PNAS NEXUS

PNAS Nexus, 2025, 4, pgaf318

https://doi.org/10.1093/pnasnexus/pgaf318
Advance access publication 4 November 2025
Research Report

Triangle weaver spiders construct spring-loaded webs using a novel set of genes for exceptionally proline-rich silk

Steven Casey [a], Sandra M. Correa-Garhwal [b,c], Richard H. Baker [b], Jay A. Stafstrom [d], Hudson Kern Reeve [d], Cheryl Y. Hayashi [b] and Jessica E. Garb [a,*]

[a]Department of Biological Sciences, University of Massachusetts Lowell, 198 Riverside Street, Lowell, MA 01854, USA
[b]Division of Invertebrate Zoology and Institute for Comparative Genomics, American Museum of Natural History, 200 Central Park West, New York, NY 10024, USA
[c]Department of Biology, Suffolk University, 8 Ashburton Place, Boston, MA 02108, USA
[d]Department of Neurobiology and Behavior, Cornell University, 215 Tower Road, Ithaca, NY 14853, USA
*To whom correspondence should be addressed: Email: jessica_garb@uml.edu
Edited By Li-Jun Ma

Abstract

Spiders amplify their physical capabilities by synthesizing multiple high performing silks. Renowned for its toughness, major ampullate (MA) silk composes the spiderweb frame, providing support and absorbing high-energy impacts. In ecribellate orb-weavers, proline-rich motifs in MaSp2 proteins of MA silk are linked to a range of mechanical properties, including extensibility, elasticity, stiffness, and supercontraction. We show a modification of these motifs outside of this clade in a spider that constructs a spring-loaded web. The triangle weaver spider *Hyptiotes cavatus* (family Uloboridae) stores energy in the support lines of its triangular web, then rapidly releases the tension to catapult forward, collapsing the web around prey. *Hyptiotes* has an expanded set of *MaSp2* genes which encode proteins with far higher proline contents than typical MaSp2. The predominant GPGPQ motifs present in *Hyptiotes* spidroins also occur abundantly in MaSp sequences of distantly related spiders that produce the most extensible dragline, implying silk protein convergence. Proline-rich MaSp2 proteins constitute half of all MA gland expression in *Hyptiotes*, and we show that the resulting fibers are the most proline-rich spider silk measured to date. This unique silk composition suggests a functional importance that may facilitate the spring-loaded prey capture mechanism of this species' web and may inspire the design of novel biomaterials using protein engineering.

Keywords: spider silk, spidroin, proline, gene family evolution, dragline

Significance Statement

Spider silks are famed for their exceptional material properties. Some spider species construct highly modified webs with unusual mechanical demands, but it is unknown whether these adaptations coincide with a modification of their silk's biochemical composition. *Hyptiotes cavatus* uses its triangle web as a "tool" for external power amplification, catapulting itself forward, jolting sticky threads onto intercepted prey. Here, we show that the spring-loaded *Hyptiotes* silk has the highest proline content of any previously characterized spider silk. Evolutionary genomic comparisons reveal that expansion and modification of silk genes drive this unique composition. Given proline's association with silk extensibility, the evolution of *Hyptiotes*' proline-rich silk may facilitate its spring-loading mechanism and provide a blueprint for novel biomaterials.

Introduction

Spiders (Araneae) are an extremely diverse order found in almost all terrestrial as well as some aquatic ecosystems. Their success as predators can be attributed to their extensive use of silk to augment their physical capabilities (1) (2). From protection to prey capture to dispersal through the air, spiders have developed a wide range of ways to use their silk. Each purpose often coincides with a dedicated silk type produced in a distinct gland; a single spider can have as many as eight different types of silk glands (3, 4). Dragline silk, also known as major ampullate (MA) silk, is produced by all true spiders (Araneomorphae) and is the most well studied of spider silks due to its extraordinary toughness (5). It is typically used as both a safety line traversing the environment and as structural support in webs, such as the radial lines in iconic orb webs.

Competing Interests: The authors declare no competing interests.
Received: February 5, 2025. **Accepted:** September 24, 2025

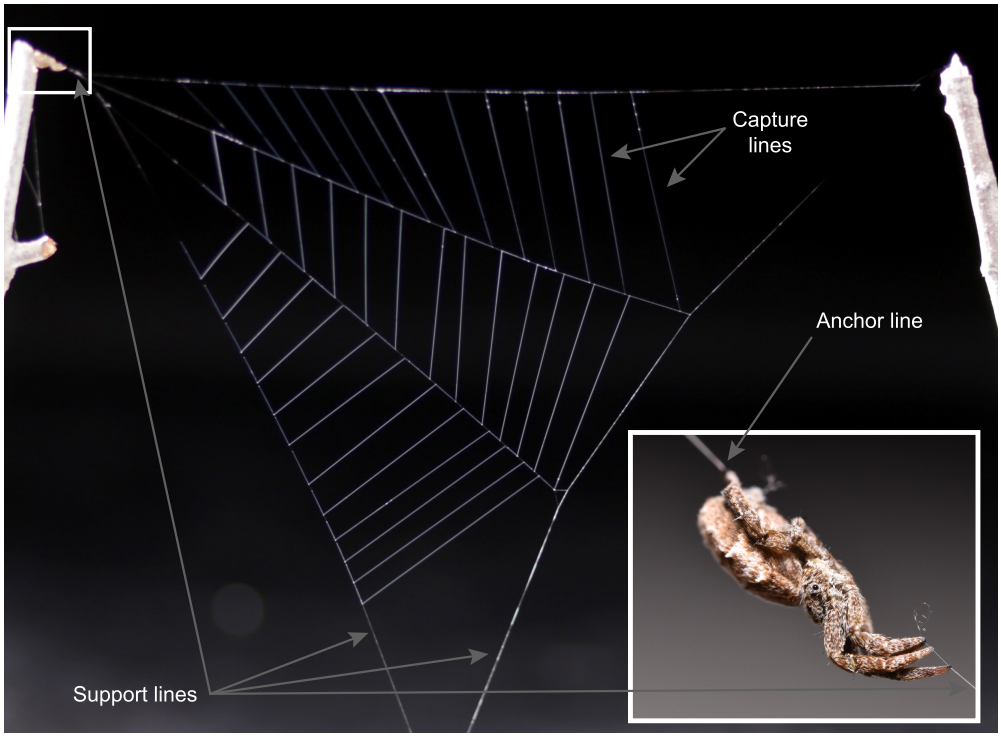

Fig. 1. *Hyptiotes cavatus* web. *Hyptiotes* builds a three-sector triangle web. The spider pulls the web taught by walking backwards on the anchor line. It quickly springs forward and releases the energy stored in its support lines in short bursts to shake and collapse the capture threads covered in wispy, Velcro-like cribellate silk onto prey. Web photo credit: Sarah Han; inset: J.E.G.

The triangle weaver spider *H. cavatus* uses its MA silk to store and release energy in a remarkable way. Species in the genus *Hyptiotes* construct a three-sided web, considered a reduced form of the orb webs produced by other genera in the same family, Uloboridae (Fig. 1) (6). *Hyptiotes* uses its body as an integral part of the web, acting as a bridge between two MA silk fibers (7, 8). With its front legs, the spider holds onto a support line attached to one vertex of the triangle. Its back legs hold a separate anchor line cemented to a supporting tree or shrub. This anchor line runs from its spinnerets (silk-extruding structures) to the supporting foliage. To increase tension in the web, *Hyptiotes* pulls on the anchor line, creating coils of loose silk by walking its fourth legs backward along the fiber, continuing this motion until the spider rests against the supporting foliage. When prey flies into the web, *Hyptiotes* loosens its grip on the anchor line, releasing the energy stored in the web in extremely quick successive bursts. These bursts cause the web to collapse, tangling the prey as it comes into contact with more of the web's puffy, Velcro-like cribellate capture silk strands (9, 10). The rates of acceleration reached using the energy stored in the MA fibers of the web are well beyond the limits of the muscles in the spider's body, illustrating a unique example of power amplification by tool use (7).

Spider silks are made of proteins called spidroins. The ability of spiders to synthesize a diverse range of silk materials and construct web architectures is due to the evolutionary expansion of a single gene family that underlies spidroin production (11–13). Spidroin sequences are highly repetitive, consisting of repeating blocks of amino acids. The repetition and composition of these amino acid motifs largely determine the various mechanical properties of the different silk types (14–16). For example, in the well-studied araneoid orb weavers, MA silk is mainly composed of spidroins MaSp1 and MaSp2 (17–19). Both MaSp1 and MaSp2 contain homopolymeric alanine (polyalanine) regions that form crystalline beta-sheet structures in the silk, imparting strength to the fibers (14, 15). Amorphous glycine-rich regions connect these beta sheets (20) and are associated with silk extensibility (21–23). MaSp2 confers extensibility to dragline through GPGX or GPGXX motifs (where X is typically G, Q, or Y) in its amorphous region, proposed to form beta turns involved in nanospring structures (11, 14, 16) (24). GPGX motifs are also found in abundance in the spidroins of orb-weaver flagelliform silk, composing the highly extensible orb capture spiral, reinforcing the link between proline and silk extensibility.

