Integrative and Comparative Biology, 2021, pp. 1–6

doi: DOI HERE
Advance Access Publication Date: Day Month Year
Paper

PAPER

Repeat variation resolves a complete aggregate silk sequence of bolas spider *Mastophora phrynosoma*

Sarah D. Stellwagen[1,*] and Mercedes Burns[2]

[1]Department of Biological Sciences, UNC Charlotte, 9201 University City Blvd, 28223, NC, USA and [2]Department of Biological Sciences, University of Maryland, Baltimore County, 1000 Hilltop Circle, 21250, MD, USA
[*]Corresponding author. stellwagen@uncc.edu
FOR PUBLISHER ONLY Received on Date Month Year; revised on Date Month Year; accepted on Date Month Year

Abstract

Many species of spider use a modified silk adhesive, called aggregate glue, to aid in prey capture. Aggregate spidroins (spider fibroins) are modified members of the spider silk family, however they are not spun into fibers as are their solid silk relatives. The genes that encode for aggregate spidroins are the largest of the known spidroin genes and are similarly highly repetitive. In this study, we used long read sequencing to discover the aggregate spidroin genes of the toad-like bolas spider, *Mastophora phrynosoma*, which employs the glue in a unique way, using only a single, large droplet to capture moths. While Aggregate Spidroin 1 (AgSp1) remains incomplete, AgSp2 is more than an extraordinary 62 kilobases of coding sequence, 20 kb longer than the longest spidroin on record. The structure of repeats from both aggregate silk proteins follows a similar pattern seen in other species, with the same strict conservation of amino acid residue number for much of the repeats' lengths. Interestingly, AgSp2 lacks the elevated number and groupings of glutamine residues seen in the other reported AgSp2 of a classic orb weaving species. The role of gene length in glue functionality remains a mystery, and thus discovering length differences across species will allow understanding and harnessing of this attribute for the next generation of bio-inspired adhesives.

Key words: spidroin, silk, long-read, bolas spider, *Mastophora*, aggregate glue

Introduction

Spiders use a remarkable set of silks throughout their lives for a variety of tasks. Each fiber type contributes unique material properties tuned for specific functions, such as forming lifelines or wrapping egg sacs. In addition to the familiar solid silk fibers, some spiders also use a specialized silk glue to retain prey that have intercepted a web. Most silk proteins belong to the spidroin (spider fibroin) family and most spidroins are transformed from a liquid dope that is produced and stored within a spider's glands, into a solid fiber as the material moves through glandular ducts and exits via spigots on the spinnerets. In contrast, aggregate spider glues, which are largely composed of modified silk proteins (Collin et al., 2016), are extruded as a liquid similarly to how they are stored. Furthermore, unlike other silks, aggregate glue proteins are extensively glycosylated, which contributes to glue stickiness (Sahni et al., 2010; Singla et al., 2018).

Spiders use aggregate glue for prey capture, however the glue is employed in a variety of web morphologies. In orb webs, the glue is extruded onto the spiral silk of the web's central region (Apstein, 1889). Cobweb weavers deposit the glue at the base of "gumfoot" lines loosely attached to the substrate that easily release when intercepted by prey, suspending the targets in mid-air until the spider can reel them in (Wiehle, 1931; Benjamin and Zschokke, 2002; Argintean et al., 2006). Bolas spiders in the genus *Mastophora* and some species from related genera use a particularly fascinating hunting technique where females produce a single droplet of glue suspended at the end of a silk strand (Fig. 1) and release pheromone mimics that draw in male moths prior to capture (Hutchinson, 1903; Eberhard, 1977). While not uncommon and relatively large with abdominal widths 1.5 cm or more (Levi, 1955), these spiders are extremely difficult to collect due to their cryptic daytime resting behavior, which resembles bird droppings on

the upper surface of a leaf, and lack of an obvious web that aids locating other spider species.

The biomechanics of spider glues have been studied extensively in orb and cobweb weavers, and many species' aggregate silks exhibit unique material properties, including different stretching, adhesion, and energy absorption capabilities, as well as differing responses to ambient conditions like humidity (Sahni et al., 2011b,a; Opell et al., 2011, 2013). Recently, research on the glue of moth-specialist species *Cyrtarachne akirai*, which is related to bolas spiders, showed that, combined with large droplet size, low viscosity allows effective glue spreading around and beneath the fragile scales of moth prey, which otherwise allows these moths to escape typical orb weaving spider webs (Diaz et al., 2018). Further research showed that the glue actually cements the scales together, effectively using them as a weapon against the moths themselves (Diaz et al., 2020).

