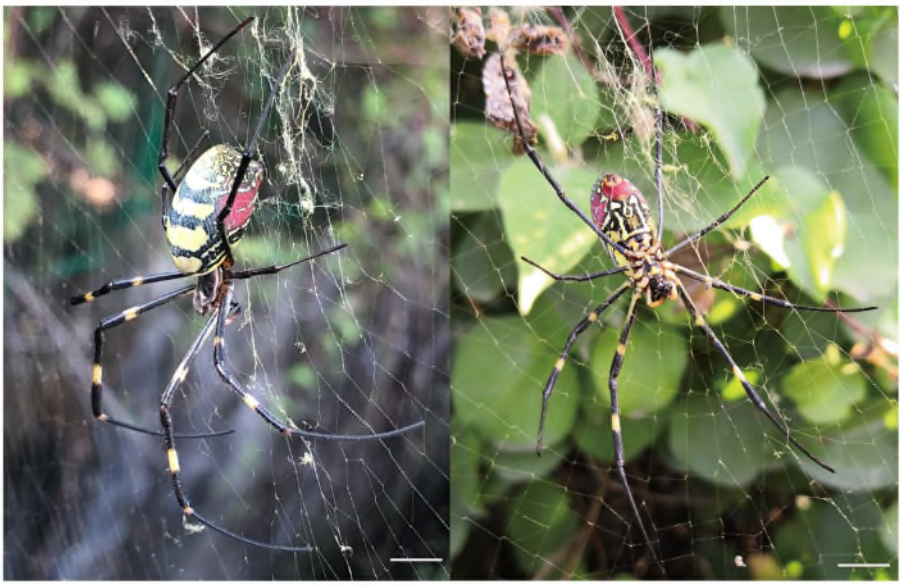
A Newly Evolved Small Secretory Peptide Enhances Mechanical Properties of Spider Silk

Anqiang Jia[1,3,#], Yudi Mao[1,#], Tianfang Yang[1,#], Guoqing Zhang[1], Qingyuan Wang[1], Wenbo Hu[1], Zhaoming Dong[1], Zhisheng Zhang[2], Sanyuan Ma[1,*], Yi Wang[1,*]

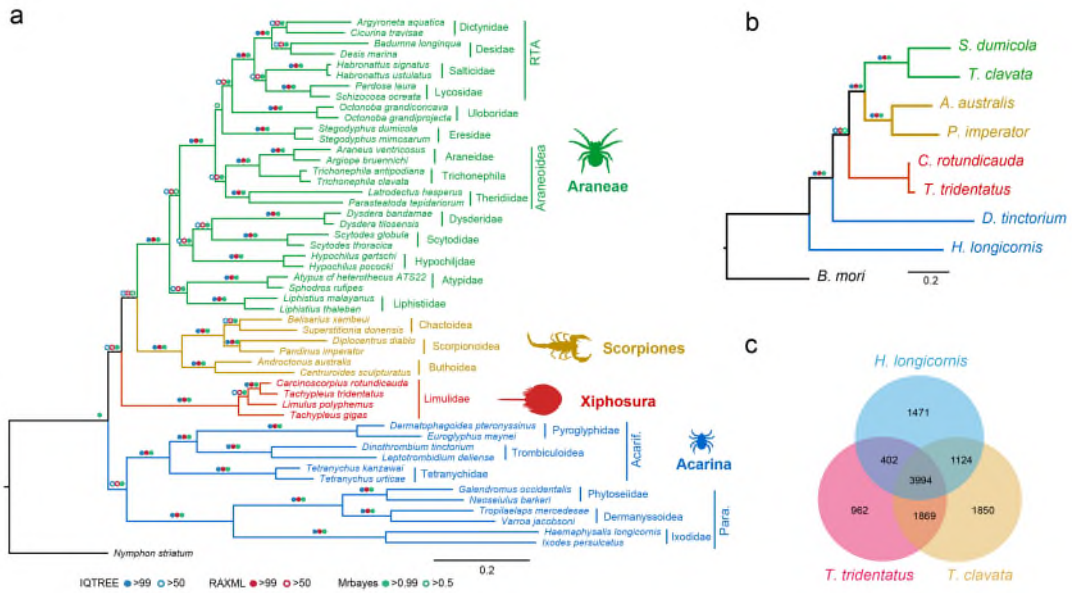
[1]Biological Science Research Center, Southwest University, Chongqing 400715, China.
[2]School of Life Sciences, Southwest University, Chongqing 400715, China.
[3]Yazhouwan National Laboratory, Sanya 572024, China.
[#]These authors contributed equally to this work.
[*]Corresponding authors.
E-mail: masy@swu.edu.cn, yiwang28@swu.edu.cn.