Proline is the only amino acid where the side chain forms a ring with the protein backbone. This pyrrolidine ring structure imparts a conformational rigidity and an inability for amide hydrogen bonding; thus, proline has been long known to have significant impacts on protein secondary structures in general (25, 26), and on spider silk properties in particular (27–30). Proline content in spider silk has been linked to fiber extensibility (21), supercontraction (27), and elasticity (28). Biophysical studies suggest that proline increases supercontraction by disrupting more ordered glycine-rich chains in the amorphous region of spider silk (28, 31) and by steric exclusion of hydrogen bonds within the molecule (32). A meta-analysis (29) of spider silk composition and mechanical performance found that proline content in MA silk is correlated with increasing toughness (via increased extensibility) and decreasing stiffness (Young's modulus). In canonical models of MA silk, proline occurs almost exclusively in the form of GPGX or GPGXX motifs (11) in MaSp2 proteins, proposed to form elastin-like beta spirals (14). The conservation of these motifs across species implies the importance of proline spacing in silk fiber mechanics and structure.

MA spidroins have diversified beyond this canonical proline-poor MaSp1 and proline-rich MaSp2 model. Recent genomic studies of orb weavers in the Araneoidea superfamily reveal that some

species have as many as 14 MaSp genes (33). This expansion suggests an investment in MA silk construction that could explain the great toughness of araneoid dragline (24, 33–35). In contrast, little is known about dragline silk composition outside of araneoids. Compared with araneoid orb weavers, MA silk proteins from other spiders were thought traditionally to have limited diversity (12, 36). These nonaraneoids include the Uloboridae family of cribellate orb weavers which contains *Hyptiotes*.

A recent genome from another species in the Uloboridae family, *Uloborus diversus*, contains one *MaSp1* and two *MaSp2* genes, supporting the view of limited MaSp protein diversity outside of araneoid spiders (37, 38). Since *U. diversus* makes a horizontal web that remains stationary, the mechanical requirements of the dragline silk in this web are markedly different than that of the confamilial triangle weaver *Hyptiotes*. The prolonged energy storage, semi-ordered coiling of the anchor line that does not tangle upon release, and cycle of rapid release of tension and subsequent resetting of the *Hyptiotes* web impose seemingly unique mechanical demands upon its MA fibers (7, 39).

Given its notable web structure and MA silk function, we sequenced the genome of *H. cavatus* to characterize its spidroins in search of molecular enhancements in silk properties that complement the web's spring-loaded mechanism. We discovered that *H. cavatus* has expanded its repertoire of *MaSp* genes, and these genes are radically modified to have increased proline content in their predicted proteins. We also found *H. cavatus* dragline fibers to have the highest proline content measured to date in spiders, consistent with the MaSp sequence characteristics, suggesting traits that may be important to the function of this web. Understanding the structure–function relationship between these proteins and the remarkable web is not only important for understanding the evolutionary novelties that enable ecological specialization but also for designing high-performance, silk-inspired biomaterials.

Results

The *H. cavatus* genome was assembled using Oxford nanopore long reads with Flye (40), resulting in a total length of 4.26 Gb and contig N50 of 107 kb. We evaluated the quality of the genome using a search for Benchmark Universal Single Copy Orthologues (BUSCOs) and found 95.4% of these genes (81.8% complete and single-copy, 4.9% complete and duplicated, 8.7% fragmented, see Fig. S1). Spidroin genes were verified as distinct loci by examining flanking sequences and other genes on spidroin-containing contigs. Twenty spidroin genes were recovered from the genome (Fig. 2), of which 17 were complete.

We classified these genes based on a phylogeny of the terminal domains (Fig. S2) as well as their repetitive structure. Spidroin genes were recovered for each expected spidroin type. The list of full-length spidroins consists of six MA genes (*MaSp*—used in dragline and web supporting lines), four minor ampullate genes (*MiSp*—possibly used in web temporary scaffolding), and one each of tubuliform (*TuSp*—egg case silk), aciniform (*AcSp*—wrapping silk), and pseudoflagelliform genes (*Pflag*—support in capture silk). Spidroin family members labeled *AmSp* and *SpvC* eluded traditional spidroin classification but were named following the same convention in *U. diversus* (37). The cribellar spidroin (*CrSp*—puffy, sticky capture silk) appeared to be incompletely assembled into two fragments, which are concatenated here and in further analyses. A small C-terminal fragment of pyriform spidroin (*PySp*) was recovered, which appears at the end of a contig. Since this is a short fragment resulting from an incomplete assembly of the gene, *PySp* was not included in subsequent amino acid composition analyses of its predicted translation. In addition, two genes proposed as candidates for paracribellar spidroins, which connect capture silk threads in uloborids, were recovered (*SpvA* and *SpvB*) (37).

One additional N-terminal fragment of an MA spidroin (*MaSp2.5*) was found. This gene appears in the center of a contig and contains a stop codon immediately after the N-terminal domain and exhibits very low expression (Dataset S1). This suggests pseudogenization, and so MaSp2.5 was not included in further analyses.

Duplicated *MaSp2* silk genes evolved proline-rich motifs

The *Hyptiotes* genome contains double to triple the number of *MaSp* genes compared with other Uloboridae genomes published to date. We recovered six distinct full-length *Hyptiotes MaSp* spidroins (Figs. 2–4), while three were found in recent *U. diversus* genomes (37, 38) and two in each genome of three *Octonoba* species (42). The six *Hyptiotes* MaSp proteins were classified as either MaSp1 or MaSp2 based on the proline content of their repetitive regions. Using this method, one protein has a low-proline content and abundant GGx motifs typically associated with MaSp1 and is labeled as such. The remaining five proline-rich sequences are designated as MaSp2 variants (Fig. 3).

To illustrate the proline-rich nature of *H. cavatus* MaSp2 proteins, we categorized the repeats present in *H. cavatus* and *U. diversus* MaSp sequences. MaSp2-like repeats contain canonical MaSp2 motifs found in other uloborids and in ecribellate orb weavers and cobweb weavers (e.g. *Argiope*, *Latrodectus*, and *Trichonephila*): GPGx or GPGxx (where x is usually Q, G, or S) and a poly-A motif. In comparison, MaSp4-like sequences (first described in the MA proteins of the distantly related Darwin's bark spider *Caerostris darwini*) (24) consist of the more proline-rich GPGPQ motif and an S/T-rich motif. The extra proline in the GPGPQ motif compared with GPGx or GPGxx has been hypothesized to contribute to the increased extensibility of *C. darwini* dragline (24). The third type of *H. cavatus* MaSp2 protein contains a mixture of MaSp4-like GPGPQ motifs and the MaSp2-like poly-A motif.

Surprisingly, there are no instances of canonical MaSp2-like repeats in *Hyptiotes* MaSp2 sequences; the repetitive regions are entirely composed of proline-rich MaSp4-like and MaSp2–MaSp4 hybrid repeats (Fig. 3). MaSp2.2 is entirely composed of MaSp4-like repeats. Representative repeats for each MaSp spidroin in *Hyptiotes* and *Uloborus* are also shown in Fig. 3. The length and amino acid composition of these repeats varies within each spidroin. Usually, the number of GPGx or GPGPQ repeats varies by one to three motifs between the poly-A- or S/T-rich domains. Full gene and protein sequences are shown in Datasets S2 and S3.

Proline-rich sequences arise in multiple MaSp lineages

The MaSp2 sequences in *Hyptiotes* are extraordinarily proline rich. To understand the relationship of these *MaSps* to the silk genes in other spiders, we mapped the proportions of proline in the repetitive regions as characters on a spidroin gene tree and performed ancestral state reconstruction (Fig. 4). This phylogeny contains sequences from genomes with complete spidroin catalogs of two additional cribellate and three ecribellate orb weavers. Compared with other uloborids, the *Hyptiotes* genome reveals a *MaSp2* gene expansion that is associated with increased proline content. *Octonoba sybotides* and *U. diversus* (representatives of two other uloborid genera) have one and two MaSp2 spidroins, respectively, containing between 15 and 20% proline. In contrast, the