Although bolas spiders use their glues in a unique mode, their phylogenetic position within the orb weaving family Araneidae, and sister to *Araneus*, one of several genera within the Araneidae that make classic orb webs (Blackledge et al., 2009), suggests that the glues of the two groups should be similarly encoded. Furthermore, while biomechanical research on glues is an important aspect of understanding glue function, the contribution of genetics to performance is far less understood. Only two full-length glue genes from a single species have thus far been published despite thousands of glue-producing spider species and a variety of unique web structures, habitats, and prey targets (Stellwagen and Renberg, 2019). As traditional short-read sequencing alone is not capable of uncovering complete repetitive spider silk sequences, it is only recently that long-read sequencing has allowed a more detailed look at the variation of repeat number and gene length, which will promote understanding of the contributions of these traits to glue functionality.

Similar to the solid fiber members from this extensive gene family, aggregate spidroins are encoded by large, highly repetitive genes that are difficult to fully sequence. The challenge of sequencing repetitive genes lies in the current technologies commonly applied to DNA sequencing. Techniques that validate polymorphisms and scaffold the full length of non-repetitive genes using short reads (i.e. less than 250 base pairs per read) are not successful in sequencing aggregate glues due to repeat size and extreme uniformity. Short reads cannot connect anchoring sequence (such as unique portions of the termini or introns) across repeats if they are shorter than the repetitive regions, nor be confidently placed within the repetitive region itself. Long-read sequencing technology has vastly improved gene and genome assemblies, aiding in the resolution of repetitive regions via scaffolding or direct assembly with appropriate depth of coverage. However, these third-generation sequencing technologies, which can consistently produce reads upwards of 40 kilobases or more, may similarly struggle to produce adequate depth of coverage to scaffold very long single or limited exon genes.

In this study, we used long reads to sequence the complete, yet unpolished Aggregate Spidroin 2 (AgSp2) from the moth specialist bolas spider, *Mastophora phrynosoma*, yielding a glue gene sequence spanning over 100 kilobases and more than 60 kb of coding sequence, the longest spidroin discovered to date. While we were successful in discovering the complete genomic sequence for one of the two main aggregate genes, Aggregate Spidroin 1 (AgSp1) remains unresolved due to the challenges of obtaining a sufficient number of reads with a larger genome size. Here, however, we demonstrate a strategy for the assembly of long repetitive reads using unique patterns formed from single base variations across repeats. We end by discussing future-minded evolutionary questions elicited by the study of these massively repetitive genes.

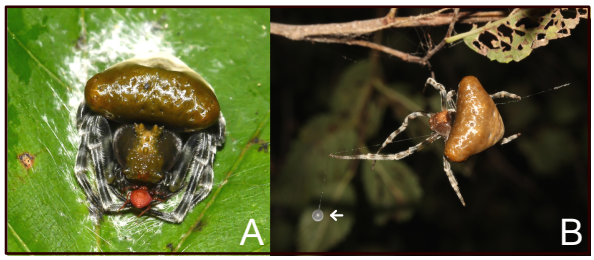

Fig. 1. Female *Mastophora phrynosoma* and smaller red male resting on the surface of a leaf during the day (A) and female hunting with a bolas (arrow points to glue droplet) at night (B). Images courtesy of Bonnie Ott (A) and MJ Hatfield (B).

Methods

Adult female *Mastophora phrynosoma* were collected from the Living Farm Heritage Museum grounds in West Friendship, Maryland and the Howard County Conservancy in Woodstock, Maryland during late September, 2019. Genomic DNA was extracted from prosomal and leg tissue using the MasterPure Complete DNA and RNA Purification Kit following the DNA Purification section protocol (cat.no. MC89010). Four samples each with a total of 10 µg of the gDNA extraction was loaded onto a Sage Science BluePippin High Pass Plus cassette (cat.no. BPLUS10) and run with a 15 kb high pass threshold overnight. The resultant elutions were cleaned with 1X AMPure XP beads (cat.no. A63880) and eluted overnight in water. The cleaned DNA was then used directly in Oxford Nanopore's 1D Genomic DNA by Ligation protocol (SQK-LSK109). A total of two runs were completed using SpotON Flow Cells (R9.4; cat.no. FLO-MIN106) and resultant fast5 files were basecalled using Oxford Nanopore's program Guppy (v3.6.1+249406c). A total of 26.7 Gbp and 10.8 million reads across two Nanopore flow cells were produced.