Supplementary Figures

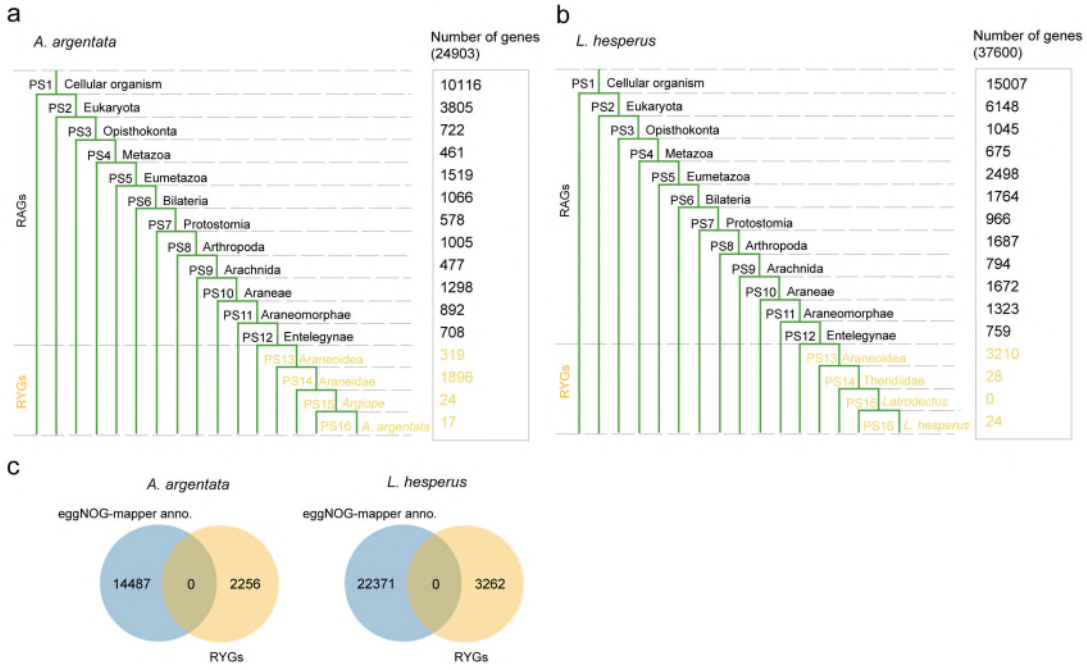
Supplementary Fig. 1 | Photograph of *T. clavata* spiders. *T. clavata* spiders were captured from the wild in Dali City, Yunnan Province, China. The white scale in the figure represents 1cm.