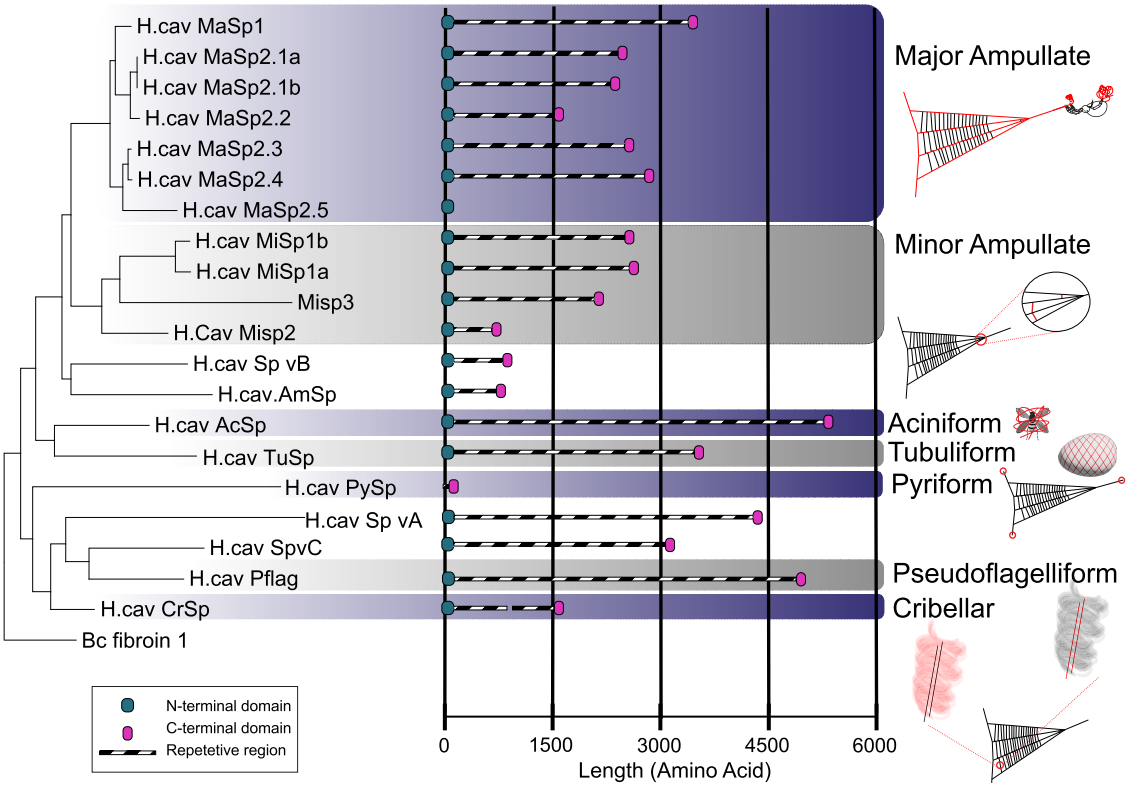

Fig. 2. Classifications and lengths of *H. cavatus* spidroins. Gene tree showing the relationships of the spidroins inferred from the *H. cavatus* genome. The tree was constructed by maximum likelihood analysis of concatenated N- and C-terminal domains, with *B. californicum* fibroin 1 as the outgroup. The length of each spidroin is shown, along with the presence of the N-terminal domain, C terminal domain, and/or repetitive region. For the pyriform spidroin (PySp), only the C-terminal fragment was recovered, and the cribellar spidroin (CrSp) appeared to be incompletely assembled into two fragments, which are concatenated here and in further analyses.

five MaSp2 sequences of *Hyptiotes* range from 23 to 30% proline. This prevalence is even higher than in any of the included araneoid flagelliform or aggregate spidroins (8 to 13%), which are known for their high-proline content and respective association with very extensible capture spiral or viscous glue silk (13, 43). To our knowledge, the proline content of the repetitive region of *H. cavatus* MaSp2.1a (~30%) is the most proline-rich spidroin in Uloboridae, and across spiders is only matched by araneid MaSp4, which is the most proline-rich spidroin reported to date.

Proline-rich MA spidroins have been characterized in other spiders, mainly ecribellate orb weavers. The phylogeny of the terminal domains (Fig. 4, detailed tree in Fig. S3) shows proline-rich MaSp sequences in different species are spread across divergent spider lineages. MaSp2.2 in *Hyptiotes* and MaSp4 in *C. darwini* contain both numerous tandem GPGPQ motifs and S/T-rich motifs, but these spidroins are not closely related based on their terminal domains. These tandem GPGPQ motifs also do not appear in other nonspidroin proteins encoded in the genome of *Hyptiotes* or in any other annotated spider genome. This suggests these extremely similar spidroin repeats in very distant species evolved convergently. High-proline spidroins in *Trichonephila* and *Argiope* also are more closely related to low-proline sequences than to each other, consistent with the rapid divergence of MaSp repetitive sequences (33).

Predominant MaSp2 expression in MA glands results in extremely proline-rich dragline silk

Using RNA-Seq analysis, we found that *MaSp2* genes are highly expressed in the major ampullate glands (Fig. 5A and B). Intriguingly, we found that, along with the cribellar spidroin gene (*CrSp*), these *MaSp2* genes were expressed to varying degrees in cribellar glands (Fig. S4). In MA glands, proline-rich *MaSp2* genes were the most highly expressed spidroin type. As an average across the three MA gland replicates, 71% of spidroin expression consists of the moderately proline-rich MaSp2.3 (23% proline) and MaSp2.4 (24% proline), while the remaining spidroin expression (19%) is mostly the even more proline-rich MaSp 2.1a, 2.1b, and 2.2 (25–28% proline). Low-proline sequences only account for a small amount of MA gland spidroin expression, split primarily between MaSp1 and MiSp3 (in total 10% of expression).

The expression data indicate that *Hyptiotes* produces an extremely proline-rich dragline. We predicted the overall dragline proline content using the amino acid composition for each spidroin and weighted transcript expression proportions (Dataset S1). This expression-based predicted estimate of *Hyptiotes* dragline proline content was calculated at 21.8%. To verify this estimate, we measured amino acid composition directly from pulled dragline silks. Dragline proline content across six individuals ranged from 20.9 to 24.3% and in the three samples that were analyzed for hydroxyproline, none was detected (Table S1, Dataset S1). This proline content coincides with our expression data and is the highest measured proline content in any spider silk (Fig. 5C detailed in Table S2).

To examine the functional implications of *H. cavatus* MaSp diversity and composition, we tested the mechanical performance of pulled dragline fibers. This testing (Table S3, Fig. S5) is within the range of previous measurements for both *H. cavatus* and *Hyptiotes affinis* (9, 51). This reinforces that when using standard

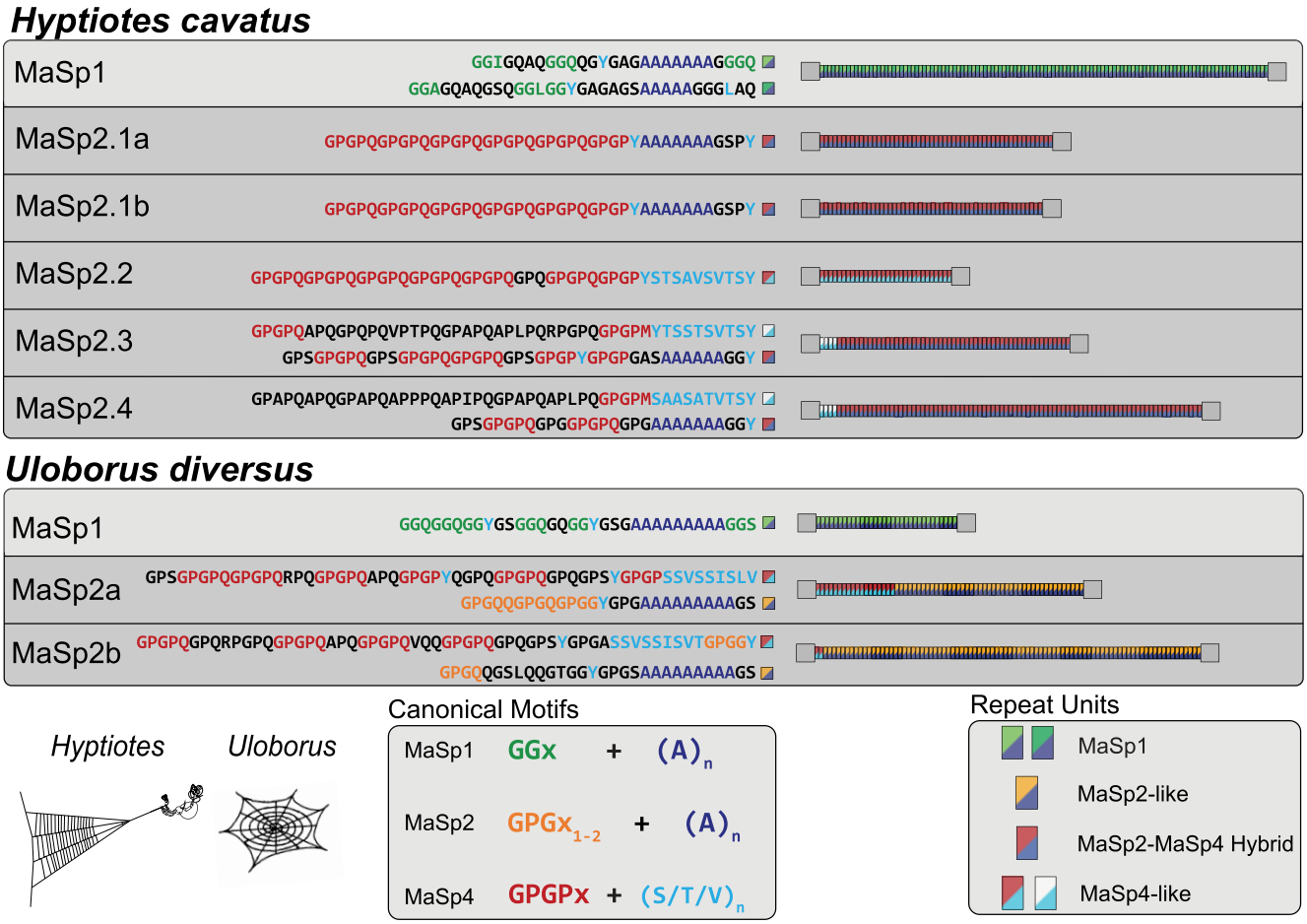

Fig. 3. Repeat structure of MA spidroin sequences in *Hyptiotes* differs from those known in other Uloboridae. The canonical motifs and their organization into repeats of the MaSp sequences in *H. cavatus* and the *U. diversus* genomes are shown. Repeats are categorized and color coded based on the presence of typical motifs found in MaSp1, MaSp2, and MaSp4. Representative repeats are also listed for each type of repeat present in each spidroin.

