Insect Biochemistry and Molecular Biology 135 (2021) 103594

Contents lists available at ScienceDirect

Insect Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/ibmb

Gene expression profiling reveals candidate genes for defining spider silk gland types

R. Crystal Chaw [a,1,*], Thomas H. Clarke III [b,2], Peter Arensburger [c], Nadia A. Ayoub [b], Cheryl Y. Hayashi [a,3]

[a] University of California, Riverside, Department of Evolution, Ecology, and Organismal Biology, 2710 Life Science Building, Riverside, CA, 92521, USA
[b] Washington and Lee University, Department of Biology, Howe Hall, Lexington, VA, 24450, USA
[c] Department of Biological Sciences, California State Polytechnic University, Pomona, CA, 91768, USA

ARTICLE INFO

Keywords:
Argiope
Spider silk
Silk glands
Transcriptome
Differential gene expression

ABSTRACT

Molecular studies of the secretory glands involved in spider silk production have revealed candidate genes for silk synthesis and a complicated history of spider silk gene evolution. However, differential gene expression profiles of the multiple silk gland types within an individual orb-web weaving spider are lacking. Each of these gland types produces a functionally distinct silk type. Comparison of gene expression among spider silk gland types would provide insight into the genes that define silk glands generally from non-silk gland tissues, and the genes that define silk glands from each other. Here, we perform 3' tag digital gene expression profiling of the seven silk gland types of the silver garden orb weaver *Argiope argentata*. Five of these gland types produce silks that are non-adhesive fibers, one silk includes both fibers and glue-like adhesives, and one silk is exclusively glue-like. We identify 1275 highly expressed, significantly upregulated, and tissue specific silk gland specific transcripts (SSTs). These SSTs include seven types of spider silk protein encoding genes known as spidroin genes. We find that the fiber-producing major ampullate and minor ampullate silk glands have more similar expression profiles than any other pair of glands. We also find that a subset of the SSTs is enriched for transmembrane transport and oxidoreductases, and that these transcripts highlight differences and similarities among the major ampullate, minor ampullate, and aggregate silk glands. Furthermore, we show that the wet glue-producing aggregate glands have the most unique SSTs, but still share some SSTs with fiber producing glands. Aciniform glands were the only gland type to share a majority of SSTs with other silk gland types, supporting previous hypotheses that duplication of aciniform glands and subsequent divergence of the duplicates gave rise to the multiple silk gland types within an individual spider.

1. Introduction

Silk production has evolved numerous times among arthropods (Sutherland et al., 2010). Spider silk has been widely studied due to the ubiquity of spider webs and the outstanding material properties of spider silk fibers (e.g., Lewis, 1992; Yarger et al., 2018). However, many of the genes underlying spider silk production remain elusive. Molecular research on spider silks has focused on the spider-specific family of structural proteins known as spidroins (a contraction of spider-fibroin; Hinman and Lewis, 1992). Spidroins are synthesized and stored inside abdominal silk glands. Ecribellate orb-web weaving spiders (Araneoidea) are prolific users of silk in the construction of iconic wagon-wheel shaped webs, and their silk glands can number into the hundreds.

The silk glands of araneoid spiders, including orb-web weavers such

Abbreviations: Aci, aciniform; Agg, aggregate; ALS, anterior lateral spinneret; Cep, cephalothorax; DGE, digital gene expression; Fla, flagelliform; Maj, major ampullate; Min, minor ampullate; PLS, posterior lateral spinneret; PMS, posterior median spinneret; Pyr, pyriform; SST, silk gland specific transcript; Tub, tubuliform.

* Corresponding author.
E-mail addresses: rcrystal@ucr.edu (R.C. Chaw), tclarke@jvci.org (T.H. Clarke), parensburger@cpp.edu (P. Arensburger), ayoubn@wlu.edu (N.A. Ayoub), chayashi@amnh.org (C.Y. Hayashi).
[1] Present addresses: Oregon Health and Science University, Department of Neurology. Portland, OR, 97239, USA.
[2] Present addresses: J. Craig Venter Institute, Rockville, MD, 20850, USA.
[3] Present addresses: Division of Invertebrate Zoology and Sackler Institute for Comparative Genomics, American Museum of Natural History, New York, NY, 10024, USA.

https://doi.org/10.1016/j.ibmb.2021.103594
Received 7 March 2021; Received in revised form 11 May 2021; Accepted 13 May 2021
Available online 27 May 2021

as *Argiope* (garden) spiders and cob-web weavers such as *Latrodectus* (widow) spiders, are abundant and can be grouped into seven different types (Fig. 1A): aciniform, aggregate, flagelliform, major ampullate, minor ampullate, pyriform, and tubuliform (also referred to as cylindrical). All spider silk glands have at least two components, a lumen where liquid silk is stored and a duct that connects the silk gland to an external spigot on the spider's spinneret (Tillinghast and Townley, 1993). However, each silk gland type is morphologically distinct and produces a signature proteinaceous silk type that has specialized mechanical properties suited to its function (e.g. Vollrath, 2000; Blackledge and Hayashi, 2006; Clarke et al., 2017; Chaw and Hayashi, 2018). Each of the silk gland types produces silks that are primarily composed of a unique spidroin or combination of spidroins, and the spidroins are named after their corresponding gland type.

The unique morphologies of spider silk glands are easily discernible under a dissecting microscope, and the location and morphology of the external silk gland spigots is another way that silk glands can be identified (e.g. Coddington, 1989; Moon, 2012). Araneoid spiders have three pairs of external spinnerets, the Anterior Lateral, Posterior Lateral, and Posterior Median spinnerets (ALS, PLS, and PMS, respectively). Aciniform glands are individually tiny but can occur by the hundreds, bunched together in grape-like clusters. Aciniform glands produce remarkably tough fibrils (Blackledge and Hayashi, 2006) that are used in prey wrapping and web decorations (e.g. Tillinghast and Townley, 1993). Aciniform spigots are found on the PMS and the PLS.

In contrast to aciniform glands, aggregate glands are very large and few in number. The four (two pairs) of aggregate glands are sea anemone-shaped, with multi-lobed lumens and wide ducts (e.g. Moon, 2018). Instead of fibers, aggregate glands produce viscous, glue-like silk used in prey capture (e.g. Townley and Tillinghast, 2013). Aggregate silk is used in combination with flagelliform (capture spiral) silk. The ducts of aggregate silk glands connect to relatively large spigots on the PLS, and form a triad with the spigot of the flagelliform silk gland.

Flagelliform silk is highly extensible (Blackledge and Hayashi, 2006) and is the fibrous silk along which aggregate silk is dotted for ensnaring prey in orb-webs (e.g. Kovoor, 1987; Tillinghast and Townley, 1993). A spider has only one pair of flagelliform silk glands and the morphology varies from species to species (Kovoor, 1987). In the silver garden spider, *Argiope argentata* (Fabricius, 1775), the flagelliform glands have a sinuous tail region that leads to cylindrical storage lumen and a zig-zagging duct. Flagelliform silk spigots are found on the PLS in a triad with the aggregate silk gland spigots.

The major ampullate gland also occurs in only one pair and features a long secretory tail at one end of a curved, bulbous lumen, and at the other end has a zig-zag duct that doubles back on itself as it approaches the spigot (e.g. Andersson et al., 2013). Major ampullate silk has high strength and toughness (e.g. Blackledge and Hayashi, 2006), and is the main component of draglines, silken lines used to prevent catastrophic falls, and the frame and radii of the orb-web. Minor ampullate glands are diminutive versions of the major ampullate glands. Minor ampullate silk fibers are used as the temporary spiral during orb-web construction and are thought to compose "bridging" lines that spiders use to travel long distances (e.g. Tillinghast and Townley, 1993). Minor ampullate silk fibers are generally less strong but more extensible than major ampullate silk fibers (Blackledge and Hayashi, 2006; Colgin and Lewis, 1998). Major ampullate silk gland spigots are found on the ALS, whereas minor ampullate silk gland spigots are found on the PMS.

Whenever any silk fiber, such as a dragline, web frame, or temporary spiral, must connect to another silk fiber or a substrate, that connection is done with pyriform silk. Pyriform silk is a mix of glue and fibrils that cements other silk types to various substrates (Kovoor and Zylberberg, 1980; Wirth et al., 2019). Pyriform glands are pear-shaped, even smaller than aciniform glands and can also number into the hundreds (Kovoor and Zylberberg, 1980). Pyriform silk gland spigots are found only on the ALS.

The final gland type in an araneoid spider is present only in mature females: tubuliform glands that produce the thick, often pigmented, fluffy fibers used to wrap egg clutches (Blackledge and Hayashi, 2006; Chaw et al., 2016). Tubuliform glands resemble undulating spaghetti noodles that grow to a relatively large size in anticipation of egg laying. Tubuliform silk gland spigots are found on the PLS and the PMS.

The differentiated silk glands within an individual orb-web weaving spider represent two levels of tissue specialization. First, is the specialization of silk glands, regardless of type, for the mass production and storage of spidroins and other silk proteins. Second, is the subspecialization of silk glands into differentiated types that are morphologically and functionally distinct from each other. Gene expression profiling of silk glands with RNA-seq has provided clues about how the glands have evolved to produce and process silk proteins (e.g., Lane et al., 2013; Clarke et al., 2017). These transcriptomic studies along with genomic studies have resulted in candidates for genes involved in silk gland function, and have revealed a complicated evolutionary history for silk gland specific genes (e.g. Babb et al., 2017; Chaw et al., 2018).