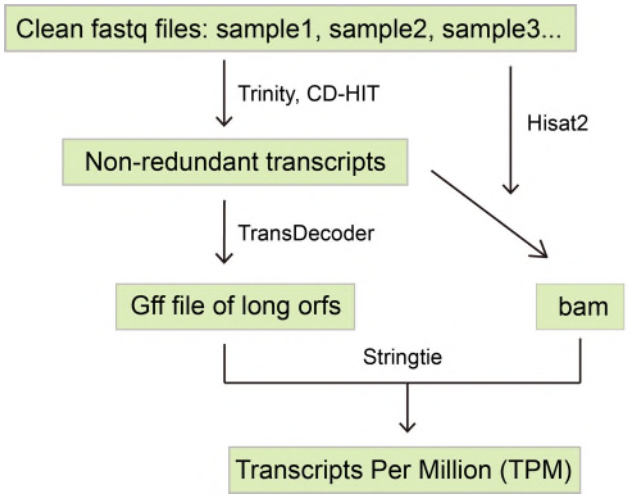
Supplementary Fig. 2 | Phylogenetic relationships of the four orders of Arachnida. **a**, ML phylogenetic trees from the concatenated orthologous genes of 50 chelicerates. Orthologous genes were identified using Blastp (E-value < $1e^{-5}$) in an all-vs-all alignment, with *N. striatum* (sea spider) designated as the outgroup. **b**, The ML phylogenetic relationships of nine arthropod species. The single-copy gene set was identified using OrthoFinder. Different colored circles on branches denote support values obtained in each approach. *B. mori* (silkworm) as outgroup. **c**, Shared and unique orthologous clusters between tick (*H. longicornis*), horseshoe crab (*T. tridentatus*) and spider (*T. clavata*), and horseshoe crab and spider share more orthologous clusters.



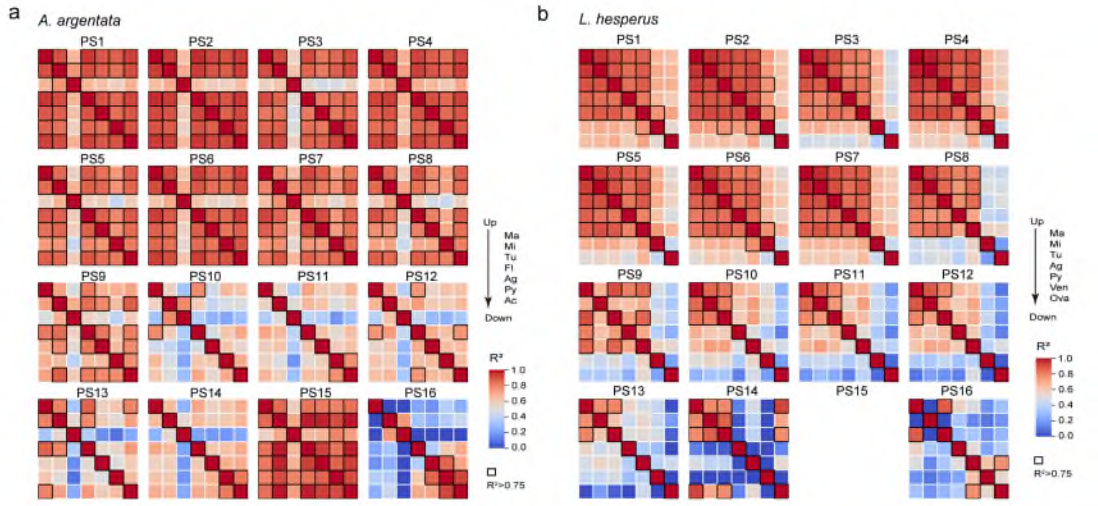
Supplementary Fig. 3 | Phylostratigraphic profiles of two Araneoidea spiders. a, The number distribution of unique transcripts from golden orb-weaving spider (*A. argentata*) across 16 PS levels. **b**, The number distribution of unique transcripts from black widow spider (*L. hesperus*) across 16 PS levels. The unique transcripts were assembled based on a nonreference genome. **c**, The Venn diagram shows that genes from PS13 to PS16 are classified as "unknown functional annotation". RAGs refer to relatively ancient genes, while RYGs refer to relatively young genes. PS refer to phylogenetic strata.



