Genomic adaptations to semi-aquatic and aquatic life in spiders

Zheng Fan[1, 2, 3, 7], Lu-Yu Wang[1, 7], Bin Luo[1], Tian-Yu Ren[1], Jia-Xin Gao[1], Piao Liu[1], Ling-Xin Cheng[1], Yu-Jun Cai[1], Bing Tan[1], Qian Huang[1], Ming-Qin Deng[1], Qing Zuo[1], Xiang-Yun Zhang[1], Jin-Zhen Lu[1], Li-Na Sun[1], Muhammad Irfan[1], Ning Liu[2, 3], Chao Tong[6]*, Ming Bai[2, 3, 4, 5]* and Zhi-Sheng Zhang[1]*

1. Key Laboratory of Eco-environments in Three Gorges Reservoir Region (Ministry of Education), School of Life Sciences, Southwest University, Chongqing 400715, China.
2. Key Laboratory of Animal Biodiversity Conservation and Integrated Pest Management, Chinese Academy of Sciences, Beijing 100101, China.
3. Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China.
4. University of Chinese Academy of Sciences, Beijing 100101, China.
5. Northeast Asia Biodiversity Research Center, Northeast Forestry University, Harbin 150040, China.
6. School of Life Sciences, Arizona State University, Tempe, AZ 85287, USA.
7. These authors contributed equally to this work.

* These authors jointly directed this work. Correspondence should be addressed to C.T. (tongchao1990@gmail.com); M.B. (baim@ioz.ac.cn); Z.Z. (zhangzs327@qq.com).

Abstract

Spiders have primarily adapted to terrestrial life, yet a number of species have made evolutionary transitions to marine and freshwater environments. While its physiological and behavioral adaptations have been characterized, the genetic basis of semi-aquatic and aquatic adaptation in spiders remains poorly understood. Here, we provide a high-quality, chromosome-level genome assembly for the aquatic spider *Argyroneta aquatica*, alongside a reference transcriptome for the semi-aquatic spider *Desis martensi*. We performed comparative genomes analyses of 22 spider species, including a unique aquatic spider, two semi-aquatic spiders and 19 terrestrial spider species, with a focus on those in the marronoid clade. By integrating morphological, phylogenomic, comparative genomic, transcriptomic, and metabolomic analyses, we explored the genomic adaptations of aquatic and semi-aquatic spiders. Phylogenomic analysis suggests that aquatic and semi-aquatic spiders have independently evolved from their terrestrial ancestors and represent divergent evolutionary routes We found hundreds of genes tend to experience relaxed selection, positive selection, and evidence of horizontal gene transfer (HGT) associated with the transition to aquatic and semi-aquatic life in spiders. These genes are associated with respiratory, osmoregulatory, fat metabolism and digestion, hypoxia, and thermal functions, putatively facilitate the adaptations to diverse underwater life. Altogether, our findings highlight the divergent evolutionary mechanisms enabling spiders to thrive in diverse aquatic environments, providing insights into the genomic basis of adaptations to semi-aquatic and aquatic habitats.

Key words: semi-aquatic and aquatic adaptation, marronoid clade, comparative genomics, molecular evolution, horizontal gene transfer (HGT)

Introduction

Most spiders have adapted to terrestrial life, while a noteworthy proportion of spider species, which were originally terrestrial, have undergone a reversal of adaptation, back to marine and freshwater habitats (Leggett, Vink, and Nelson, 2024). This evolutionary transition has occurred either in semi-aquatic or fully aquatic spider species. One exemplar group is the marronoid clade, encompassing approximately nine families including Agelenidae, Amaurobiidae, Cybaeidae, Cycloctenidae, Desidae, Dictynidae, Hahniidae, Stiphidiidae, and Toxopidae, which exhibit diverse life-history strategies, particularly adapted to a range of extreme environments (Gorneau et al., 2023). For instance, spiders that traditionally inhabited terrestrial environments have demonstrated semi-aquatic lifestyle, such as those in the intertidal spider genus *Desis* (Desidae), which are found on rocky coasts worldwide. Notably, the unique water spider *Argyroneta aquatica* (Dictynidae) has undergone complete adaptation to an aquatic lifestyle (Stefano et al., 2016; Schaber et al., 2023).

Various physiological adaptations to semi-aquatic and aquatic environments have been revealed in spiders. Respiration is most challenging for semi-aquatic and aquatic spiders, because their respiratory systems are the same as these in terrestrial spiders, with book lungs and trachea. The semi-aquatic spider *Desis* live in the spaces among the rocks and crevices and/or shells with available air space in and around the nest (McQueen et al., 2012). While the water spider can construction a "diving bell" for breath underwater, which is filled with air transported from the water surface (Schaber et al., 2023). Due to limited air availability, hypoxic adaptation is especially important for aquatic spider (Seymour and Hetz, 2011). For semi-aquatic spiders *Desis* living in marine habitats, osmoregulation represents one of the most important physiological processes. In semi-aquatic spiders, a variety of physiological, behavioral, or morphological adaptations are caused by osmotic pressure. Unlike terrestrial spiders which prey on insects, fish may represent one of the important food sources for semi-aquatic and aquatic spiders (Nyffeler and Pusey, 2014). Thus, many morphological, physiological, and behavioral characteristics differ among terrestrial, semi-aquatic, and aquatic spiders. However, the molecular and genomic differences underlying adaptations to semi-aquatic and aquatic life in spiders remain poorly understood due to the long-term lack of relevant genomic resources.

The divergent adaptations of spiders to different habitats are hypothesized to be associated with several evolutionary mechanisms, including relaxed selection, positive selection, and likely horizontal gene transfer (HGT). Specifically, species tended to undergo diverse selections on key phenotypic traits during the adaptation to a new environment. Once these key phenotypic traits reach a new optimal value, they may tend to experience weaker selection, leading to a relaxation of selective pressures (Hughes, 2012; Yoder et al., 2010; Hunt et al., 2011). For instance, cetaceans (whales and dolphins) exhibit dramatic morphological changes from their terrestrial ancestors, including the development of flippers and tail flukes, driven by rapid evolutionary processes (Yim et al., 2014; Tian et al., 2016; Noh et al., 2022). In addition, positive selection can drive the fixation of beneficial mutations that enhance fitness in a particular habitat, as seen in a cold-adapted ant (*Tetramorium alpestre*) (Cicconardi et al., 2020). Another example, bdelloid rotifers are microscopic aquatic animals that have acquired a significant portion of their genome through HGT from bacteria, fungi, and plants, which involved in various metabolic processes, including the degradation of complex carbohydrates

and the synthesis of essential amino acids, which facilitate the rotifers surviving in nutrient-poor aquatic environments (Gladyshev et al., 2008). Increasingly, HGT has been documented in insects, where it can contribute to the acquisition of genes that offer survival advantages in specific environments (Neumann et al., 2013). In *Pyrrhocoris apterus* (Pyrrhocoroidea), salivary proteins putatively derived from HGT are critical for salivary sheath formation and other feeding processes (Huang et al., 2024). The horizontally acquired pectin-digesting polygalacturonases (PGs) of the leaf beetle *Phaedon cochleariae* from fungi, which enhance nutrient accessibility, development, and survival (Kirsch et al., 2022). In silkworms and ladybird beetles, HGT-acquired genes have been identified that contribute to pathogen resistance (Zhu et al., 2011; Li et al., 2021). In lepidopteran insects, evidence suggested that HGT-acquired genes is linked to male courtship behavior (Li et al., 2022). Moreover, in whitefly and leaf beetle, HGT-acquiredgenes are essential for survival (Gilbert and Maumus, 2022). The SMaseD enzyme, a distinctive toxin enzyme isolated from the spider *Loxosceles*, has been suggested to evolve through HGT, providing the initial evidence of this phenomenon in spiders (Binford et al., 2005). However, little is known about whether these evolutionary mechanisms contribute to the divergent adaptation to semi-aquatic and fully aquatic life in spiders.

In this study, we generated a high-quality chromosome-scale genome assembly for the aquatic spider *Argyroneta aquatica* and a reference transcriptome for the semi-aquatic spider *Desis martensi*, and integrated publicly available genomes and transcriptomes of their closely related spider species. We performed comparative genomics and molecular evolution analyses of above 22 spider species, including the unique aquatic spider, two semi-aquatic spiders and 19 terrestrial spider species within the in the marronoid clade. Moreover, we presented the use of comparative transcriptomics and metabolomics with the focus on hypoxia response in water spiders. Altogether, we sought to identify genomic signatures associated with semi-aquatic and aquatic adaptations in spiders.

Results

Spider genomic features and phylogenomic analyses

To investigate the genomic adaptations to semi-aquatic and aquatic life in spiders, we selected 22 species, including the unique aquatic spider species, *Argyroneta aquatica*, two semi-aquatic spider species (*Desis martensi* and *Desis japonica*), and 19 terrestrial spider species (*Uloborus diversus* as outgroup) (Fig. 1A and 1B, Table S1) for comparative analyses, with the focus on those in the marronoid clade. In this study, we sequenced the genome of water spider *A. aquatica* and a reference transcriptome of one semi-aquatic species *D. martensi*. We also re-annotated the publicly available genomes of two closely related terrestrial spider species that does not have RefSeq annotations, and *de novo* assembled 17 species based on raw RNA sequencing data from NCBI database (Table S1). Specifically, we reported a high-quality and chromosome-level genome of water spider *A. aquatica*, with an N50 of 2.44 Mbp and total assembly length of 753.77 Mbp that distributed in 11 chromosomes (Fig. 1C, Fig. S1, Table S2), detecting 98.0% arthropod conserved genes (1,013 genes) by BUSCO (Table S2). In addition, we found that approximately 41.53% of the whole genome contains repetitive

121 regions (Fig. 1C, Table S3). A combination of homolog search, transcriptome-based and *ab initio*
122 prediction for searching gene models resulted in a total of 17,632 protein-coding genes (Table S4).
123 Moreover, building on the existing transcriptome of a semi-aquatic spider, *D. japonica*, we
124 supplemented a reference transcriptome of another semi-aquatic species, *D. martensi*, detecting 97.2%
125 arthropod conserved genes by BUSCO. We performed genome re-annotation for two closely related
126 terrestrial spider species, *Amaurobius ferox* (GCA_951213105.1) and *Dolomedes plantarius*
127 (GCA_907164885.2), resulting in 14,258 proteins for *A. ferox* with BUSCO completeness scores of
128 74% and 76,921 proteins for *D. plantarius* with 80.6% arthropod conserved genes (Table S4).
129 The genome-scale phylogeny of aquatic spider, *A. aquatica*, semi-aquatic spiders *D. japonica*,
130 *D. martensi*, and other 19 terrestrial spiders were reconstructed based on a set of 941 single copy
131 orthologous genes. Molecular dating analysis indicated that the divergence time of marronoid spider
132 clade was split approximately 100.07 million years ago (Fig. 1D). The aquatic spider *A. aquatica* is a
133 sister group to Dictynid spiders (i.e. *Adenodictyna kudoae*, *Dictyna felis* and *Lathys maculosa*) with the
134 divergence time at 43.79 million years ago. The *Desis* (*D. japonica* and *D. martensi*) as members of the
135 family Desidae with semi-aquatic lifestyle, while species within the other genera of this family (i.e.
136 *Metaltella simoni* and *Badumna longinqua*) with terrestrial lifestyle, with the divergence time at 55.82
137 million years ago (Fig. 1D).
138
139 **Proposed Semi-aquatic and aquatic transitions**
140 We proposed the hypotheses of the divergent evolutionary routes for semi-aquatic and aquatic spiders
141 (Fig. 1E). Semi-aquatic and aquatic spiders have each independent route evolved from ancestral
142 terrestrial spiders. First, water spiders inhabit freshwater lakes or ponds, while semi-aquatic spiders (*M.*
143 *simoni* and *B. longinqua*) live in intertidal zones along the seashore (Fig. 1A, 1B). Phylogenetic
144 analysis revealed that semi-aquatic spiders, along with closely related species (e.g. *M. simoni* and *B.*
145 *longinqua*) within the same family, exhibit a terrestrial lifestyle. Similarly, aquatic spider, along with
146 its closely related species (e.g. *A. kudoae*, *D. felis*, and *L. maculosa*) within the same family, also
147 exhibit a terrestrial lifestyle (Fig. 2A, 2B).
148
149 **Relaxed selection associated with spider underwater transitions**
150 We estimated the gene-wide ratios of non-synonymous to synonymous substitution (dN/dS, ω) across
151 the spider phylogeny for 4,744 (semi-aquatic transition) and 4,871 (aquatic transition) of shared
152 orthologs, and detected the selection acting on two divergent underwater transitions in spider by
153 calculating three discrete categories of ω ratios (purifying selection: ω1, neutral evolution: ω2, positive
154 selection: ω3) using RELAX. The parameter K-value indicated overall relaxation ($K < 1$) or
155 intensification ($K > 1$) of selection.
156 For semi-aquatic transition, we identified genes that experienced relaxed or intensified
157 selection in semi-aquatic species relative to the other species involved in "evolutionary route 1" (Fig.
158 1E). We found 58 genes under intensification of selection ($K > 1$, q-value <0.05), and 67 genes under
159 relaxation of selection ($K < 1$, q-value <0.05) (Table S5). The sets of genes under intensification and
160 relaxation were not associated with statistically significant enrichment of specific biological functions

