Relaxed selection and horizontal gene transfer fuel underwater adaptation in a water spider

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), Institute of Zoology, 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

The water spider (*Argyroneta aquatica*) exhibits remarkable adaptability to thrive in aquatic environments and represents the only spider species capable of spending its entire life underwater. While its physiological and behavioral adaptations had been characterized, the genetic basis of underwater adaptation in spiders remains a fascinating topic that is poorly understood at multi-omics levels. Here, we *de novo* assemble a chromosome-level, high-quality genome of the water spider. A comparative analysis of spider genomes indicates that water spider has elevated genome-wide rates of molecular evolution associated with relaxed selection, which caused by reduced effective population size during the transition from territorial to underwater lifestyle. Remarkably, we found that horizontal gene transfer (HGT) emerges as a pivotal evolutionary force facilitating the underwater adaptation in water spider. We observed significant differentiation between water spiders under normal oxygenic and hypoxic conditions by utilizing comparative transcriptomics and metabolomics, and identified genes associated with hypoxia response pathways, especially hypoxia-inducible factor 1 (HIF-1) signaling pathway. Notably, we determined that genes within regulation HIF-1 pathway underwent positive selection (such as *MARK1*, *Glut1*, *Hkdc1*, and *Hsp90b1*) or gene expansions through HGT-acquired (such as ABC and ACAD gene families), enhancing the hypoxia adaptation of the water spider. Altogether, our study provides insights into the genomic signature of adaptive evolution and specific genetic changes that underpin the transition to water life in spiders.

Introduction

The water spider, *Argyroneta aquatica* (Fig. 1A), represents a remarkable example of morphological and physiological adaptation to an aquatic lifestyle[1]. This species, unique among over 51,000 described spiders worldwide[2], has evolved to spend its entire life cycle submerged in freshwater environments. Morphologically, the water spider is distinguished by its specialized hydrophobic hairs covering its abdomen, which facilitate the creation of a thin layer of air, essential for respiration underwater[3,4]. Additionally, its ability to construct air-filled 'diving bells' using silk webs is a key adaptation for both respiration and reproduction[5,6]. Remarkably, the water spider also possesses what can be termed as 'physical gills', a rare adaptation in spiders, which allow it to accumulate and store oxygen. This adaptation is particularly important for surviving in hypoxia conditions often found in aquatic environments. These 'gills' enable the spider to extend its time submerged, a critical advantage for an organism whose life cycle is entirely water-bound. These morphological features are not just mere physical adaptations but are integral to the spider's unique underwater lifestyle, enabling it to efficiently navigate, hunt, and breed in an aquatic environment[3,7].

Untangling the evolutionary process including rapid evolution and relaxed selection is crucial to understand the mechanisms behind morphological innovations and adaptation to unique environments[8,9]. Rapid evolution occurs when species adapt quickly to new environments, often driven by changes in selective pressures. Relaxed selection, on the other hand, refers to a reduction in these pressures, allowing for greater genetic variability and potentially fostering novel traits[10]. This interplay between rapid evolution and relaxed selection is particularly significant in understanding how organisms adapt morphologically and genetically to specialized environments, such as underwater habitats. Several case studies illustrate this phenomenon in underwater adaptations. 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[11-13]. Similarly, the evolution of fish from their ancestral forms showcases various adaptations like streamlined bodies and gill modifications, facilitated by relaxed selective pressures in aquatic environments[14-17]. These examples highlight how rapid evolution and relaxed selection can lead to significant morphological innovations, enabling species to exploit new ecological niches. Despite its fascinating underwater lifestyle and evident morphological adaptations in water spider, the long lack of corresponding genomics data has impeded a deeper understanding of the genomic changes associated with this unique underwater lifestyle.

Recent studies have shed light on another fascinating aspect of specialized adaptation in animals: horizontal gene transfer (HGT)[18-22]. HGT, typically observed in prokaryotes, involves the transfer of genetic material between different species, bypassing the traditional parent-offspring lineage. This phenomenon is increasingly being recognized in various animals and is postulated to play a significant role in their adaptation to unique environments. For example, in silkworms and ladybird beetles, certain HGT candidate genes have been identified that contribute to pathogen resistance[18,19]. In lepidopterans, some HGT candidate genes have been linked to male courtship behavior[20]. Additionally,

in whitefly and leaf beetle, HGT candidate genes are found to be essential for survival[21,22]. However, the presence of HGT in spiders, particularly in relation to their adaptation to aquatic environments, has not been reported yet. In the context of the water spider, the investigation into HGT may present an intriguing avenue for understanding how genes from other organisms might have been incorporated into its genome, potentially facilitating its unique adaptations.

Hypoxia, characterized by reduced oxygen availability, is a prevalent challenge for animals in extreme environments such as the plateau or deep sea. Animals dwelling in these areas have evolved a variety of unique adaptations to cope with low-oxygen conditions. For instance, enhanced oxygen storage capacity is achieved through specialized oxygen transport proteins, as observed in species like the Minke whale and the Arctic gyrfalcon[11,23]. Additionally, animals including the naked mole-rat and *Ips nitidus*, exhibit increased energy metabolism by augmenting the activity of key enzymes within the tricarboxylic acid (TCA) cycle[24,25]. Another common adaptation involves the regulation of the hypoxia-inducible factor 1 (HIF-1) signaling pathway, exemplified by the plateau pika, *Ochotona curzoniae*[26], which plays a significant role in managing cellular responses to hypoxia. Insects in functionally hypoxic environments also demonstrate similar adaptations[27]. The water spider provides a compelling case study, because its unique adaptation involves the creation of "diving bells"[3], exhibiting a low partial pressure of oxygen (PO_2). This indicates the water spider's ability to thrive in hypoxic conditions for extended periods. This adaptation is crucial for its survival and reproductive success in underwater environments, highlighting the spider's specialized evolutionary response to the challenges posed by low oxygen availability.

In this study, we provide a chromosome-level and high-quality genome of the water spider and present comparative genomics and molecular evolutionary analyses. Our research primarily tests three hypotheses: the relaxed selection during spider adaptation to underwater lifestyle; the presence and impact of horizontal gene transfer (HGT) in the water spider's genome; and the water spider's regulatory mechanism of adaptation to hypoxia. We conducted hypoxia experiments to explore these aspects, utilizing comparative transcriptomics and metabolomics to reveal the water spider's adaptations at a molecular level. These experiments are designed to verify how relaxed selection and potential HGT have contributed to spider's unique adaptations, including its ability to withstand low-oxygen environments. Furthermore, our study addresses a pressing conservation concern. The water spider population has seen a significant decline due to habitat fragmentation and deterioration. This decline is noted in several Regional Red Lists and Red Books across Europe[28]. By analyzing the genome of the water spider, we seek to unveil the genetic basis underlying its extraordinary adaptations to an underwater lifestyle. This study will also provide genomic insights into effective conservation and management of water spider populations, helping to secure their future in the face of environmental challenges.

Results

Genome assembly, annotation, and phylogeny

We reported a high-quality, 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. S1), detecting 98.0% arthropod conserved genes by BUSCO (Table S2). Further genome annotation suggested that approximately 41.53% of whole genome are repetitive regions (Fig. 1B, Table S3). A combination of homolog search, transcriptome-based and ab initio prediction for searching gene models resulted in a total of 17,632 protein-coding genes. Based on the single-copy ortholog dataset, we reconstructed the genome-scale phylogeny for these 14 spider species, which was strongly supported (bootstrap value = 100) for all nodes (Fig. 1C). The estimated divergence ages of the most recent common ancestors between *A. aquatica* and *Uloborus diversus* was dated at 117.29 Ma (Fig. 1C). The comparative analysis among water spider and other spiders based on whole genome reflected a lot of genes changes in water spider associated with underwater transition. We also identified the expansion gene families of water spiders (Fig. 1C). In the GO terms and KEGG pathways of expansion gene families enrichment analysis, we found some genes which functions related to water spider underwater life, such as the GO terms of ATP hydrolysis activity and response to heat, the KEGG pathways of purine metabolism, TCA cycle, ABC transporters and MAPK signal pathway (Fig. S2).