A custom *M. phrynosoma* BLAST database was created from the sequence data using Geneious v11.1.5 (Kearse et al., 2012), and termini and repeat-encoding sequences from *A. trifasciata* aggregate spidroins were used as query sequences to initially identify aggregate reads from the *M. phrynosoma* data set. After AgSp reads were discovered, repeat motifs were used as queries against the data set to extract all repeat-containing reads. In addition to searching manually for short matching sequences using command line, and adjusting BLAST parameters within Geneious to ensure discovery of all AgSp reads, using the default parameters for Megablast recovered all applicable reads gathered via other methods. Once the longest reads available were selected, the repeats of each read were identified and annotated. Repeats from a single read were then extracted and aligned, and the resultant alignment visually inspected to identify single base variations across repeats, which could be separated from

errors due to their consistency from repeat to repeat (Fig. 2A). Repeat annotations from each read were then color coded to reflect the base variation at the chosen position (Fig. 2B). Using the pattern formed by variations across repeats, reads were manually combined to approximate the complete genetic sequence length (Fig. 2C), as short read data is currently unavailable for long read correction. Additionally, a limited number of repeats are either significantly shorter or longer than the standard repeat size, which provided further evidence for alignment precision. Sequence alignments and analysis was conducted using Geneious with the Geneious alignment tools.

Results

We sequenced and assembled the complete Aggregate Spidroin 2 (AgSp2) and partial AgSp1 from the toad-like bolas spider species *Mastophora phrynosoma*. AgSp1 and AgSp2 are members of the spider silk (spidroin) family (Collin et al., 2016) and conform to the same N-terminus/Repeat(n)/C-terminus structure. AgSp2 was manually assembled from six long reads (Supplementary File 1) and is encoded by nearly 63 kilobases, dwarfing the previous spidroin record holder from *Argiope trifasciata*'s AgSp1, which is just over

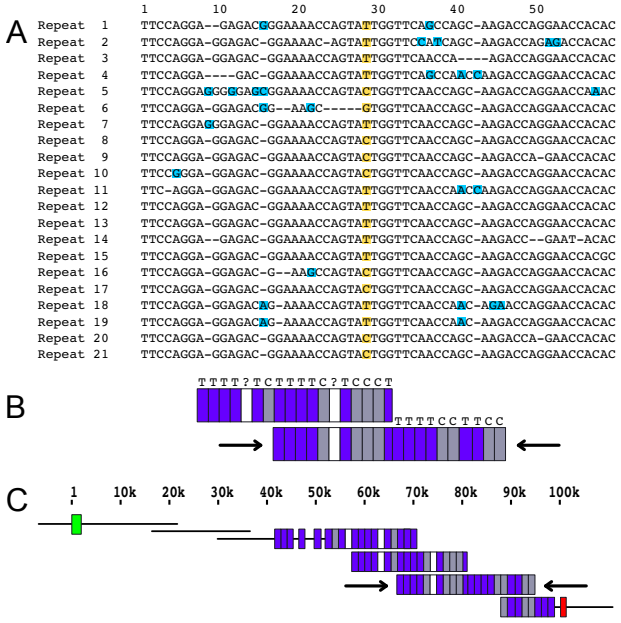

Fig. 2. Long-read alignment strategy used to resolve the complete AgSp2 spidroin from the bolas spider *Mastophora phrynosoma*. Alignment of the relevant portion from all 21 repeats of an AgSp2 read that contains a single nucleotide variant (schematic of this read is between arrows in B and C). Random highlighted bases generally represent sequencing errors (blue highlight), however the base at position 29 is variable and can be either C or T, or occasionally, another base or sequencing error (yellow highlight)(A). After annotating each repeat as containing a T (purple), C (grey), or alternate/unknown (white) at the variable base location, the pattern between these bases across repeats can be used to align long repetitive reads (B). The resultant gene after alignment of all relevant reads, including non-coding sequence (black line) and coding sequence (N-terminal encoding region in green and C-terminal encoding region in red), is just over 100 kb (C). The coding sequence of AgSp2 is approximately 62.7 kb, however as the gene contains sequencing errors, exact length and polished sequence is still unclear.