161 (Fig. S2-S5). In semi-aquatic spiders, among these genes under intensified selection, we found a set of
162 genes associated with respiratory chain and energy metabolism, such as *SdhD* (Succinate
163 dehydrogenase [ubiquinone] cytochrome b small subunit), *Pfk-2* (phosphofructokinase-2/fructose-2,6-
164 bisphosphatase), *SLC2A3* (Solute Carrier Family 2, Facilitated Glucose Transporter Member 3), with
165 trachea devolopment, such as *Hdac6* (Histone deacetylase 6), with transmembrane ion transport, such
166 as *Slc24a5* (Sodium/potassium/calcium exchanger 5), with fat metabolism and digestion, such as *boca*
167 (LDLR chaperone boca), *Pnliprp2* (Pancreatic lipase-related protein 2), with DNA repair such as
168 *Polr1b* (DNA-directed RNA polymerase I subunit RPA2) and *Rad23a* (UV excision repair protein
169 RAD23 homolog A) (Fig. 2C). Among these genes under relaxation of selection, we found genes
170 associated with fat metabolism and digestion, such as *apmap* (Adipocyte plasma membrane-associated
171 protein), with respond to heat stress such as *HSBP1* (Heat shock factor-binding protein 1), with DNA
172 repair such as *brcc3* (Lys-63-specific deubiquitinase BRCC36) and *Mes4* (DNA polymerase epsilon
173 subunit 4) (Fig. 2C).
174 For aquatic transition, we identified genes that experienced relaxed or intensified selection in
175 water spider species relative to the other species involved in "evolutionary route 2" (Fig. 1E). We
176 identified 92 genes under intensification of selection (K > 1, q-value <0.05), and 59 genes under
177 relaxation of selection (K < 1, q-value <0.05) (Table S6). The sets of intensification genes were
178 associated with statistically significant enrichment of nucleolus, obsolete chromosomal part, and small
179 molecule catabolic process (Fig. S6, S7). Among the set of genes under intensification of selection, we
180 found genes may be associated with the mitochondrial electron transport chain, such as *Uqcrfs1*
181 (Cytochrome b-c1 complex subunit Rieske), *COX5B* (Cytochrome c oxidase subunit 5B), *UQCRC2*
182 (Cytochrome b-c1 complex subunit 2), *FUNDC1* (FUN14 domain-containing protein 1), *CA2*
183 (Carbonic anhydrase 2), and *Uqcrb* (Cytochrome b-c1 complex subunit 7), with the development of the
184 tracheae such as *TMEM234* (Transmembrane protein 234 homolog), *Arp3* (Actin-related protein 3) and
185 *ECE1* (Endothelin-converting enzyme 1), *STX5* (Syntaxin-5), *VEGF* (Vascular endothelial growth
186 factor A), with energy metabolism such as *GPI* (Glucose-6-phosphate isomerase), *GFPT2* (Glutamine--
187 fructose-6-phosphate aminotransferase) and *UQCRC2* (Cytochrome b-c1 complex subunit 2,
188 mitochondrial), with cytochrome P450 enzymes such as *Cyp2j6* (Cytochrome P450 2J6) and *Cyp3a9*
189 (Cytochrome P450 3A9), with DNA repair such as *ppp4cb* (Serine/threonine-protein phosphatase 4
190 catalytic subunit B), with fat metabolism and digestion such as *Pnliprp2* (Pancreatic lipase-related
191 protein 2) and *PNLIP* (Pancreatic triacylglycerol lipase). While the sets of genes under relaxation were
192 not associated with statistically significant enrichment of specific biological functions (Fig. S8, S9).
193 Among the set of genes under relaxation of selection, we found genes may associated with the
194 development of lung and trachea such as *Wnt-1, Catenin beta, TMEM59* (Transmembrane protein 59),
195 with energy metabolism such as *GlyP* (Glycogen phosphorylase), *slc37a3* (Sugar phosphate exchanger
196 3), *ETFA* (Electron transfer flavoprotein subunit alpha, mitochondrial), with DNA repair such as
197 *ALKBH2* (DNA oxidative demethylase ALKBH2), *PGBD4* (PiggyBac transposable element-derived
198 protein 4) (Fig. 2D).
199
200 **Positive selection facilitates spider adaptations to underwater life**

201 We identified 127 positively selected genes (PSGs) in semi-aquatic spiders (Table S7). Among these
202 genes, we found some genes associated with transmembrane ion transport such as *Slc24a5* and *slc37a3,*
203 with DNA repair such as *KIN* (DNA/RNA-binding protein KIN17) and *Rad23a* (UV excision repair
204 protein RAD23 homolog A), with fat metabolism and digestion such as *Pnliprp2, boca, ACADM*
205 (Medium-chain specific acyl-CoA dehydrogenase, mitochondrial), *ABHD17B* (Alpha/beta hydrolase
206 domain-containing protein 17B)*, Acot13* (Acyl-coenzyme A thioesterase 13), with respond to heat
207 stress such as *DNAJB4* (DnaJ homolog subfamily B member 4), with respiratory chain and energy
208 metabolism such as *SdhD* (Succinate dehydrogenase), *SLC2A3*, with DNA repair such as *KIN*
209 (DNA/RNA-binding protein KIN17), with karyopherin gene *Kpna3* (Importin subunit alpha-4) (Fig.
210 2E). While, a total of 26 PSGs were identified in aquatic spider (Table S8), which were associated with
211 fat metabolism and digestion such as *ND-ACP* (*Acyl carrier protein, mitochondrial*)*,* or with
212 karyopherin gene *Kpnb1* (Importin subunit beta-1) (Fig. 2F).
213
214 **Divergent morphology of respiratory systems**
215 The book lung and trachea of four spider species were dissected, respectively. Significant
216 morphological differences between semi-aquatic spider (*D. martensi*) and its terrestrial relative
217 (*Badumna tangae*) within semi-aquatic transition, aquatic spider species (*A. aquatica*) and its terrestrial
218 relative (*D. felis*) within aquatic transition had been observed in their book lung (Fig. 2B). Specifically,
219 compared to its closely related terrestrial species *B. tangae*, the semi-aquatic *D. martensi* has more
220 tracheoles and thicker book lungs. The respiratory system of the water spider (*A. aquatica*) shows
221 significant variation. Compared to its terrestrial relative *D. felis*, water spider has spiracles located
222 closer to the front of its body, along with more tracheoles and thicker book lungs (Fig. 2B).
223
224 **Expansion of the ABC and ACAD gene families through putative horizontal gene transfer (HGT)**
225 Horizontal gene transfer (HGT) has been hypothesized as one of evolutionary mechanisms may
226 facilitate the underwater adaptation in water spider. The ABC and ACAD gene families in water spider
227 genome were found to be significantly expanded in comparison to terrestrial spider *A. ferox* and *D.*
228 *plantarius* genomes (Fig. 3A, 3B). In addition, a set of genes belong to ABC and ACAD gene families
229 that are unique to water spiders showed highly homologous with bacteria. We further validated the
230 reliability of inference of these putative HGT-acquired genes (i.e. ABC and ACAD gene families)
231 using a combination of approaches including Blast-search, phylogenetic mapping, accessing RNA
232 expression, whole genome resequencing reads mapping, and PCR cloning (see methods for details). In
233 total, we identified 75 ABC genes in water spider genome, including 39 genes acquired through HGT.
234 While, its closely related terrestrial species *A. ferox* has 26 ABC genes, and *D. plantarius* has 41 ABC
235 genes (Table S9). For ACAD genes, we identified 35 ACAD genes in water spider genome, including
236 21 genes acquired through HGT. While its closely related species *A. ferox* has 9 ACAD genes, and *D.*
237 *plantarius* has 12 ACAD genes (Table S10). To further understand the characteristics of HGT-acquired
238 genes in ABC and ACAD gene families of water spider, we analyzed the number of Exons and the GC
239 content of each gene family member (Fig 3A, 3B, 3C, Table S9, S10). Intriguingly, we found that the
240 HGT genes have only 1-2 exons, which is significantly fewer than the non-HGT genes. The GC

content of the HGT genes varies widely, differing from that of the non-HGT genes. This variation may be because the GC content of the HGT genes is related to the GC content of the source bacteria. The distribution of HGT genes was on scaffolds that have not been assembled into chromosomes. Additionally, the distribution of these HGT genes is relatively scattered (Fig. 3D and 3E). We mapped the HGT genes of ABC and ACAD gene families to the genome resequencing data of water spider. Except for ABCF-15 (gene ID number: Argyroneta_aquatica_00016198), most genes can be mapped to the genome re-sequencing reads (Fig. 3F).

We pinpointed that these ABC and ACAD HGT-acquired genes were from bacteria, including 14 (29%) genes from Alcaligenaceae, 10 (20.5%) genes from Comamonadaceae, 10 (20.5%) genes from Moraxellaceae, 5 (10%) from Flavobacteriaceae, 4 (8%) from Rhodocyclaceae, 2 (4%) from Chitinophagaceae, and 4 (8%) from other family or order (Fig. S19A). And the main genus of the bacteria was *Acinetobacter*, *Alkanindiges*, and *Emticicia*. Notably, in the freshwater, evidence showed that all these bacteria can be found and confirmed (Fig. S19B).

Comparative transcriptomics and metabolomics in water spider response to hypoxia stress

The water spider exhibits a remarkable ability to survive for over 48 hours with their spiracles closed using paraffin. We conducted comparative transcriptomics and metabolomics analyses to further investigate the molecular basis of hypoxia stress response in water spider.

Under hypoxia stress, comparative transcriptomic analysis of hypoxia group (LO) showed a total of 7 genes up-regulated and 17 genes down-regulated compared to control group (CK) (Fig. 4A). Among these genes, the significantly enriched functions were mainly related to energy metabolism, including *Gld* (Glucose dehydrogenase), *HAO1* (Hydroxyacid oxidase 1), *Enox1* (Ecto-NOX disulfide-thiol exchanger 1), *Slc13a5* (Solute carrier family 13 member 5), *AACS* (Acetoacetyl-CoA synthetase), and *SLC1A2* (Excitatory amino acid transporter 2). Notably, *HAO1* showed intensification selection, while *SLC1A2* showed relaxation selection.

By widely targeted metabolomics methods, in water spider, a total of 579 metabolites were identified and 95 differentiated metabolites (including 57 upregulated and 58 downregulated) were screened under hypoxia stress (Fig. 4B). Under hypoxia stress, the content of intermediate metabolites associated with major energy supply pathways such as glycolysis and TCA cycle changed significantly, for example, Pyridoxal 5'-phosphate, O-Acetylcarnitine, Acetylphosphate, L-Homoserine, GTP and 4-Hydroxyphenylpyruvic acid were up-regulation, while the D-Erythrose 4-phosphate, Glutathione, and Succinic acid were down-regulation in hypoxia stress treatment group, which involved in metabolism of glucide, lipid, protein and nucleic acid.

Comparative tissue-specific and developmental-specific transcriptomes

Finally, we sought to investigate the spatial and temporal expression patterns of genes under selection or HGT genes in water spider. We performed multiple transcriptomic profiling of five tissues (venom, trachea, muscle, testis and ovary), and four development stages (3, 5, 6 and 19 days post fertilization). The expression patterns among different tissues or development stages of genes under relaxation, intensification and positive selection showed no significance (Fig. 5A, 5B).Interestingly, we observed

most of HGT and non-HGT genes within ABC or ACAD gene family differentially expressed in all five tissues, while no HGT genes are expressed during the developmental stages.