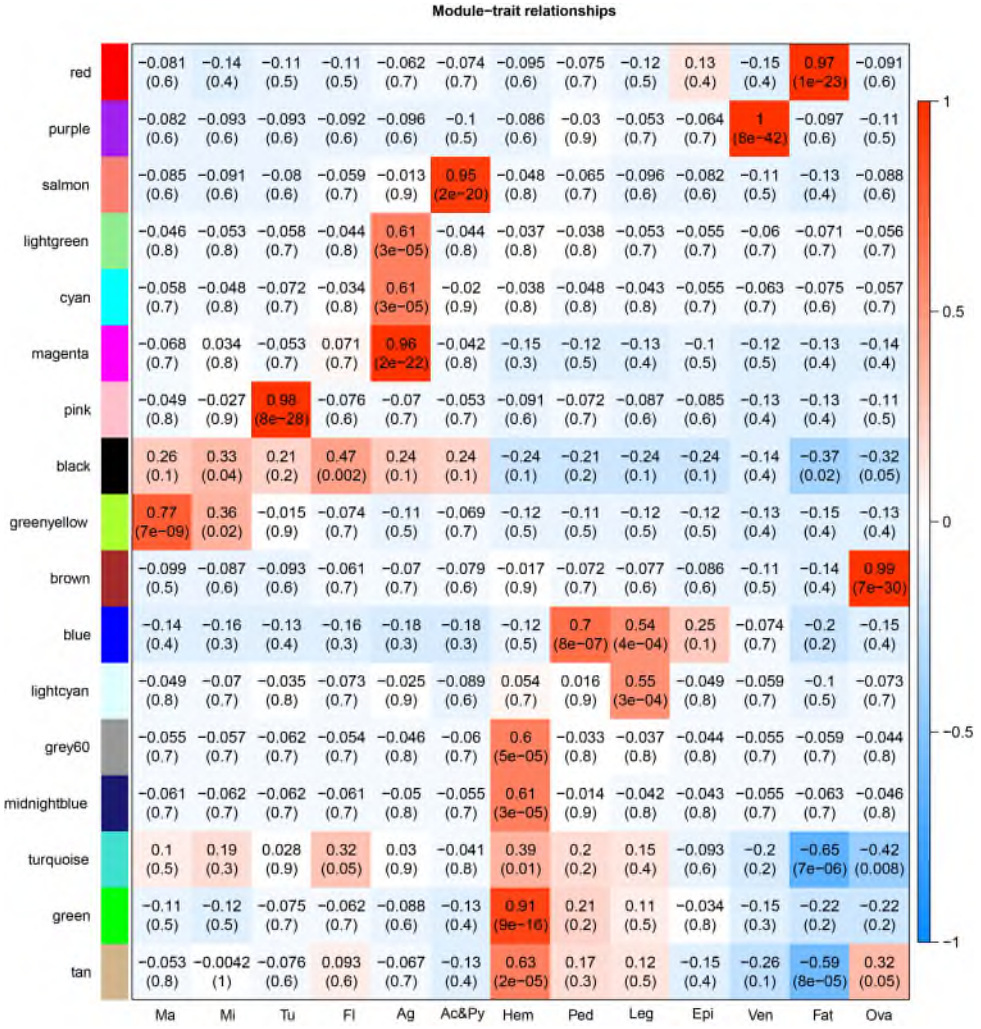
Supplementary Fig. 4 | A custom nonreference transcriptome assembly and quantification pipeline. Detailed descriptions refer to the "Data acquisition" section of "Materials and methods".