42 kb. Interestingly, the AgSp2 ortholog in *A. trifasciata* is approximately 20.5 kb, only a third of the size of that of the bolas spider. Similar to AgSp2 of *A. trifasciata*, the N-terminal encoding region of AgSp2 of *M. phrynosoma* is separated from the rest of the gene by a massive intron. The intron in *M. phrynosoma* is approximately 37.5 kb, similar to that of *A. trifasciata*'s 31.5 kb intron. Unlike *A. trifasciata*, the bolas spider AgSp2 does not have the same glutamine-rich regions between sections of repeats.

M. phrynosoma's AgSp2 is comprised of approximately 47 repeats, however, repeats near the termini, particularly towards the N-terminus, tend to be less strictly conserved as the central repeats, which can result in difficulties assigning consistent beginnings and ends. Central repeats are formed from 1,326 bp encoding 442 amino acids, which can be further divided into four sub-repeat units, each highly similar to the main repeat found in AgSp2 of *A. trifasciata* (Fig. 3). As the full length of AgSp1 from *M. phrynosoma* has not been resolved, it is not possible to determine the number of repeat units within the gene, however there are more than 70 repeats within the acquired overlapping reads, totalling more than 42 kb. The gene is at least somewhat larger than this, as there are repeats not included in the count total within the N- and C-terminal encoding reads, and, based on unique repeat structure of the terminal-adjacent repeats, there are additional reads needed to bridge these terminal reads to the remainder of the central repeats.

Both AgSp1 and AgSp2 had variable bases that allowed proper alignment of the long repetitive regions. They varied, however, in the frequency of synonymous changes to the translated amino acids. While AgSp2 repeat variants were differentiated by a single base change (Fig. 3B) that led to a synonymous substitution in a tyrosine residue, AgSp1 repeat variants differed at a total of six bases. These included three synonymous substitutions at histidine and valine residues, and three non-synonymous substitutions convert serine, glutamine, and leucine to arginine, arginine, and phenylalanine residues, respectively.

Consistent with previous reports for AgSp1 and AgSp2 of both orb weavers and cobweb weavers (Stellwagen and Renberg, 2019), the number of amino acids for much of *M. phrynosoma* repeats is highly conserved, and variable within the tail region (Fig. 3). *M. phrynosoma* repeats, however, are formed from larger, variable sub-repeats that are each similar to the full repeat units of other species. Sequence deviations across sub-repeat units, particularly in the tail region, increases the size of the overall repetitive pattern of amino acids.

Discussion

The toad-like bolas spider *Mastophora phrynosoma* produces a large droplet of aggregate glue that it uses to capture moth prey. In other closely related orb-weaving species, the main glue proteins are encoded by two extremely large, repetitive genes Aggregate Spidroin 1 and 2 (AgSp1 and AgSp2) (Choresh et al., 2009; Collin et al., 2016; Stellwagen and Renberg, 2019). We sequenced part of AgSp1 and discovered the complete, though unpolished, AgSp2 from *M. phrynosoma*. AgSp2 has a generally similar structure to that of *A. trifasciata*, except it is three times larger with more than 60 kilobases of coding sequence, and lacks the glutamine

rich regions. The repeats of both glue proteins are similar to those reported in other species.

Repeat variant structure of bolas spider aggregate spidroins

Due to the extreme repetition of the aggregate spidroin genes in the bolas spider, we relied on the presence of variable DNA bases in repeating sequence units of the gene for alignment. The translational impact of these base variations is minimal for AgSp2, which had only one variant that led to a synonymous substitution. This result speaks to the overwhelmingly conserved structure of the *M. phrynosoma* AgSp2, even in comparison with the AgSp2 gene of *A. trifasciata*, which lacks the second order, sub-repetition found in *M. phrynosoma*.

We found six variable bases that were well-characterized by reads of repeat variants in the *M. phrynosoma* AgSp1. Half of these resulted in synonymous substitutions. The other half produced nonsynonymous substitutions that involved transitions of amino acids with either polar (Ser, Gln) or hydrophobic (Leu) side chains to those with positively-charged (Arg) or hydrophobic side chains (Phe). That the charge and polarity of these residues is not maintained in repeat variants suggests that the these nonsynonymous substitutions may have functional significance for AgSp1 protein. This is particularly true for residues that vary in post-translational modification, particularly *O*-glycosylation of serine, which confers stickiness (Tillinghast et al., 1993; Dreesbach et al., 1983). The functional distinguishment of the two aggregate spidroin products remains unclear, but it seems apparent that mechanisms of purifying selection and/or gene conversion (Garb et al., 2007) for repeat conservancy are not equally applied to the separate gene copies.