Gene expression profiling of spider silk glands has also informed an understanding of what makes each silk gland type unique. Silks from a cob- or orb-web weaving spider can be exclusively made of fibers, glue, or a combination of a fiber and glue. In cob-web weaving spiders, the expression profiles of fiber-forming glands, (aciniform, flagelliform, major ampullate, minor ampullate, pyriform, and tubuliform) were found to be divergent from the exclusively glue-forming aggregate

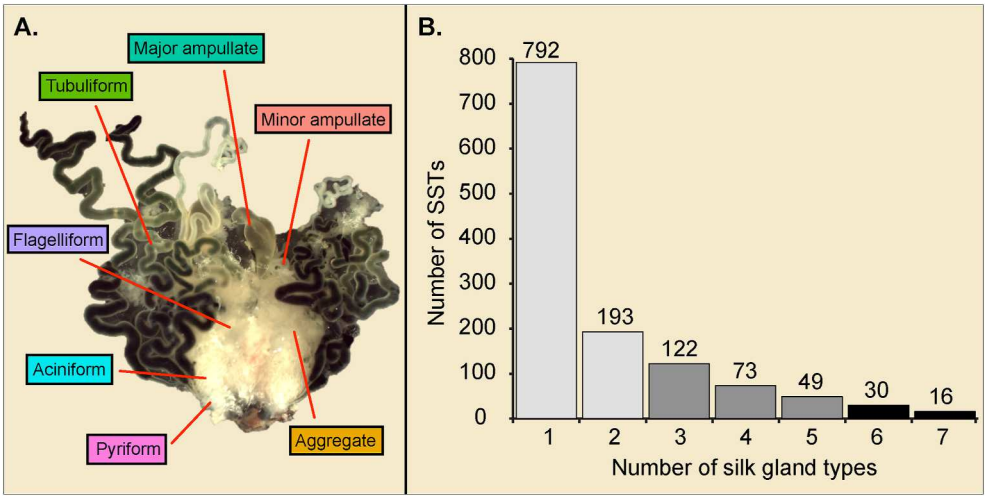

Fig. 1. A) All silk glands dissected from a sexually mature female *Argiope argentata.* Posterior external spinnerets are attached (bottom of glands). Approximate locations of each silk gland type are indicated. Colors correspond to Figs. 2 and 3. B) Number of silk gland specific transcripts (SSTs) that are present in 1–7 gland types. Light gray: SSTs found in one or two gland types, narrow; Gray: SSTs found in 3–5 gland types, moderate; Dark gray: SSTs found in 6–7 gland types, broad. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

glands (Clarke et al., 2017). However, Clarke et al. (2017) pooled the gene expression data from all silk gland types for a single differential expression analysis with non-silk gland tissue, and therefore did not compare the differentially expressed profiles in the individual silk gland types. Other transcriptome studies also pooled silk gland types or investigated only a subset of silk glands. In general, these studies focused on how silk glands differ from non-silk gland tissues, or how silk gland types evolved across species (e.g. Lane et al., 2013; Clarke et al., 2014; Chaw et al., 2016; Whaite et al., 2018; Correa-Garhwal et al., 2018). Direct comparison of gene expression profiles among all silk gland types within a species (i.e., reflecting differential expression of genes within a common genome) is therefore lacking.

In this study we profiled the gene expression of all seven silk gland types in the silver garden orb-weaver *A. argentata*. We hypothesized that the silk gland expression profiles would reflect the differentiation of silk glands from non-silk gland tissues and silk gland types from each other. For example, the major ampullate silk glands and minor ampullate silk glands are more similar to each other in shape and in the functions of the silks they produce than any other gland type. Given our hypothesis, the major ampullate and minor ampullate silk glands would have more similar expression profiles to each other than any other gland type. By contrast, the expression profile of the glue-forming aggregate silk glands would be the most divergent because the aggregate glands are the only gland type that exclusively produces glues. To investigate our hypotheses, we assembled a transcriptome for *A. argentata*, and performed 3' digital gene expression (DGE) tag profiling to obtain transcript counts for differential gene expression analyses among all seven silk gland types. We identify a subset of highly expressed and highly differentially expressed silk gland specific transcripts that include spidroin genes and non-spidroin genes with essential functions in silk production such as protein transport molecules and peptidases (e.g. Clarke et al., 2014; Clarke et al., 2017). In addition to providing candidates for further silk studies, the expression patterns of these genes among silk gland types support our hypothesis and other, existing hypotheses about the evolution of spider silk glands.

2. Methods and materials

2.1. Tissue collection, RNA extraction, library construction

Mature female *Argiope argentata* spiders were collected in San Diego County, CA, USA. Spiders were starved for at least two days, anesthetized using CO_2 gas, and sacrificed by separating the abdomen from the cephalothorax. From some spiders, the seven different silk gland types were collected and stored separately: aciniform, aggregate, flagelliform, major ampullate, minor ampullate, pyriform, and tubuliform. Silk gland types were identified by their distinctive morphologies (described above), confirmed by tracing ducts to spinnerets. From other spiders, the total set of silk glands was collected. Cephalothoraxes (head-body, non-silk gland control) were also collected. Each tissue type was placed into individual tubes, snap frozen in liquid N_2, and stored at −80 °C.

RNA extraction was described in Chaw et al. (2015). In total, seven individual spiders were used in this study. All tissues for the first replicate came from two spiders (ID: 2, 15), and the tissues for the second replicate came from two different spiders (ID: 10, 18). The only exception was the second replicate RNA extraction of the pyriform gland. This replicate was from a different set of two individual spiders (ID: 16, 17). The total set of silk glands was used from another (the seventh) individual spider (Chaw et al., 2015).

Construction of cDNA libraries was described in Chaw et al. (2015). Briefly, extracted and purified total RNA from each individual silk gland type and cephalothorax tissue was processed into cDNA using the Ovation 3′-DGE System (Digital Gene Expression; NuGen, San Carlos, CA, USA). The resulting libraries were purified and then Illumina compatible adaptors were attached with the Encore NGS Multiplex System I (NuGen). Silk gland libraries were named according to tissue type (aciniform, aggregate, cephalothorax, flagelliform, major ampullate, minor ampullate, pyriform, and tubuliform were abbreviated Aci, Agg, Cep, Fla, Maj, Min, Pyr, and Tub, respectively) and replicate (replicate 1 and 2, R1 and R2, respectively). Total RNA from the cephalothorax, pyriform, and flagelliform tissue replicate with the higher RNA yield (Cep_R2, Fla_R1, Pyr_R1) and two tubuliform silk gland replicates (Tub_R1 and _R2), were made into five RNA-seq libraries with the TruSeq RNA Sample Prep Kit v2 (Illumina) by the Johns Hopkins University School of Medicine Deep Sequencing and Microarray Core. Total RNA extracted from the total set of silk glands was made into an RNA-seq library using the TruSeq RNA Sample Prep Kit v1 (Illumina, San Diego, CA, USA; Supplementary Table S1).

2.2. Sequencing

Library sequencing was described in Chaw et al. (2015). Libraries were sequenced on a HiSeq 2500 platform (Illumina) at the University of California, Riverside Institute for Integrative Genome Biology. The RNA-seq libraries (total silk, Cep_R2, Fla_R1, Pyr_R1, and Tub_R1 and R2) were paired-end sequenced for 100 cycles. The 3′ DGE libraries (Aci, Agg, Fla, Cep, Maj, Min, Pyr, Tub, with R1 and R2 for each) were single-end sequenced for 50 cycles. The biological replicates of the 3′DGE libraries were run in separate lanes, and each library was run twice, generating technical replicates (Supplementary Table S1). All sequencing reads are available at the Sequence Read Archive (BioProject accession PRJNA322068).

2.3. Transcriptome de novo assembly

A reference transcriptome was assembled with six RNA-seq paired-end read libraries (Supplementary Table S1; total silk Cep_R2, Fla_R1, Pyr_R1 and Tub_R1 and R2) as in Clarke et al. (2014) with minor modifications. The five tissue-specific libraries were added to the total silk gland libraries to provide sampling from a non-silk gland control tissue (Cep_R2) and representative samples from the individual spiders used for the 3' DGE libraries. The flagelliform and tubuliform gland libraries (Fla_R1, Tub_R1 and R2) had relatively high RNA-yields. We included the pyriform gland library (Pyr_R1) because the pyriform glands are the smallest glands in size and the additional sampling from the library would decrease bias against transcripts specific to pyriform glands. Prior to assembly, FASTQ results from library sequencing were filtered using Trimmomatic and SortMeRNA to remove adaptors, sequences with low quality scores, and ribosomal RNA sequences (Bolger et al., 2014; Kopylova et al., 2012).