Accelerated evolution and relaxed selection associated with underwater transition

We estimated the genome-wide ratio of non-synonymous to synonymous substitution (dN/dS, ω) across the spider phylogeny-based thousands of shared orthologs dataset. We found that water spider experienced elevated genome-wide rates of molecular evolution compared to the other spiders at both genome-wide (Likelihood Ratio Test, LRT, p-value <0.00001) and gene-wide levels Wilcoxon rank-sum test, p-value = 0.0014, Fig. S3). Next, we sought to determine the selection acting on the underwater transition in spider by calculating three discrete categories of ω ratios (purifying selection: ω1, neutral evolution: ω2, positive selection: ω3) for concatenated and coalescent genesets using RELAX. The parameter K-valuewas used to indicate overall relaxation (K < 1) or intensification (K > 1). We found that the uplift of genome-wide rate of molecular evolution in water spider was primarily caused by relaxed selection (K = 0.95, LRT, p-value <0.001, Fig. 2A).

Further, we identified genes that experienced relaxed or intensified selection in water spider branch relative to all other spider branches. We found that increasing number of genes under relaxed selection (151 genes, K < 1, q-value <0.05,Table S5) compared to intensified selection (64 genes, K > 1, q-value <0.05, Table S5). Genes experiencing significant relaxation of selection in water spider were enriched for 30 GO terms (q-value <0.05), including eight main categories (Fig. S4a): development (e.g., imaginal disc morphogenesis; imaginal disc fusion; imaginal disc fusion, thorax closure), morphogenesis (e.g., post-embryonic animal organ morphogenesis), cytoskeletal protein binding and polarity (e.g., cytoskeletal protein; actin binding; establishment or maintenance of cytoskeleton polarity), ubiquitin ligase complex (e.g., ubiquitin ligase complex; cullin-RING ubiquitin ligase complex; Cul3-RING ubiquitin ligase complex), protein modification and catabolic (e.g., protein modification by small protein conjugation or removal, regulation of proteasomal ubiquitin-dependent protein catabolic process; positive regulation of proteolysis), cell component generation and regulation (e.g., regulation of cellular component biogenesis; chromosome condensation), regulation of Wnt

signaling pathway, and neuron remodeling. In addition, genes experiencing intensification of selection in social branches were enriched for seven GO terms (q-value <0.05, Fig. S4b), including five main categories: cell population proliferation, calcium ion channel activity (e.g., glutamate-gated calcium ion channel activity; ligand-gated calcium channel activity), cobalamin transport and binding, regulation of synaptic activity, and kainate selective glutamate receptor activity.

Finally, we inferred demographic history and effective population size (Ne) based on the whole genome resequencing data with > 20 x coverage of 20 individuals. Interestingly, we observed a pattern with initial increase, then followed by a decline around which started after the Kunlun-Yellow River movement (~ 0.8800,000 million years ago, Mya) (Fig. 2B). The demographic history analysis of batik golden web spider *Trichonephila antipodiana* (living in the forest) and lampshade web spider *Ectatosticta davidi* (living in the stone gap) showed a decline at the same time (after kunlun ice age) with water spider (Fig. S5). Compared with other spider species, the demographic history of water spider showed a continuous decline, indicating that water spider suvived under the same environmental pressures.

Positive selection associated with underwater adaptation

We detected the signal of positive selection in water spider compared with other spider species across the phylogeny, and identified a total of 12 positively selected genes (PSGs) in water spider (Fig. S6). The most representative genes were HSPs, these gene were both under positive selection (e.g., hspb11), relaxation selection (e.g., *hspb1*, *cdc37*, *hsp90b1*) and intensification selection (e.g., *hsf1*). In addition, a set of *P450* genes were under both positive selection (e.g., *cyp18a1*) and intensified selection (e.g., *cyp3a9*, *cyp4v2*, *cox5b*).

Hemocyanins gene repertoire

We conducted a comprehensive survey and identified seven hemocyanin genes in water spider genome (Fig. S7-S8, Table S4). The structural properties of water spider hemocyanin protein were predicted using Swiss-Model database, revealing a 12-mer (2 x 6-mer) multiple hexametric configuration. A total of six hemocyanin genes were expressed between multiple tissues, with high expression in muscle and trachea tissues (Fig. S8). In addition, we detected the signal of positive selection in these genes. Notably, specific positions within the hemocyanin genes were determined under positive selection, suggesting potential importance in oxygen transportation (Fig. S8).

HGT-acquiced gene repertoire

Horizontal gene transfer (HGT) is hypothesized as one of evolutionary factors facilitating the underwater adaptation in water spider genome. We identified a total of 424 genes specific genes (Fig. S9-S10) with 131 putative evidences of HGT (Table S6). We found that a set of genes that are unique to water spiders showing highly homologous relationship with bacteria. We further validated the reliability of inference of these HGT-acquired genes (i.e. ABC and ACAD gene families) using a set of approaches such as Blast-Search, phylogenetic mapping, accessing RNA expression, whole genome resequencing reads mapping, and PCR cloning (Fig. S11-S14, see methods for details).

We pinpointed that these HGT-acquired genes (e.g. genes in ABC and ACAD gene families) 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. 3a). And the main genus of the bacteria were *Acinetobacter*, *Alkanindiges*, and *Emticicia*. Notably, in the freshwater, evidences showed that all these bacteria can be found and confirmed (Fig. 3b). Moreover, near the loci of most HGT genes, transposons were observation (Fig. 3c).

Besides, we identified a total of 35 ACAD (acyl-CoA dehydrogenase) genes associated with transport and oxidative decomposition of fatty acids functions water spider genome, including 18 HGT-acquiced genes. Out of 75 ABC genes that have been identified in water spider, 32 genes were from bacteria. Intriguingly, amongthese HGT-acquired ABC genes, a set of genes had characterized functions, such as urea ABC transporter ATP-binding subuni, lipid A export ATP-binding/permease protein MsbA, lipoprotein-releasing system ATP-binding protein LolD, glucose import ATP-binding protein TsgD13, and high-affinity branched-chain amino acid transport ATP-binding protein LivG. Collectively, these HGT-acquiced genes may fuel the underwater adaptation in water spider (Table S6).

Comparative transcriptomics and metabolomics in response to hypoxia stress

We identified a substantial set of genes involved in HIF-1 (hypoxia-inducible factor 1) signaling pathway with significant changes in evolutionary shift during underwater adaptation, such as the glucose transporter type 1 (*Glut1*) gene and hexokinase domain containing 1 (*Hkdc1*) gene related to anaerobic metabolism underwent relaxed selection. MAP3K7 gene (belong to MAPK signal pathway), represents one of the indirect upstream regulates of HIF-1α and experienced relaxed selection(Fig. 4, S15-S16). To further elucidate the regulatory mechanism in water spider adaptation to underwater life, we sought to trace the changes in both transcriptional and post-transcriptional level by comparative transcriptomics and metabolomics following the hypoxia stress treatment (Fig. S17). Specifically, transcriptomic analysis showed that a total of 45 up-regulated and 66 down-regulated genes triggered by hypoxia stress (Fig. S18). The significantly enriched ontologies for up-regulated genes (GO terms, q-value < 0.05) were mainly related to enzyme activity (such as argininosuccinate synthase activity, carnitine O-palmitoyltransferase activity, glutamate decarboxylase activity, O-palmitoyltransferase activity, and palmitoleoyltransferase activity), regulation of biological processes (such as hormone metabolic process, wybutosine metabolic process, and wybutosine biosynthetic process, and argininosuccinate metabolic process), response to stimulus (such as cellular response to interleukin-6, response to interleukin-6, cellular response to fatty acid, cellular response to oleic acid), positive regulation of signal transduction (positive regulation of phosphatidylinositol 3-kinase signaling, positive regulation of platelet activation), transport (long-chain fatty acid transport, high-affinity glutamate transmembrane transporter activity), and so on (Fig. S18). While, the enriched GO terms for down-regulated genes were associated with oxidoreductase activity, metabolic process (tyrosine metabolic process, isoprenoid metabolic process, choline metabolic process, L−phenylalanine metabolic process, L−phenylalanine catabolic process, drug catabolic process, tyrosine catabolic process, isoprenoid biosynthetic process) and so on (Fig. S18). In addition, we performed additional

KEGG pathway enrichment analysis and found that, MAPK signaling pathway, lipid and atherosclerosis pathway and longevity regulating pathway were related with hypoxia stress (Fig. S18). For these down-regulated genes, several pathways including spliceosome, phagosome, tyrosine metabolism, terpenoid backbone biosynthesis, valine, leucine and isoleucine degradation were shown up.