Sequence comparison of aggregate spidroins

Interestingly, we found that the AgSp2 of *M. phrynosoma* lacked the distinctive glutamine-rich regions found in AgSp2 of *A. trifasciata*. It has been hypothesized that glutamine promotes self-aggregation of silk proteins into fibers (Geurts et al., 2010). However, unlike the aggregate glues of orb weavers that adhere and then stretch to help dissipate energy and retain fast-moving prey, *M. phrynosoma* glue must spread rapidly under moth scales before solidifying (Diaz et al., 2018, 2020), making protein aggregation a potential hindrance to the accelerated glue flow. Additionally, protein aggregation is highly concentration dependent (Poulson et al., 2020), whether by the action of catalysts (Braun et al., 2013; Sandberg et al., 2008) or via self-assembly under particular temperatures (Bansil and Turner, 2006). The aqueous solution that surrounds the bulk of the central protein core of a droplet of spider aggregate glue contains low molecular weight compounds and salts, which are known for their role in ambient water absorption and glue protein lubrication (Vollrath et al., 1990; Sahni et al., 2014). Furthermore, salt concentration has been shown to be an important factor for both solubility and aggregation for other spidroins (Knight and Vollrath, 2001; Eisoldt et al., 2010), however their contribution to aggregation of spidroin glues has not been investigated. While *M. phrynosoma* aggregate spidroins are extremely large, it it not known how the density of the glue's proteins or concentration of other compounds in the aqueous material compares between species, or the role this variation might play in functional differences. Longer protein length may provide the protein linkage necessary for adhesion, while limiting concentrations that would result in detrimental viscous aggregations, inhibiting flow on prey. In order to evaluate this hypothesis, glycoprotein concentrations of *M. phrynosoma* could be estimated and compared to those from more traditional orb weaving glue droplets. After size correction, we would expect *M. phrynosoma* glues to be less concentrated than droplets of orb weaver spiders. There is likely a delicate balance between protein size and concentration, as well as the concentration of other small molecules within glue droplets that help regulate protein behavior. Future focus on the role of aggregation in regulating solubility and viscosity would be important for evaluating the suitability of spidroins for medical application (Abdelrahman et al., 2020) or as models of fibril behavior (Kenney et al., 2002).

Repeats across *M. phrynosoma* aggregate spidroin genes contain variable bases at specific positions (Fig. 2). The unique pattern of these variations from repeat to repeat allows alignment of long reads that is not otherwise possible. However, likely due to the error rate within long reads and the length of these repetitive sequences, the alignment algorithms we applied were unable to identify and utilize nucleotide variations across repeats for alignment. Furthermore, using short sequence reads mapped to long reads for error correction can lead to mis-mapping. Any repeats with a few non-typical bases (which are common near spidroin termini and can be difficult to discern from sequencing errors) become overwhelmed by the large number of reads from the central repeats that otherwise map successfully, resulting in an incorrect consensus (Fig. 4).

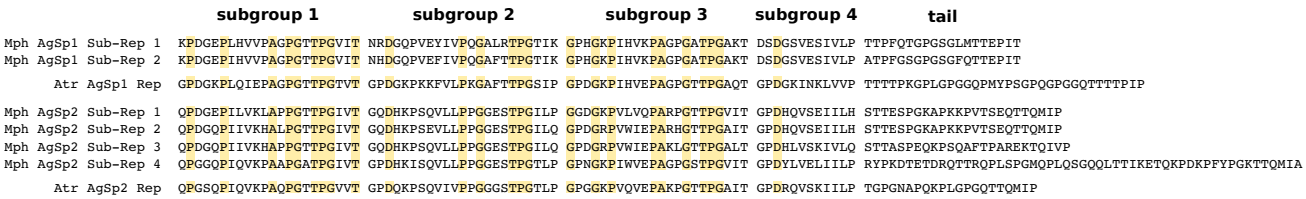

Fig. 3. Translated AgSp1 and AgSp2 repeat motifs from the bolas spider *Mastophora phryosoma* (Mph) and orb weaving spider *Argiope trifasciata* (Atr). *M. phrynosoma* repeat motifs are longer and consist of sub-repeats that are each similar to entire repeats of *A. trifasciata*. Variations from sub-repeat to sub-repeat result in a longer overall repeating unit, however *M. phrynosoma* sub-repeats are still highly similar to full repeats of *A. trifasciata*. Each sub-repeat or repeat is separated into four subgroups and a tail region as in (Stellwagen and Renberg, 2019). The number of amino acid residues within the subgroups is highly conserved across species, and this pattern in maintained in *M. phrynosoma*. The tail region is highly variable, and there are several different tail organizations across sub-repeats of *M. phrynosoma* aggregate spidroins. Yellow highlight indicates conserved residues across both species and genes.