Tissue specific assemblies were generated from 16 tissue-specific single-end read libraries (7 silk gland types plus Cep and their technical replicates; Supplementary Table S1) using the Trinity v. r2012-06-08 de novo transcriptome assembly program (Grabherr et al., 2011; Haas et al., 2013). To obtain a comprehensive reference assembly, tissue-specific assemblies were combined with CAP3 using default settings (Huang and Madan, 1999). The combined assembly was further improved by incorporating all *A. argentata* transcript sequences from GenBank (retrieved December 2018). Overlapping and identical entries (100% identity) from GenBank were made into consensus sequences using Sequencher 4.2 (Gene Codes, Ann Arbor MI). Further, we sequenced a full-length *A. argentata* minor ampullate spidroin (MiSp) from a previously constructed BAC library (GenBank accession MT977119). Library construction was described in Chaw et al. (2017 and 2018). Briefly, BAC transformations were performed using DNA from purified, intact nuclei extracted from a single virgin *A. argentata* spider. Colonies were grown on agar plates and screened with standard PCR conditions (Luo and Wing, 2003; Sambrook et al., 1989; forward primer AGTTGGACGAGGAATTACGTATG, reverse primer CAACATAACCAATGGAGGAATTG). DNA was extracted from colonies that were positive for MiSp. Equimolar pools were created from the extracted

DNAs and the pools were sequenced with a Pacific Biosciences RSII + single molecule sequencer with P6–C4 chemistry. Short reads (<1000 bp) were filtered from the raw sequence data. Filtered reads were de novo assembled with CANU, using the self-correction module and default settings (Saski et al., 2015; Koren et al., 2017). Manual editing to form a consensus sequence was performed in Consed (Gordon et al., 1998). The unique GenBank entries, resulting consensus sequences from overlapping GenBank entries, and full-length MiSp were combined with the assembly using CAP3.

The reference assembly was further curated by the following procedures. Proteins were predicted based on longest open reading frame (ORF) and frame of best hit from a BLASTx search (default settings, e-value <1e-5) to NCBI's non-redundant protein (nr) database (downloaded October 2015) using a custom written Perl script. We kept predicted proteins that were longer than 30 amino acids and removed redundant protein encoding transcripts by identifying protein sequences that formed near-identical clusters (98% identity over 98% of their length) using the BLASTClust program (ftp://ftp.ncbi.nih.gov/blast/documents/blastclust.html). For each cluster, the transcript with either the longest amino acid sequence or, in cases with sequences of identical length, the first transcript of longest length was chosen.

Chimeric transcripts and potential contaminants such as bacterial sequences were identified with a custom Python script and removed from the assembly (Clarke et al., 2014). Assembly completeness was assessed with a tBLASTn search using 2274 Benchmarking Universal Single Copy Orthologs (BUSCO v1.2) from the red deer tick *Ixodes scapularis* (Simão et al., 2015). This Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the accession GIWY00000000. The version described in this paper is the first version, GIWY01000000.

2.4. Read mapping and identification of silk gland specific transcripts (SSTs)

The filtered 3′ DGE library reads were mapped to the reference assembly using Bowtie v. 1.1.1 with parameters for finding the alignment with the fewest mismatches for each read (–best; Langmead et al., 2009). Read counts from technical replicates (replicate sequencing lanes originating from the same RNA extraction) were plotted against each other to verify correlation. As expected, technical replicates displayed strong correlation ($R > 0.99$), and raw read counts from technical replicates were summed. Summed read counts from the 3′ DGE libraries were normalized by counts per million (CPM); 3′ DGE reads are anchored to the 3' end of transcripts and generate a single read per transcript, thereby avoiding potential overestimates of transcript abundance from multiple reads mapping to the repetitive region of a single silk protein (Chaw et al., 2016; Supplementary Table S2).

Normalized counts were used in a pairwise comparison of each silk gland tissue type versus non-silk gland (cephalothorax) tissue in differential gene expression analyses with the R package DESeq v1.38.0 (Anders and Huber, 2010). We chose DESeq v1.38.0 instead of DESeq2 because DESeq2 considered spidroin gene sequences to be outliers due to their unusually high counts in a small number of replicates. With outlier filters turned off, DESeq2 still did not consider spidroin sequences significantly differentially expressed. This may be because spidroin sequence expression can vary depending on whether silk is being actively synthesized, resulting in higher dispersion of counts between replicates (e.g., Casem et al., 2010). DESeq v1.38.0 is slightly more stable when there are no outliers in the data but there is high dispersion of data (Lin and Pang, 2019). DESeq results from the pairwise comparisons between each silk gland type and cephalothorax were concatenated and silk gland type specificity was assessed with the Tau factor for tissue specificity (Yanai et al., 2005). The Tau factor ranges from 0 to 1, with 0 indicating genes that have identical expression levels in all tissues (e.g., "housekeeping" genes), and 1 indicating genes that are completely tissue specific (expression restricted to a single gland type). We defined SSTs for each silk gland type as transcripts that were highly expressed (CPM > 5), significantly upregulated in that gland type relative to cephalothorax (DESeq results: positive log 2 fold change and p-adjusted value < 0.05), and tissue specific (Tau value > 0.85, across 8 tissues: the 7 silk gland types and the non-silk gland control). The p-adjusted value was given by DESeq, and it was the p-value corrected by the default DESeq multiple test correction of benjamini-hochberg.

2.5. SST analyses

SSTs were split into three groups according to the number of gland types in which they passed all of our cutoffs for SSTs: 1–2 (narrow), 3–5 (moderate), 6–7 (broad). Relationships between the SSTs that were in 1–2 gland types were visualized using Cytoscape v. 3.5.1 (Shannon et al., 2003). SSTs in multiple gland types, here considered all SSTs from 2 to 7 gland types, were grouped with hierarchical clustering using the default R function and the Euclidean distances between the different SSTs from the proportion of the total expression mean CPM of the transcripts in each of the seven silk gland types. The hierarchical clusters were split into the seven largest clusters (number of silk gland types assayed). The clusters were further assigned to an individual gland type if the summed total expression of all the SSTs in the cluster was >50% in that gland type. All SSTs were searched for spidroins based on top BLASTX match (e-value <1e-5).

2.6. GO analyses

Gene ontology (GO) Slim term analyses were performed as in Clarke et al. (2014). Briefly, GO terms were assigned to all transcripts according to the best UniProt hits by E-score via a custom Perl script (Young et al., 2010; Clarke et al., 2014). GO Slim terms were then obtained for those transcripts with GO terms, using the program GO Slim viewer (McCarthy et al., 2006). GO and GO Slim terms significantly enriched in subsets of transcripts compared to the entire set were identified using the GOSeq R package with the Wallenius and the HyperGeometric tests (Young et al., 2012).

3. Results and discussion

3.1. Silk gland specific transcripts (SSTs) have narrow or broad distribution among silk gland types

Previous work has compared the gene expression profiles of silk gland types across spider species (e.g., Babb et al., 2017; Clarke et al., 2014; Clarke et al., 2017). We sought to compare gene expression profiles of every silk gland type within a single species. Thus, we constructed a high quality *A. argentata* reference transcriptome. We assembled over 100 million 100 base pair (bp) RNA-seq reads and over 410 million 50 bp DGE reads into 120,571 contigs. We then added complete silk sequences from NCBI nr and BAC sequencing to our assembly, and curated the contigs to avoid redundancy (see Methods). Our final reference assembly had 115,487 contigs (N50 = 1277 bp). tBLASTn analyses showed that our reference assembly was 94.5% complete using arthropod Benchmarking Universal Single Copy Orthologs (BUSCO), with the genome of the black deer tick, *Ixodes scapularis*, as a baseline (Simão et al., 2015). We mapped >317 million 3' DGE tag reads to the reference assembly (Supplementary Tables S2–S4). Comparison of normalized read counts (CPM) in each of the seven silk gland types versus the non-silk gland control (cephalothorax) resulted in an average of ~3000 significantly differentially expressed (DE) transcripts in each gland type. We used a series of cutoffs to further narrow the DE transcripts down to 165–501 highly expressed and significantly differentially expressed silk gland specific transcripts in each gland type (Table 1). Collating the SSTs and removing duplicates resulted in 1275 total unique SSTs.

Of the 1275 SSTs, 581 had BLASTx matches. These included matches

Table 1
Total number of Silk gland Specific Transcripts (SSTs) by gland type after successive cutoffs were applied. Pairwise DESeq results were winnowed according to significance (padj<.05), upregulation (positive log2fold change), and high expression (CPM>5). Remaining transcripts were then compared for tissue specificity across all 7 silk gland types and non-silk gland control tissue (Tau value > 0.85).

Gland type	padj <.05	positive log2fold change	CPM >5	Tau value > 0.85
Aci	686	225	178	165
Agg	3429	1307	877	501
Fla	6131	1737	1254	385
Maj	3216	999	644	321
Min	4048	1252	894	363
Pyr	1322	586	362	224
Tub	3724	1398	880	414

to spidroin encoding genes. The remaining 694 SSTs with no BLASTx match, and those with matches to hypothetical, predicted, or uncharacterized proteins, are candidates for novel transcripts involved in silk gland functions including the production of specialized, high performance silks (Supplementary Table S5). We anticipated that some SSTs would be in all gland types, broadly defining silk glands from non-silk gland tissues. We also expected that some SSTs would only be present in one or two gland types, these SSTs would be candidates for specializing gland types from other gland types. Indeed, SST presence was variable among gland types. We were able to establish three categories of SSTs: "narrow" SSTs were in 1–2 gland types (985), "moderate" SSTs were in 3–5 gland types (244), and "broad" SSTs were in 6–7 gland types (46; Fig. 1B). The majority of SSTs were in the narrow category, which is consistent with how we identified SSTs: transcripts that are highly expressed, significantly differentially expressed, and tissue specific among silk glands.