Similarily, comparative metabolomics analysis identified 37 and 58 metabolites showing upregulated and downregulated patterns in response to hypoxia stress (Fig. S19). Not surprisingly, most of differentially expressed metabolic products induced by hypoxia were significantly enriched to various metabolism pathways (Fig. S19). We found that a set of changing metabolites were annotated to TCA cycle of mitochondrion (Fig. 4). Specifically, succinyl-CoA and GDP were up-regulation while the succinate was down-regulation in hypoxia stress treatment group. While, the Acetyl-carnitine was up-regulated (Fig. 4). Thus, hypoxia stress may lead to diminished oxidative phosphorylation and to a depletion of cellular ATP.

Discussion

Transition from land to underwater environment poses several challenges for water spiders, such as temperature changes, changes in food sources, hypoxia. This study untangles the genomic underpinnings of water spider species during underwater adaptation, focusing on three key aspects: relaxed selection, horizontal gene transfer (HGT), and hypoxia response. Specifically, we conducted a comparative genomic analysis between water spider genomes and those of 13 other spider species, mined HGT-acquired genes, integrated comparative transcriptomic and metabolomic profiling following hypoxia stress, to uncover the genetic changes facilitating its unique underwater lifestyle.

The adaptation of water spiders to underwater environments is highlighted by various selection processes across the whole genome. Our study reveals that the heat shock protein (HSP) gene family plays a critical role in the water spider's response to environmental stressors such as temperature changes, similar to findings in high-elevation ants like *Tetramorium alpestre* adapted to cold climates[29]. The analysis of the water spider's genome shows an active presence of the HSP gene family, with indications of extended gene families, positive selection, relaxed selection, and intensified selection. Additionally, the water spider exhibits higher rates of non-synonymous to synonymous substitutions (dN/dS) compared to other spiders at both genome-wide and gene-wide levels (Fig. 2). This suggests a relaxation in purifying selection on its genes, underscoring the importance of relaxed selection in underwater adaptation.

The study also presents groundbreaking evidence of horizontal gene transfer (HGT) in water spiders, a significant finding considering its previous documentation primarily in insects[5]. The water spider's genome exhibits the acquisition of genes through HGT, enhancing its adaptive capacity to specific environmental challenges. For instance, the HGT-acquired ABC transporter genes are expanded and positively selected, enabling the transport of a wide range of substrates crucial for survival in diverse aquatic settings[30]. Similarly, the HGT-acquired ACAD genes target long-chain acyl-CoA substrates, vital for energy production through mitochondrial fatty acid beta-oxidation[31]. These adaptations are particularly advantageous in environments where high-energy food sources, such as

fish rich in long-chain fatty acids, are a primary diet[32].

Furthermore, our findings explore the gene expression and regulatory mechanisms that facilitate the water spider's adaptation to hypoxic environments. Key adaptations include the positive evolution of hemocyanin, vital for oxygen transport, and the role of the HIF-1 signaling pathway in regulating carbohydrate metabolism under low-oxygen conditions, as indicated by genes like Hkdc1 and Tigar (Table S5). This is consistent with adaptations seen in other species adapted to hypoxia, such as the bar-headed goose and hypoxia-tolerant fish species[33,34]. Moreover, the study reveals that hypoxia impacts lipid metabolism, with genes like Elovl and Lrp12 showing intensified selection, and the expansion of the ACAD gene family, key in fatty acid and amino acid catabolism[35,36]. These insights open avenues for future research on the mechanisms and selective pressures driving these adaptations in water spiders.

In summary, this study elucidates the water spider *Argyroneta aquatica*'s adaptation to underwater life, highlighting three key genetic mechanisms: relaxed selection, particularly in the heat shock protein gene family and the overall genome-wide or gene-wide evolutionary shift; the presence of horizontal gene transfer, enhancing its adaptability; and specialized gene expression and regulatory mechanisms in response to hypoxia. Collectively, these findings not only enhance our understanding of the unique evolutionary path of the water spider but also contribute to the broader knowledge of adaptive strategies in aquatic organisms. This research paves the way for future investigations into the complex interplay of genetic, environmental, and evolutionary factors that shape the life of this remarkable species.

Methods

Sample collection and hypoxia stress treatment

Water spider *A. aquatica individuals were* collected from Xilingol league of InnerMongolia. The whole body without abdomen and stomach was used for whole genome sequencing. After further dissection of water spider individual, tissues including venom, muscle, trachea, testis, and ovary were used for total RNA extraction and transcriptome sequencing on a single lane of an Illumina NovaSeq platform (Berry Genomics, Beijing, China) yielding 150-bp paired-end reads.

The lung slit and spiracle was sealed by paraffin, then raised it in boxes for hypoxia treatment stress experiment. 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. The detailed was shown in Figure S17. Finally, the sample was used for Liquid chromatography–mass spectrometry (LC-MC) assay (Suzhou PANOMIX Biomedical Tech, China) and transcriptome sequencing (Berry Genomics, Beijing, China)

Whole genome sequencing and de novo assembly

A combination of two sequencing approaches including short reads (400 bp) 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 of water spider. 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.

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[37]. The resulting genome assembly was further polished for two rounds using NGS data with NextPolish v1.0.5[38]. The Heterozygous regions were removed by Purge Haplotigs v1.1.0[39] with the haplotigs identifying of 50% cut-off. The chromosome-level assembly was generate using Hi-C reads with Juicer v1.6.2[40]. 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[41]. Finally,the quality and completeness of genome assemblies was assessed with BUSCO v5.4.5[42] with arthropoda_odb10 data set (n= 1,013).

Genome annotation

For the annotation of repetitive elements in genome, we searched it by means of a combination of ab initio and homology. Firstly, the species-specific repeat database (Ab-initio database) was constructed using RepeatModeler v2.0.1[43,] and further combined the known repeat library (Repbase) and Ab-initio database as the reference repeat database. The repetitive elements in the genome was finally identified by RepeatMasker[44].

Non-coding RNAs were annotated by Infernal v1.1.2[45] and tRNAscan-SE v2.0.6[46]. In addition, a tool “EukHighConfidenceFilter” in tRNAscanSE was used for further filtering transfer RNAs (tRNAs) with high confidence.

For the annotation of gene structure, we integrated the ab initio, transcriptome-based, and protein homology-based evidence by Maker v2.31.10[47]. Gene structure prediction of ab inito was identified by Augustus v3.3.3[48] and GeneMark-ES/ET/EP v4.48 3.60 lic[49]. To model the sequence properties accurately, both gene finders (Augustus and GeneMark-ES/ET/EP) were initially trained using the BRAKER v2.1.5 pipline[50]. The transcriptome data for whole body water spider was mapped to the genome by the software HISAT2 v2.2.0[51]. BRAKER pipeline was run with default parameters. Then further assembled RNA-Seq alignments into potential transcripts by Stringtie v2.1.3[52], and the results was provided as input for Maker via the “est” option. The protein homology-based evidence was download from NCBI database, which including the protein sequences of *Stegodyphus mimosarum* (GCA 000611955.2), *Parasteatoda tepidariorum* (GCA 000365465.3), *Trichonephila clavipes* (GCA 002102615.1), *Drosophila melanogaster* (GCA 000001215.4), *Ixodes scapularis* (GCA 002892825.2), *Daphnia pulex* (GCA 900092285.2) and *Strigamia maritima* (GCA 000239455.1), which were combined and fed to Maker pipline.

Gene function annotation was predicted by Diamond v0.9.24[53], InterProScan v5.41–78.0[54] and

eggNOG-mapper v2.0[55]. The UniProtKB/Swissprot database was against by the software Diamond with the mode of more sensitive, maximum number for target sequences of 1, and e-value of 1e−5. The following 5 databases including Pfam[56], Panther[57], Gene3D[58], Superfamily[59], and Conserved Domain Database (CDD)[60] was against by the software InterProScan. The the eggNOG v5.0 database[61] was against by the software eggNOG-mapper.

Gene clusters analysis

We download publically available genomes and proteins of spider species including *Trichonephila antipodiana* (http://dx.doi.org/10.5524/100868), *T. clavipes* (GCA_002102615.1), *Argiope bruennichi* (GCA_015342795.1), *Oedothorax gibbosus* (GCA_019343175.1), *P. tepidariorum* (GCA_000365465.3), *Stegodyphus dumicola* (GCA_010614865.1), *S. mimosarum* (GCA_000611955.2), *Caerostris darwini* (GCA_021605075.1), *Ectatosticta davidi* (https://doi.org/10.57760/sciencedb.06872), and an outgroup *Tetranychus urticae* (GCA 000239435.1). All orthologous gene clusters were classified by OrthoFinder v2.5.4[62].