metrics such as Young's modulus (9.2–10.7 Gpa) and toughness (84–164 MPa) to evaluate *H. cavatus* dragline, it performs within the range of other species in the family Uloboridae as well as across all spiders (5, 9, 51). Contrary to correlations between proline content and extensibility (52), the proline-rich *Hyptiotes* silk does not outperform the silk of other spiders in the same family in this measure. Nevertheless, the MA silk of *Hyptiotes* is distinct in that it contains a higher proportion of proline than silk characterized from any other spider.

Discussion

Proline-rich MA silk evolved with unusual web architecture and mechanics

Hyptiotes cavatus employs a unique prey capture strategy and produces unusually proline-rich MA silk to match. MA silk has evolved extraordinary mechanical properties, and because of this, it has been the focus of considerable research to both understand (53–55) and replicate (56–58) these attributes. Previous studies have shown that *MaSp* genes in ecribellate orb-web weavers have undergone significant expansion and diversification (24, 33, 59). Our *Hyptiotes* genome shows that expansion of *MaSp* genes through recent duplication events is not unique to ecribellate orb weavers (33). Understanding the relationship between function and molecular structure is imperative to understanding silk as a biomaterial (14). *Hyptiotes* actively stretches its triangle web taut, storing energy in its MA silk, and quickly releases it in short, controlled bursts (Fig. 1). Thus, markedly different mechanics and stresses operate on the support lines of *Hyptiotes* webs than exist in those of static webs of other cribellate and ecribellate orb weavers (60). It follows that the primary and secondary structures of *Hyptiotes* spidroins likely evolved in tandem with its novel prey capture strategy.

Many studies have examined the relationships of silk protein primary and secondary structure to its mechanical properties: polyalanine motifs form antiparallel beta sheets that contribute to toughness, connected by amorphous regions (14, 15). The presence of proline in these amorphous regions contributes to the fiber's extensibility (21, 27, 28, 61). Traditionally, MA silk is understood to be composed of two spidroins that are categorized as MaSp1 or MaSp2 depending in part on whether they are low or high in proline, respectively (11, 62). Recent studies have shown that the MaSp clade is more complex than previously known (24, 33, 59). Our phylogenetic analysis of the spidroins of various cribellate and ecribellate orb weavers shows that there is no single evolutionary pathway from MaSp1 and MaSp2. Rather, it appears that MaSp sequences evolve rapidly between high- and low-proline content sequences and repeatedly altered motifs, even converging on some in very distantly related species. For example,

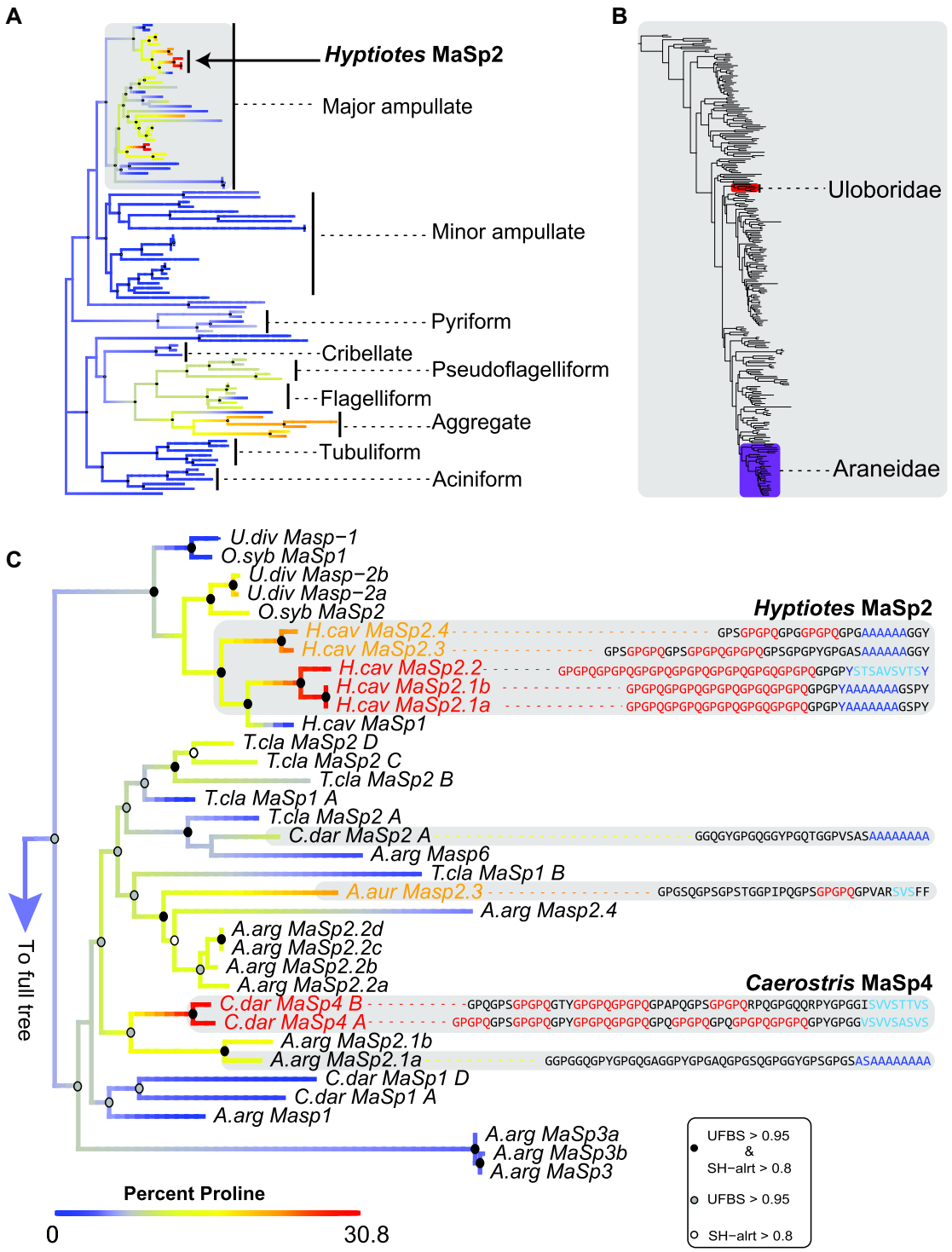

Fig. 4. Ancestral state reconstruction of spidroin repetitive region proline content. A) Spidroin gene tree from concatenated N- and C-terminal sequences from Uloboridae and Araneidae genomes. Colors indicate percentage of proline in the repetitive domain of each spidroin sequence with major spidroin types labeled (full tree in Fig. S3). B) A full spider species phylogeny adapted from Kallal et al. (41) highlighting included families. C) Enlarged view of MA clade in (a). Representative repeats are shown for proline-rich MaSp2 sequences with MaSp4-like motifs highlighted. Included species are *U. diversus* (*U. div*), *C. darwini* (*C.dar*), *A. argentata* and *A. aurantia* (*A.arg*, *A.aur*), *Trichonephila clavipes* (*T.cla*), and *Octonoba sybotides* (*O.syb*). Support values were calculated using ultrafast bootstrapping and SH-alrt.

MaSp2.2 in *Hyptiotes* is remarkably similar in repetitive primary sequence structure to MaSp4 sequences in *C. darwini*, despite their phylogenetic distance as indicated by nonrepetitive terminal domains and the ~214 million years since the two species shared their most recent common ancestor (63). This suggests potential functional convergence since it is the spidroin repetitive sequence that contributes to the mechanical function of spider silk fibers and can rapidly diversify, whereas the terminal domains are relatively conserved and phylogenetically informative (18).