SSTs in the broad category are expressed in the largest number of gland types, and are therefore likely candidates for genes that define silk glands from non-silk gland tissues. Four of the SSTs in the broad category had top BLASTX matches to spidroin genes, which we will discuss below. Of the remaining forty-two, twenty-two had BLASTX matches to hypothetical or uncharacterized proteins, or had no BLASTX match. These transcripts may be unannotated in the NCBI database, and may contain novel transcripts that differentiate silk glands from non-silk gland tissues. As expected, the other transcripts included BLASTX matches to genes thought to be important to silk gland function such as proteins important to secretory vesicle maintenance, protein transport, and peptidase activity and inhibition (Supplementary Table S5; Clarke et al., 2017). Secretory vesicle maintenance, protein transport, peptidases, and peptidase inhibitors are also a part of normal cellular functions, and future studies could investigate whether the transcripts present in the broad SSTs subset represent specialized paralogs that are used only in silk glands.

3.2. Spidroins have variable expression levels across gland types

Spidroin genes are a defining feature of spider silk glands. We expected that spidroins would have expression across multiple gland types as previously shown in other species (e.g., Garb et al., 2010; Babb et al., 2017; Clarke et al., 2017). Thirty-four transcripts with top BLASTX matches to spidroin genes were among the 1275 SSTs that we identified, and the BLASTX matches included all seven spidroin gene types (Supplementary Table S5).

Five out of seven spidroin gene types were found in multiple gland types. Major ampullate, minor ampullate, aciniform, and aggregate spidroin gene types were found in our narrow, moderate, and broad SST categories. Major ampullate and minor ampullate spidroin genes were shared among 2–5 gland types (narrow and moderate SST categories). Aciniform and aggregate spidroin genes were found shared among 2–6 gland types (narrow, moderate, and broad SST categories). Pyriform spidroin transcripts were shared among 5 gland types (moderate SST category). No spidroins were found across all 7 gland types (Supplementary Table S5).

The pyriform transcript was expressed in five gland types: aciniform, aggregate, minor ampullate, pyriform, and tubuliform glands. Pyriform silk is a composite of a fiber and a glue that forms a cement for anchoring other silk types to substrates (Kovoor and Zylberberg, 1980; Wirth et al., 2019). One explanation for the lack of pyriform gland exclusive transcripts in our analysis is that pyriform glands are relatively small and, in order to achieve enough RNA yield, the replicates for our pyriform glands differed from the replicates of all other gland types (see methods for details). This difference may be related to our lack of a gland-exclusive transcript. In addition, *A. argentata* pyriform spidroin is currently thought to be a single-copy gene (Collin et al., 2018) and thus may not have evolved a transcript that is limited in expression to a single gland type.

For all but one transcript, spidroin genes in the moderate and broad categories had the highest CPM in the gland type known to extrude the corresponding silk type. On average, spidroin transcripts for a given silk type were 20% of total SST expression in the corresponding gland type. By contrast, these spidroin transcripts averaged only 6% of total SST expression in the gland type with the next highest expression. For example, transcripts with top BLASTx matches to major ampullate spidroin-1 were 10% of total SST expression in the major ampullate gland and <1% of total SST expression in the pyriform gland (Supplementary Table S6).

The exception was Contig1075, which had a top BLASTx match to the major ampullate spidroin-2 gene *MaSp2*. Contig1075 did not have the highest CPM in major ampullate glands (25,606 CPM; 24% of total SST expression). Instead, Contig1075 had a higher CPM in pyriform glands (27,336; 29% of total SST expression). In pyriform silk glands, expression of Contig1075 was even higher than expression of the pyriform spidroin gene, which accounted for 11% of total SST expression (Supplementary Table S6). This high expression of Contig1075 in pyriform silk glands suggests that major ampullate spidroin-2 is incorporated in orb-web weaving pyriform silks. In dragline silk, major ampullate spidroin-2 is associated with extensibility (e.g., Brooks et al., 2008). Hence, the incorporation of major ampullate spidroin-2 into pyriform silk may provide the attachment disks with flexibility that would prevent breakage (Wolff et al., 2015; Wirth et al., 2019).

Flagelliform and tubuliform spidroin transcripts were only in the narrow category, and were found in only flagelliform silk glands and tubuliform silk glands, respectively. This exclusive expression pattern may exist for a number of reasons. Flagelliform silk is an extensible silk used in prey capture (Vollrath and Edmonds, 1989; Hayashi and Lewis, 2001). An analysis of the genome of the golden orb-web weaver *Trichonephila clavipes* (formerly *Nephila clavipes;* Kuntner et al., 2019) revealed two flagelliform spidroin genes. One was highly expressed in flagelliform silk glands, and the other was highly expressed in venom glands, which are found in the cephalothorax of spiders (Babb et al., 2017). In *A. argentata*, however, we found extremely few flagelliform transcripts in the cephalothorax (0.4 CPM; Supplementary Table S5). For comparison, the expression level in flagelliform silk glands was 11, 830 CPM (Supplementary Table S5). The exclusive expression of flagelliform spidroin transcripts in *A. argentata* reflects a difference in transcript specialization from *T. clavipes*.

Tubuliform silk is used in egg-case wrapping (Tillinghast and Townley, 1993). In previous work with cob-web weaving spiders, tubuliform spidroin transcripts were detected in major ampullate glands, but tubuliform spidroin proteins were not found in major ampullate glands or silks (Lane et al., 2013; Chaw et al., 2015). The exclusive expression of tubuliform spidroin genes in *A. argentata* tubuliform glands in our study may reflect a difference between *Argiope* (orb-web) and cob-web weaving spiders. Overall, our results are consistent with spidroin genes as defining features of silk glands. We also

found that the spidroin genes can be broadly expressed, but are most abundantly expressed in a single gland type, which is consistent with studies in other spider species (e.g. Garb et al., 2010; Babb et al., 2017; Clarke et al., 2017).

It should also be noted that spidroin genes feature a lengthy, highly repetitive central region flanked by shorter, non-repetitive N- and C-terminal encoding regions (e.g. Lewis, 1992; Gatesy et al., 2001). Thus, transcriptome assemblies result in fragmented spidroin sequences that do not accurately capture all possible variants, such as allelic or splice variants. Moreover, the repetitive sequence of spidroins can confound expression estimates from traditional RNA-seq reads, which is why we chose 3′DGE tag profiling because it generates more accurate read counts for spidroin genes (Chaw et al., 2016). However, 3′DGE tag reads are from the 100 to 150 base pairs proximal to the poly-A tail of mRNA, which again will not accurately capture all possible variants of a gene. Previous work has identified alternatively spliced spidroin sequences in *Trichonephila clavipes* (Babb et al., 2017). In addition, many spidroins have multiple loci (e.g. Ayoub and Hayashi, 2008; Collin et al., 2018). Future work into the genome of *A. argentata* could provide insight into how variants of SSTs are functionally distinct among silk glands.

3.3. Major ampullate and minor ampullate silk glands have the most similarities in non-spidroin SST expression, and aggregate silk glands have the most divergent non-spidroin SST expression

The majority of previous spider silk research has focused on spidroin silk genes, but we wanted to identify other SSTs that also contribute to spider silk production and to ensure that the high spidroin expression in some glands would not mask the differential expression of other, relatively lowly expressed transcripts. We therefore investigated differences in non-spidroin SST expression among spider silk gland types. In cob-web weaving spiders, the expression profiles of the fiber-forming glands were found to be divergent from the exclusively glue-forming aggregate glands (Clarke et al., 2017). As in cob-web weaving spiders, the aciniform; flagelliform; major ampullate; minor ampullate; pyriform; and tubuliform glands in *Argiope* all produce silk types that are either entirely fibrous or include silk fibers. The aggregate silk glands produce only glue-like adhesive silk.

Of all silk gland types, the major ampullate and the minor ampullate are the most similar in function and morphology (see description in Introduction). We thus expected that the SST expression in major ampullate and minor ampullate silk glands to be more similar to each other than any other pairing of silk gland types. By contrast, we expected that aggregate silk glands would be most divergent in expression profile from the other silk gland types. Specifically, we expected that aggregate silk glands would have the most unique transcripts and the fewest correlations in SST expression when compared to the other silk gland types in *A. argentata.*

To test our hypotheses, we clustered the non-spidroin SSTs present in more than one (2–7) gland type by expression level in two ways: by expression of the SSTs across all gland types and expression of the SSTs within a gland type (Fig. 2, top and left, respectively). When clustering SST expression level across all gland types, the analysis resulted in seven clusters. Expression of the SSTs in each cluster made up the majority of expression in a single gland type as follows: cluster number, gland type with highest expression (percent of expression in that gland type); 1, flagelliform (68%); 2, minor ampullate (67%); 3, major ampullate (57%); 4, aggregate (95%); 5, tubuliform (50%); 6, pyriform (58%); 7, aciniform (86%; Fig. 2, top; Supplementary Table S7). We will refer to each cluster according to the gland type with highest expression.