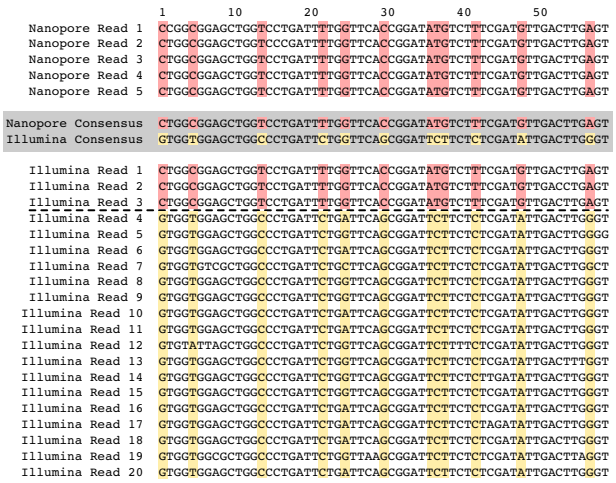

Fig. 4. Nanopore long reads and corresponding alignment consensus (Nanopore Consensus), Illumina short reads mapped to the Nanopore Consensus, and the consensus formed from Illumina read mapping (Illumina Consensus). An initial consensus sequence (Nanopore Consensus, grey box) for a section of a terminal-adjacent repeat (which typically contain a few base variations compared to the bulk of central repeats) from a spidroin of a species that was not part of this study, is obtained by aligning Nanopore reads 1-5. Depth of coverage using Nanopore long reads is not sufficient to resolve all errors with confidence and are therefore often corrected using Illumina data. After mapping RNA-seq Illumina short reads (GenBank accession: SRR5131057) to the Nanopore Consensus to correct sequencing errors (for example, Fig. 2A, blue highlighted bases), Illumina reads 1-3 (above dotted line) map correctly. However, Illumina reads 4-20 (below dotted line) incorrectly map due to general repeat similarity, overwhelming the few accurate reads, and producing a new incorrect consensus (Illumina Consensus, grey box). Red highlights over read bases indicate consistency with the Nanopore Consensus, while yellow highlights indicate consistency with the mis-mapped Illumina reads 4-20.

Base-to-base confidence in extreme-length sequences like aggregate spidroins can currently only be derived from adequate long-read sequencing coverage of 30-40x. Moreover, the reads at this depth of coverage must be long enough to be placed confidently, as we were able to do only for AgSp2 in this study. This methodology is therefore currently out of the realm of standard sequencing projects for organisms with larger genomes like those of spiders (usually larger than 1-2 GBp; (Gregory and Shorthouse, 2003)). Even though sufficient depth of coverage is possible for most regions of the genome, achieving this depth with sufficiently long reads necessary to assemble extreme-length repetitive spidroins is currently a great challenge. Furthermore, while the technology is available for tedious, by-hand reconstruction and correction of these genes with a combination of long and short reads, building a substantial catalog of high quality, full-length glue genes will come only after obtaining consistently long reads either of higher quality or higher depth becomes more feasible.

Conclusion

Aggregate spidroins are the largest members of the spider silk gene family, and the bolas spider has now extended the known length of these genes to an incredible >60 kb of coding sequence. Using long read sequencing and single nucleotide variations across repeats, the massive AgSp2 gene size was discovered and the repeat units of both AgSp1 and AgSp2 were resolved. While sequencing technology still struggles to allow for consistent resolution of extreme-length repetitive sequences from many species, we are able to begin investigating repeat number and size, and, subsequently, their potential contributions to glue functionality. However, until either long-read error rates improve or greater depth of coverage is more cost effective, broad-scale data sets of these large, repetitive spidroins from many species will remain out of reach, hampering efforts to uncover the fundamental adhesive gene structures for bio-inspired material applications. Moreover, a broader question remains: why are aggregate spidroin repeats so conserved, and yet so repetitive? The order of evolutionary events leading to the genomic duplication and functional specialization remain in question. Furthermore, the means by which repetitive units are serially modified invites more substantial research into the mechanisms of gene conversion or unequal crossing-over (Garb et al., 2007) that could be capable of exerting such extreme homogenization.