A total of 409 single-copy orthologous genes were used to examine the phylogenetic relationships among these spiderspecies. First, the genes were separately aligned by MAFFT v7.471[63] with L-INS-I strategy. The resulting alignments were then trimmed by trimAl v1.4[64] with method “automated1”. Then resulting were concatenated by FASconCATG v1.04[65]. Finally, the maximum likelihood (ML) tree was constructed by IQ-TREE v2.0.3[66] with the optimization number of rcluster algorithm, replicates for ultrafast bootstrap, and replicates for Shimodaira–Hasegawa (SH) approximate likelihood ratio tests being 10, 1000, and 1000, respectively.

The divergence time calibration was performed with the reference from Magalhaes[67], paleobiodb database (https://paleobiodb.org/) and TimeTree database (http://www.timetree.org/), with Nephilinae stem (43-47 Mya), Orchestininae stem (98.17-99.41 Mya), Synspermiata stem (164-175.1 Mya), and Araneae stem (374-382.7 Mya). The divergence time was calculated using package MCMC Tree in PAML v4.10.0[68]. Finally, the analysis of gene family expansion and contraction was performed with the use of Cafe v5.1.04[69].

Codon substitution rate estimation and comparison

By using thousands of orthologs that shared by 14 spider species, we prepared the codon alignments of these orthologs which derived from amino acid alignments and corresponding DNA sequences using PAL2NAL v.14[70]. The discrete codon substitution rate (dN/dS, ω) representing water spider and other species were estimated by HyPhy[71] following the pipeline as previously described[72]. Further, the pattern of selection acting on the transition from of territorial to underwater in water spider was detected by RELAX[73] at genome-wide and gene-wide levels. Genes are considered to experience significantly relaxed selection with a K-value < 1 (q-value < 0.05). Conversely, when K-value > 1 (q-value < 0.05), the genes are considered to experience significantly intensified selection. The q-values represent an estimate of the false discovery rate (FDR) for each gene. Finally, the signal of positive selection in genes of water spider was detected by BUSTED (Branch-site Unrestricted Statistical Test for Episodic Diversification)[74].

Hemocyanins and bHLH-PAS gene annotation

We sought to analyze the hemocyanins genes in water spider, and download the amino acid sequences for hemocyanins genes of several species from NCBI database as references, including *Araneus ventricosus*, *Eurypelma californicum*, *Austrarchaea mainae*, and *Physocyclus mexicanus*. First, we performed a blastp-like search by MMseqs2 v13.45111[75] with four rounds of iteration. Then we used Interproscan (https://www.ebi.ac.uk/interpro/search/sequence/) to confirm the specific conserved domains including all-alpha domain (IPR005203, PF03722), ig-like domain (IPR005203, PF03723), copper containing domain (IPR000896, PF00372). The invalid matches of proteins were deleted according to the output of MMseqs2 with a tblatn-like search, and the intron/exon boundarie was also examined. The protein structure of hemocyanins protein of water spider was predicted by SWISS-MODEL[76]. Finally, we detected the selective pressure of hemocyanins genes between water spider and other spider species using BUSTED.

We next focused on the hypoxia inducible factor (HIF-α) genes and aryl hydrocarbon receptor nuclear translocator (ARNT/HIF-β), which both werebelonged to the bHLH-PAS gene family. Similarily, we analyzed the HIF genes in water spider and downloaded the bHLH-PAS genes of selected species from NCBI as references, including *Monodelphis domestica*, *Anas platyrhynchos*, *Ficedula albicollis*, *Anolis carolinensis*, *Apis mellifera*, *Bombyx mori*, *Anopheles gambiae*, *Drosophila melanogaster*, and *Acyrthosiphon pisum.* Further, all the sequences were identified by MMseqs2 with 4 rounds of iteration. Then we used Interproscan to confirm the specific conserved domains including basic helix-loop-helix (bHLH) domain (IPR011598, PF00010) PAS fold-3 domain (IPR013655, PF08447), and PAS fold domain (IPR013767, PF00989). The invalid matches of the proteins were deleted according to the results of software MMseqs2 with a tblatn-like search, and the intron/exon boundarie was also examined. Altogether,the phylogenetic trees for the hemocyanins and bHLH-PAS genes based on protein sequences were constructed by integrating the tools of MAFFT, trimAl and IQ-TREE.

Demographic history analysis

We sought to infer the demographic history and effective population size (Ne) of water spiders. For this, we performed additional whole genome resequencing of 20 water spider individuals (10 female and 10 male). The raw whole genome resequencing data of *T. antipodiana* was download from NCBI with Bioproject ID of PRJNA627506[77]. And the *E. davidi* was download from NCBI with BioProject ID PRJNA853523[78]. We ultilized a pairwise sequentially Markovian coalescence (PSMC) model[79] to finalize the analysis.

Horizontal gene transfer (HGT) gene identification and validation

To determine if particular genes in water spider that may have been horizontally acquired, we employed a robust and conservative phylogeny-based approach[20]. For ABC and ACAD genes in water spiders, we first blast these genes with corresponding protein sequence by BLASTP in NCBI database. To evaluate the reliability of putative HGT-acquired genes, we first compared our putative HGT-

acquired genes with previously published genes across the phylogenetic tree. Then we calculated these gene expression among different tissues. In addition, we randomly sampled 9 putative HGT-acquired genes, including 6 ABC genes and 3 ACAD genes. For these genes, we performed additional PCR assays to amplify upstream or downstream regions that flanked the foreign gene. Each PCR target size was 700 ~ 1,500 bp. The agarose gel electrophoretic analysis was used to determine whether PCR products were expected. The PCR product was sequenced using Sanger sequence technology. If the upstream and downstream regions of HGT-acquired gene was successfully amplified and their Sanger sequences were nearly identical to DNA sequences with identity of R over 98% in reference genome, the HGT-acquired gene was considered successfully validated. The primer sequences were shown in Table S7.

Differential gene expression analysis

To explore the changes in gene expression response to hypoxia stress, we performed comparative transcriptomic analysis between spider samples under hypoxia stress and normal oxygene condition as control. For this, we aligned the transcriptome sequencing reads of whole body to the whole genome of water spider by HISAT2. The read counts were counted by featureCounts v1.6.4[80] and then normalized among samples FPKM normalization method by DESeq2 package v1.24.0[81]. Genes with an adjusted p value <0.05 and $|\log2 \text{ ratio}| \geq 1$ in at least one temperature condition were considered to be differentially expressed.

Differential metabolites analysis

To investigate the differential metabolites induced by hypoxia treatment, we performed metabolomics by comparing spider samples under hypoxia and nornal oxygene condition using non-targeted LC-MC (Suzhou PANOMIX Biomedical Tech, China). Specifically, the samples were prepared following the procedures: first accurately weighed 1g sample, 1000 μL tissue extract (75% 9:1 methanol: chloroform, 25% H_2O), and 3 steel balls into a 2mL centrifuge tube. Then, put it into the tissue grinder, grind it at 50 Hz for 60 s, and repeat the above operation twice. Next, ultrasonic shaking for 30 min and then ice bath for 30 minutes. After centrifuged for 10 min at 12,000 rpm in 4 ℃, take all the supernatant, dry it. The 200 μL 50% acetonitrile solution was used to prepare 2-Amino-3-(2-chloro- phenyl)-propionic acid (4 ppm), then re-dissolve the sample, filter the supernatant by 0.22 μm membrane and transfer into the detection bottle for LC-MS detection.