Discussion

Understanding the genomic changes that enable spiders to adapt to underwater life is crucial, especially given their terrestrial ancestry. In this study, we conducted comparative genomic analyses of spiders from terrestrial, semi-aquatic, and aquatic habitats to address these questions. The marronoid spider clade includes species that inhabit a variety of specialized microhabitats, such as tundras, deserts, caves, intertidal areas, salt flats, and water, encompassing terrestrial, semi-aquatic, and aquatic environments (Gorneau et al., 2023; Leggett, Vink, and Nelson, 2024). Our study focused on the unique water spider *A. aquatica*, the semi-aquatic *Desis* genus, and various closely related terrestrial spiders within the marronoid clade. Notably, *A. aquatica* is the only spider that spends its entire life underwater. While previous studies have predominantly primarily explored the behavioral and physiological adaptations of water spiders to their aquatic lifestyle (Seymour and Hetz, 2011; Kang et al., 2014; Schaber et al., 2023), we provide a high-quality chromosome-level genome assembly of the water spider *A. aquatica*. To address the long-term lack of reference genomic resource for the semi-aquatic *Desis* genus and various closely related terrestrial spiders within in marronoid clade, we assembled and annotated the transcriptomic data of these spiders.

Genomic evidence supports divergent evolutionary trajectories in semi-aquatic or aquatic adaptation

The evolution of semi-aquatic and aquatic spiders from terrestrial ancestors has been a topic of significant interest. Current taxonomy disperses these semi-aquatic and aquatic taxa across various families (Crews et al., 2020), suggesting that both types have independently evolved from terrestrial spiders (Leggett, Vink, and Nelson, 2024). Our comparative genomics analysis, integrated with life history and predation strategies, reveals distinct evolutionary routes for semi-aquatic and aquatic spiders adapting to underwater life. We observed a higher number of genes under intensified selection associated with the transition to aquatic life compared to semi-aquatic life. Specifically, the ratio of genes under intensified selection to those under relaxed selection was greater in the aquatic transition (92 intensification vs. 59 relaxation) than in the semi-aquatic transition (58 intensification vs. 67 relaxation). This relative abundance of intensified selection indicates that the evolutionary pressures to adapt to a fully aquatic environment are stronger and more urgent than those for a semi-aquatic environment. Adapting to an aquatic habitat requires comprehensive and often drastic genetic changes to survive and reproduce in water-dominated conditions, resulting in a higher number of genes under intensified selection. In contrast, the semi-aquatic transition involves organisms spending time both in water and on land. This dual lifestyle may not necessitate extensive genetic changes, or the selective pressures might be less intense because these organisms can still rely on terrestrial capabilities. Interestingly, there were more positively selected genes (PSGs) in semi-aquatic spiders (127 genes) compared to aquatic spiders (67 genes). This suggests that semi-aquatic spiders experience diverse and

321 dynamic environmental pressures, requiring a greater number of genes to undergo positive selection
322 during the adaptation to both terrestrial and aquatic environments. Aquatic spiders may have already
323 undergone significant evolutionary adaptations to their environment, resulting in fewer genes currently
324 under positive selection. Alternatively, the aquatic environment might present more stable or less
325 varied selective pressures compared to the dual habitats of semi-aquatic spiders. Collectively, semi-
326 aquatic spiders exhibit a higher number of positively selected genes due to the need to adapt to diverse
327 conditions, while aquatic spiders display a higher number of genes under intensified selection,
328 reflecting strong selective pressures for comprehensive adaptations to a fully aquatic environment.
329
330 **Respiratory adaptation in semi-aquatic and aquatic spiders**
331 In aquatic and semi-aquatic environments, spiders encounter significant risks of respiratory distress and
332 drowning, necessitating crucial adaptations to prevent asphyxia for their survival. Spider respiration
333 depends on internal branching tubes and sacs connected to the atmosphere through spiracles on the
334 abdominal wall (Foelix, 2011). These spiracles have a valvular mechanism that can remain open or
335 partially closed to minimize water loss. Our study reveals significant advancements in the respiratory
336 systems of semi-aquatic and aquatic spiders compared to their terrestrial counterparts. This evolution is
337 particularly pronounced in fully aquatic spiders, which exhibit a highly refined gaseous exchange
338 system (Fig. 2B). Among the genes with intensified or relaxed selection in water spiders, eight genes
339 are associated with the development of the tracheae, as their functions have been reported in other
340 animals. In water spiders, the *Wnt-1* and *Catenin beta* genes exhibited relaxed selection, with the
341 Wnt/β-catenin pathway playing a pivotal role in trachea development (Llimargas and Lawrence, 2001;
342 Gerhardt et al., 2018; Brechbuhl et al., 2011). The *TMEM234* gene shows intensified selection, while
343 *TMEM59* exhibits relaxed selection. Although information regarding their structures, functions, and
344 mechanisms within the TMEM family is limited, there is evidence suggesting that these gene family
345 proteins are required for normal trachea development (Rock et al., 2008; Scholl et al., 2019; Sébastien
346 et al., 2020). The *Arp3* and *ECE1* genes may regulate the diameter of the trachea and its contraction
347 and expansion during respiration by participating in actin polymerization and contraction in smooth
348 muscle (Zhang et al., 2005; De Campo et al., 2002). These genes also show intensified selection in
349 water spiders. The *Syntaxin 5* gene, essential for the seamless growth and maintenance of tubular
350 structures (Bourne et al., 2022), exhibits heightened selection pressure in water spiders. Additionally,
351 *VEGF*, which controls cell migration during trachea development (Kipryushina et al., 2015), shows
352 heightened selection pressure in water spiders. In semi-aquatic spiders, the gene *Hdac6* has been
353 identified as under intensified selection and may function in the cilia of the trachea (Ran et al., 2016).
354 This specific gene potentially plays a functional role in the respiratory adaptations required for the dual
355 aquatic and terrestrial lifestyle of semi-aquatic spiders.
356
357 **Osmoregulatory adaptation in semi-aquatic and aquatic spiders**
358 Osmoregulation is a critical physiological process for spiders living in aquatic or semi-aquatic
359 environments. While cuticular structure and reduced metabolic rates help combat some osmotic stress,
360 they are not sufficient on their own (Renault et al., 2016). Ion regulation is crucial for maintaining

osmotic pressure. Under various experimental saline conditions, the concentrations of metal ions (Ca^{2+}, K^{+}, Mg^{2+}, Na^{+}) in spiders were reported to vary (Renault et al., 2016). The gene *Slc24a5*, which encodes a sodium/calcium/potassium exchanger involved in transmembrane ion transport, showed intensified positive selection in semi-aquatic spiders but exhibited relaxed selection in water spiders. Semi-aquatic spiders, such as those in the *Desis* genus, inhabit environments with high osmotic pressure, whereas water spiders face low osmotic pressure. In water spiders, horizontally acquired ABC genes, such as the high-affinity branched-chain amino acid transport ATP-binding proteins LivG and LivF, were identified. These genes provide the energy required to drive the transmembrane transport of branched-chain amino acids (Albers et al., 2004), potentially indirectly affecting osmotic pressure and regulation. Elevated salt loads may provoke DNA breaks, protein denaturation, and uncontrolled molecular interactions. Many genes related to DNA repair showed signs of intensified, relaxed, or positive selection among water and semi-aquatic spiders. In water spiders, genes such as *ALKBH2* and *PGBD4* exhibited relaxed selection, while *ppp4cb* showed intensified selection. In semi-aquatic spiders, genes such as *brcc3* and *Mes4* exhibited relaxed selection, while *Polr1b* and *Rad23*a showed intensified selection. Additionally, *KIN* and *Rad23a* genes exhibited positive selection. These findings underscore the complex and varied genetic adaptations that enable spiders to thrive in aquatic and semi-aquatic environments.

Genetic adaptation in fat metabolism and digestion in semi-aquatic and aquatic spiders

In both aquatic and semi-aquatic spiders, various genes related to fat metabolism and digestion exhibit intensified or positive selection. For example, in semi-aquatic spiders, the genes *boca*, *ACADM*, *ABHD17B*, and *Acot13* show positive selection. The gene *Pnliprp2* exhibits both intensified and positive selection. In water spiders, the genes *Pnliprp2* and *PNLIP* display intensified selection, while *ND-ACP* shows positive selection. Additionally, many horizontally acquired ACAD genes were identified in water spiders, which are essential for energy production through mitochondrial fatty acid beta-oxidation (Shen et al., 2009). These genetic adaptations are particularly advantageous in environments where high-energy food sources, such as fish rich in long-chain fatty acids, constitute a primary diet (Nyffeler and Pusey, 2014; Shen et al., 2009).

Hypoxia adaptation in aquatic spiders

Adaptation to hypoxia conditions is crucial for the successful transition of spiders from terrestrial back to aquatic environments. In water spiders, a variety of hypoxia-regulated genes involved in various aspects of energy metabolism have been identified as being under selection (Fuhrmann and Brune, 2017). Genes under intensified selection include *GPI*, *GFPT2*, and *UQCRC2*, while those under relaxed selection include *GlyP*, *slc37a3*, and *ETFA*. Under hypoxic conditions, enzyme reactions involving oxygen consumption may be influenced (Huwiler and Pfeilschifter, 2006). Notably, two genes encoding cytochrome P450, specifically *Cyp2j6* and *Cyp3a9*, have been identified as undergoing intensified selection. This finding aligns with the adaptive processes required for hypoxia tolerance in aquatic environments.

Thermal adaptation in semi-aquatic Spiders

As poikilotherms, the semi-aquatic *Desis* spiders, being intertidal arthropods, require diverse strategies to cope with extreme temperature variations. The genes Heat shock factor-binding protein 1 (under relaxed selection) and *DNAJB4* (under positive selection) may play crucial roles in the cellular response to heat stress in these spiders.

In conclusion, these results provide critical insights into the genomic adaptations that enable spiders to thrive in underwater environments. By examining the roles of relaxed selection, intensified selection, positive selection, and HGT, we elucidate the genomic changes involved in the adaptation of spiders from terrestrial to semi-aquatic or aquatic habitats. Collectively, these findings not only enhance our understanding of the evolutionary paths of the semi-aquatic and aquatic spiders but also contribute to the broader knowledge of adaptive strategies in aquatic organisms.

Methods

Sample collection, sequencing and draft data collection

Water spider *A. aquatica* individuals were collected from Xilingol league of InnerMongolia (Figure 1A). Ten whole body of female adult water spiders without abdomen and stomach were used for whole genome sequencing. And an additional eight adults water spiders (four females and four males) were processed for whole genome resequencing. The tissues (including venom, muscle, trachea, testis, and ovary) and the embryo among different stages (3, 5, 6 and 19 days post fertilization) of water spider were used for transcriptome sequencing, each group contains three replicates. In addition, a female *Desis martensi* was collected from Sanya, Hainan Province (Figure 1A), followed by Starvation treatment for three days and used for transcriptome sequencing.

A combination of two sequencing approaches including short reads on MGISEQ 2000 platform and long reads on Oxford Nanapore Technology (ONT) PromethION platform (Nextomics Bioscience, Wuhan, China) had been performed for the whole genome sequencing. Specifically, the total genomic DNA for genome sequencing was extracted using Qiagen Blood & Cell Culture DNA Mini Kit. For short reads, the DNA libraries were prepared with the MGIEasy FS DNA Library Prep Set (BGI, Shenzhen, China) following the MGISEQ protocol, a total of 212.12 Gb from 2.13 billion clean reads were obtained. For long reads, the ONT libraries were prepared following the protocol of Ligation Sequencing Kit 1D (SQK-LSK108, ONT), the N50 of the raw nanopore reads length was 25 kb. An additional Hi-C sequencing was performed (Berry Genomics, Beijing, China). After cross-linked long-distance physical interactions and digested with MboI, the 300bp paired-end sequencing libraries were prepared for sequencing on an Illumina HiSeq PE150 platform. Finally, the transcriptome sequencing was performed on an Illumina NovaSeq 6000 platform (Berry Genomics, Beijing, China) yielding 150-bp paired-end reads. The details were shown in Table S11.