Competing interests

The authors declare no competing interests.

Author contributions statement

S.D.S. conceived and conducted the experiment(s), S.D.S and M.B. analysed the results, wrote, and reviewed the manuscript.

Funding

This research was funded in part by the Herb Levi Memorial Fund for Arachnological Research administered by the American Arachnological Society.

Acknowledgments

We sincerely thank Bonnie Ott for her help locating *M. phrynosoma* specimens. Additionally we thank the Living Farm Heritage Museum and the Howard County Conservancy in Maryland for generous permission to collect specimens on their premises. We also wish to acknowledge SICB for supporting our symposium and three anonymous reviewers who provided helpful comments to improve this manuscript.

References

S. Abdelrahman, M. Alghrably, J. I. Lachowicz, A. H. Emwas, C. A. E. Hauser, and M. Jaremko. What doesn't kill you makes you stronger: Future applications of amyloid aggregates in biomedicine. *Molecules*, 25(22):5245, 2020.

C. H. Apstein. Bau und Function der Spinnendrusen der Araneida. *Archiv fur Naturgeschichte*, page 29, 1889.

S. Argintean, J. Chen, M. Kim, and A. M. F. Moore. Resilient silk captures prey in black widow cobwebs. *Applied Physics A*, 82(2):235–241, 2006.

R. Bansil and B. S. Turner. Mucin structure, aggregation, physiological functions and biomedical applications. *Current Opinion in Colloid Interface Science*, 11(2-3): 164–170, 2006.

S. P. Benjamin and S. Zschokke. Untangling the tangle-web: Web construction behavior of the comb-footed spider Steatoda triangulosa and comments on phylogenetic implications (Araneae: Theridiidae). *Journal of Insect Behavior*, 15(6):791–809, 2002.

T. Blackledge, N. Scharff, J. A. Coddington, T. Szuts, J. W. Wenzel, C. Y. Hayashi, and I. Agnarsson. Reconstructing web evolution and spider diversification in the molecular era. *PNAS*, 106(13):5229–5234, 2009.

M. Braun, M. Menges, F. Opoku, and A. Smith. The relative contribution of calcium, zinc and oxidation-based cross-links to the stiffness of *Arion subfuscus* glue. *Journal of Experimental Biology*, 216:1475–1483, 2013.

O. Choresh, B. Bayarmagnai, and R. V. Lewis. Spider Web Glue: Two proteins expressed from opposite strands of the same DNA sequence. *Biomacromolecules*, 10(10): 2852–2856, 2009.

M. A. Collin, T. H. Clarke, N. A. Ayoub, and C. Y. Hayashi. Evidence from multiple species that spider silk glue component ASG2 is a spidroin. *Scientific Reports*, 6: 21589, 2016.

C. Diaz, A. Tanikawa, T. Miyashita, G. Amarpuri, D. Jain, A. Dhinojwala, and T. A. Blackledge. Supersaturation with water explains the unusual adhesion of aggregate glue in the webs of the moth-specialist spider, *Cyrtarachne akirai*. *Royal Society Open Science*, 5(11):181296, 2018.

C. Diaz, D. Maksuta, G. Amarpuri, A. Tanikawa, T. Miyashita, A. Dhinojwala, and T. A. Blackledge. The moth specialist spider *Cyrtarachne akirai* uses prey scales to increase adhesion. *Journal of The Royal Society Interface*, 17(162):20190792, 2020.

K. Dreesbach, G. Uhlenbruck, and E. K. Tillinghast. Carbohydrates of the trypsin soluble fraction of the orb web of *Argiope trifasciata*. *Insect Biochemistry*, 13(6):627–631, 1983.

W. G. Eberhard. Aggressive chemical mimicry by a bolas spider. *Science*, 198(4322):1173–1175, 1977.

L. Eisoldt, J. G. Hardy, M. Heim, and T. R. Scheibel. The role of salt and shear on the storage and assembly of spider silk proteins. *Journal of Structural Biology*, 170(2):413–419, 2010.

