure 1. Location of the different silk glands in the opisthosoma of a female *Trichonephila clavipes* spider. (1) Major ampullate silk gland, (2) Minor ampullate silk gland, (3) Flagelliform silk gland, and (4) Tubuliform silk gland. The same illustration without labelling is provided as Supporting Information Figure 1.

Figure 2. Schematic peptide coverage of the Major ampullate gland silk type-1 spidroins (a) MaSp-c, and (b) MaSp-b. The sequences of the proteins follow the work by Babb et al. (Babb et al., 2017). Representative peptides were chosen for the alignment with the protein sequence, and the corresponding C-terminal amino acid residues from the chymotryptic cleavage are underlined. The whole set of identified peptides is presented in Supporting Information Table 3. The following color code is used to indicate the gland in which the peptide was found. Red- major ampullate gland; Green- flagelliform gland, Yellow- minor ampullate gland, and Violet- tubuliform gland. Peptides identified in more than one gland are labelled correspondingly.

Figure 3. Schematic peptide coverage of the Major ampullate gland silk type-2 spidroins (a) MaSp-g, and (b) MaSp-h. The sequences of the proteins follow the work by Babb et al. (Babb et al., 2017). Representative peptides were chosen for the alignment with the protein sequence, and the corresponding C-terminal amino acid residues from the chymotryptic cleavage are underlined. The whole set of identified peptides is presented in Supporting Information Table 3. The following color code is used to indicate the gland in which the peptide was found. Red- major ampullate gland, Yellow- minor ampullate gland. Peptides identified in more than one gland are labelled correspondingly.

Figure 4. Schematic peptide coverage of the Flagelliform gland silk spidroin-a. The sequence of the protein follows the work by Babb et al. (Babb et al., 2017). Representative peptides were chosen for the alignment with the protein sequence, and the corresponding C-terminal amino acid residues from the chymotryptic cleavage are underlined. The whole set of identified peptides is presented in Supporting Information Table 3. All peptides were found in the flagelliform gland exclusively.

Figure 5. Schematic peptide coverage of the minor ampullate gland silk spidroins (a) MiSp-d, (b) MiSp-b and (c) MiSp-c. The sequences of the proteins follow the work by Babb et al. (Babb et al., 2017). Representative peptides were chosen for the alignment with the protein sequence, and the corresponding C-terminal amino acid residues from the chymotryptic cleavage are underlined. The whole set of identified peptides is presented in Supporting Information Table 3. The following color code is used to indicate the gland in which the peptide was found. Red- major ampullate gland; Green- flagelliform gland, Yellow- minor ampullate gland, and Violet- tubuliform gland. Peptides identified in more than one gland are labelled correspondingly.

Figure 6. Schematic peptide coverage of the Tubuliform gland silk spidroin. The sequence of the protein follows the work by Babb et al. (Babb et al., 2017). Representative peptides were chosen for the alignment with the protein sequence, and the corresponding C-terminal amino acid residues from the chymotryptic cleavage are underlined. The whole set of identified peptides is presented in Supporting Information Table 3.   All peptides were found in the tubuliform gland exclusively.

Supporting Information Figure 1. Location of the different silk glands in the opisthosoma of a female *T. clavipes* spider.

Supporting Information Figure 2. Sequence match spectra of the peptides with an identity score between 95%-100%.

Supporting Information Table 1. List of assigned peptide-spectrum-match (PSM) from spidroins at 1% false discovery rate (FDR).

Supporting Information Table 2. List of spidroins identified from the genomics work by Babb et al. (Babb et al., 2017).

Supporting Information Table 3. Complete set of peptides assigned to different spidroin proteins found in the four silk glands included in this study.