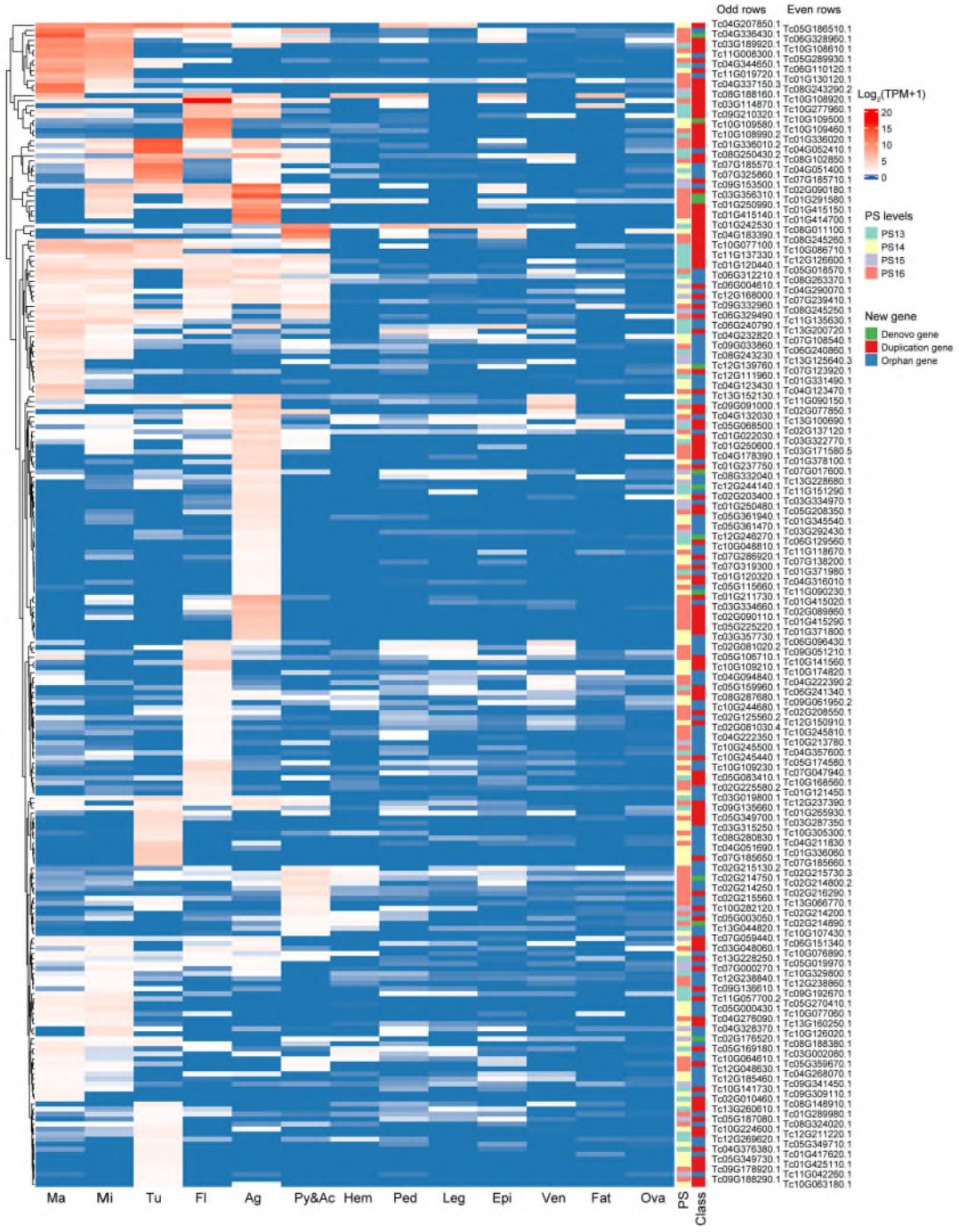
Supplementary Fig. 5 | The heatmap of gene expression correlations across the 16 PS levels. a, The expression correlation of transcripts from seven tissues of *A. argentata* across the 16 PS levels. Tissue order in each PS heatmap (top-to-bottom, left-to-right) follows the arrow-indicated order in the legend. **b**, The expression correlation of transcripts from seven tissues of *L. hesperus* across the 16 PS levels. PS15 did not match any transcripts, indicating the limitations of nonreference transcriptome assembly. However, the expression correlation trends across tissues are similar to those observed in the reference genome (*T. clavata*). Ma: major ampullate gland; Mi: minor ampullate gland; Tu: tubuliform gland; Fl: flagelliform gland; Ag: aggregate gland; Ac: aciniform gland; Py: pyriform gland; Ven: venom gland; Ova: ovary.