Besides the newly sequenced omics data, we download available genomes of terrestrial spider species including *Amaurobius ferox* (GCA_951213105.1), *Dolomedes plantarius* (GCA_907164885.2) and *Uloborus diversus* (GCF_026930045.1), and existing transcriptome data of spider species

including a semi-aquatic spider *Desis japonica* (DRR296513), and terrestrial spiders *Metaltella simoni* (DRR296491), *Badumna longinqua* (DRR297554), *Adenodictyna kudoae* (DRR297439), *Dictyna felis* (DRR297156), *Brommella punctosparsa* (DRR296607), *Lathys maculosa* (DRR297435), *Cybaeus nipponicus* (DRR297040), *Cicurina japonica* (DRR297872), *Hahnia corticicola* (DRR296643), *Neoantistea quelpartensis* (DRR297193), *Agelena silvatica* (DRR297956), *Pireneitega luctuosa* (DRR297239), *Tegenaria domestica* (DRR297377), *Ambicodamus* sp. (DRR297731), *Pandava laminata* (DRR297807), *Nurscia albofasciata* (DRR296579) from NCBI Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra).

Hypoxia stress treatment

For hypoxia treatment stress experiment, the lung slit and spiracle of adult male water spider was sealed by paraffin, then raised it in boxes at room temperature. After 48 hours, the spider was freezed with liquid nitrogen, removed the paraffin, and stored at −80 °C. The control group was not treated with paraffin seal, and raised it in boxes at room temperature. The details were shown in Figure S20. Finally, the samples were used for Liquid chromatography–mass spectrometry (LC-MC) assay and transcriptome sequencing (Suzhou PANOMIX Biomedical Tech, China). For LC-MC assay, each group contains eight replicates. For transcriptome sequencing, each group contains three replicates.

The morphological analysis of respiratory systems

The four spiders including *Desis martensi*, *Badumna tangae*, water spider and *Dictyna felis* were used for morphological analysis of respiratory systems. All specimens were preserved in 75% ethanol in in the School of Life Sciences, Southwest University, Chongqing, China (SWUC). After washed by ddH_2O, the specimen was then cleared by immersing them in pancreatin (Álvarez-Padilla and Hormiga 2007). The respiratory systems including book lungs and trachea were examined and measured using a Leica M205A stereomicroscope equipped with a drawing tube, a Leica DFC450 camera, and LAS software v4.6. Photos were taken with a Kuy Nice CCD mounted on an Olympus BX53 compound microscope. Compound focus images were generated using Helicon Focus 6.7.1 software.

Genome assembly

The genome of water spider was assembled with one road of polishing iterations and the minimum overlap between reads of 3,000 using Flye v2.5 (Kolmogorov et al., 2019). The resulting genome assembly was further polished for two rounds using NGS data with NextPolish v1.0.5 (Hu et al., 2020). The Heterozygous regions were removed by Purge Haplotigs v1.1.0 (Roach et al., 2018) with the haplotigs identifying of 50% cut-off. The chromosome-level assembly was generated using Hi-C reads with Juicer v1.6.2 (Durand et al., 2016). To remove the potential contaminant sequences, the assembly genome was inspected against the NCBI nucleotide (nt) and UniVec databases by the software of HS BLASTN and BLAST+ (blastn) v2.7.1 (Chen et al., 2015). Finally, the quality and completeness of genome assemblies was assessed with BUSCO v5.4.5 (Simao et al., 2015) with arthropoda_odb10 data set (n= 1,013).

481 **Genome annotation**

482 For the annotation of repetitive elements in genome, we searched it by means of a combination of ab
483 initio and homology. Firstly, the species-specific repeat database (Ab-initio database) was constructed
484 using RepeatModeler v2.0.1 (Flynn et al., 2020), and further combined the known repeat library
485 (Repbase) and Ab-initio database as the reference repeat database. The repetitive elements in the
486 genome was finally identified by RepeatMasker (Tarailo-Graovac and Chen, 2009).

487 Non-coding RNAs were annotated by Infernal v1.1.2 (Nawrocki and Eddy, 2013) and tRNAscan-
488 SE v2.0.6 (Chan et al., 2021). In addition, a tool "EukHighConfidenceFilter" in tRNAscanSE was used
489 for further filtering transfer RNAs (tRNAs) with high confidence.

490 For the annotation of gene structure, we integrated the ab initio, transcriptome-based, and protein
491 homology-based evidence by Maker v2.31.10 (Holt and Yandell, 2011). Gene structure prediction of
492 ab inito was identified by Augustus v3.3.3 (Hoff and Stanke, 2019) and GeneMark-ES/ET/EP v4.48
493 3.60 lic (Bruna et al., 2020). To model the sequence properties accurately, both gene finders (Augustus
494 and GeneMark-ES/ET/EP) were initially trained using the BRAKER v2.1.5 pipline (Bruna et al., 2021).
495 The transcriptome data for whole body water spider was mapped to the genome by the software
496 HISAT2 v2.2.0 (Kim et al., 2015). BRAKER pipeline was run with default parameters. Then further
497 assembled RNA-Seq alignments into potential transcripts by Stringtie v2.1.3 (Pertea et al., 2015), and
498 the results was provided as input for Maker via the "est" option. The protein homology-based evidence
499 was download from NCBI database, which including the protein sequences of *Stegodyphus mimosarum*
500 (GCA 000611955.2), *Parasteatoda tepidariorum* (GCA 000365465.3), *Trichonephila clavipes* (GCA
501 002102615.1), *Drosophila melanogaster* (GCA 000001215.4), *Ixodes scapularis* (GCA 002892825.2),
502 *Daphnia pulex* (GCA 900092285.2) and *Strigamia maritima* (GCA 000239455.1), which were
503 combined and fed to Maker pipline.

504 Gene function annotation was predicted by Diamond v0.9.24 (Buchfink et al., 2015),
505 InterProScan v5.41–78.0 (Mulder and Apweiler, 2007) and eggNOG-mapper v2.0 (Cantalapiedra et al.,
506 2021). The UniProtKB/Swissprot database was against by the software Diamond with the mode of
507 more sensitive, maximum number for target sequences of 1, and e-value of 1e-5. The following 5
508 databases including Pfam (Mistry et al., 2021), Panther (Mi et al., 2016), Gene3D (Lewis et al., 2018),
509 Superfamily (Pandurangan et al., 2019), and Conserved Domain Database (CDD) (Marchler-Bauer et
510 al., 2017) was against by the software InterProScan. The the eggNOG v5.0 database (Huerta-Cepas et
511 al., 2019) was against by the software eggNOG-mapper.

512 Due to the annotation files for *A. ferox* and *D. plantarius* are not publicly available, we also
513 performed separate annotations for these two species in this project.

514

515 **RNA-seq assembly and annotation**

516 We assembled and annotated the transcriptomes data in these study. Quality control of the raw data was
517 performed using BBTools suite v38.67 of "bbduk.sh" for trimming the reads' ends to Q20 with reads
518 shorter than 15 bp or with >5 Ns. Then the transcript sequences was assembled with de novo assembly
519 using Trinity v.25.1 (Grabherr et al., 2011). The redundancy sequences were removed by cd-hit-est
520 v4.8.1. And the assembled transcript sequences were annotated by Trinotate v3.0.1 (Bryant et al., 2017).

521

522 **Species phylogeny and gene orthology**

523 A total of eight families in marronoid spider clade including Amaurobiidae (*A. ferox*), Agelenidae (*A.*
524 *silvatica*, *P. luctuosa*, *T. domestica*), Desidae (*D. japonica*, *D. martensi*, *M. simoni*, *B. longinqua*),
525 Hahniidae (*H. corticicola*, *N. quelpartensis*), Cybaeidae (*C. nipponicus*), Cicurinidae (*C. japonica*),
526 Dictynidae (*A. kudoae*, *D. felis*, *L. maculosa*), and five closely related families spidcies including
527 Cicurinidae (*B. punctosparsa*), Nicodamidae (*Ambicodamus* sp.), Titanoecidae (*P. laminata*, *N.*
528 *albofasciata*), Pisauridae (*D. plantarius*), and the Uloboridae (*U. diversus*) as outgroup were combined
529 with water spider *A. aquatica* for Gene clusters analysis. The proteins of these species were fed to
530 OrthoFinder v2.5.4 (Emms et al., 2019) for orthologous gene clusters analysis.

531 For phylogenetic analysis, a total of 166 single copy orthologous genes were used to examine the
532 phylogenetic relationships. First, the genes were separately aligned by MAFFT v7.471 (Katoh and
533 Standley, 2013) with L-INS-I strategy. The resulting alignments were then trimmed by trimAl v1.4
534 (Capella-Gutierrez et al., 2009) with method "automated1". Then the results were concatenated by
535 FASconCATG v1.04 (Kuck and Longo, 2014). Finally, the maximum likelihood (ML) tree was
536 constructed by IQ-TREE v2.0.3 (Minh et al., 2020) with the optimization number of rcluster algorithm,
537 replicates for ultrafast bootstrap, and replicates for Shimodaira–Hasegawa (SH) approximate likelihood
538 ratio tests being 10, 1000, and 1000, respectively. The divergence time calibration was performed with
539 the reference from Magalhaes (Magalhaes et al., 2020; Moradmand, Schonhofer and Jager, 2014),
540 paleobiodb database (https://paleobiodb.org/) and TimeTree database (http://www.timetree.org/), with
541 Hahniinae (43-47.8 Mya), Dictynidae (100-189.7 Mya), and Entelegynae (132.9-140 Mya). The
542 divergence time was calculated using package MCMC Tree in PAML v4.10.0 (Yang, 2007).

543

544 **Pattern of molecular evolution in shared orthologs**

545 To determine the patterns of molecular evolution of terrestrial, semi-aquatic and aquatic spiders in the
546 set of semi-aquatic and aquatic branches across the phylogeny, we characterized the rates of
547 nonsynonymous to synonymous rate (dN/dS) in each shared ortholog. The protein sequence alignment
548 was performed by MAFFT v7.471 (Katoh and Standley, 2013). Then the codon alignments of these
549 orthologs which derived from amino acid alignments and corresponding DNA sequences using
550 PAL2NAL v.14 (Suyama et al., 2006). And the shared ortholog alignments with length of at least 50
551 codons was executed filtration. The discrete codon substitution rate (dN/dS, ω) representing water
552 spider, semi-aquatic spider and terrestrial spider species were estimated by HyPhy (Pond et al., 2005)
553 following the pipeline as previously described (Tong et al., 2022).

554 Further, the pattern of selection acting on the semi-aquatic transition and aquatic
555 transitionfrom was detected by RELAX (Wertheim et al., 2015) at gene-wide level. Genes are
556 considered to experience significantly relaxed selection with a K-value < 1 (q-value < 0.05).
557 Conversely, when K-value > 1 (q-value < 0.05), the genes are considered to experience significantly
558 intensified selection. The q-values represent an estimate of the false discovery rate (FDR) for each gene
559 which were calculated using the Benjamini–Hochberg method.