With the exception of the pyriform and tubuliform clusters, percentage of total SST expression was at least three fold higher in the gland with the highest expression compared to the gland type with the next highest expression. Pyriform cluster SSTs were 42% of total expression in aciniform glands, and tubuliform cluster SSTs were 30% of total expression in minor ampullate glands (Fig. 2, Supplementary Table S7). These results indicate that SSTs with expression in multiple gland types are nevertheless most highly expressed in a single gland type.

Both clustering methods place the major ampullate and minor ampullate glands sister to each other, indicating that these silk glands share more similarities in SST expression with each other than with other gland types. When clustering by expression within a gland type, aggregate gland SST expression places the aggregate gland outside of the other gland types, which is consistent with SSTs in the aggregate gland being the most divergent. Clustering by SST expression level across silk glands places the aggregate cluster sister to the tubuliform cluster, suggesting that the aggregate and tubuliform silk glands have similar expression levels of the SSTs identified in multiple gland types (Fig. 2).

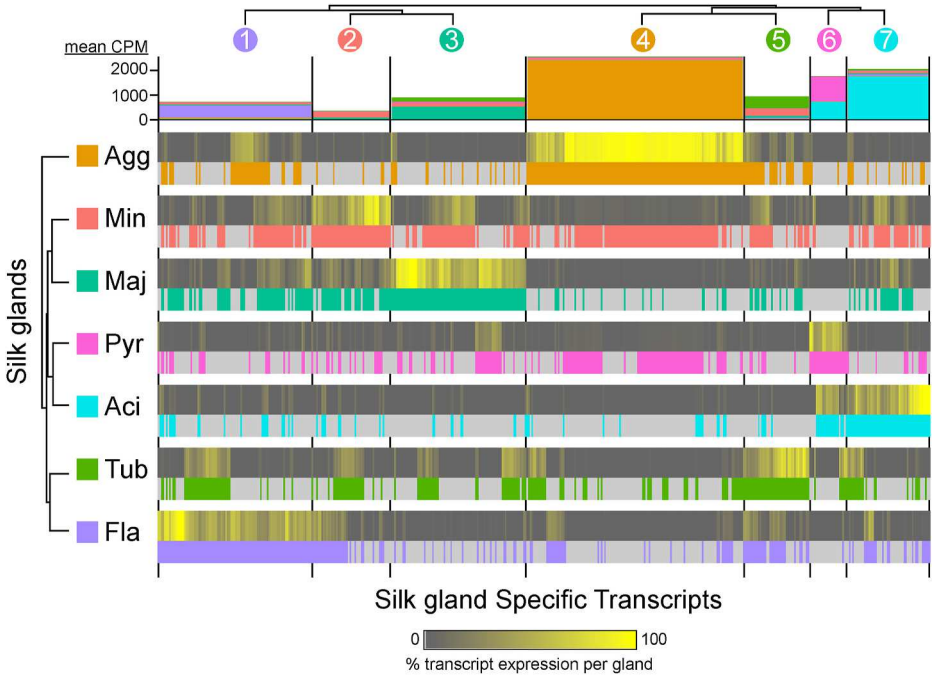

Fig. 2. Silk specific transcripts shared among 2–7 gland types, with spidroin transcripts removed. Gland types on Y axis, SSTs on X axis. Two rows next to each gland type indicate proportion of transcript expression in the gland type (top) and whether the contig was identified as significantly differentially expressed in a pairwise analysis with cephalothorax tissue (bottom, colored according to tissue type). Based on expression, SSTs group into seven clusters (numbers, clusters at top; colors correspond to gland with majority expression). Mean CPM expression from each gland type shown at top, below numbered circles. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

However, the SSTs in the aggregate cluster account for 95% of the expression in the aggregate gland, and only 0.4% of the expression in the tubuliform gland. Conversely, the SSTs in the tubuliform cluster are 50% of total expression in the tubuliform gland and 2% of expression in the aggregate gland (Supplementary Table S7). The clustering relationship between the aggregate and tubuliform gland SSTs is therefore likely to be driven by SSTs that are lowly expressed. Overall, our clustering analyses were consistent with our predictions that the major ampullate and minor ampullate silk glands likely employ similar transcripts to accomplish gland functions, and the aggregate silk glands possess the most unique SSTs and the most divergent expression profile of SSTs.

3.4. Positive correlations in SST expression suggest coordinated up-regulation of the transcripts in the fiber forming silk gland types and a close relationship between major ampullate and minor ampullate silk glands

Another way to elucidate similarities in silk gland transcript expression is to compare SST expression levels between clusters. We expected that the expression levels of the SSTs associated with fiber-forming glands would positively correlate, and that glue-forming SSTs would show no positive correlations with the fiber-forming SSTs. We performed pairwise comparisons of raw SST expression in each cluster, which revealed some significant positive correlations (Pearson's coefficient >0.1; Supplementary Table S8). Minor ampullate cluster SST expression levels were positively correlated with tubuliform, major ampullate, flagelliform, and aggregate cluster SST expression levels (0.65, 0.48, 0.27, and 0.14, respectively). Major ampullate cluster SST expression levels were also positively correlated with tubuliform cluster SST expression levels (0.50). Pyriform cluster SST expression levels were positively correlated with aciniform cluster SST expression levels (0.27; Supplementary Table S8). The positive correlations suggest coordinated up-regulation of the transcripts in the major ampullate gland with the minor ampullate gland, the minor ampullate gland with the flagelliform gland, the pyriform gland with the aciniform gland, and the major ampullate and minor ampullate glands with the tubuliform gland. Each of these silk glands produce a silk type that is either entirely fibrous or includes a fiber component (Tillinghast and Townley, 1993), as expected.

The major ampullate and minor ampullate clusters had strong positive correlation (0.48). To support this correlation, we also considered SSTs in 1 or 2 gland types, the narrow category of SSTs, because the narrow SSTs includes those transcripts that are most likely to define a silk gland type from all other silk gland types (single gland SSTs) and to reveal silk glands that are closely related (2 gland SSTs). We found that the major ampullate glands and minor ampullate glands shared the most SSTs out of all pairings of gland types (32; Fig. 3, Supplementary Table S9). The proportion of SSTs shared between the major ampullate and minor ampullate glands relative to total SSTs found in the gland was similar (10% and 9%, respectively; Supplementary Table S9). These results support the hypothesis that major ampullate glands and minor ampullate glands use similar transcripts.

3.5. Aggregate silk glands have a unique expression profile, but still have some shared SSTs with other glands

We did not expect the positive correlation (0.14) between minor ampullate and aggregate silk gland SST expression levels (Supplementary Table S8). We also found shared SSTs between the minor ampullate and aggregate silk glands in the narrow category of SSTs. Minor ampullate glands shared 23 SSTs with aggregate glands, second only to major ampullate glands, and 23 SSTs was the most SSTs that the aggregate gland shared with any other single gland type. Proportionately, 23 SSTs was a similar percent of the total SSTs in the minor ampullate and aggregate glands (6% and 5%, respectively; Supplementary Table S9).

Previous research on cob-web weaving spiders found that the expression of silk-gland specific transcripts was highest in aggregate silk glands (Clarke et al., 2017). This discrepancy between cob-web weaving spiders and *A. argentata* may be due to how the aggregate glands were analyzed in each study. In our analyses, the aggregate silk glands in *A. argentata* had the highest proportion of SSTs found only in a single gland (501 SSTs were found in aggregate silk glands, 255 (51%) of them were only in the aggregate gland) and a low number of SSTs shared with one other gland type (65; 12%; Supplementary Table S9). The high proportion of SSTs found only in the aggregate gland indicates that the *A. argentata* aggregate glands have a unique complement of SSTs. If the aggregate glands in cob-web weaving spiders also have many unique transcripts, then these may have separated the aggregate gland from other gland types in the hierarchical clustering of the over expressed silk gland transcripts in Clarke et al. (2017). In our study, we excluded these aggregate gland restricted transcripts from our clustering analyses, which revealed the relationship between aggregate and minor ampullate SST expression. Taken together, our results indicate that the aggregate glands in *A. argentata* may have evolved a set of SSTs to accomplish aggregate gland-specific tasks, such as keeping aggregate silk glue-like upon extrusion. However, the process of making aggregate silk may still rely on molecules that are shared with fiber-forming gland types. We further investigate this with GO term annotations, discussed below.