Data availability

The genome sequence of the *Argyroneta aquatica* is available at the scienceDB[82] (https://www.scidb.cn/s/Urai2a). 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. The raw Illumina data[83] (https://www.scidb.cn/s/Urai2a), the metabolites and transcriptomic raw data[84] (https://www.scidb.cn/s/amuIva) are available on scienceDB. All other relevant data are available upon request.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Reference

1. Schutz, D. & Taborsky, M. Adaptations to an aquatic life may be responsible for the reversed sexual size dimorphism in the water spider, *Argyroneta aquatica*. *Evol Ecol Res* **5**, 105-117 (2003).
2. WSC. World Spider Catalog. Version 24.5 Natural History Museum Bern. http://wsc.nmbe.ch (2024).
3. Seymour, R. S. & Hetz, S. K. The diving bell and the spider: the physical gill of *Argyroneta aquatica*. *J Exp Biol* **214**, 2175-2181, doi:10.1242/jeb.056093 (2011).
4. Schmitz, A. Respiration in spiders (Araneae). *J Comp Physiol B* **186**, 403-415, doi:10.1007/s00360-016-0962-8 (2016).
5. Neumann, D. & Kureck, A. Composite structure of silken threads and a proteinaceous hydrogel which form the diving bell wall of the water spider *Agyroneta aquatica*. *Springerplus* **2**, 223, doi:10.1186/2193-1801-2-223 (2013).
6. Neumann, D. & Kureck, A. Composite structure of silken threads and a proteinaceous hydrogel which form the diving bell wall of the water spider *Agyroneta aquatica*. *Springerplus* **2**, 223, doi:10.1186/2193-1801-2-223 (2013).
7. Schaber, C. F., Grawe, I. & Gorb, S. N. Attachment discs of the diving bell spider *Argyroneta aquatica*. *Commun Biol* **6**, 1232, doi:10.1038/s42003-023-05575-7 (2023).
8. Hughes, A. L. Evolution of adaptive phenotypic traits without positive Darwinian selection. *Heredity* **108**, 347-353, doi:10.1038/hdy.2011.97 (2012).
9. Yoder, J. B. *et al.* Ecological opportunity and the origin of adaptive radiations. *J Evol Biol* **23**, 1581-1596, doi:10.1111/j.1420-9101.2010.02029.x (2010).
10. Hunt, B. G. *et al.* Relaxed selection is a precursor to the evolution of phenotypic plasticity. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 15936-15941, doi:10.1073/pnas.1104825108 (2011).
11. Yim, H. S. *et al.* Minke whale genome and aquatic adaptation in cetaceans. *Nature genetics* **46**, 88-92, doi:10.1038/ng.2835 (2014).
12. Tian, R. *et al.* Evolutionary Genetics of Hypoxia Tolerance in Cetaceans during Diving. *Genome Biology and Evolution* **8**, 827-839, doi:10.1093/gbe/evw037 (2016).
13. Noh, H. J. *et al.* The Antarctic Weddell seal genome reveals evidence of selection on cardiovascular phenotype and lipid handling. *Commun Biol* **5**, 140, doi:10.1038/s42003-022-03089-2 (2022).
14. Fuselli, S. *et al.* Relaxation of Natural Selection in the Evolution of the Giant Lungfish Genomes. *Molecular biology and evolution* **40**, doi:10.1093/molbev/msad193 (2023).
15. Lu, Y., Xing, H. & Zhang, D. Evidence for relaxed selection of mitogenome in rapid-flow cyprinids. *Genes Genomics* **41**, 863-869, doi:10.1007/s13258-019-00817-7 (2019).

16. Velotta, J. P., McCormick, S. D., O'Neill, R. J. & Schultz, E. T. Relaxed selection causes microevolution of seawater osmoregulation and gene expression in landlocked Alewives. *Oecologia* **175**, 1081-1092, doi:10.1007/s00442-014-2961-3 (2014).

17. Messler, A., Wund, M. A., Baker, J. A. & Foster, S. A. The effects of relaxed and reversed selection by predators on the antipredator behavior of the threespine stickleback,. *Ethology* **113**, 953-963, doi:10.1111/j.1439-0310.2007.01416.x (2007).

18. Zhu, B. *et al.* Horizontal gene transfer in silkworm, *Bombyx mori*. *BMC genomics* **12**, 248, doi:10.1186/1471-2164-12-248 (2011).

19. Li, H. S. et al. Horizontally acquired antibacterial genes associated with adaptive radiation of ladybird beetles. *BMC biology* **19**, doi:ARTN 710.1186/s12915-020-00945-7 (2021).

20. Li, Y. *et al.* HGT is widespread in insects and contributes to male courtship in lepidopterans. *Cell* **185**, 2975-2987 e2910, doi:10.1016/j.cell.2022.06.014 (2022).

21. Gilbert, C. & Maumus, F. Multiple Horizontal Acquisitions of Plant Genes in the Whitefly Bemisia tabaci. *Genome Biol Evol* **14**, doi:10.1093/gbe/evac141 (2022).

22. Kirsch, R. et al. Metabolic novelty originating from horizontal gene transfer is essential for leaf beetle survival. *Proceedings of the National Academy of Sciences of the United States of America* **119**, doi:ARTN e220585711910.1073/pnas.2205857119 (2022).

23. Hu, L. *et al.* Arctic introgression and chromatin regulation facilitated rapid Qinghai-Tibet Plateau colonization by an avian predator. *Nature communications* **13**, 6413, doi:10.1038/s41467-022-34138-3 (2022).

24. Park, T. J. *et al.* Fructose-driven glycolysis supports anoxia resistance in the naked mole-rat. *Science* **356**, 307-311, doi:10.1126/science.aab3896 (2017).

25. Wang, Z. et al. Genome and transcriptome of provide insights into high-altitude hypoxia adaptation and symbiosis. *Iscience* 26, doi:ARTN 10779310.1016/j.isci.2023.107793 (2023).

26. Zhang, X. Z. *et al.* Lung transcriptome analysis for the identification of genes involved in the hypoxic adaptation of plateau pika (*Ochotona curzoniae*). *Comp Biochem Physiol Part D Genomics Proteomics* **41**, 100943, doi:10.1016/j.cbd.2021.100943 (2022).

27. Harrison, J. F., Greenlee, K. J. & Verberk, W. Functional Hypoxia in Insects: Definition, Assessment, and Consequences for Physiology, Ecology, and Evolution. *Annu Rev Entomol* **63**, 303-325, doi:10.1146/annurev-ento-020117-043145 (2018).

28. Milano, F. *et al.* Spider conservation in Europe: a review. *Biol Conserv* **256**, doi: ARTN 10902010.1016/j.biocon.2021.109020 (2021).

29. Cicconardi, F. *et al.* Genomic Signature of Shifts in Selection in a Subalpine Ant and Its Physiological Adaptations. *Molecular biology and evolution* **37**, 2211-2227, doi:10.1093/molbev/msaa076 (2020).

30. Dermauw, W. & Van Leeuwen, T. The ABC gene family in arthropods: comparative genomics and role in insecticide transport and resistance. *Insect Biochem Mol Biol* **45**, 89-110, doi:10.1016/j.ibmb.2013.11.001 (2014).

31. Shen, Y. Q., Lang, B. F. & Burger, G. Diversity and dispersal of a ubiquitous protein family: acyl-CoA dehydrogenases. *Nucleic Acids Res* **37**, 5619-5631, doi:10.1093/nar/gkp566 (2009).

32. Nyffeler, M. & Pusey, B. J. Fish predation by semi-aquatic spiders: a global pattern. *Plos One* **9**, e99459, doi:10.1371/journal.pone.0099459 (2014).

33. Wang, W. *et al.* First de novo whole genome sequencing and assembly of the bar-headed goose. *PeerJ* **8**, e8914, doi:10.7717/peerj.8914 (2020).

34. Rytkonen, K. T., Vuori, K. A., Primmer, C. R. & Nikinmaa, M. Comparison of hypoxia-inducible factor-1 alpha in hypoxia-sensitive and hypoxia-tolerant fish species. *Comp Biochem Physiol Part D Genomics Proteomics* **2**, 177-186, doi:10.1016/j.cbd.2007.03.001 (2007).

35. Huwiler, A. & Pfeilschifter, J. Hypoxia and lipid signaling. *Biol Chem* **387**, 1321-1328, doi:10.1515/BC.2006.165 (2006).

36. Shen, Y. Q., Lang, B. F. & Burger, G. Diversity and dispersal of a ubiquitous protein family: acyl-CoA dehydrogenases. *Nucleic Acids Res* **37**, 5619-5631, doi:10.1093/nar/gkp566 (2009).

37. Kolmogorov, M., Yuan, J., Lin, Y. & Pevzner, P. A. Assembly of long, error-prone reads using repeat graphs. *Nature biotechnology* **37**, 540-546, doi:10.1038/s41587-019-0072-8 (2019).

38. Hu, J., Fan, J., Sun, Z. & Liu, S. NextPolish: a fast and efficient genome polishing tool for long-read assembly. *Bioinformatics* **36**, 2253-2255, doi:10.1093/bioinformatics/btz891 (2020).