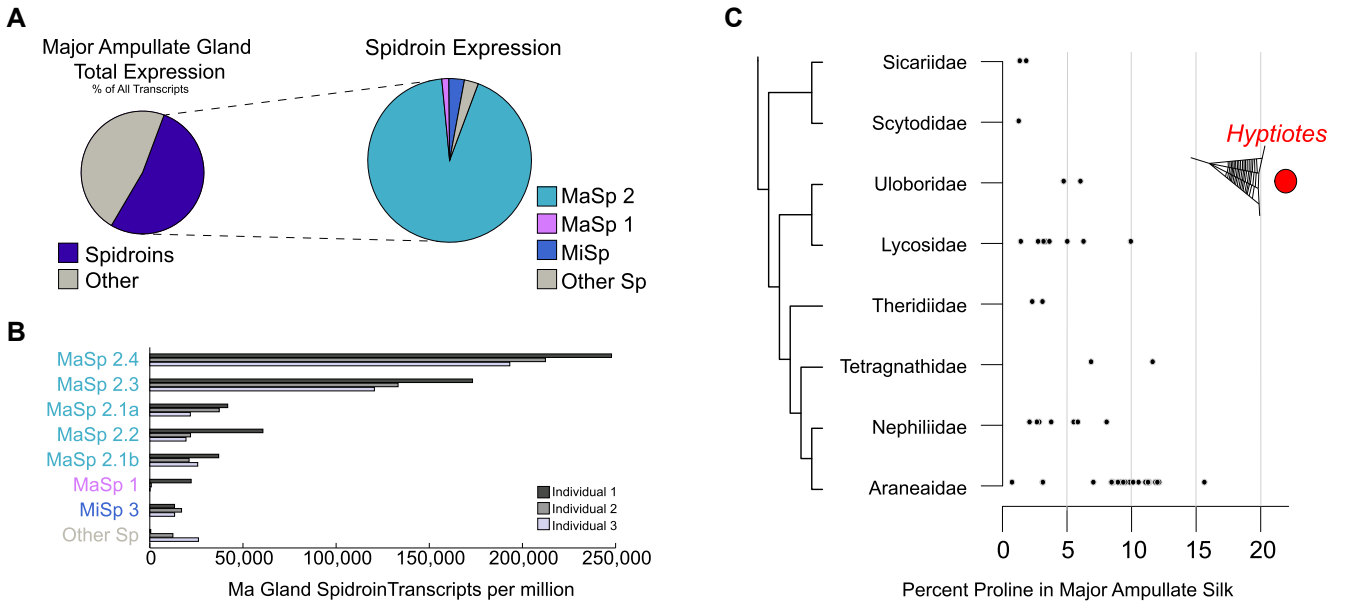

Fig. 5. MA gland expression and measured proline content in dragline. A) Spidroin expression in *H. cavatus* MA glands visualized as a proportion of all expression across three individuals (share of total TPM). B) Detailed MA gland spidroin expression data for each of three individuals. C) The proline content of collected MA silk for spiders where composition data are known (detailed in Table S2) (27, 29, 44–50). Each point represents a separate species, arranged by family and according to the current understanding of their phylogenetic relationships (41).

Caerostris darwini is known to produce the toughest measured dragline of any spider, largely due to increased extensibility, used in webs reaching 2.8 m^2 in area that are suspended by bridge lines as long as 25 m across rivers (64). GPGPQ-rich motifs are present in all highly expressed MA spidroins in *Hyptiotes*, while *C. darwini* blends its MaSp4 sequence with more conventional MaSp1 and MaSp2 (24, 54). The differing proportions of these proteins in the silk of the two species, along with spinning conditions, likely explain why the silk of *Hyptiotes* does not match the extreme toughness of that of *C. darwini*. Nevertheless, the potential convergent development of these motifs in spiders with novel web structures strongly suggests the functional importance of these similar silk proteins in web mechanics.

The expansion of GPGPQ motifs in *Hyptiotes* MaSp2 spidroins, in comparison with much lower frequencies in closely related uloborid MaSp2 proteins, raises the question of how such rapid molecular diversification occurs. We surveyed all predicted proteins from the *Hyptiotes* genome and from every spider genome for which there exists a protein annotation and found there are no instances of tandem GPGPQ motifs outside of MaSp spidroins. Tandem GPGPQ motifs do occur with some frequency in the MaSp spidroins of other uloborids. These motifs appear in the first two to three ensemble repeats of one MaSp in *U. diversus* and in *Octonoba sinensis*, and they occupy the first third of a second *U. diversus* MaSp (Fig. 3). This suggests that the GPGPQ motif existed in low frequencies in the MaSp spidroins of ancestral uloborids and likely expanded in *Hyptiotes*, rather than being recruited from nonspidroin genes. Similarly, in other spider genomes where GPGPQ appears as a single instance within a spidroin, it is most often in the repetitive sequence flanking the terminal domains, which are much less homogenized than intervening repeats, providing a source of novel motifs (65). Studies have suggested that replication slippage, intragenic exchange, and unequal crossing over contribute to spidroin diversification and repeat homogenization (33). It is possible that through such molecular mechanisms the initial, more variable repeats near terminal domains can sweep through the rest of the gene and homogenize through selection (66, 67). Gene duplication can increase the likelihood of this differentiation by either relaxing selection on any given gene or facilitating intergenic exchange of repeat motifs. In other words, GPGPQ motifs may occur sporadically through mutation in the more variable repeats immediately flanking terminal domains but only propagate and homogenize throughout the entire repetitive region when selectively advantageous for silk mechanics. Similar evolutionary mechanisms independently occurring in distant spider lineages could also explain the convergence of the *Caerostris* MaSp4 sequences with *Hyptiotes* GPGPQ-rich MaSp2.

Hyptiotes dragline composition remains unusual when compared with other animal biomaterials

Proline-rich silks and other biomaterials have evolved multiple times among diverse animal lineages. Known to disrupt alpha helices and beta sheets, proline is frequently found in amorphous regions of proteins, where it is essential to the formation and function of fibrous and elastomeric proteins (68). One well-known proline-rich animal protein is collagen, where numerous proline- and hydroxyproline-rich residues stabilize the formation of its characteristic triple helices (69). There are also numerous examples of the evolution of proline-rich biomaterials across invertebrates. For instance, willow sawflies secrete a silk (10.0% proline) of small collagen-like proteins which can form the typical triple helices but without hydroxyproline. Some paper wasps (70) and caddisflies (71) secrete proline-rich proteins to build their nests and retreats (~10% proline). There are also a number of proline-rich secretions in animals that form gel-like substances, either through heavily *O*-glycosylated proteins (spider aggregate silk (72) and human mucin (73) or through proteins with cross-linked short hydrophobic or complementary charged domains (velvet worm slime) (74). While *Hyptiotes* is not unique in producing a proline-rich biomaterial, its dragline remains

extraordinarily proline rich through comparison across metazoa; its 20.9–24.3% measured proline content is only rivaled by the combined total of proline and hydroxyproline in collagen (~23%) (75).

Many elastomeric proteins have variations of proline-rich motifs similar to the GPGPQ motif in *Hyptiotes* silk. Collagen (75) and willow sawfly silk (76) contain GXY (where X is often P) motifs which form triple helixes. Both elastin's VPGXG pentapeptide motifs (77) and gluten's QPGQ motifs in the HMM subunit of glutenin (78) have been proposed to form β-spiral structures which contribute to the elasticity of these materials. The glutamine residues in glutenin likely explain its altered elastic characteristics when compared with elastin, as glutamine forms many hydrogen bonds between these β-spirals and across chains (79) Glutamine is also in the GPGPQ motifs of *Hyptiotes* MaSp2, and the uniformity with which it appears strongly suggests its functional importance. Glutamine is also common in MaSp2 motifs of other spiders, and glutamine hydrogen bonding may similarly stabilize secondary structures within and between silk proteins in dragline. Thus, while proline-rich proteins tend to exhibit elastomeric properties, the function and mechanics of the material they compose are dictated by the frequency of specific proline-rich motifs. While the GPGPQ throughout *Hyptiotes* dragline are similar to those of other elastomeric proteins, they remain distinct in proline spacing, which may result in unique secondary structures and mechanics.

Proline content has multifaceted effect on spider silk mechanics

Mechanical testing of *Hyptiotes* dragline yielded values for extensibility and tensile strength that were within the range of other cribellate orb weavers that produce MA silk with much lower proline contents (9, 51). This contradicts multiple studies correlating increased proline content to increases in extensibility and elasticity and decreases in the stiffness (Young's modulus) of dragline silk (28, 29, 52). Structural models (16, 30, 32) of proline in dragline focus on the predominant GPGX or GPGXX motifs in typical MaSp2 proteins. The increased proximity of prolines in the GPGPQ motifs throughout *Hyptiotes* MaSp sequences may impart additional steric constraints that may lead to modified molecular dynamics, including different secondary structures, which do not follow previously established correlations between proline content and multiple mechanical properties.

Since the *H. cavatus* MaSp motifs and amino acid composition are dissimilar to closely related spider species, an untested property conferred to the silk by this extra proline may facilitate its prey capture strategy of storing and quickly releasing energy in its dragline. One such metric may be minimizing energy loss through the cycle of storage and rapid release. Studies of dragline silk have shown that the amount of energy lost during this cycle is dependent on both the strain rate and the proline content (80). Higher proline content has also been associated with higher work recovery and better strain recovery through multiple cycles (52). *Hyptiotes* releases the tension in its web incredibly quickly and then cycles tension into the same support line. Since it would be advantageous to lose as little energy as possible during this process, there may be selective pressure for higher proline content to conserve the spider's energy in this unique system. The expression of these proline-rich sequences in cribellate glands also suggests their importance in the mechanics of *Hyptiotes* cribellate capture silk. However, because this is the first measure of cribellate gland expression in a spider and there is wide variability across individuals (Fig. S4), the role of MaSps in the structure of *Hyptiotes* cribellate silk and how this contrasts to other cribellate silk remains an intriguing open question.