560 Finally, the signal of positive selection in genes involved in semi-aquatic and aquatic

561 transitions were detected by BUSTED-PH (https://github.com/veg/hyphy-
562 analyses/blob/master/BUSTED-PH/README.md) (Murrell et al., 2015) and followed by the
563 correcting for multiple testing (Benjamini–Hochberg). We defined the genes under positive selection
564 with q-value < 0.05 for foreground branches (i.e. aquatic spider), q-value > 0.05 for background
565 branches (i.e. territorial spiders), and q-value < 0.05 for the difference between foreground and
566 background branches.
567
568 **HGT gene identification and validation**
569 To determine if specific genes in water spider that may have been horizontally acquired, we employed
570 conservative phylogeny-based approach following the methodology (Li et al., 2022). For ABC and
571 ACAD genes in water spiders, we first blast these genes with corresponding protein sequence by
572 BLASTP in NCBI database.
573 To further validate the putative HGT-acquired genes, we compared them against previously
574 published genes across the phylogenetic tree. We then analyzed gene expression across different tissues
575 to assess their biological relevance. Additionally, we selected nine putative HGT-acquired genes—six
576 ABC genes and three ACAD genes—for further experimental validation. For these genes, we
577 performed PCR assays to amplify regions flanking the foreign gene, targeting sequences between 700
578 and 1,500 bp. The PCR products were analyzed via agarose gel electrophoresis to confirm the expected
579 sizes, and successful amplifications were sequenced using Sanger sequencing. If the flanking regions of
580 the HGT-acquired genes were successfully amplified and the Sanger sequences showed over 98%
581 identity with the reference genome, the gene was considered validated. Primer sequences for these PCR
582 assays are provided in Table S12. Additionally, we visualized the distribution of HGT-acquired genes
583 on scaffold 1 of the water spider genome using the online tool MG2C v2.1 (Chao et al., 2021).
584 To confirm the HGT-acquired genes available in genome, the kallisto v0.48.0 (Bray et al.,
585 2016) was used for counts the reads for HGT-acquired genes among the six water spider re-sequencing
586 genome reads.
587
588 **Differential gene expression analysis**
589 To explore the changes in gene expression response to hypoxia stress, we performed comparative
590 transcriptomic analysis between spider samples under hypoxia stress and normal oxygene condition as
591 control. For this, we aligned the transcriptome sequencing reads of whole body to the whole genome of
592 water spider by HISAT2. The read counts were counted by featureCounts v1.6.4 (Liao et al., 2014) and
593 then normalized among samples FPKM normalization method by DESeq2 package v1.24.0 (Love et al.,
594 2014). Genes with an adjusted p value <0.05 and |log2 ratio| ≥ 1 were considered to be differentially
595 expressed. The volcano plot was generated using the "ggplot2" package in R.
596
597 **Differential metabolites analysis**
598 To investigate differential metabolites induced by hypoxia treatment, we conducted non-targeted
599 metabolomics using LC-MS (Suzhou PANOMIX Biomedical Tech, China) to compare spider samples
600 under hypoxic and normal oxygen conditions. The sample preparation followed these steps: first, 1g of

sample was accurately weighed and mixed with 1000 μL of tissue extract (75% methanol [9:1], 25% H2O) and three steel balls in a 2mL centrifuge tube. The samples were then ground in a tissue grinder at 50 Hz for 60 seconds, repeating the process twice. After grinding, the samples underwent ultrasonic shaking for 30 minutes, followed by an ice bath for 30 minutes. The mixture was centrifuged at 12,000 rpm for 10 minutes at 4°C, and the supernatant was collected and dried. To prepare the samples for LC-MS detection, 200 μL of 50% acetonitrile solution containing 2-Amino-3-(2-chlorophenyl)-propionic acid (4 ppm) was used to re-dissolve the dried samples. The supernatant was filtered through a 0.22 μm membrane and transferred to a detection vial for LC-MS analysis. Various statistical methods, including principal component analysis, clustering, and univariate analysis, were applied to compare metabolite profiles between the different groups. The volcano plot illustrating significant metabolite changes was generated using the "ggplot2" package in R.

Data availability

The raw Nanopore data are available on NCBI with the project of PRJNA559680 with the SRA accession numbers of SRR9994193. And the whole-body RNA accession number is SRR9994194. For water spider *Argyroneta aquatica,* the genome sequence (https://cstr.cn/31253.11.sciencedb.14858), the raw Illumina data (https://cstr.cn/31253.11.sciencedb.11012), the Hi-C data (https://doi.org/10.57760/sciencedb.15431), the metabolites and transcriptomic raw data (https://cstr.cn/31253.11.sciencedb.15069), the RNA data for different stage of embryo (https://cstr.cn/31253.11.sciencedb.10923), and the RNA data of different tissues (https://cstr.cn/31253.11.sciencedb.15029) are available on scienceDB. And the RNA data of a female *Desis martensi* (https://cstr.cn/31253.11.sciencedb.09032 is also available on scienceDB.

Reference

Albers SV, Koning SM, Konings WN, and Driessen AJM. 2004. Insights into ABC transport in archaea. *J Bioenerg Biomembr.* 36:5-15.

Álvarez-Padilla F., Hormiga G. 2007. A protocol for digesting internal soft tissues and mounting spiders for scanning electron microscopy. *The Journal of Arachnology*. 35: 538-542.

Binford GJ, Cordes MH, and Wells MA. 2005. Sphingomyelinase D from venoms of *Loxosceles* spiders: evolutionary insights from cDNA sequences and gene structure. *Toxicon.* 45:547-560.

Bourne CM, Lai DC, and Schottenfeld-Roames J. 2022. Regulators of the secretory pathway have distinct inputs into single-celled branching morphogenesis and seamless tube formation in the *Drosophila* trachea. *Dev Biol.* 490:100-109.

Bray NL, Pimentel H, Melsted P, and Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nature biotechnology.* 34:525-527.

Brechbuhl HM, Ghosh M, Smith MK, Smith RW, Li B, Hicks DA, Cole BB, Reynolds PR, and Reynolds SD. 2011. Beta-catenin dosage is a critical determinant of tracheal basal cell fate determination. *Am J Pathol.* 179:367-379.

641 Bruna T, Hoff KJ, Lomsadze A, Stanke M, and Borodovsky M. 2021. BRAKER2: automatic
642 eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein
643 database. *NAR genomics and bioinformatics.* 3:lqaa108-lqaa108.
644 Bruna T, Lomsadze A, and Borodovsky M. 2020. GeneMark-EP+: eukaryotic gene prediction with
645 self-training in the space of genes and proteins. *NAR Genom Bioinform.* 2:lqaa026.
646 Bryant DM, Johnson K, DiTommaso T, Tickle T, Couger MB, Payzin-Dogru D, Lee TJ, Leigh ND,
647 Kuo TH, Davis FG, Bateman J, Bryant S, Guzikowski AR, Tsai SL, Coyne S, Ye WW, Freeman
648 RM, Jr., Peshkin L, Tabin CJ, Regev A, Haas BJ, and Whited JL. 2017. A Tissue-Mapped Axolotl
649 De Novo Transcriptome Enables Identification of Limb Regeneration Factors. *Cell Rep.* 18:762-
650 776.
651 Buchfink B, Xie C, and Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat*
652 *Methods.* 12:59-60.
653 Cantalapiedra CP, Hernandez-Plaza A, Letunic I, Bork P, and Huerta-Cepas J. 2021. eggNOG-mapper
654 v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic
655 Scale. *Molecular biology and evolution.* 38:5825-5829.
656 Capella-Gutierrez S, Silla-Martinez JM, and Gabaldon T. 2009. TrimAl: a tool for automated alignment
657 trimming in large-scale phylogenetic analyses. *Bioinformatics.* 25:1972-1973.
658 Chan PP, Lin BY, Mak AJ, and Lowe TM. 2021. tRNAscan-SE 2.0: improved detection and functional
659 classification of transfer RNA genes. *Nucleic Acids Res.* 49:9077-9096.
660 Chen Y, Ye WC, Zhang YD, and Xu YS. 2015. High speed BLASTN: an accelerated MegaBLAST
661 search tool. *Nucleic Acids Res.* 43:7762-7768.
662 Chao J, Li Z, Sun Y, Aluko OO, Wu X, Wang Q, and Liu G. 2021. MG2C: a user-friendly online tool
663 for drawing genetic maps. *Molecular Horticulture.* 1.
664 Cicconardi F, Krapf P, D'Annessa I, Gamisch A, Wagner HC, Nguyen AD, Economo EP, Mikheyev AS,
665 Guenard B, Grabherr R, Andesner P, Wolfgang A, Di Marino D, Steiner FM, and Schlick-Steiner
666 BC. 2020. Genomic Signature of Shifts in Selection in a Subalpine Ant and Its Physiological
667 Adaptations. *Molecular biology and evolution.* 37:2211-2227.
668 Crews SC, Garcia EL, Spagna JC, Van Dam MH, and Esposito LA. 2020. The life aquatic with spiders
669 (Araneae): repeated evolution of aquatic habitat association in Dictynidae and allied taxa.
670 *Zoological Journal of the Linnean Society.* 189:862-920.
671 De Campo BA, Goldie RG, Jeng AY, and Henry PJ. 2002. Role of endothelin-converting enzyme,
672 chymase and neutral endopeptidase in the processing of big ET-1, ET-1(1-21) and ET-1(1-31) in
673 the trachea of allergic mice. *Clin Sci (Lond).* 103 Suppl 48:353S-356S.
674 Durand NC, Shamim MS, Machol I, Rao SSP, Huntley MH, Lander ES, and Aiden EL. 2016. Juicer
675 Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst.* 3:95-
676 98.
677 Emms DM, and Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative
678 genomics. *Genome Biol.* 20:238.
679 Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, and Smit AF. 2020. RepeatModeler2
680 for automated genomic discovery of transposable element families. *Proceedings of the National*

681 *Academy of Sciences of the United States of America.* 117:9451-9457.
682 Foelix RF. 2011. Biology of Spiders 3rd. New York: Oxford University Press.
683 Fuhrmann DC, and Brune B. 2017. Mitochondrial composition and function under the control of
684 hypoxia. *Redox Biol.* 12:208-215.
685 Gerhardt B, Leesman L, Burra K, Snowball J, Rosenzweig R, Guzman N, Ambalavanan M, and Sinner
686 D. 2018. Notum attenuates Wnt/beta-catenin signaling to promote tracheal cartilage patterning.
687 *Dev Biol.* 436:14-27.
688 Gilbert C, and Maumus F. 2022. Multiple Horizontal Acquisitions of Plant Genes in the Whitefly
689 *Bemisia tabaci*. *Genome Biol Evol.* 14.
690 Gorneau JA, Crews SC, Cala-Riquelme F, Montana KO, Spagna JC, Ballarin F, Almeida-Silva LM, and
691 Esposito LA. 2023. Webs of intrigue: museum genomics elucidate relationships of the marronoid
692 spider clade (Araneae). *Insect Syst Diver.* 7.
693 Gladyshev EA, Meselson M, and Arkhipova IR. 2008. Massive horizontal gene transfer in Bdelloid
694 rotifers. *Science.* 320:1210-1213.
695 Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L,
696 Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F,
697 Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, and Regev A. 2011. Full-length
698 transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology.*
699 29:644-652.
700 Hoff KJ, and Stanke M. 2019. Predicting Genes in Single Genomes with AUGUSTUS. *Curr Protoc*
701 *Bioinformatics.* 65:e57.
702 Holt C, and Yandell M. 2011. MAKER2: an annotation pipeline and genome-database management
703 tool for second-generation genome projects. *BMC Bioinformatics.* 12:491.
704 Hu J, Fan J, Sun Z, and Liu S. 2020. NextPolish: a fast and efficient genome polishing tool for long-
705 read assembly. *Bioinformatics.* 36:2253-2255.
706 Huerta-Cepas J, Szklarczyk D, Heller D, Hernandez-Plaza A, Forslund SK, Cook H, Mende DR,
707 Letunic I, Rattei T, Jensen LJ, von Mering C, and Bork P. 2019. eggNOG 5.0: a hierarchical,
708 functionally and phylogenetically annotated orthology resource based on 5090 organisms and
709 2502 viruses. *Nucleic Acids Res.* 47:D309-D314.
710 Hughes AL. 2012. Evolution of adaptive phenotypic traits without positive Darwinian selection.
711 *Heredity.* 108:347-353.
712 Hunt BG, Ometto L, Wurm Y, Shoemaker D, Yi SV, Keller L, and Goodisman MAD. 2011. Relaxed
713 selection is a precursor to the evolution of phenotypic plasticity. *Proceedings of the National*
714 *Academy of Sciences of the United States of America.* 108:15936-15941.
715 Huang HJ, Li LL, Ye ZX, Lu JB, Lou YH, Wei ZY, Sun ZT, Chen JP, Li JM, and Zhang CX. 2024.
716 Salivary proteins potentially derived from horizontal gene transfer are critical for salivary sheath
717 formation and other feeding processes. *Commun Biol.* 7:257.
718 Huwiler A, and Pfeilschifter J. 2006. Hypoxia and lipid signaling. *Biol Chem.* 387:1321-1328.
719 Kang J-H, Lee Y-J, Oh B-K, Lee S-K, Hyun B-R, Lee B-W, Choi Y-G, Nam K-S, and Lim J-D. 2014.
720 Microstructure of the water spider (*Argyroneta aquatica*) using the scanning electron microscope.