39. Roach, M. J., Schmidt, S. A. & Borneman, A. R. Purge Haplotigs: allelic contig reassignment for third-gen diploid genome assemblies. *BMC Bioinformatics* **19**, 460, doi:10.1186/s12859-018-2485-7 (2018).

40. Durand, N. C. *et al.* Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst.* **3**, 95-98, doi:10.1016/j.cels.2016.07.002 (2016).

41. Chen, Y., Ye, W. C., Zhang, Y. D. & Xu, Y. S. High speed BLASTN: an accelerated MegaBLAST search tool. *Nucleic Acids Res* **43**, 7762-7768, doi:10.1093/nar/gkv784 (2015).

42. Simao, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210-3212, doi:10.1093/bioinformatics/btv351 (2015).

43. Flynn, J. M. *et al.* RepeatModeler2 for automated genomic discovery of transposable element families. *Proceedings of the National Academy of Sciences of the United States of America* **117**, 9451-9457, doi:10.1073/pnas.1921046117 (2020).

44. Tarailo-Graovac, M. & Chen, N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr Protoc Bioinformatics* **Chapter 4**, 4 10 11-14 10 14, doi:10.1002/0471250953.bi0410s25 (2009).

45. Nawrocki, E. P. & Eddy, S. R. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* **29**, 2933-2935, doi:10.1093/bioinformatics/btt509 (2013).

46. Chan, P. P., Lin, B. Y., Mak, A. J. & Lowe, T. M. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res* **49**, 9077-9096, doi:10.1093/nar/gkab688 (2021).

47. Holt, C. & Yandell, M. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* **12**, 491, doi:10.1186/1471-2105-12-491 (2011).

48. Hoff, K. J. & Stanke, M. Predicting Genes in Single Genomes with AUGUSTUS. *Curr Protoc*

Bioinformatics **65**, e57, doi:10.1002/cpbi.57 (2019).

49. Bruna, T., Lomsadze, A. & Borodovsky, M. GeneMark-EP plus: eukaryotic gene prediction with self-training in the space of genes and proteins. *Nar Genomics and Bioinformatics* **2**, doi:10.1093/nargab/lqaa026 (2020).

50. Bruna, T., Hoff, K. J., Lomsadze, A., Stanke, M. & Borodovsky, M. BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR genomics and bioinformatics* **3**, lqaa108-lqaa108, doi:10.1093/nargab/lqaa108 (2021).

51. Kim, D., Landmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**, 357-U121, doi:10.1038/Nmeth.3317 (2015).

52. Pertea, M. *et al*. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature biotechnology* **33**, 290-295, doi:10.1038/nbt.3122 (2015).

53. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* **12**, 59-60, doi:10.1038/nmeth.3176 (2015).

54. Mulder, N. & Apweiler, R. InterPro and InterProScan: tools for protein sequence classification and comparison. *Methods Mol Biol* **396**, 59-70, doi:10.1007/978-1-59745-515-2_5 (2007).

55. Cantalapiedra, C. P., Hernandez-Plaza, A., Letunic, I., Bork, P. & Huerta-Cepas, J. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Molecular biology and evolution* **38**, 5825-5829 (2021).

56. Mistry, J. *et al*. Pfam: The protein families database in 2021. *Nucleic Acids Res* **49**, D412-D419, doi:10.1093/nar/gkaa913 (2021).

57. Mi, H., Poudel, S., Muruganujan, A., Casagrande, J. T. & Thomas, P. D. PANTHER version 10: expanded protein families and functions, and analysis tools. *Nucleic Acids Res* **44**, D336-342, doi:10.1093/nar/gkv1194 (2016).

58. Lewis, T. E. *et al.* Gene3D: Extensive prediction of globular domains in proteins. *Nucleic Acids Res* **46**, D435-D439, doi:10.1093/nar/gkx1069 (2018).

59. Pandurangan, A. P., Stahlhacke, J., Oates, M. E., Smithers, B. & Gough, J. The SUPERFAMILY 2.0 database: a significant proteome update and a new webserver. *Nucleic Acids Res* **47**, D490-D494, doi:10.1093/nar/gky1130 (2019).

60. Marchler-Bauer, A. *et al.* CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* **45**, D200-D203, doi:10.1093/nar/gkw1129 (2017).

61. Huerta-Cepas, J. *et al*. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* **47**, D309-D314, doi:10.1093/nar/gky1085 (2019).

62. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* **20**, 238, doi:10.1186/s13059-019-1832-y (2019).

63. Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular biology and evolution* **30**, 772-780, doi:10.1093/molbev/mst010 (2013).

64. Capella-Gutierrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972-1973,

doi:10.1093/bioinformatics/btp348 (2009).

65. Kuck, P. & Longo, G. C. FASconCAT-G: extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. *Front Zool* **11**, 81, doi:10.1186/s12983-014-0081-x (2014).

66. Minh, B. Q. *et al.* IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular biology and evolution* **37**, 1530-1534, doi:10.1093/molbev/msaa015 (2020).

67. Magalhaes, I. L. F., Azevedo, G. H. F., Michalik, P. & Ramirez, M. J. The fossil record of spiders revisited: implications for calibrating trees and evidence for a major faunal turnover since the Mesozoic. *Biol Rev Camb Philos Soc* **95**, 184-217, doi:10.1111/brv.12559 (2020).

68. Yang, Z. H. PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular biology and evolution* **24**, 1586-1591, doi:10.1093/molbev/msm088 (2007).

69. De Bie, T., Cristianini, N., Demuth, J. P. & Hahn, M. W. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* **22**, 1269-1271, doi:10.1093/bioinformatics/btl097 (2006).

70. Suyama, M., Torrents, D. & Bork, P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res* **34**, W609-W612, doi:10.1093/nar/gkl315 (2006).

71. Pond, S. L., Frost, S. D. & Muse, S. V. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* **21**, 676-679, doi:10.1093/bioinformatics/bti079 (2005).

72. Tong, C., Aviles, L., Rayor, L. S., Mikheyev, A. S. & Linksvayer, T. A. Genomic signatures of recent convergent transitions to social life in spiders. *Nature communications* **13**, 6967, doi:10.1038/s41467-022-34446-8 (2022).

73. Wertheim, J. O., Murrell, B., Smith, M. D., Kosakovsky Pond, S. L. & Scheffler, K. RELAX: detecting relaxed selection in a phylogenetic framework. *Molecular biology and evolution* **32**, 820-832, doi:10.1093/molbev/msu400 (2015).

74. Murrell, B. *et al.* Gene-wide identification of episodic selection. *Molecular biology and evolution* **32**, 1365-1371, doi:10.1093/molbev/msv035 (2015).

75. Steinegger, M. & Soding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature biotechnology* **35**, 1026-1028, doi:10.1038/nbt.3988 (2017).

76. Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296-W303, doi:10.1093/nar/gky427 (2018).

77. Fan, Z. *et al.* A chromosome-level genome of the spider *Trichonephila antipodiana* reveals the genetic basis of its polyphagy and evidence of an ancient whole-genome duplication event. *GigaScience* **10**, doi:10.1093/gigascience/giab016 (2021).

78. Fan, Z. et al. Lampshade web spider *Ectatosticta davidi* chromosome-level genome assembly provides evidence for its phylogenetic position. *Communications Biology* **6**, doi:ARTN 74810.1038/s42003-023-05129-x (2023).

79. Liu, J., Ji, X. & Chen, H. Beta-PSMC: uncovering more detailed population history using beta distribution. *BMC genomics* **23**, 785, doi:10.1186/s12864-022-09021-6 (2022).

80. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930, doi:10.1093/bioinformatics/btt656 (2014).

81. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).

82. Fan, Z. & Zhang Z. S. The genome sequence of the water spider *Argyroneta aquatica*. V1. Science Data Bank, https://cstr.cn/31253.11.sciencedb.14858. CSTR:31253.11.sciencedb.14858. (2024).

83. Fan, Z., Wang, L.Y. & Zhang Z. S. The short reads of whole genome sequencing in water spider *Argyroneta aquatica*. V1. *Science Data Bank*, https://cstr.cn/31253.11.sciencedb.11012. CSTR:31253.11.sciencedb.11012. (2023).

84. Fan, Z. & Zhang Z. S. The metabolites and transcriptomic data of water spider under hypoxia stress and normal oxygene condition. V1. *Science Data Bank*, https://cstr.cn/31253.11.sciencedb.15069. CSTR:31253.11.sciencedb.15069. (2024).