The rapid release of tension by *Hyptiotes* in its support line is achieved in multiple short bursts that cause the spider and the entire web to oscillate (7). These oscillations allow more capture threads to contact and tangle prey. Since proline introduces turns in protein structures, the increased proline in the amorphous region of MaSp sequences may act as nano-springs (16) that allow for increased web oscillations. Furthermore, given that the GPGPQ motifs in *Hyptiotes* dragline arise convergently across spider lineages and in species that make webs with atypical mechanical demands, it is highly likely that these motifs have functional importance to the silk. Further study into *Hyptiotes* silk mechanics and secondary structure is needed to illuminate the effect of its predominant GPGPQ motifs on performance, which may also provide insight into their functional importance in the extraordinarily extensible silk of distantly related species. Understanding these relationships will contribute to designing biomaterials for a wide range of prospective applications.

Methods

Genome sequencing and assembly

Genome sequencing was carried out using Oxford Nanopore PromethION sequence data and Illumina reads for error correction. DNA was extracted from a combined sample of silk glands from ten individuals using the Gentra Puregene tissue kit (Qiagen cat. 158667) and treated with the Circulomics short-read eliminator kit (PacBio cat. SS-100-101-01). This DNA prep was used for both the Nanopore and Illumina sequencing. The Nanopore sequencing was conducted on two PromethION flowcells, producing a total of 96 Gb of sequence data with an N50 of 9.8 kb. The Illumina sequencing produced 104 Gb of sequence data.

Spidroin annotation

Nanopore reads were filtered to a minimum length of 5,000 base pairs and assembled using Flye (40). This assembly produced a genome of size 4.26 Gb comprised of 91,457 contigs with an N50 of 107 kb. Spidroin sequences were discovered using blastx searches of the genome against a private database of a curated list of known spidroin protein terminal domains. Full contigs for each spidroin hit were separated and corrected using Illumina DNA short reads with Pilon (81), which was set to only correct frameshifts to minimize artificial repeat homogenization from short-read correction. Both short and long reads were then mapped to the contigs using minimap2 (82). These mappings were viewed in Integrative Genomics Viewer (IGV) (83) and misassembled frameshifts in spidroin sequences were manually corrected.

Spidroin sequences were categorized by both examining the repeat structure of each sequence and by creating a phylogeny of the concatenated N- and C-terminal domains including spidroins from *U. diversus*. Flanking regions and other genes on spidroin-containing contigs were examined to verify that each spidroin gene was a distinct locus and did not represent allelic variants. MaSp2.1a and Masp2.1b, the two most similar sequences, were located on the same contig. In this case, their validity as distinct loci was determined by locating multiple individual nanopore reads that spanned portions of both genes. This was repeated for the only other spidroin pair that occurred on a single contig, MiSp1a and MiSp1b.

MaSp sequences were separated into MaSp1 and MaSp2 clades based on the repeat motifs and proportion of proline in the repetitive region, as MaSp2 is typically differentiated from MaSp1 by

having a significant number of GPGx repeats. Repeat motifs were both manually defined and computationally defined using the program RADAR (84).

Phylogenetic analysis

Phylogenetic trees were created by isolating and concatenating the N-terminal and C-terminal domains of each spidroin protein sequence. Sequences were aligned using Clustalw (85), and maximum likelihood trees were created using iqtree2 (86). The substitution model Q.insect + R5 was used, selected by using the built-in model finder which selects for the model with the lowest BIC. Support values were calculated using Ultrafast Bootstrap (87) and Shimodaira–Hasegawa approximate likelihood ratio test (sh-alrt) (88), each using 5,000 replicates. Strong support cutoffs of 95 and 80% for UFbootstrap and sh-alrt, respectively, were used, as suggested in the iqtree manual.

To map proline content as a continuous character onto the phylogeny, the function contMap in the R package phytools (89) was used. Species within Uloboridae and Araneoidea with high-quality genomes, full spidroin catalogs, and notably high-proline MA spidroins were included in this analysis. *Argiope argentata* does not have a full-length *MaSp2.3* gene present in other *Argiope* species, so the MaSp2.3 sequence from *Argiope aurantia* was included in its place. Proline content was determined by isolating the repetitive region of each protein sequence and determining the amino acid composition of each sequence. N- or C-terminal fragments without repetitive region were excluded. The mygalomorph spidroin sequence *Bothriocyrtum californicum* spidroin 1 was selected as an outgroup given its available N- and C-terminal domain.

RNA-seq and expression

RNA was extracted from six *H. cavatus* individuals to create three replicates, for each of four tissue types: total silk glands, cephalothorax, cribellate glands, and MA glands. The cribellate and MA gland replicates were created from the same three individuals, while the cephalothorax and total silk glands replicates originated from the remaining three. RNA was extracted using the PureLink RNA Mini kit with on-column DNAse treatment (Thermo Fisher Scientific) and quantification was done with a Qubit Fluorometer. All samples were sent to Novogene for library preparation and sequencing. Each library was paired-end 150 bp sequenced on an Illumina HiSeq System.

The RNA-seq reads from all libraries were mapped to the genome assembly using the splice-aware alignment program GSNAP (90). A reference-based transcriptome using these mapped reads was assembled using Stringtie (91). Differential gene expression was determined using only uniquely mapping reads from alignment with STAR (92). The program Deseq2 (93) was then used to make pairwise comparisons across each tissue type. Transcripts for MA glands were then separated into four groups: genes upregulated in MA glands compared with cephalothorax tissue, genes upregulated in MA compared with both cephalothorax and cribellate glands, spidroins, and all other genes. This process was repeated for cribellate gland libraries. Transcripts per million (TPM) values were calculated using Salmon (94) to map reads to the genome-guided transcriptome created by Stringtie.

Amino acid composition and mechanical testing

Pulled silk was reeled from six *Hyptiotes*, one *Zosis*, and two *Miagrammopes* individuals for amino acid composition. Composition was analyzed using Hitachi L-8900 Amino Acid Analyzer. Three of the *Hyptiotes* silk samples were analyzed using a Hitachi LA8080 to detect hydroxyproline. For mechanical testing, single-fiber MA silk was forcibly pulled from two individuals. Fiber diameter was measured using the SEM JEOL JSM 6390. Seven replicates for each individual were tested using the Tensile Tester Nano Single Fiber KLA T150 UTM with a strain rate of 1/s, harmonic frequency of 20 Hz and harmonic force of 4.5 μN. Fiber lengths were 21 mm.

Supplementary Material

Supplementary material is available at *PNAS Nexus* online.

Funding

This work was supported by the U.S. National Science Foundation (award numbers: IOS-2128028 to S.M.C.-G., C.Y.H., and R.H.B., IOS-2128029 to J.E.G., and IOS-2128027 to H.K.R.).

Author Contributions

Steven Casey (Conceptualization, Formal analysis, Investigation, Visualization, Writing—original draft, Writing—review & editing), Sandra M. Correa-Garhwal (Conceptualization, Formal analysis, Funding acquisition, Investigation, Visualization, Writing—original draft, Writing—review & editing), Richard H. Baker (Conceptualization, Formal analysis, Funding acquisition, Investigation, Writing—review & editing), Jay A. Stafstrom (Conceptualization, Funding acquisition, Writing—review & editing), Hudson Kern Reeve (Conceptualization, Funding acquisition, Writing—review & editing), Cheryl Y. Hayashi (Conceptualization, Funding acquisition, Writing—review & editing), and Jessica E. Garb (Conceptualization, Formal analysis, Funding acquisition, Investigation, Visualization, Writing—original draft, Writing—review & editing)

Data Availability

Genome assembly, genomic reads, and RNA-seq reads are found under BioProject PRJNA1125421 at NCBI. Raw RNA-seq reads and genomic reads can be found in the SRA database under SRR29456179–SRR29456193 silk gene and protein sequences can be found in Datasets S2 and S3, respectively.