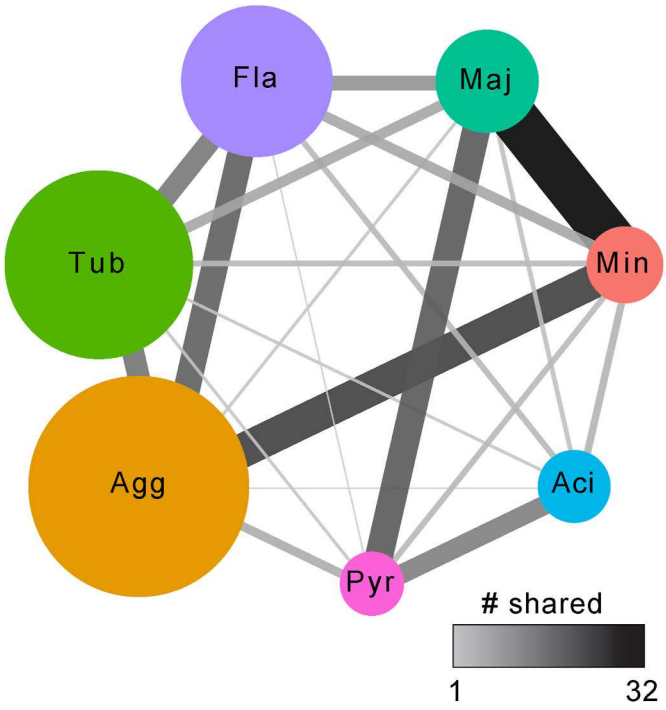

Fig. 3. Network diagram of unique SSTs and SSTs shared between two gland types. Gland types are the nodes, lines connecting nodes indicate shared contigs. Nodes are sized according to the number of unique contigs in each gland type. Line darkness and width correspond to the number of contigs shared between glands (Supplementary Table S9).

3.6. SSTs that match to the GO functions oxidoreductase or transmembrane transport differ among silk glands

Transcripts in the narrow category underlie differences between silk gland types and functional relationships between pairs of glands. We therefore functionally annotated our transcriptome with GO (Gene Ontology) terms and asked if any GO terms were significantly enriched in the narrow SSTs. A GOSLIM analysis of the 985 narrow SSTs mapped 139 transcripts to GOSLIM terms. The GOSLIM terms GO:0016491 (oxidoreductase activity) and GO:0055085 (biological process: transmembrane transport) were significantly enriched in the narrow SSTs

relative to the entire transcriptome (false discovery rate adjusted p-value <1e-5).

Both of these GOSLIM terms were also found in cob-web weaving silk gland specific transcripts (Clarke et al., 2014, 2017). Oxidation-reduction activity can facilitate pH changes, which are known to be part of the process that major ampullate spider silk undergoes as it transitions from a liquid to a fiber in the silk gland duct (Dicko et al., 2004). Transmembrane transport molecules are likely to play a role in moving spidroins from secretory cells into the lumen of the silk gland (Kovoor, 1987; Andersson et al., 2013).

Thirty-nine contigs mapped to oxidoreductase activity and twenty-three mapped to transmembrane transport (Fig. 4; Supplementary Table S10) among narrow SSTs. Other GO terms were enriched in the silk gland specific transcripts of cob-web weavers such as translation, peptidase activity and inhibition, proteinase, and proteinase inhibitors (Clarke et al., 2017). We did not find significant enrichment in the narrow SSTs of these GO terms, however, transcripts with top BLASTX matches to genes in all of these categories were found throughout our 1275 SSTs (Supplementary Table S5). The most likely explanation for the difference in our results compared to previous work in cob-web weaving spiders is our focus on the narrow SSTs within our already stringent SST criteria.

Most of the narrow SST contigs that mapped to the significantly enriched GO terms were found only in one gland type (Fig. 4). No contig among the narrow SSTs that mapped to GO:0016491 (oxidoreductase activity) or GO:0055085 (transmembrane transport) was shared between major ampullate glands and minor ampullate glands. Sixteen of the contigs that mapped to oxidoreductase activity were found in either the major ampullate gland or minor ampullate gland. Ten of these were unique to major ampullate glands, three were unique to minor ampullate glands, and three were found in minor ampullate and aggregate silk glands. The three that were found in both minor ampullate and aggregate silk glands have top BLASTX (nr) hits to two coenzymeA genes and glucose dehydrogenase.

Coenzyme A is necessary for hundreds of metabolic processes including lipid synthesis and energy generation (Leonardi et al., 2005). Glucose dehydrogenase is an enzyme that transfers electrons without the use of oxygen (Ferri et al., 2011). Of note, these genes are also enriched in the silk gland transcriptomes of the silkworms *Bombyx mandarina* and *Bombyx mori.* Fang et al. (2015) identified highly expressed and highly differentially expressed genes in the silk glands of *B. mandarina* and *B. mori.* Among these genes were coenzymeA and glucose dehydrogenase. In contrast, in the muga silkworm *Antheraea assamensis,* genes with oxidoreductase functions were not significantly enriched in silk glands relative to non-silk gland tissue (Chetia et al., 2017). In the fruit fly *Drosophila melanogaster,* glucose dehydrogenase is expressed in a tissue-specific manner in the reproductive tract (Schiff et al., 1992).

The exact function of coenzyme A and glucose dehydrogenase in spider silk glands remains unknown, but their broad importance in the metabolic processes and silk glands of other organisms suggests that they could be related to basic functions such as cellular energy generation. Enrichment of contigs that match the oxidoreductase GO term could therefore reflect the higher energy needs of silk producing glands as they synthesize silk proteins. Indeed, contigs that mapped to oxidoreductase were found in six of the seven spider silk gland types (Fig. 4). That minor ampullate and aggregate silk glands share expression of specific coenzymeA and glucose dehydrogenase transcripts is consistent with our results indicating that aggregate silk glands may retain some of the same functions as fiber-forming glands.

None of the significantly enriched narrow SSTs that map to transmembrane transport were found in major ampullate silk glands or minor ampullate silk glands. Ten of the SSTs that map to transmembrane transport were found in aggregate silk glands. Most of these have top BLASTX matches to inorganic phosphate cotransporter isoforms. Inorganic phosphate is an essential nutrient that is key to cellular metabolic processes. Transporters move inorganic phosphate across cell

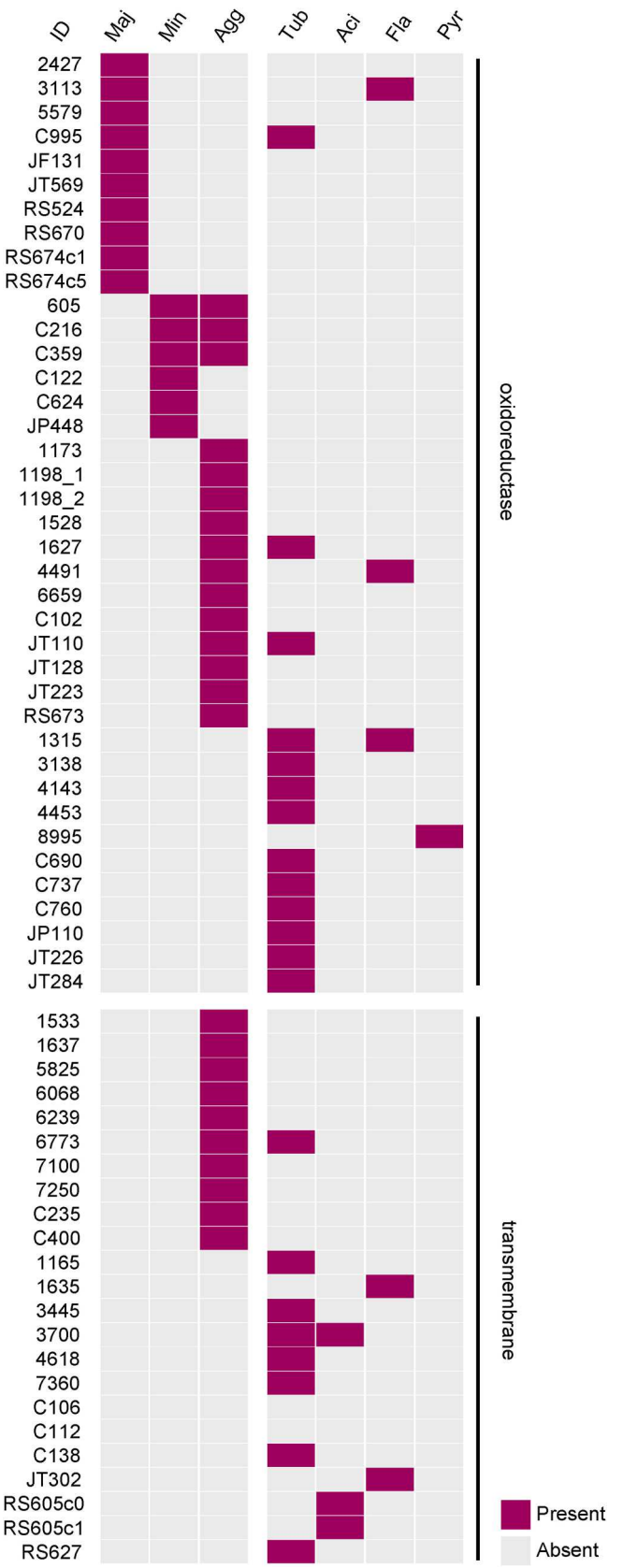

Fig. 4. SSTs in the narrow category (found in only 1 or 2 gland types) that matched the GOSLIM term for oxidoreductase (top) or transmembrane transport (bottom) function. Results are organized by gland type with major ampullate, minor ampullate, and aggregate glands at the left. Full BLASTX matches are available in Supplementary Table S10.

boundaries (Tenenhouse, 2007). Genes involved in protein excretion and translocation were also discovered in the highly expressed and highly differentially expressed silk gland genes from *B. mandarina* and *B. mori* (Fang et al., 2015), although these genes did not map to the same transmembrane transport GO term. Similarly, in *A. assamensis,* Chetia et al. (2017) found enrichment of genes with transport function in the silk gland as compared to non-silk gland tissue. Oxidoreductase activity and transmembrane transport are both functions that could be considered common among all spider silk glands, but our results suggest that transcripts with these functions may be influential in defining spider silk gland types from each other because they are enriched in our narrow SSTs. Moreover, most of the transcripts in the narrow category that mapped to these enriched GO terms have gland-specific expression.