721 *Journal of Asia-Pacific Biodiversity.* 7:484-488.
722 Katoh K, and Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7:
723 Improvements in Performance and Usability. *Molecular biology and evolution.* 30:772-780.
724 Kim D, Landmead B, and Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory
725 requirements. *Nat Methods.* 12:357-U121.
726 Kipryushina YO, Yakovlev KV, and Odintsova NA. 2015. Vascular endothelial growth factors: A
727 comparison between invertebrates and vertebrates. *Cytokine Growth F R.* 26:687-695.
728 Kirsch R, Okamura Y, Haeger W, Vogel H, Kunert G, and Pauchet Y. 2022. Metabolic novelty
729 originating from horizontal gene transfer is essential for leaf beetle survival. *Proceedings of the*
730 *National Academy of Sciences of the United States of America.* 119.
731 Kolmogorov M, Yuan J, Lin Y, and Pevzner PA. 2019. Assembly of long, error-prone reads using repeat
732 graphs. *Nature biotechnology.* 37:540-546.
733 Kuck P, and Longo GC. 2014. FASconCAT-G: extensive functions for multiple sequence alignment
734 preparations concerning phylogenetic studies. *Front Zool.* 11:81.
735 Leggett MA, Vink CJ, and Nelson XJ. 2024. Adaptation and Survival of Marine-Associated Spiders
736 (Araneae). *Annu Rev Entomol.* 69:481-501.
737 Lewis TE, Sillitoe I, Dawson N, Lam SD, Clarke T, Lee D, Orengo C, and Lees J. 2018. Gene3D:
738 Extensive prediction of globular domains in proteins. *Nucleic Acids Res.* 46:D1282.
739 Li HS, Tang XF, Huang YH, Xu ZY, Chen ML, Du XY, Qiu BY, Chen PT, Zhang W, Slipinski A, et al.
740 2021. Horizontally acquired antibacterial genes associated with adaptive radiation of ladybird
741 beetles. *BMC biology.* 19:7.
742 Li Y, Liu Z, Liu C, Shi Z, Pang L, Chen C, Chen Y, Pan R, Zhou W, Chen XX, Rokas A, Huang J, and
743 Shen XX. 2022. HGT is widespread in insects and contributes to male courtship in lepidopterans.
744 *Cell.* 185:2975-2987 e2910.
745 Liao Y, Smyth GK, and Shi W. 2014. FeatureCounts: an efficient general purpose program for
746 assigning sequence reads to genomic features. *Bioinformatics.* 30:923-930.
747 Llimargas M, and Lawrence PA. 2001. Seven Wnt homologues in *Drosophila*: a case study of the
748 developing tracheae. *Proceedings of the National Academy of Sciences of the United States of*
749 *America.* 98:14487-14492.
750 Love MI, Huber W, and Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-
751 seq data with DESeq2. *Genome Biol.* 15:550.
752 Magalhaes ILF, Azevedo GHF, Michalik P, and Ramirez MJ. 2020. The fossil record of spiders
753 revisited: implications for calibrating trees and evidence for a major faunal turnover since the
754 Mesozoic. *Biol Rev Camb Philos Soc.* 95:184-217.
755 Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer RC,
756 Gonzales NR, et al. 2017. CDD/SPARCLE: functional classification of proteins via subfamily
757 domain architectures. *Nucleic Acids Res.* 45:D200-D203.
758 McQueen DJ, and McLay CL. 2012. How does the intertidal spider *Desis marina* (Hector) remain
759 under water for such a long time? *New Zealand Journal of Zoology.* 10:383-391.
760 Mi H, Poudel S, Muruganujan A, Casagrande JT, and Thomas PD. 2016. PANTHER version 10:

761 expanded protein families and functions, and analysis tools. *Nucleic Acids Res.* 44:D336-342.
762 Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, and Lanfear R.
763 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic
764 Era. *Molecular biology and evolution.* 37:1530-1534.
765 Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE,
766 Paladin L, Raj S, Richardson LJ, et al. 2021. Pfam: The protein families database in 2021. *Nucleic*
767 *Acids Res.* 49:D412-D419.
768 Moradmand M, Schonhofer AL, and Jager P. 2014. Molecular phylogeny of the spider family
769 Sparassidae with focus on the genus Eusparassus and notes on the RTA-clade and 'Laterigradae'.
770 *Molecular phylogenetics and evolution.* 74:48-65.
771 Mulder N, and Apweiler R. 2007. InterPro and InterProScan: tools for protein sequence classification
772 and comparison. *Methods Mol Biol.* 396:59-70.
773 Murrell B, Weaver S, Smith MD, Wertheim JO, Murrell S, Aylward A, Eren K, Pollner T, Martin DP,
774 Smith DM, et al. 2015. Gene-wide identification of episodic selection. *Molecular biology and*
775 *evolution.* 32:1365-1371.
776 Nawrocki EP, and Eddy SR. 2013. Infernal 1.1: 100-fold faster RNA homology searches.
777 *Bioinformatics.* 29:2933-2935.
778 Neumann D, and Kureck A. 2013. Composite structure of silken threads and a proteinaceous hydrogel
779 which form the diving bell wall of the water spider *Agyroneta aquatica*. *Springerplus.* 2:223.
780 Noh HJ, Turner-Maier J, Schulberg SA, Fitzgerald ML, Johnson J, Allen KN, Huckstadt LA, Batten AJ,
781 Alfoldi J, Costa DP, et al. 2022. The Antarctic Weddell seal genome reveals evidence of selection
782 on cardiovascular phenotype and lipid handling. *Commun Biol.* 5:140.
783 Nyffeler M, and Pusey BJ. 2014. Fish predation by semi-aquatic spiders: a global pattern. *Plos One.*
784 9:e99459.
785 ONO H. 2009. The spiders of Japan with keys to the families and genera and illustrations of the species.
786 *Tokai University Press.*
787 Pandurangan AP, Stahlhacke J, Oates ME, Smithers B, and Gough J. 2019. The SUPERFAMILY 2.0
788 database: a significant proteome update and a new webserver. *Nucleic Acids Res.* 47:D490-D494.
789 Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, and Salzberg SL. 2015. StringTie
790 enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature biotechnology.*
791 33:290-295.
792 Pond SL, Frost SD, and Muse SV. 2005. HyPhy: hypothesis testing using phylogenies. *Bioinformatics.*
793 21:676-679.
794 Roach MJ, Schmidt SA, and Borneman AR. 2018. Purge Haplotigs: allelic contig reassignment for
795 third-gen diploid genome assemblies. *BMC Bioinformatics.* 19:460.
796 Rock JR, Futtner CR, and Harfe BD. 2008. The transmembrane protein TMEM16A is required for
797 normal development of the murine trachea. *Dev Biol.* 321:141-149.
798 Ran J, Yu F, Qin J, Zhang YJ, Yang YF, Li DW, Zhou J, and Liu M. 2016. Functional interplay between
799 cylindromatosis and histone deacetylase 6 in ciliary homeostasis revealed by phenotypic analysis
800 of double knockout mice. *Oncotarget.* 7:27527-27537.

801 Renault D, Puzin C, Foucreau N, Bouchereau A, and Petillon J. 2016. Chronic exposure to soil salinity
802 in terrestrial species: Does plasticity and underlying physiology differ among specialized ground-
803 dwelling spiders? *J Insect Physiol.* 90:49-58.
804 Schaber CF, Grawe I, and Gorb SN. 2023. Attachment discs of the diving bell spider Argyroneta
805 aquatica. *Commun Biol.* 6:1232.
806 Scholl A, O'Brien MJ, Chandran RR, and Jiang L. 2019. The novel gene apnoia regulates Drosophila
807 tracheal tube size. *Dev Dyn.* 248:477-487.
808 Sébastien M, Thomas D, Chen JW, Marina B, Johan W, Carine M, and Benjamin L. 2020.
809 Transmembrane (TMEM) protein family members: Poorly characterized even if essential for the
810 metastatic process. *Semin Cancer Biol.* 60:96-106.
811 Seymour RS, and Hetz SK. 2011. The diving bell and the spider: the physical gill of *Argyroneta*
812 *aquatica*. *J Exp Biol.* 214:2175-2181.
813 Shen YQ, Lang BF, and Burger G. 2009. Diversity and dispersal of a ubiquitous protein family: acyl-
814 CoA dehydrogenases. *Nucleic Acids Res.* 37:5619-5631.
815 Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, and Zdobnov EM. 2015. BUSCO: assessing
816 genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics.*
817 31:3210-3212.
818 Stefano M, Riccardo C, and Marco I. 2016. Ecological preference of the diving bell spider *Argyroneta*
819 *aquatica* in a resurgence of the Po plain (Northern Italy) (Araneae: Cybaeidae). *Fragmenta*
820 *entomologica.* 48:9-16
821 Suyama M, Torrents D, and Bork P. 2006. PAL2NAL: robust conversion of protein sequence
822 alignments into the corresponding codon alignments. *Nucleic Acids Res.* 34:W609-W612.
823 Tarailo-Graovac M, and Chen N. 2009. Using RepeatMasker to identify repetitive elements in genomic
824 sequences. *Curr Protoc Bioinformatics.* Chapter 4:4 10 11-14 10 14.
825 Tian R, Wang ZF, Niu X, Zhou KY, Xu SX, and Yang G. 2016. Evolutionary Genetics of Hypoxia
826 Tolerance in Cetaceans during Diving. *Genome Biology and Evolution.* 8:827-839.
827 Tong C, Aviles L, Rayor LS, Mikheyev AS, and Linksvayer TA. 2022. Genomic signatures of recent
828 convergent transitions to social life in spiders. *Nature communications.* 13:6967.
829 Wertheim JO, Murrell B, Smith MD, Kosakovsky Pond SL, and Scheffler K. 2015. RELAX: detecting
830 relaxed selection in a phylogenetic framework. *Molecular biology and evolution.* 32:820-832.
831 Yang ZH. 2007. PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular biology and*
832 *evolution.* 24:1586-1591.
833 Yim HS, Cho YS, Guang X, Kang SG, Jeong JY, Cha SS, Oh HM, Lee JH, Yang EC, Kwon KK, et al.
834 2014. Minke whale genome and aquatic adaptation in cetaceans. *Nature genetics.* 46:88-92.
835 Yoder JB, Clancey E, Des Roches S, Eastman JM, Gentry L, Godsoe W, Hagey TJ, Jochimsen D,
836 Oswald BP, Robertson J, et al. 2010. Ecological opportunity and the origin of adaptive radiations.
837 *J Evol Biol.* 23:1581-1596.
838 Zhu B, Lou MM, Xie GL, Zhang GQ, Zhou XP, Li B, and Jin GL. 2011. Horizontal gene transfer in
839 silkworm, Bombyx mori. *BMC genomics.* 12:248.
840 Zhang WW, Wu YD, Du LP, Tang DLD, and Gunst SJ. 2005. Activation of the Arp2/3 complex by N-

WASp is required for actin polymerization and contraction in smooth muscle. *Am J Physiol-Cell Ph.* 288:C1145-C1160.

Competing interests

The authors declare no competing interests.

Acknowledgements

We thanked Kun Yu and Feng Lu for provide photos of water spider. This work was supported by National Key R&D Program of China (Grant No. 2022YFC2601200, 2023YFC2604904), the National Science & Technology Fundamental Resources Investigation Program of China (Nos. 2023FY100301,2022FY100500), National Science Foundation of China (Nos. 32270468,32261143728), the project of the Northeast Asia Biodiversity Research Center (NABRI202203) by Ming Bai. This research is also supported by the Science Foundation of School of Life Sciences SWU (20232008071901 and 20212020110501) and the Natural Science Foundation of Chongqing (No. cstc2019jcyj-zdxmX0006) by Zhi-Sheng Zhang.

Author contributions

C.T., M.B. and Z.Z. designed the original research. Z.F. and C.T. performed data analysis and drafted the manuscript. B.L., L.S., T.R. and P.L. performed sample preparation and data analysis. L.W., J.G., L.C., and B.T. collected the samples and performed sample preparation. Y.C. drew the illustration. Q.H., M.D., Q.Z., X.Z., J.L., N.L. and M.I. performed formal analyses and visualization of the data. All authors revised and approved the final manuscript.

Figure Legends

Figure 1. The genome structure analysis and phylogenetic analysis of water spider *Argyroneta aquatica* with other spiders.