References

1 Eberhard W. *Spider webs: behavior, function, and evolution*. University of Chicago Press, 2020.
2 Vollrath F. 2000. Strength and structure of spiders' silks. *J Biotechnol*. 74:67–83.
3 Peters HM. On the spinning apparatus and the structure of the capture threads of *Deinopis subrufus* Araneae Deinopidae. *Zoomorphology*. 112:27–37. https://doi.org/10.1007/BF01632992
4 Coddington JA. 1989. Spinneret silk spigot morphology: evidence for the monophyly of orbweaving spiders, cyrtophorinae (Araneidae), and the group Theridiidae plus Nesticidae. *J Arachnol*. 17:71–95.
5 Agnarsson I, Kuntner M, Blackledge TA. 2010. Bioprospecting finds the toughest biological material: extraordinary silk from a giant riverine orb spider. *PLoS One*. 5:e11234. https://doi.org/10.1371/journal.pone.0011234
6 Eberhard WG, Opell BD. 2022. Orb web traits typical of Uloboridae (Araneae). *J Arachnol*. 50:351–384. https://doi.org/10.1636/JoA-S-21-050

7 Han SI, Astley HC, Maksuta DD, Blackledge TA. 2019. External power amplification drives prey capture in a spider web. *Proc Natl Acad Sci U S A*. 116:12060–12065. https://doi.org/10.1073/pnas.1821419116

8 Han SI, Htut KZ, Blackledge TA. 2021. Permanent deformation of triangle weaver silk enables ultrafast tangle-free release of spider webs. *Sci Nat*. 108:60. https://doi.org/10.1007/s00114-021-01769-3

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

10 Opell BD. Cribellar thread. In: Nentwig W, editors. *Spider ecophysiology*. Springer, 2013. p. 303–315.

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

12 Garb J. Spider silk: an ancient biomaterial for 21st century research. In: Penney D, editors. *Spider research in the 21st century: trends & perspectives*. Siri Scientific Press, 2013. p. 252.

13 Babb PL, *et al*. 2017. The *Nephila clavipes* genome highlights the diversity of spider silk genes and their complex expression. *Nat Genet*. 49:895–903. https://doi.org/10.1038/ng.3852

14 Hayashi CY, Shipley NH, Lewis RV. 1999. Hypotheses that correlate the sequence, structure, and mechanical properties of spider silk proteins. *Int J Biol Macromol*. 24:271–275.

15 Yarger JL, Cherry BR, van der Vaart A. 2018. Uncovering the structure–function relationship in spider silk. *Nat Rev Mater*. 3:1–11. https://doi.org/10.1038/natrevmats.2018.8

16 Becker N, *et al*. 2003. Molecular nanosprings in spider capture-silk threads. *Nature Mater*. 2:278–283. https://doi.org/10.1038/nmat858

17 Ayoub NA, Garb JE, Tinghitella RM, Collin MA, Hayashi CY. 2007. Blueprint for a high-performance biomaterial: full-length spider dragline silk genes. *PLoS One*. 2:e514. https://doi.org/10.1371/journal.pone.0000514

18 Garb JE, Ayoub NA, Hayashi CY. 2010. Untangling spider silk evolution with spidroin terminal domains. *BMC Evol Biol*. 10:243. https://doi.org/10.1186/1471-2148-10-243

19 Blackledge TA. Spider silk: molecular structure and function in webs. In: Nentwig W, editors. *Spider Ecophysiology*. Springer, 2013. p. 267–281.

20 Sampath S, *et al*. 2012. X-ray diffraction study of nanocrystalline and amorphous structure within major and minor ampullate dragline spider silks. *Soft Matter*. 8:6713–6722. https://doi.org/10.1039/C2SM25373A

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

22 Termonia Y. 1994. Molecular modeling of spider silk elasticity. *Macromolecules*. 27:7378–7381. https://doi.org/10.1021/ma00103a018

23 Nova A, Keten S, Pugno N, Redaelli A, Buehler M. 2010. Molecular and nanostructural mechanisms of deformation, strength and toughness of spider silk fibrils. *Nano Lett*. 10(7):2626–2634. https://pubs.acs.org/doi/10.1021/nl101341w

24 Garb JE, *et al*. 2019. The transcriptome of Darwin's bark spider silk glands predicts proteins contributing to dragline silk toughness. *Commun Biol*. 2:1–8. https://doi.org/10.1038/s42003-019-0496-1

25 Ananthanarayanan VS, Attah-poku SK, Mukkamala PL, Rehse PH. 1985. Structural and functional importance of the β-turn in proteins. Studies on proline-containing peptides. *J Biosci*. 8:209–221. https://doi.org/10.1007/BF02703977

26 MacArthur MW, Thornton JM. 1991. Influence of proline residues on protein conformation. *J Mol Biol*. 218:397–412. https://doi.org/10.1016/0022-2836(91)90721-H

27 Liu Y, Sponner A, Porter D, Vollrath F. 2008. Proline and processing of spider silks. *Biomacromolecules*. 9:116–121. https://doi.org/10.1021/bm700877g

28 Savage KN, Gosline JM. 2008. The role of proline in the elastic mechanism of hydrated spider silks. *J Exp Biol*. 211:1948–1957. https://doi.org/10.1242/jeb.014225

29 Craig HC, Piorkowski D, Nakagawa S, Kasumovic MM, Blamires SJ. 2020. Meta-analysis reveals materiomic relationships in major ampullate silk across the spider phylogeny. *J R Soc Interface*. 17:20200471. https://doi.org/10.1098/rsif.2020.0471

30 Shi X, Yarger JL, Holland GP. 2014. Elucidating proline dynamics in spider dragline silk fibre using ^{2}H–^{13}C HETCOR MAS NMR. *Chem Commun*. 50:4856–4859. https://doi.org/10.1039/C4CC00971A

31 Guan J, Vollrath F, Porter D. 2011. Two mechanisms for supercontraction in *Nephila* spider dragline silk. *Biomacromolecules*. 12:4030–4035. https://doi.org/10.1021/bm201032v

32 Liu D, *et al*. 2019. Spider dragline silk as torsional actuator driven by humidity. *Sci Adv*. 5:eaau9183. https://doi.org/10.1126/sciadv.aau9183

33 Baker RH, Corvelo A, Hayashi CY. 2022. Rapid molecular diversification and homogenization of clustered major ampullate silk genes in *Argiope* garden spiders. *PLoS Genet*. 18:e1010537. https://doi.org/10.1371/journal.pgen.1010537

34 Malay AD, Craig HC, Chen J, Oktaviani NA, Numata K. 2022. Complexity of spider dragline silk. *Biomacromolecules*. 23:1827–1840. https://doi.org/10.1021/acs.biomac.1c01682

35 Swanson BO, Blackledge TA, Summers AP, Hayashi CY. 2006. Spider dragline silk: correlated and mosaic evolution in high-performance biological materials. *Evolution*. 60:2539–2551. https://doi.org/10.1111/j.0014-3820.2006.tb01888.x

36 Blackledge TA. 2012. Spider silk: a brief review and prospectus on research linking biomechanics and ecology in draglines and orb webs. *J Arachnol*. 40:1–12.

37 Correa-Garhwal SM, Baker RH, Clarke TH, Ayoub NA, Hayashi CY. 2022. The evolutionary history of cribellate orb-weaver capture thread spidroins. *BMC Ecol Evol*. 22:89. https://doi.org/10.1186/s12862-022-02042-5

38 Miller J, Zimin AV, Gordus A. 2023. Chromosome-level genome and the identification of sex chromosomes in *Uloborus diversus*. *GigaScience*. 12:giad002. https://doi.org/10.1093/gigascience/giad002

39 Opell BD. 1982. Post-hatching development and web production of *Hyptiotes cavatus* (Hentz) (Araneae, Uloboridae). *J Arachnol*. 10:185–191.

40 Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol*. 37:540–546. https://doi.org/10.1038/s41587-019-0072-8

41 Kallal RJ, *et al*. 2021. Converging on the orb: denser taxon sampling elucidates spider phylogeny and new analytical methods support repeated evolution of the orb web. *Cladistics*. 37:298–316. https://doi.org/10.1111/cla.12439

42 Kono N, Nakamura H, Mori M, Tomita M, Arakawa K. 2020. Spidroin profiling of cribellate spiders provides insight into the evolution of spider prey capture strategies. *Sci Rep*. 10:15721. https://doi.org/10.1038/s41598-020-72888-6

43 Hayashi CY, Blackledge TA, Lewis RV. 2004. Molecular and mechanical characterization of aciniform silk: uniformity of iterated sequence modules in a novel member of the spider silk fibroin gene family. *Mol Biol Evol*. 21:1950–1959. https://doi.org/10.1093/molbev/msh204

44 Correa-Garhwal SM, Garb JE. 2014. Diverse formulas for spider dragline fibers demonstrated by molecular and mechanical characterization of spitting spider silk. *Biomacromolecules*. 15: 4598–4605. https://doi.org/10.1021/bm501409n

45 Blamires SJ, Wu C-L, Blackledge TA, Tso I-M. 2012. Post-secretion processing influences spider silk performance. *J R Soc Interface*. 9: 2479–2487. https://doi.org/10.1098/rsif.2012.0277

46 Blamires SJ, *et al.* 2018. Multiscale mechanisms of nutritionally induced property variation in spider silks. *PLoS One*. 13: e0192005. https://doi.org/10.1371/journal.pone.0192005

47 Lacava M, *et al.* 2018. Web building and silk properties functionally covary among species of wolf spider. *J Evol Biol*. 31:968–978. https://doi.org/10.1111/jeb.13278

48 Coddington JA, Chanzy HD, Jackson CL, Raty G, Gardner KH. 2002. The unique ribbon morphology of the major ampullate silk of spiders from the genus *Loxosceles* (recluse spiders). *NIST*. 3:5–8.