Competing interests

The authors declare no competing interests.

Acknowledgements

We thanked for 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 of the original research. Z.F. 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. Y.C., 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 the phylogenetic analysis of water spider *A. aquatica* with other spiders.

(a) Images of an adult female water spider.

(b) 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 elements, DNA: DNA transposable elements, Gene: distribution of genes, GC: GC content.

(c) Phylogenetic and expansion gene family analyses of water spider. The divergence times among different species are shown at the bottom. Node values indicate gene families showing expansion (red), contraction (blue), and rapid evolution (black in brackets).

Figure 2. The genome wide dN/dS analysis and the demographic history analysis of water spider and other spiders.

(a) The genome wide dN/dS analysis of water spider and other spiders. RELAX modeled the distribution of three categories of dN/dS, and the parameter K was used to indicates overall relaxation ($K < 1$) or intensification ($K > 1$).

(b) The demographic history analysis of water spider. The point from right to left respectively were humman migrated out of Africa, MPT, Shishapangma Ice Age, Kunlun-Yellow River, and Kunlun Ice Age.

Figure 3. The source of HGT-acquired ABC and ACAD genes in water spider.

(a) The genus of bacteria of the source for HGT-acquired ABC and ACAD genes.

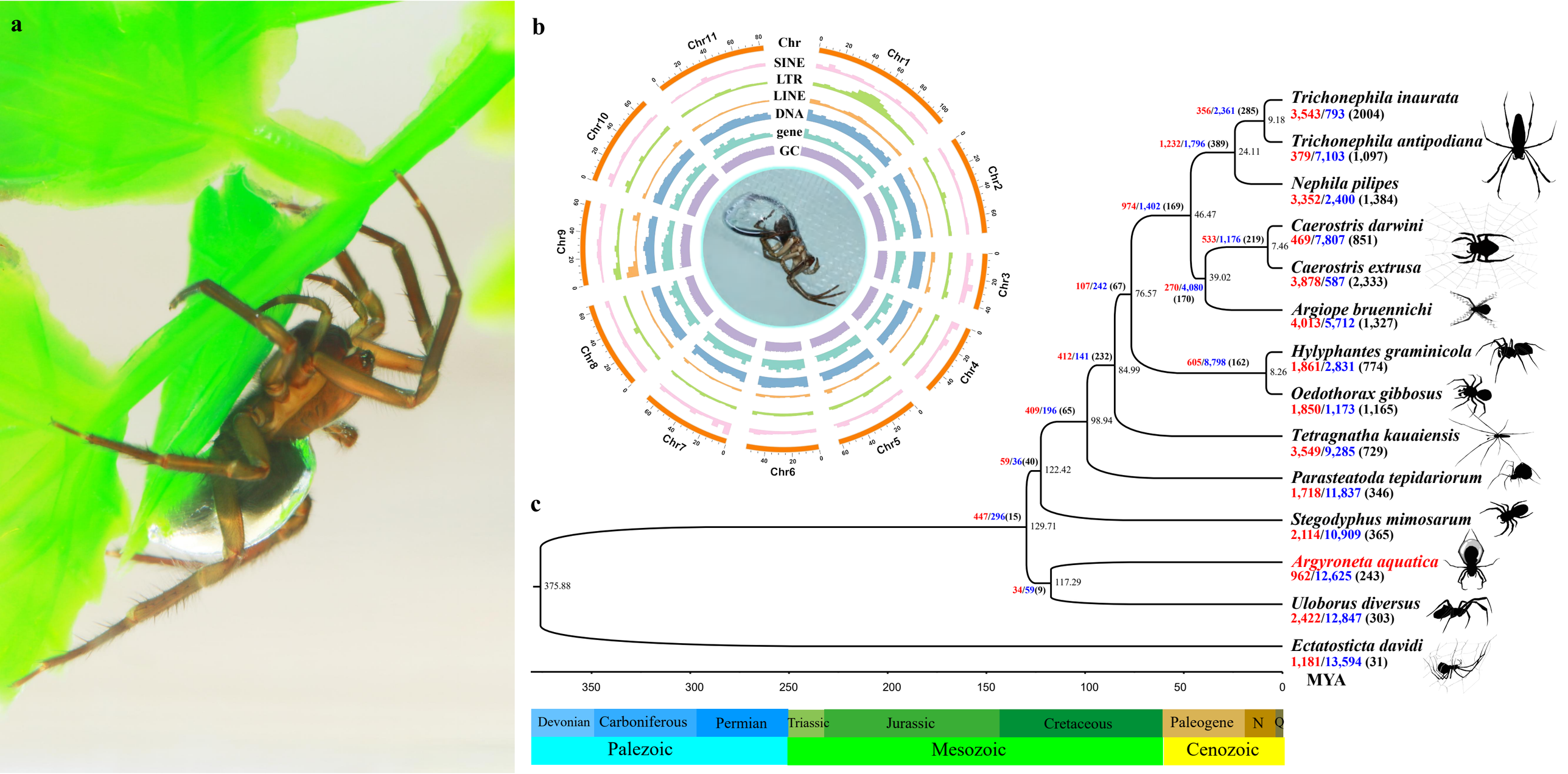
(b) The main living environment which transferred for water spider HGT-acquired ABC and ACAD genes.

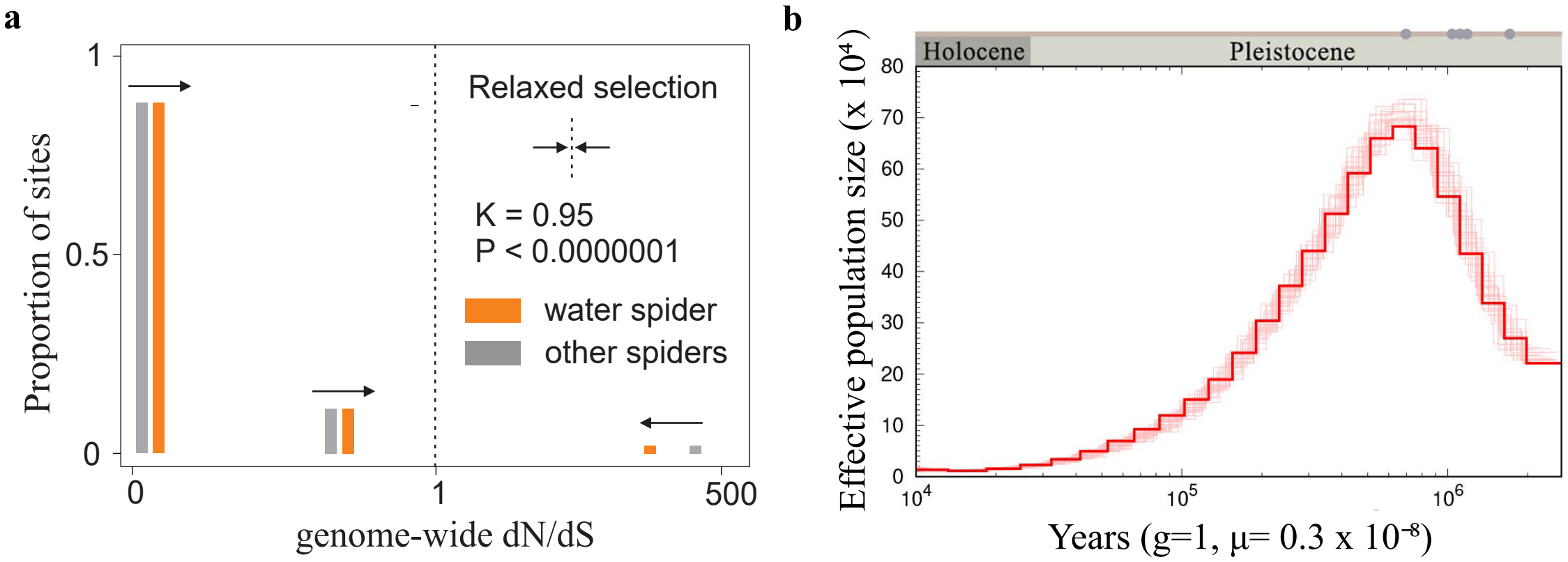
(c) The type of TEs near these HGT-acquired ABC and ACAD genes.