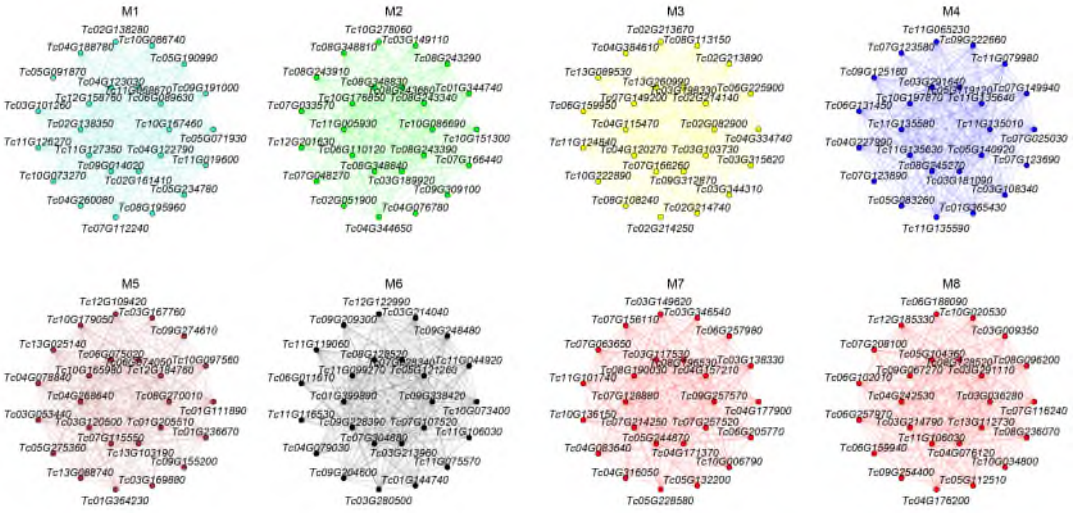
Supplementary Fig. 6 | Module-trait relationships of *T. clavata* genes. Each row corresponds to a module, and each column corresponds to a tissue. The correlation and p-value are shown within each box. The correlation intensity increases from blue to red.



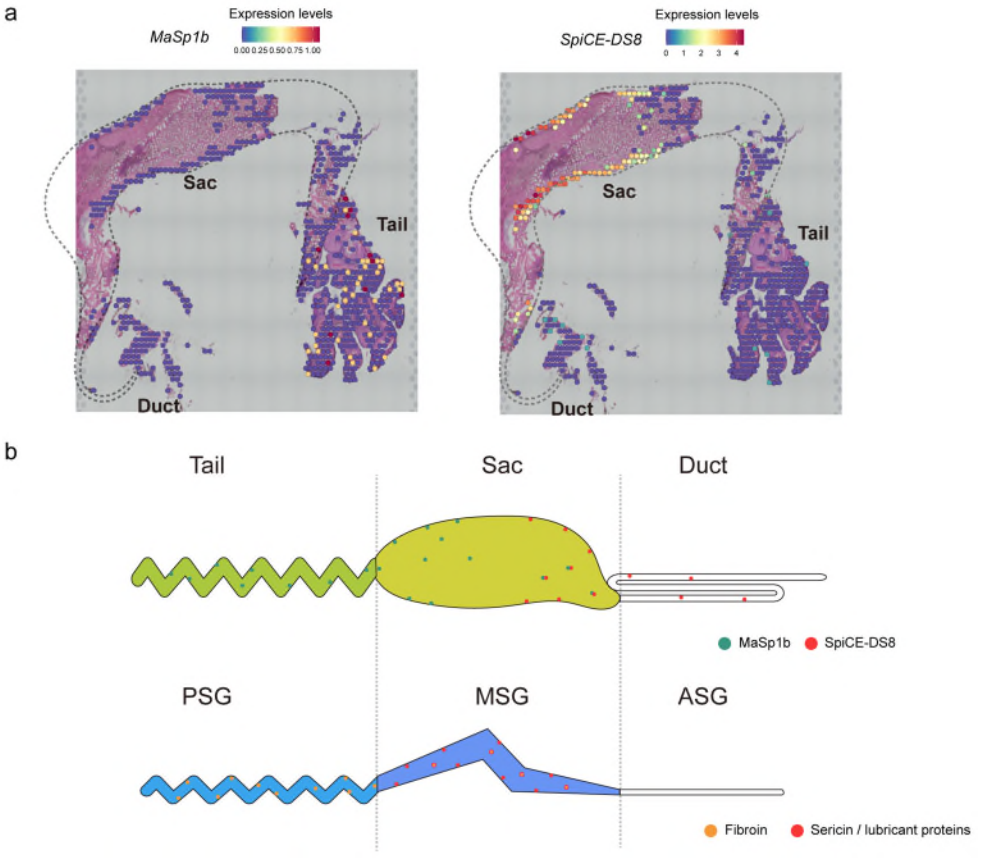
Supplementary Fig. 7 | Expression patterns of 232 Araneoidea-specific genes across 13 tissues of *T. clavata*.