J. E. Garb, T. DiMauro, R. V. Lewis, and C. Y. Hayashi. Expansion and intragenic homogenization of spider silk genes since the triassic: Evidence from mygalomorphae (tarantulas and their kin) spidroins. *Molecular Biology and Evolution*, 24(11):2454–2464, 2007.

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

T. R. Gregory and D. P. Shorthouse. Genome sizes of spiders. *Journal of Heredity*, 94(4):285–290, 2003.

C. E. Hutchinson. A BOLAS-THROWING SPIDER. *Scientific American*, 89(10):172, 1903.

M. Kearse, R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes, and A. Drummond. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(12):1647–1649, 2012.

S. Kenney, D. Knight, M. J. Wise, and F. Vollrath. Amyloidogenic nature of spider silk. *European Journal of Biochemistry*, 269(16):4159–4163, 2002.

D. P. Knight and F. Vollrath. Changes in element composition along the spinning duct in a Nephila spider. *Naturwissenschaften*, 88(4):179–182, 2001.

H. W. Levi. The bolas spiders of the genus Mastophora (Araneae: Araneidae). *Bulletin of the Museum of Comparative Zoology*, 106(4):221–254, 1955.

B. D. Opell, S. E. Karinshak, and M. A. Sigler. Humidity affects the extensibility of an orb-weaving spider's viscous thread droplets. *Journal of Experimental Biology*, 214(17): 2988–2993, 2011.

B. D. Opell, S. E. Karinshak, and M. A. Sigler. Environmental response and adaptation of glycoprotein glue within the droplets of viscous prey capture threads from araneoid spider orb-webs. *Journal of Experimental Biology*, 216: 3023–3034, 2013.

B. G. Poulson, K. Szczepski, J. I. Lachowicz, L. Jaremko, A.-H. Emwas, and M. Jaremko. Aggregation of biologically important peptides and proteins: inhibition or acceleration depending on protein and metal ion concentrations. *Royal Society of Chemistry Advances*, 10:215–227, 2020.

V. Sahni, T. A. Blackledge, and A. Dhinojwala. Viscoelastic solids explain spider web stickiness. *Nature Communications*, 1(19):1–4, 2010.

V. Sahni, T. A. Blackledge, and A. Dhinojwala. *A review on spider silk adhesion*, volume 87. 2011a. ISBN 6.

V. Sahni, T. A. Blackledge, and A. Dhinojwala. Changes in the adhesive properties of spider aggregate glue during the evolution of cobwebs. *Scientific Reports*, 1:41, 2011b.

V. Sahni, T. Miyoshi, K. Chen, D. Jain, S. J. Blamires, T. A. Blackledge, and A. Dhinojwala. Direct solvation of glycoproteins by salts in spider silk glues enhances adhesion and helps to explain the evolution of modern spider orb webs. *Biomacromolecules*, 15(4):1225–1232, 2014.

T. Sandberg, H. Blom, and K. D. Caldwell. Potential use of mucins as biomaterial coatings. i. fractionation, characterization, and model adsorption of bovine, porcine, and human mucins. *Journal of Biomedical Materials Research*, 91A(3):762–772, 2008.

S. Singla, G. Amarpuri, N. Dhopatkar, T. A. Blackledge, and A. Dhinojwala. Hygroscopic compounds in spider aggregate glue remove interfacial water to maintain adhesion in humid conditions. *Nature Communications*, 9:1890, 2018.

S. D. Stellwagen and R. L. Renberg. Toward spider glue: Long read scaffolding for extreme length and repetitious silk family genes agsp1 and agsp2 with insights into functional adaptation. *G3: Genes, Genomes, Genetics*, 9(6):1909–1919, 2019.

E. K. Tillinghast, M. A. Townley, T. N. Wight, G. Uhlenbrock, and E. Janssen. The adhesive glycoprotein of the orb web of *Argiope aurantia* (araneae, araneidae). *Materials Research Society Symposium Proceedings*, 292: 9–23, 1993.

F. Vollrath, W. J. Fairbrother, R. J. P. Williams, E. K. Tillinghast, D. T. Bernstein, K. S. Gallagher, and M. A. Townley. Compounds in the droplets of the orb spider's viscid spiral. *Nature*, 345:526–528, 1990.

H. Wiehle. Neue beiträge zur kenntnis des Fanggewebes der Spinnen aus den familien argiopidae, uloboridae und theridiidae. *Zeitschrift für Morphologie und Ökologie der Tiere*, 22(2-3):349–400, 1931.