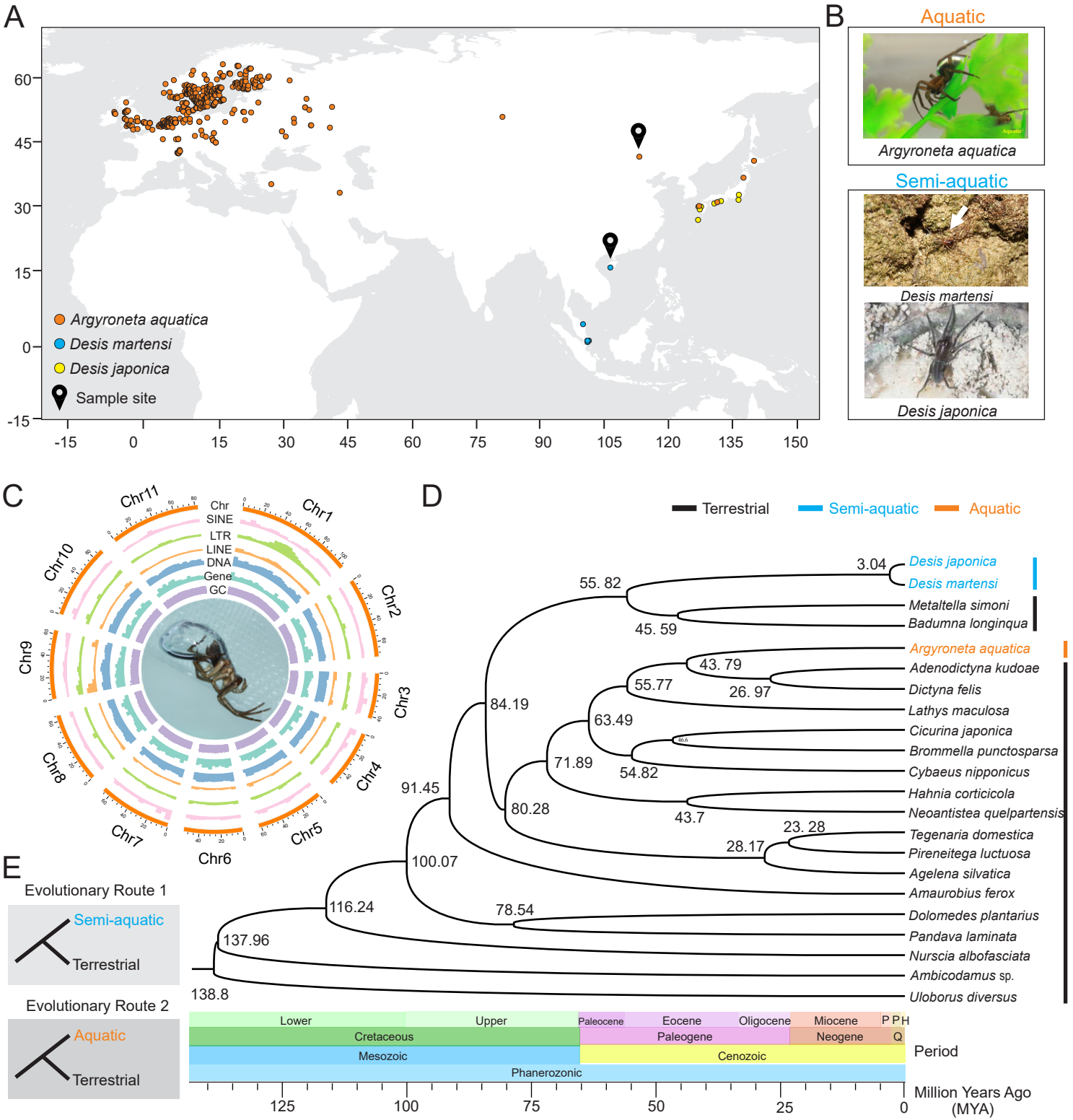
A. The global distribution of *A. aquatica* and *Desis martensi* and *Desis japonic*. The distribution data for aquatic (*A. aquatica*) and semi-aquatic (*D. martensi* and *D. japonica*) spiders from the World Spider Catalog (https://wsc.nmbe.ch/). Based on the latitude and longitude coordinates information, the point map for global distribution of aquatic and semi-aquatic spiders drawing by the online website of SimpleMappr (https://www.simplemappr.net/#tabs=0).

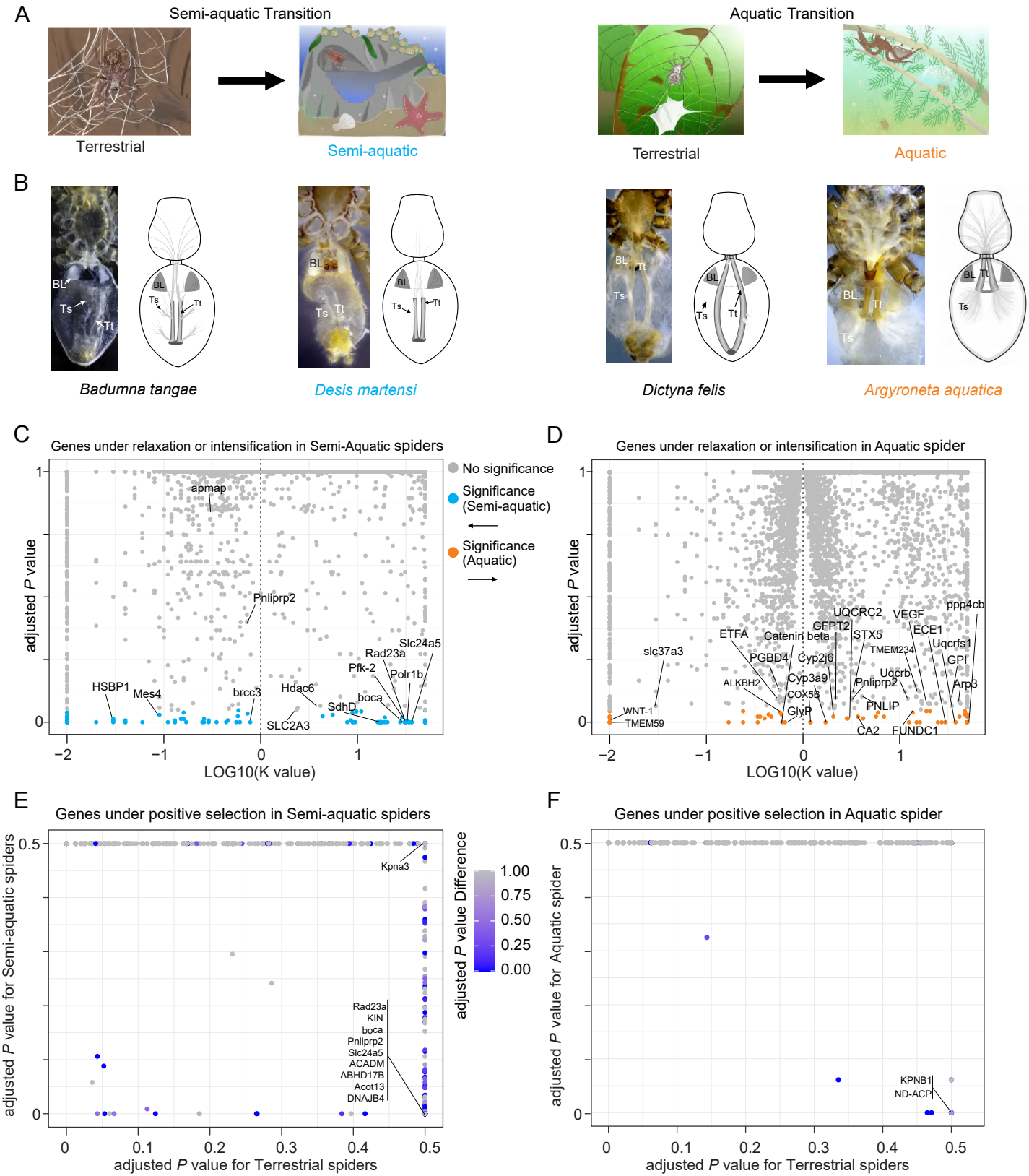
B. Images of *A. aquatica, D. martensi* and *D. japonic.* The Photograph of *A. aquatica* by Qian-Le Lu, *D. martensi by* Lu-Yu Wang. And the Photograph of *D. japonic* was cited from Ono (Ono, 2009).

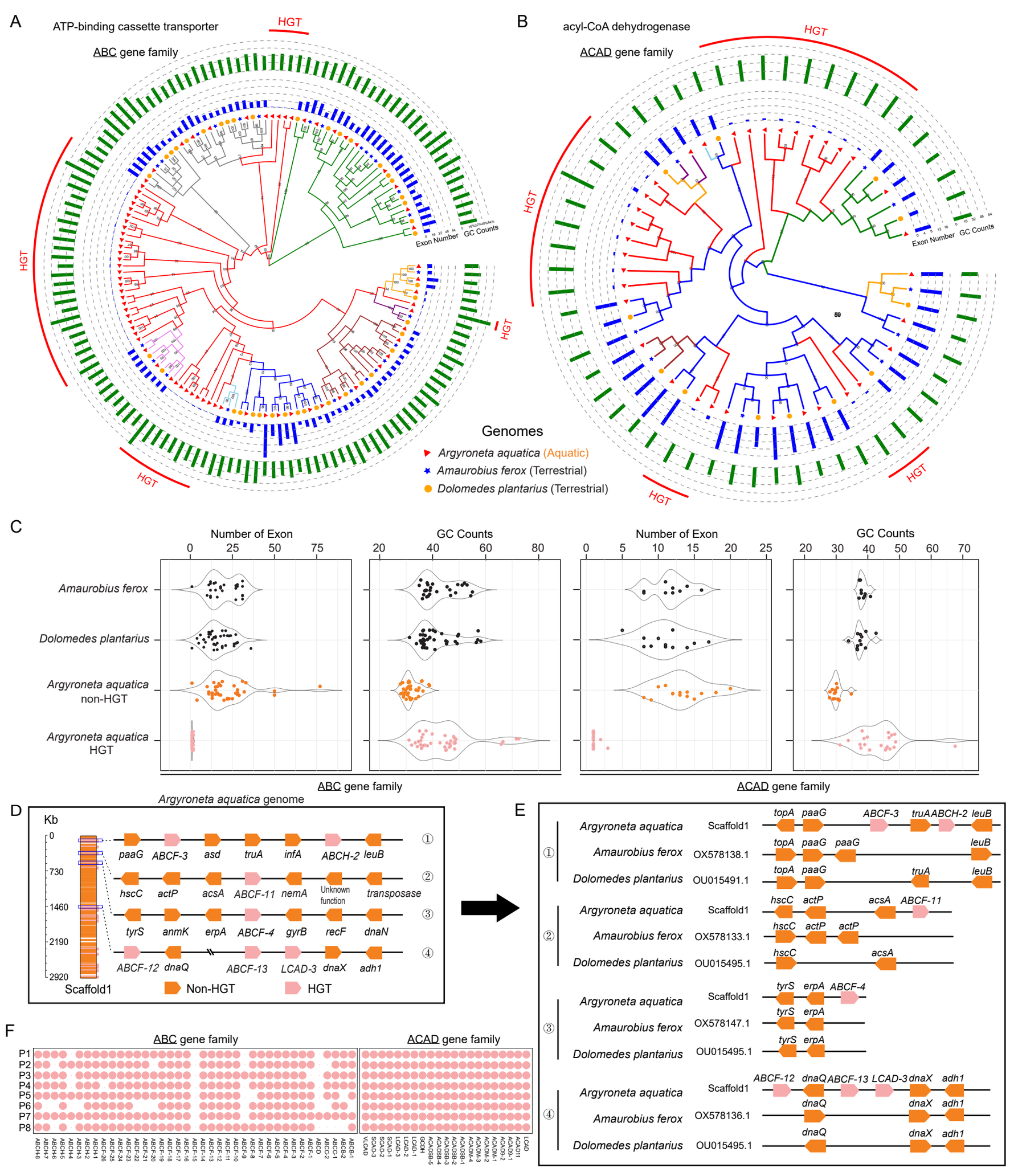
C. Distribution of chromosomal elements of water spider. The inner ring contains a picture of water spider. The outer rings of the circle represent means bellow, respectively: Chr: chromosomes, SINE: short interspersed nuclear element, LTR: long terminal repeat, LINE: long interspersed nuclear