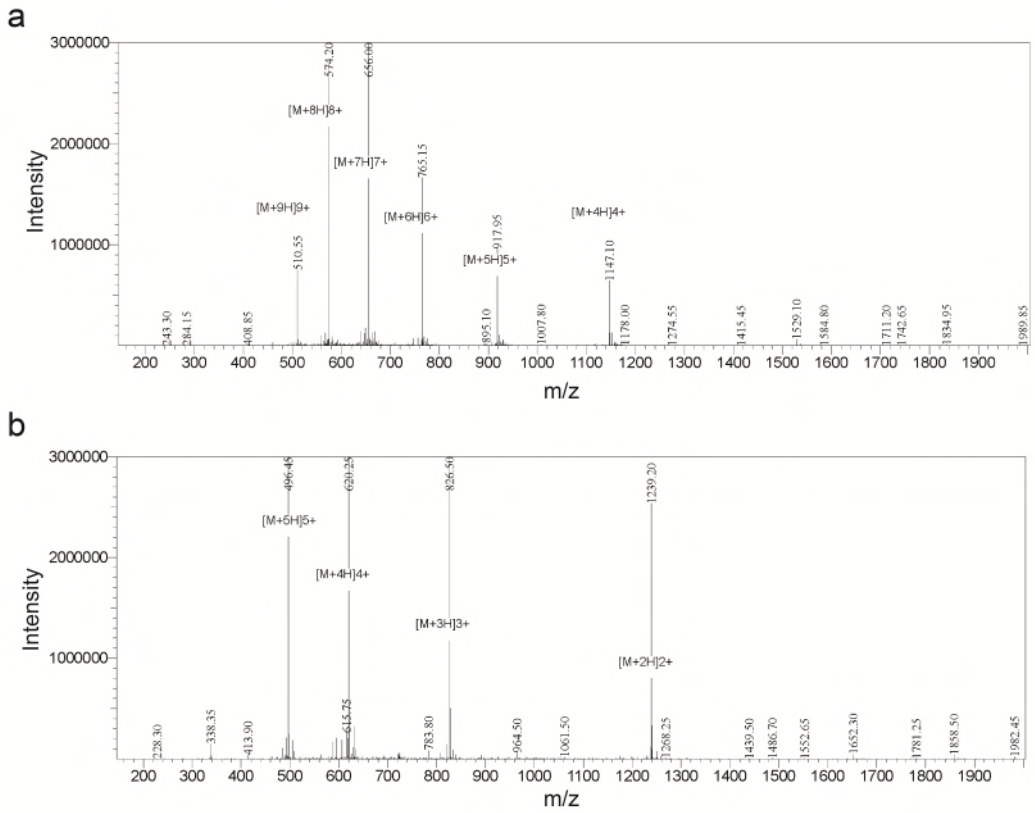
Supplementary Fig. 8 | The eight modules of the Ma single-cell coexpression network. The top 25 genes from each module are displayed.