3.7. Tubuliform and aciniform silk glands have SST expression profiles that may correlate with their function and evolutionary history, respectively

Our SST analyses revealed expression profiles for the tubuliform and aciniform silk glands that may correlate with the function of tubuliform silk, and the evolutionary history of aciniform silk glands. The non-spidroin SSTs in the narrow category accounted for 76% of total SST expression in the tubuliform glands, which was the highest narrow category percentage for all the gland types (ranged from 11% to 76%, Supplementary Table 11). Meanwhile, the non-spidroin SSTs in the moderate category had the highest proportion in aciniform silk glands, 72% of total SST expression, much greater than the percentage in the other gland types (6%–41%; Supplementary Table S11). We then further subdivided the narrow, moderate, and broad categories of non-spidroin SSTs into those expressed in all seven gland types, in six gland types, and so forth, down to only one gland type. Looking at non-spidroin SST expression this way, 75% of total non-spidroin SST expression in tubuliform glands is from SSTs that are found in only one gland type. By contrast, in aciniform glands, 52% of non-spidroin SST expression is from SSTs shared among five gland types (Fig. 5; Supplementary Table S11).

The high percentage of gland-specific SSTs in tubuliform silk glands may be related to the unique function of tubuliform silk. Tubuliform silk is used only in egg case wrapping, and tubuliform silk glands only develop in female spiders when the spider reaches sexual maturity (Tillinghast and Townley, 1993). The specialized transcripts in tubuliform silk glands may be necessary for developing silk glands at a late stage in the life cycle of the spider or related to the uniqueness of tubuliform silk. Unlike the other silk types, tubuliform silk must be stored for long periods of time until the moment a female commences egg case construction.

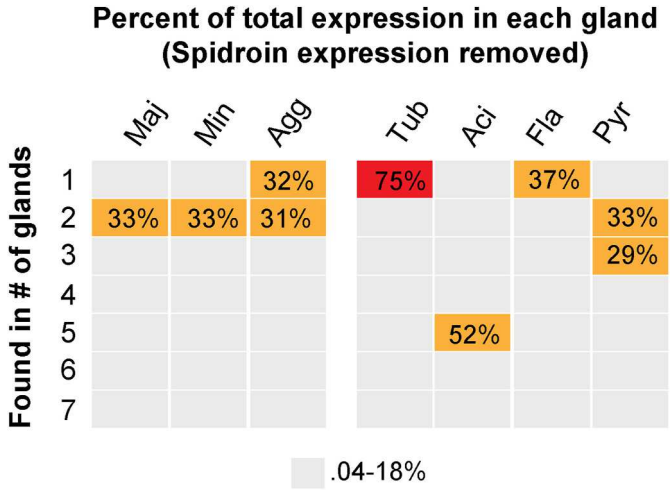

Fig. 5. Percent of total SST expression for the SSTs that are unique to one gland type or shared with other gland types. Spidroin gene expression has been removed.

The broad expression of aciniform gland SSTs may be related to spidroin evolution. Aciniform silk glands are thought to have been present in the last common ancestor of true spiders (Coddington and Levi, 1991), and aciniform spidroin genes likely duplicated and diversified prior to the morphological diversification of silk glands (Starrett et al., 2012; Clarke et al., 2015). Aciniform glands may therefore express non-spidroin genes that also duplicated, diversified, and were then used in other spider silk gland types.

4. Conclusions

Our expression profiling of individual silk glands in *A. argentata* provided the resolution necessary to categorize silk gland specific genes as either narrowly or broadly expressed. From our comprehensive *A. argentata* transcriptome, we identified 1275 transcripts that are abundantly and significantly differentially expressed in *A. argentata* silk glands (Fig. 1). We found that different spidroin types have different expression patterns. The flagelliform and tubuliform spidroin genes were narrowly confined to their respective silk gland types. By contrast, aciniform, aggregate, and pyriform spidroin genes were broadly expressed in as many as six silk gland types (Fig. 2, Supplementary Tables S5 and S6).

Expression level clustering and correlation analyses indicated that the major ampullate and minor ampullate silk glands had the most similar non-spidroin SST expression profiles and the most SSTs in common out of any pair of gland types. Aggregate silk glands had the most unique non-spidroin SST expression profile. However, contrary to our expectations, aggregate silk gland SST expression was positively correlated with the minor ampullate silk glands and aggregate and minor ampullate silk glands had some SSTs in common (Figs. 2 and 3; Supplementary Tables S8 and S9). The relationship between aggregate and minor ampullate silk glands was further supported by our GO term enrichment analysis. The aggregate and minor ampullate silk glands shared three contigs that mapped to the enriched GO term for oxidoreductase activity (Fig. 4, Supplementary Table S10). We also found that the tubuliform silk glands had the most narrowly expressed non-spidroin SSTs, and that the aciniform silk glands had the most broadly expressed non-spidroin SSTs (Fig. 5; Supplementary Table S11).

Our results are consistent with a glandular affiliation hypothesis of silk protein evolution, which implies that silk protein relationships should follow silk gland relationships (Hayashi and Lewis, 1998; Ayoub et al., 2012). Based on silk gland spigots and silk spinning behavior, systematic studies suggest that the major ampullate gland evolved before the minor ampullate gland and that the aggregate gland is a much more recent innovation (e.g. Coddington and Levi, 1991; Hormiga and Griswold, 2014). Our results reinforce the affiliation of the major ampullate and minor ampullate silk glands with each other, and that the aggregate gland is the most divergent. Furthermore, aciniform silk glands are thought to be one of the earliest silk gland types to evolve (Shultz, 1987). A high percentage (52%, Fig. 5) of the SSTs found in aciniform silk glands were shared among multiple gland types. If aciniform glands were among the first to evolve, then the proteins associated with aciniform glands have had the most time to duplicate and diversify, and to be co-opted for use in other glands. Tubuliform glands are found in a wide variety of spider species, and it is most likely that the specificity of the SSTs in tubuliform glands is linked to their highly specialized use for egg case wrapping silk production during the reproductive cycle of adult spiders.

Our hypotheses focused on the major ampullate, minor ampullate, and aggregate silk glands, and on highly expressed transcripts. Future research could explore the connection that we found between aggregate silk glands and minor ampullate silk glands. The SSTs shared between the aggregate and minor ampullate silk glands appear to be necessary for

cellular metabolism, but why these were shared only between these two gland types remains unknown. We also cannot discount the potential importance of silk gland specific transcripts that are lowly expressed. For example, the differentially expressed genes in each gland that did not make our cutoffs for SST designation may contain genes that must be downregulated in order for silk production to proceed. Future work could explore lowly expressed genes in spider silk glands as compared to the cephalothorax.

We were able to characterize shared and divergent transcripts among fiber forming and glue forming silk glands in *A. argentata,* and we identified novel candidate genes for further study of how spiders accomplish silk production. Comparison of all silk gland types in a common genome provides insight into the individual transcripts that underlie the production of unique spider silk glands and silk types. Our work advances knowledge about individual spider silk glands and the transcripts that they use to produce different silk types.

Funding

This work was supported by the Army Research Office (W911NF-11-1-0299, W911NF-15-1-0099) and by the National Science Foundation (IOS-0951086, IOS-0951061, IOS-1755142, IOS-1754979).

Acknowledgements

The authors thank S. Correa-Garhwal, R. Baker, and J. Gatesy for improving this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ibmb.2021.103594.

References

Anders, S., Huber, W., 2010. Differential expression analysis for sequence count data. Nat. Precedings. 1–1 https://doi.org/10.1038/npre.2010.4282.1.

Andersson, M., Holm, L., Ridderstrale, Y., Johansson, J., Rising, A., 2013. Morphology and composition of the spider major ampullate gland and dragline silk. Biomacromolecules 14, 2945–2952.

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

Ayoub, N.A., Garb, J.E., Kuelbs, A., Hayashi, C.Y., 2012. Ancient properties of spider silks revealed by the complete gene sequence of the prey-wrapping silk protein (AcSp1). Mol. Biol. Evol. 30, 589–601.

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

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

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformation 30, 2114–2120.

Brooks, A.E., Nelson, S.R., Jones, J.A., Koenig, C., Hinman, M., Stricker, S., Lewis, R.V., 2008. Distinct contributions of model MaSp1 and MaSp2 like peptides to the mechanical properties of synthetic major ampullate silk fibers as revealed in silico. Nanotechnol. Sci. Appl. 1, 9–15.