49 Blamires SJ, Wu C-L, Tso I-M. 2012. Variation in protein intake induces variation in spider silk expression. *PLoS One*. 7:e31626. https://doi.org/10.1371/journal.pone.0031626

50 Work RW, Young CT. 1987. The amino acid compositions of major and minor ampullate silks of certain orb-web-building spiders (Araneae, Araneidae). *J Arachnol*. 15:65–80.

51 Arakawa K, *et al.* 2022. 1000 spider silkomes: linking sequences to silk physical properties. *Sci Adv*. 8:eabo6043. https://doi.org/10.1126/sciadv.abo6043

52 Liu Y, Shao Z, Vollrath F. 2008. Elasticity of spider silks. *Biomacromolecules*. 9:1782–1786. https://doi.org/10.1021/bm7014174

53 Chalek K, Soni A, Lorenz CD, Holland GP. 2024. Proline–tyrosine ring interactions in black widow dragline silk revealed by solid-state nuclear magnetic resonance and molecular dynamics simulations. *Biomacromolecules*. 25:1916–1922. https://doi.org/10.1021/acs.biomac.3c01351

54 Htut KZ, *et al.* 2021. Correlation between protein secondary structure and mechanical performance for the ultra-tough dragline silk of Darwin's bark spider. *J R Soc Interface*. 18:20210320. https://doi.org/10.1098/rsif.2021.0320

55 Hu W, *et al.* 2023. A molecular atlas reveals the tri-sectional spinning mechanism of spider dragline silk. *Nat Commun*. 14:837. https://doi.org/10.1038/s41467-023-36545-6

56 Jaleel Z, *et al.* 2020. Expanding canonical spider silk properties through a DNA combinatorial approach. *Materials*. 13:3596. https://doi.org/10.3390/ma13163596

57 Salehi S, Koeck K, Scheibel T. 2020. Spider silk for tissue engineering applications. *Molecules*. 25:737. https://doi.org/10.3390/molecules25030737

58 Ramezaniaghdam M, Nahdi ND, Reski R. 2022. Recombinant spider silk: promises and bottlenecks. *Front Bioeng Biotechnol*. 10: 835637. https://doi.org/10.3389/fbioe.2022.835637

59 Kono N, *et al.* 2019. Orb-weaving spider *Araneus ventricosus* genome elucidates the spidroin gene catalogue. *Sci Rep*. 9:8380. https://doi.org/10.1038/s41598-019-44775-2

60 Opell BD. 1987. Changes in web-monitoring forces associated with web reduction in the spider family Uloboridae. *Can J Zool*. 65:1028–1034. https://doi.org/10.1139/z87-163

61 Marhabaie M, Leeper TC, Blackledge TA. 2014. Protein composition correlates with the mechanical properties of spider (*Argiope trifasciata*) dragline silk. *Biomacromolecules*. 15:20–29. https://doi.org/10.1021/bm401110b

62 Xu M, Lewis RV. 1990. Structure of a protein superfiber: spider dragline silk. *Proc Natl Acad Sci U S A*. 87:7120–7124. https://doi.org/10.1073/pnas.87.18.7120

63 Garrison NL, *et al.* 2016. Spider phylogenomics: untangling the spider tree of life. *PeerJ*. 4:e1719. https://doi.org/10.7717/peerj.1719

64 Kuntner M. 2010. Web gigantism in Darwin's bark spider, a new species from Madagascar (Araneidae: *Caerostris*). *J Arachnol*. 38: 346–356. https://doi.org/10.1636/B09-113.1

65 Chaw RC, *et al.* 2014. Intragenic homogenization and multiple copies of prey-wrapping silk genes in *Argiope* garden spiders. *BMC Evol Biol*. 14:31. https://doi.org/10.1186/1471-2148-14-31

66 Garb JE, DiMauro T, Lewis RV, Hayashi CY. 2007. Expansion and intragenic homogenization of spider silk genes since the triassic: evidence from mygalomorphae (tarantulas and their kin) spidroins. *Mol Biol Evol*. 24:2454–2464. https://doi.org/10.1093/molbev/msm179

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

68 Rauscher S, Baud S, Miao M, Keeley FW, Pomès R. 2006. Proline and glycine control protein self-organization into elastomeric or amyloid fibrils. *Structure*. 14(11):1667–1676. https://doi.org/10.1016/j.str.2006.09.008

69 Shoulders MD, Raines RT. 2009. Collagen structure and stability. *Annu Rev Biochem*. 78(1):929–958. https://doi.org/10.1146/annurev.biochem.77.032207.120833

70 Espelie KE, Himmelsbach DS. 1990. Characterization of pedicel, paper, and larval silk from nest of Polistes annularis (L.). *J Chem Ecol*. 16:3467–3477. https://doi.org/10.1007/BF00982111

71 Heckenhauer J, *et al.* 2023. Characterization of the primary structure of the major silk gene, h-fibroin, across caddisfly (Trichoptera) suborders. *iScience*. 26:107253. https://doi.org/10.1016/j.isci.2023.107253

72 Townley MA, Tillinghast EK. Aggregate silk gland secretions of araneoid spiders. In: Nentwig W, editors. *Spider Ecophysiology*. Springer, 2013. p. 283–302.

73 Pajic P, *et al.* 2022. A mechanism of gene evolution generating mucin function. *Sci Adv*. 8:eabm8757. https://doi.org/10.1126/sciadv.abm8757

74 Haritos VS, *et al.* 2010. Harnessing disorder: onychophorans use highly unstructured proteins, not silks, for prey capture. *Proc Biol Sci*. 277:3255–3263. https://doi.org/10.1098/rspb.2010.0604

75 Krane SM. 2008. The importance of proline residues in the structure, stability and susceptibility to proteolytic degradation of collagens. *Amino Acids*. 35:703–710. https://doi.org/10.1007/s00726-008-0073-2

76 Sutherland TD, *et al.* 2013. A new class of animal collagen masquerading as an insect silk. *Sci Rep*. 3:2864. https://doi.org/10.1038/srep02864

77 Roberts S, Dzuricky M, Chilkoti A. 2015. Elastin-like polypeptides as models of intrinsically disordered proteins. *FEBS Lett*. 589: 2477–2486. https://doi.org/10.1016/j.febslet.2015.08.029

78 Anjum FM, *et al.* 2007. Wheat gluten: high molecular weight glutenin subunits—structure, genetics, and relation to dough elasticity. *J Food Sci*. 72:R56–R63. https://doi.org/10.1111/j.1750-3841.2007.00292.x

79 Shewry PR, Halford NG, Belton PS, Tatham AS. 2002. The structure and properties of gluten: an elastic protein from wheat grain. *Philos Trans R Soc Lond B Biol Sci*. 357:133–142. https://doi.org/10.1098/rstb.2001.1024

80 Patil SP, Markert B, Gräter F. 2014. Rate-dependent behavior of the amorphous phase of spider dragline silk. *Biophys J*. 106: 2511–2518. https://doi.org/10.1016/j.bpj.2014.04.033

81 Walker BJ, *et al.* 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One*. 9:e112963. https://doi.org/10.1371/journal.pone.0112963

82 Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 34:3094–3100. https://doi.org/10.1093/bioinformatics/bty191
83 Robinson JT, *et al.* 2011. Integrative genomics viewer. *Nat Biotechnol*. 29:24–26. https://doi.org/10.1038/nbt.1754
84 Heger A, Holm L. 2000. Rapid automatic detection and alignment of repeats in protein sequences. *Proteins: Struct Funct Bioinf*. 41:224–237. https://doi.org/10.1002/1097-0134(20001101)41:2<224::AID-PROT70>3.0.CO;2-Z
85 Larkin MA, *et al.* 2007. Clustal W and clustal X version 2.0. *Bioinformatics*. 23:2947–2948. https://doi.org/10.1093/bioinformatics/btm404
86 Minh BQ, *et al.* 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol*. 37:1530–1534. https://doi.org/10.1093/molbev/msaa015
87 Minh BQ, Nguyen MAT, von Haeseler A. 2013. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol*. 30:1188–1195. https://doi.org/10.1093/molbev/mst024
88 Guindon S, *et al.* 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol*. 59:307–321. https://doi.org/10.1093/sysbio/syq010
89 Revell LJ. 2012. Phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol*. 3:217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x
90 Wu TD, Watanabe CK. 2005. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics*. 21:1859–1875. https://doi.org/10.1093/bioinformatics/bti310
91 Pertea M, *et al.* 2015. StringTie enables improved reconstruction of a transcriptome from RNA-Seq reads. *Nat Biotechnol*. 33:290–295. https://doi.org/10.1038/nbt.3122
92 Dobin A, *et al.* 2013. STAR: ultrafast universal RNA-Seq aligner. *Bioinformatics*. 29:15–21. https://doi.org/10.1093/bioinformatics/bts635
93 Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biol*. 15:550. https://doi.org/10.1186/s13059-014-0550-8
94 Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*. 14:417–419. https://doi.org/10.1038/nmeth.4197