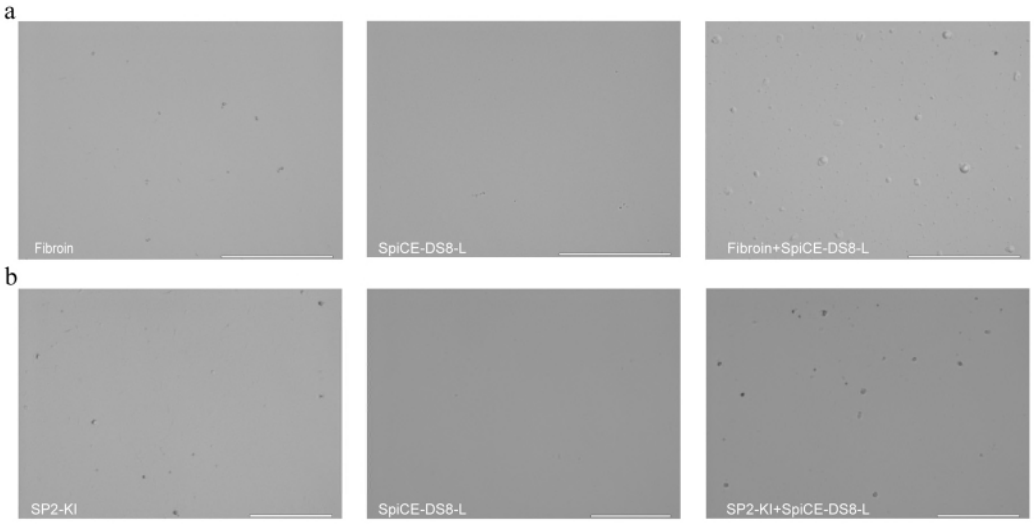
Supplementary Fig. 9 | Spatial expression of *MaSp1b* and *SpiCE-DS8* and their proposed interaction schematic in the silk gland. a, Spatial transcriptomic (ST) spots of *MaSp1b* and *SpiCE-DS8*. The dotted lines depict the outline of the Ma gland. **b**, Comparative secretion models of silk components in spider (top) and silkworm (bottom) glands. Top: MaSp1b-SpiCE-DS8 interaction in the major ampullate lumen. Bottom: Fibroin and sericin/lubricant protein interactions in the silk gland lumen.



Supplementary Fig. 10 | Mass spectrometry analysis of peptide synthesis for SpiCE-DS8-L and SpiCE-DS8-S. a, Mass spectrometry analysis of peptide synthesis for SpiCE-DS8-L. The y-axis represents intensity, and the x-axis represents the mass-to-charge ratio (m/z). The peaks correspond to different adducts. The molecular ion peak, identified at $[M+7H]^{7+}$ (m/z 656.00), corresponds to an observed molecular weight of 4585.00, which closely matches the theoretical molecular weight (4585.375), confirming that the sample's molecular weight aligns with expectations. The mass-to-charge ratio is *m*, the charge number is *n*, and the molecular weight is *M*.
The equations are $M = nm - n$. **b,** Mass spectrometry analysis of peptide synthesis for SpiCE-DS8-S. The molecular ion peak, observed at $[M+4H]^{4+}$ (m/z: 620.25), corresponds to an observed molecular weight of 2477.00, which closely matches the theoretical molecular weight (2476.915), confirming that the sample's molecular weight aligns with expectations.



Supplementary Fig. 11 | Protein-solidification assay by bright-field microscopy. a, A mixture of fibroin and SpiCE-DS8-L incubated at 23℃ resulted in extensive formation of protein aggregates, as observed by bright-field microscopy. **b**, SP2-KI combined with SpiCE-DS8-L under identical conditions similarly results in protein aggregation. The white scale bars represent 100 μm in (a) and 50 μm in (b).



Supplementary Fig. 12 | Schematic representation and predicted interaction models of FibH and MaSp1b with SpiCE-DS8-L. **a**, Sequence architecture of FibH and MaSp1b. Schematic diagrams display structural composition (not to scale). **b**, Sequence alignment of N-terminal domains from FibH and MaSp1b. **c**, Predicted 3D interaction interfaces between SpiCE-DS8-L and N-terminal domains of FibH (left) and MaSp1b (right). **d**, Interaction heatmaps of SpiCE-DS8-L with N-terminal domains of FibH (left) and MaSp1b (right). **e**, Predicted 3D interfaces and interaction heatmaps between SpiCE-DS8-L and truncated N-terminal domains, with corresponding sequence alignment.