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

Chaw, R.C., Hayashi, C.Y., 2018. Dissection of silk glands in the Western black widow *Latrodectus hesperus*. J. Am. Arachnol. 46, 159–161.

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

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

Chaw, R.C., Saski, C.A., Hayashi, C.Y., 2017. Complete gene sequence of spider attachment silk protein (PySp1) reveals novel linker regions and extreme repeat homogenization. Insect Biochem. Mol. Biol. 1, 80–90.

Chaw, R.C., Collin, M.A., Wimmer, M., Helmrick, K.L., Hayashi, C.Y., 2018. Egg case silk gene sequences from *Argiope* spiders: evidence for multiple loci and a loss of function between paralogs. G3- Genes Genom. Genet. 8, 231–238.

Chetia, H., Debajyoti, K., Singh, D., Mosahari, P.V., Das, S., Sharma, P., Neog, K., Sharma, S., Jayaprakash, P., Bora, U., 2017. De novo transcriptome of the muga silkworm, *Antheraea assamensis* (Helfer). Gene 611, 54–65.

Clarke III, T.H., Garb, J.E., Hayashi, C.Y., Haney, R.A., Lancaster, A.K., Corbett, S., Ayoub, N.A., 2014. Multi-tissue transcriptomics of the black widow spider reveals expansions, co-options, and functional processes of the silk gland gene toolkit. BMC Genom. 15, 1–17.

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

Clarke III, T.H., Garb, J.E., Haney, R.A., Chaw, R.C., Hayashi, C.Y., Ayoub, N.A., 2017. Evolutionary shifts in gene expression decoupled from gene duplication across functionally distinct spider silk glands. Sci. Rep. 7, 8393. https://doi.org/10.1038/s41598-017-07388-1.

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

Coddington, J.A., Levi, H.W., 1991. Systematics and evolution of spiders (Araneae). Annu. Rev. Ecol. Systemat. 22, 565–592.

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

Collin, M.A., Clarke III, 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. Int. J. Biol. Macromol. 113, 829–840.

Correa-Garhwal, S.M., Chaw, R.C., Clarke III, T.H., Alaniz, L.G., Chan, F.S., Alfaro, R.E., Hayashi, C.Y., 2018. Silk genes and silk gene expression in the spider *Tengella perfuga* (Zoropsidae), including a potential cribellar spidroin (CrSp). PLoS One 9. https://doi.org/10.1371/journal.pone.0203563.

Dicko, C., Vollrath, F., Kenney, J.M., 2004. Spider silk protein refolding is controlled by changing pH. Biomacromolecules 5, 704–710.

Fang, S.-M., Hu, B.-L., Zhou, Q.-Z., Yu, Q.-Y., Zhang, Z., 2015. Comparative analysis of the silk gland transcriptome between the domestic and wild silkworms. BMC Genom. 16, 60. https://doi.org/10.1186/s12864-015-1287-9.

Ferri, S., Kojima, K., Sode, K., 2011. Review of glucose oxidases and glucose dehydrogenases: a bird's eye view of glucose sensing enzymes. J. Diabetes Sci. Technol. 5, 1068–1076.

Garb, J.E., Ayoub, N.A., Hayashi, C.Y., 2010. Untangling spider silk evolution with spidroin terminal domains. BMC Evol. Biol. 10, 1–16.

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

Gordon, D., Abajian, C., Green, P., 1998. Consed: a graphical tool for sequence finishing. Genome Res. 8, 195–202, 1998.

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., 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat. Biotechnol. 29, 644.

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., 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8, 1494–1512.

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

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

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

Hormiga, G., Griswold, C.E., 2014. Systematics, phylogeny, and evolution of orb-weaving spiders. Annu. Rev. Entomol. 59, 487–512.

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

Kopylova, E., Noé, L., Touzet, H., 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformation 28, 3211–3217.

Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., Phillippy, A.M., 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 27, 722–736.

Kovoor, J., 1987. Comparative structure and histochemistry of silk-producing organs in arachnids. In: Nentwig, W. (Ed.), Ecophysiology of Spiders. Springer, Berlin, Heidelberg, pp. 160–186.

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

Kuntner, M., Hamilton, C.A., Cheng, R.C., Gregorič, M., Lupše, N., Lokovšek, T., Lemmon, E.M., Lemmon, A.R., Agnarsson, I., Coddington, J.A., Bond, J.E., 2019. Golden orbweavers ignore biological rules: phylogenomic and comparative analyses unravel a complex evolution of sexual size dimorphism. Syst. Biol. 68, 555–572.

Lane, A.K., Hayashi, C.Y., Whitworth, G.B., Ayoub, N.A., 2013. Complex gene expression in the dragline silk producing glands of the Western black widow (*Latrodectus hesperus*). BMC Genom. 14, 1–12.

Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, 1–10.

Leonardi, R., Zhang, Y.M., Rock, C.O., Jackowski, S., 2005. Coenzyme A: back in action. Prog. Lipid Res. 44, 125–153.

Lewis, R.V., 1992. Spider silk: the unraveling of a mystery. Acc. Chem. Res. 25, 392–398.
Lin, B., Pang, Z., 2019. Stability of methods for differential expression analysis of RNA-seq data. BMC Genom. 20, 1–13.
Luo, M., Wing, R.A., 2003. An improved method for plant BAC library construction. In: Grotewold, E. (Ed.), Plant Functional Genomics. Humana Press, New Jersey, pp. 3–19.
McCarthy, F.M., Wang, N., Magee, G.B., Nanduri, B., Lawrence, M.L., Camon, E.B., Barrell, D.G., Hill, D.P., Dolan, M.E., Williams, W.P., Luthe, D.S., 2006. AgBase: a functional genomics resource for agriculture. BMC Genom. 7, 1–13.
Moon, M.J., 2012. Organization of the spinnerets and spigots in the orb web spider, *Argiope bruennichi* (Araneae: Araneidae). Entomol. Res. 42, 85–93.
Moon, M.J., 2018. Fine structure of the aggregate silk nodules in the orb-web spider *Nephila clavata*. Anim. Cell Syst. 22, 421–428.
Sambrook, J., Fritsch, E.R., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual, second ed. Cold Spring Harbor Press, New York.
Saski, C.A., Feltus, F.A., Parida, L., Haiminen, N., 2015. BAC sequencing using pooled methods. In: Narayanan, K. (Ed.), Bacterial Artificial Chromosomes. Humana Press, New York, pp. 55–67.
Schiff, N.M., Feng, Y., Quine, J.A., Krasney, P.A., Cavener, D.R., 1992. Evolution of the expression of the Gld gene in the reproductive tract of Drosophila. Mol. Biol. Evol. 9, 1029–1049.
Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504.
Shultz, J.W., 1987. The origin of the spinning apparatus in spiders. Biol. Rev. 62, 89–113.
Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformation 31, 3210–3212.
Starrett, J., Garb, J.E., Kuelbs, A., Azubuike, U.O., Hayashi, C.Y., 2012. Early events in the evolution of spider silk genes. PLoS One 7, e38084.
Sutherland, T.D., Young, J.H., Weisman, S., Hayashi, C.Y., Merritt, D.J., 2010. Insect silk: one name, many materials. Annu. Rev. Entomol. 55, 171–188.
Tenenhouse, H.S., 2007. Phosphate transport: molecular basis, regulation and pathophysiology. J. Steroid Biochem. Mol. Biol. 103, 572–577.
Tillinghast, E.K., Townley, M.A., 1993. Silk glands of araneid spiders: selected morphological and physiological aspects. In: Kaplan, D., Adams, W.W., Farmer, B., Viney, C. (Eds.), Silk Polymers. Am. Chem. Soc. Press, Washington, D.C., pp. 29–44
Townley, M.A., Tillinghast, E.K., 2013. Aggregate silk gland secretions of araneoid spiders. In: Nentwig, W. (Ed.), Spider Ecophysiology. Springer, Berlin, Heidelberg, pp. 283–302.
Vollrath, F., 2000. Strength and structure of spiders' silks. Rev. Mol. Biotechnol. 74, 67–83.
Vollrath, F., Edmonds, D.T., 1989. Modulation of the mechanical properties of spider silk by coating with water. Nature 340, 305–307.
Whaite, A.D., Wang, T., Macdonald, 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, e0204243.
Wirth, M., Wolff, J.O., Appel, E., Gorb, S.N., 2019. Ultrastructure of spider thread anchorages. J. Morphol. 280, 534–543.
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.
Yanai, I., Benjamin, H., Shmoish, M., Chalifa-Caspi, V., Shklar, M., Ophir, R., Bar-Even, A., Horn-Saban, S., Safran, M., Domany, E., Lancet, D., 2005. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. Bioinformation 21, 650–659.
Yarger, J.L., Cherry, B.R., Van Der Vaart, A., 2018. Uncovering the structure–function relationship in spider silk. Nat. Rev. Mat. 3, 18008, 2018.
Young, M.D., Wakefield, M.J., Smyth, G.K., Oshlack, A., 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol. 11, R14, 2010.
Young, M.D., Wakefield, M.J., Smyth, G.K., Oshlack, A., 2012. goseq: gene Ontology testing for RNA-seq datasets. R Bioconductor 8, 1–25.