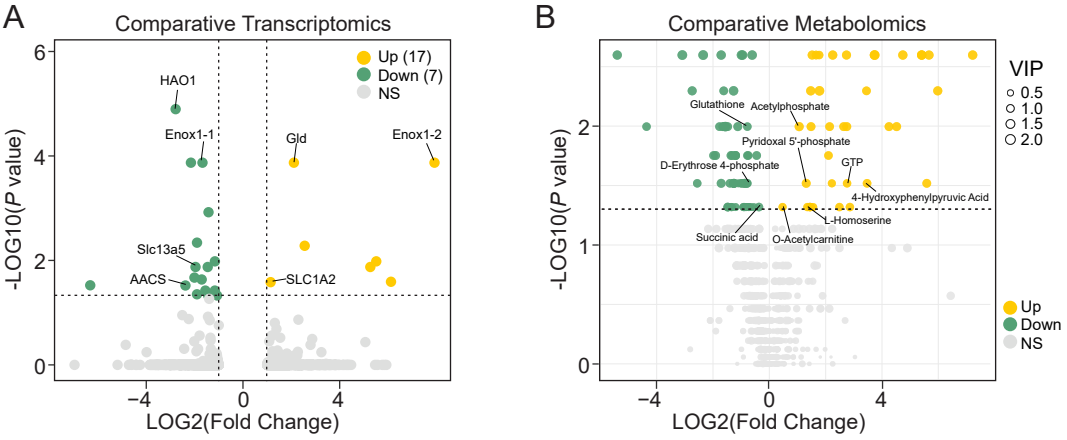
881 elements, DNA: DNA transposable elements, Gene: distribution of genes, GC: GC content. The
882 Photograph of *A. aquatica* by Kun Yu.
883 D. Phylogenetic analysis of *A. aquatica* and most "marronoid" spider species with the Uloboridae (*U.*
884 *diversus*) as outgroups. The divergence times among different species are shown at the bottom and
885 branch node. The black line represents the spiders' terrestrial living habits, blue line with semi-aquatic
886 living habits, and orange line with aquatic living habits.
887 E. The hypothesis of evolutionary route of terrestrial, Semi-aquatic and aquatic.
888
889 **Figure 2. The environment, respiratory systems, and genome-wide dN/dS changes of spiders**
890 **transitioning from terrestrial to semi-aquatic and terrestrial to aquatic environments.**
891 A. Diagram illustrating the transition of spiders from terrestrial to semi-aquatic and terrestrial to
892 aquatic environments.
893 B. The morphological of respiratory systems changes of spiders transitioning from terrestrial to semi-
894 aquatic and terrestrial to aquatic environments. Tt: tracheal tube; Ts: tracheoles; BL: book lungs.
895 Tracheoles are extremely fine tubes, often less than 1 micron in diameter.
896 C. The volcano plot of genes under relaxation or intensification selection of semi-aquatic spiders
897 compared with terrestrial spiders.
898 D. The volcano plot of genes under relaxation or intensification selection of aquatic spiders compared
899 with terrestrial spiders.
900 E. The volcano plot of genes under positive selection of semi-aquatic spiders compared with terrestrial
901 spiders.
902 F. The volcano plot of genes under positive selection of aquatic spiders compared with terrestrial
903 spiders.
904
905 **Figure 3. The feature of HGT-acquired ABC and ACAD genes in *Argyroneta aquatica*.**
906 A. Phylogenetic analysis of ABC gene family of *A. aquatica* and other two spiders *Amaurobius ferox*
907 and *Dolomedes plantarius.* Red line represent HGT-acquired ABC gene.
908 B. Phylogenetic analysis of ACAD gene family of *A. aquatica* and other two spiders *A. ferox* and *D.*
909 *plantarius.* Red line represent HGT-acquired ACAD gene.
910 C. The features of the number of Exon and the GC counts of ABC and ACAD gene family among *A.*
911 *aquatica, A. ferox* and *D. plantarius.*
912 D. The exhibition of HGT genes in the scaffold 1 genome of *A. aquatica.*
913 E. Collinearity analyses of some conserved genes near HGT-acquired genes in three species (*A.*
914 *aquatica*, *A. ferox*, and *D. plantarius*)
915 F. The heatmap of the HGT-acquired ABC and ACAD genes mapped to eight re-sequencing data of *A.*
916 *aquatica* (P1-4: female; P5-8: male).
917
918 **Figure 4. The compare transcriptome and metabonomics analysis of the hypoxia treatment group**
919 **(LO) and the control group (CK).**
920 A. The volcano plots of the results of transcriptome for up- and down-regulation genes of CK vs LO groups.

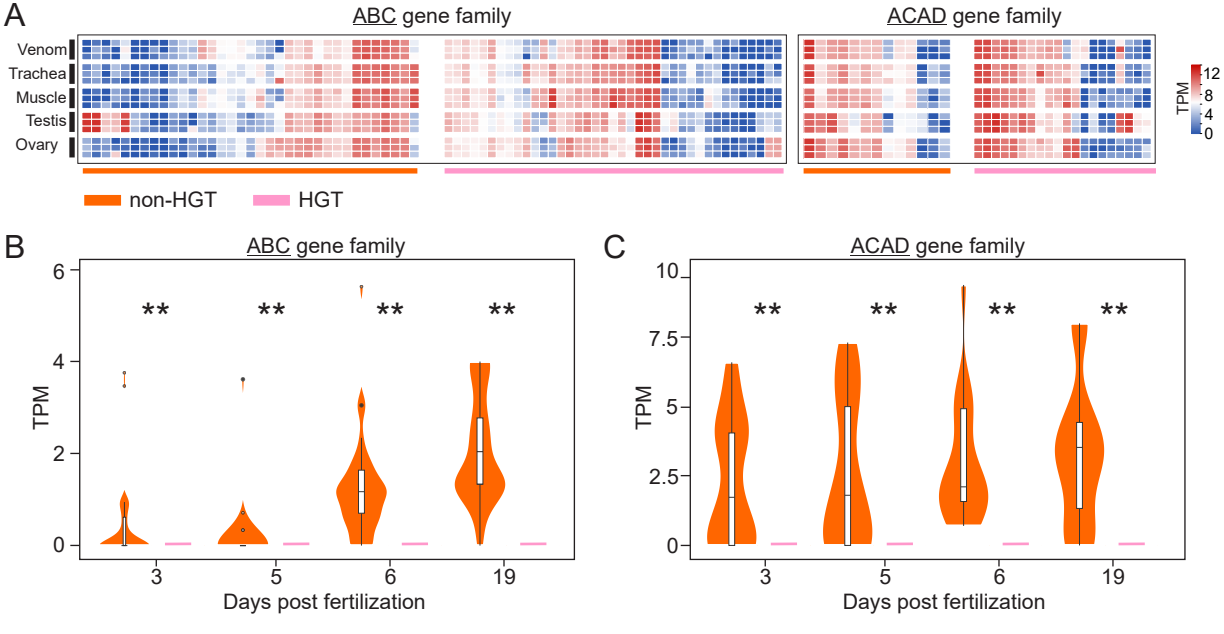
921 B. The volcano plots of the results of metabonomics for up- and down-regulation genes of CK vs LO
922 groups.
923
924 **Figure 5. The expression of ABC and ACAD gene family among different tissues and develop**
925 **stages.**
926 A. The heatmaps of non-HGT and HGT-acquired ABC and ACAD genes expression among
927 different tissues including venom, trachea, muscle, testis and ovary.
928 B. The expression of non-HGT and HGT ABC and ACAD gene family among 3, 5, 6 and 19 days
929 post fertilization.


A
Argyroneta aquatica
Desis martensi
Desis japonica
Sample site
B
Aquatic
Argyroneta aquatica
Semi-aquatic
Desis martensi
Desis japonica
C
Chr
SINE
LTR
LINE
DNA
Gene
GC
Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7
Chr8
Chr9
Chr10
Chr11
D
Terrestrial
Semi-aquatic
Aquatic
Desis japonica
Desis martensi
Metaltella simoni
Badumna longinqua
Argyroneta aquatica
Adenodictyna kudoae
Dictyna felis
Lathys maculosa
Cicurina japonica
Brommella punctosparsa
Cybaeus nipponicus
Hahnia corticicola
Neoantistea quelpartensis
Tegenaria domestica
Pireneitega luctuosa
Agelena silvatica
Amaurobius ferox
Dolomedes plantarius
Pandava laminata
Nurscia albofasciata
Ambicodamus sp.
Uloborus diversus
3.04
55. 82
45. 59
43. 79
55.77
26. 97
84.19
63.49
71.89
54.82
91.45
80.28
43.7
23. 28
28.17
100.07
116.24
78.54
137.96
138.8
E
Evolutionary Route 1
Semi-aquatic
Terrestrial
Evolutionary Route 2
Aquatic
Terrestrial
Lower
Upper
Paleocene
Eocene
Oligocene
Miocene
P P H
Cretaceous
Paleogene
Neogene
Q
Mesozoic
Cenozoic
Phanerozonic
Period
125
100
75
50
25
0
Million Years Ago (MYA)


A
Semi-aquatic Transition
Terrestrial
Semi-aquatic
Aquatic Transition
Terrestrial
Aquatic
B
BL
Ts
Tt
Badumna tangae
Desis martensi
Dictyna felis
Argyroneta aquatica
C
Genes under relaxation or intensification in Semi-Aquatic spiders
No significance
Significance (Semi-aquatic)
Significance (Aquatic)
adjusted P value
LOG10(K value)
apmap
Pnliprp2
Slc24a5
Rad23a
Pfk-2
Polr1b
HSBP1
Mes4
brcc3
Hdac6
boca
SdhD
SLC2A3
D
Genes under relaxation or intensification in Aquatic spider
adjusted P value
LOG10(K value)
UQCRC2
VEGF
ppp4cb
GFPT2
ETFA
Catenin beta
STX5
ECE1
Uqcrfs1
slc37a3
TMEM234
PGBD4
Cyp2j6
GPI
Uqcrb
Cyp3a9
Pnliprp2
Arp3
ALKBH2
COX5B
PNLIP
WNT-1
TMEM59
GlyR
CA2
FUNDC1
E
Genes under positive selection in Semi-aquatic spiders
adjusted P value for Semi-aquatic spiders
adjusted P value for Terrestrial spiders
adjusted P value Difference
Kpna3
Rad23a
KIN
boca
Pnliprp2
Slc24a5
ACADM
ABHD17B
Acot13
DNAJB4
F
Genes under positive selection in Aquatic spider
adjusted P value for Aquatic spider
adjusted P value for Terrestrial spiders
KPNB1
ND-ACP


A
ATP-binding cassette transporter
ABC gene family
HGT
Exon Number
GC Counts
B
acyl-CoA dehydrogenase
ACAD gene family
Genomes
Argyroneta aquatica (Aquatic)
Amaurobius ferox (Terrestrial)
Dolomedes plantarius (Terrestrial)
C
Number of Exon
GC Counts
Amaurobius ferox
Dolomedes plantarius
Argyroneta aquatica non-HGT
Argyroneta aquatica HGT
ABC gene family
ACAD gene family
D
Argyroneta aquatica genome
Kb
Scaffold1
paaG ABCF-3 asd truA infA ABCH-2 leuB
hscC actP acsA ABCF-11 nemA Unknown function transposase
tyrS anmK erpA ABCF-4 gyrB recF dnaN
ABCF-12 dnaQ ABCF-13 LCAD-3 dnaX adh1
Non-HGT
HGT
E
Argyroneta aquatica
Amaurobius ferox
Dolomedes plantarius
Scaffold1
OX578138.1
OU015491.1
OX578133.1
OU015495.1
OX578147.1
OX578136.1
F
ABC gene family
ACAD gene family
P1
P2
P3
P4
P5
P6
P7
P8


A
Comparative Transcriptomics
Up (17)
Down (7)
NS
HAO1
Enox1-1
Gld
Enox1-2
Slc13a5
AACS
SLC1A2
-LOG10(P value)
LOG2(Fold Change)
B
Comparative Metabolomics
VIP
0.5
1.0
1.5
2.0
Glutathione
Acetylphosphate
Pyridoxal 5'-phosphate
D-Erythrose 4-phosphate
GTP
4-Hydroxyphenylpyruvic Acid
L-Homoserine
Succinic acid
O-Acetylcarnitine
Up
Down
NS
-LOG10(P value)
LOG2(Fold Change)


A
ABC gene family
ACAD gene family
Venom
Trachea
Muscle
Testis
Ovary
TPM
12
8
4
0
non-HGT
HGT
B
ABC gene family
TPM
6
4
2
0
**
3
5
6
19
Days post fertilization
C
ACAD gene family
TPM
10
7.5
5
2.5
0
**
3
5
6
19
Days post fertilization