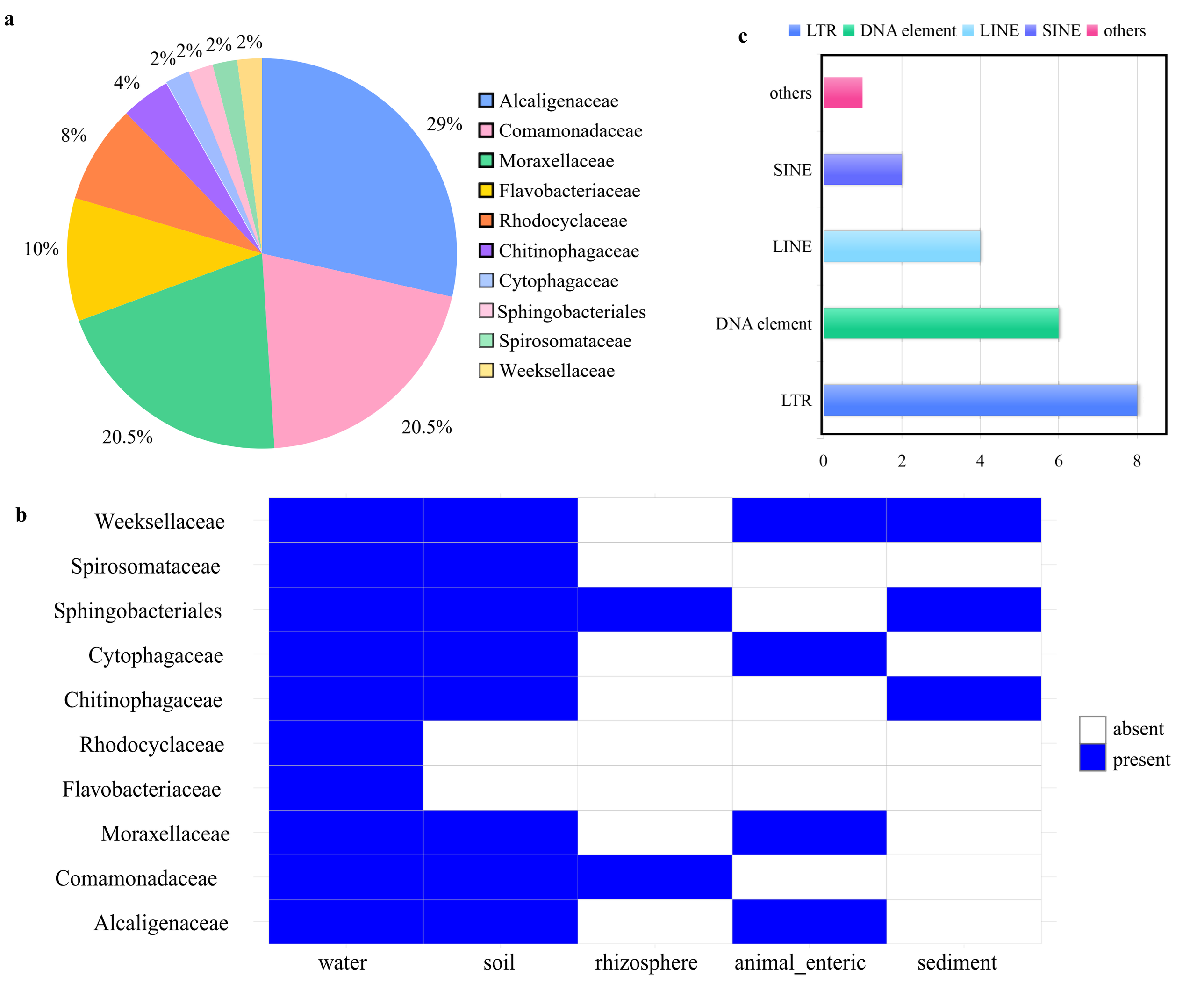
Figure 4. Schematic summary of adaptive strategies of water spider to hypoxia adaptation.

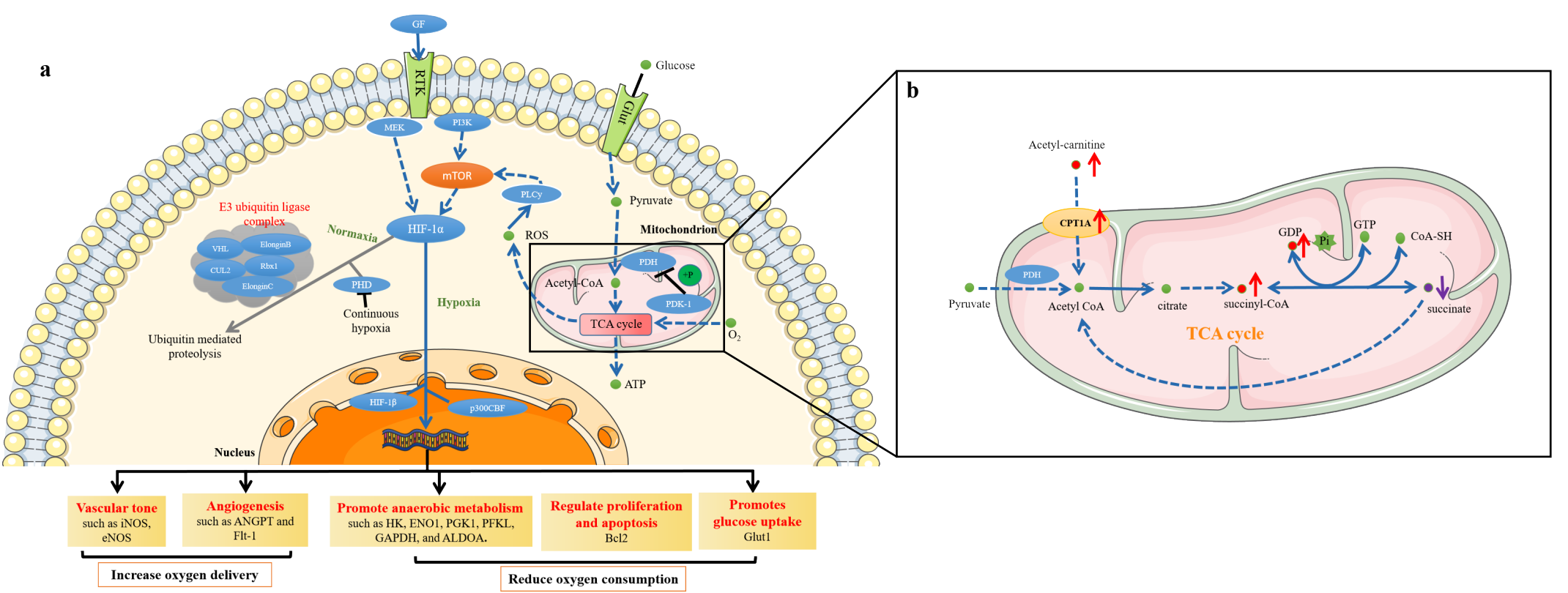
(a) HIF-1 signaling pathway of water spider.

(b) Changes of TCA cycle when facing hypoxia in water spider.


a
b
Chr
SINE
LTR
LINE
DNA
gene
GC
Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7
Chr8
Chr9
Chr10
Chr11
c
Trichonephila inaurata
3,543/793 (2004)
Trichonephila antipodiana
379/7,103 (1,097)
Nephila pilipes
3,352/2,400 (1,384)
Caerostris darwini
469/7,807 (851)
Caerostris extrusa
3,878/587 (2,333)
Argiope bruennichi
4,013/5,712 (1,327)
Hylyphantes graminicola
1,861/2,831 (774)
Oedothorax gibbosus
1,850/1,173 (1,165)
Tetragnatha kauaiensis
3,549/9,285 (729)
Parasteatoda tepidariorum
1,718/11,837 (346)
Stegodyphus mimosarum
2,114/10,909 (365)
Argyroneta aquatica
962/12,625 (243)
Uloborus diversus
2,422/12,847 (303)
Ectatosticta davidi
1,181/13,594 (31)
356/2,361 (285)
9.18
1,232/1,796 (389)
24.11
974/1,402 (169)
46.47
533/1,176 (219)
7.46
39.02
107/242 (67)
76.57
270/4,080 (170)
412/141 (232)
84.99
605/8,798 (162)
8.26
409/196 (65)
98.94
59/36(40)
122.42
447/296(15)
129.71
34/59(9)
117.29
375.88
350
300
250
200
150
100
50
0
MYA
Devonian
Carboniferous
Permian
Triassic
Jurassic
Cretaceous
Paleogene
N
Q
Palezoic
Mesozoic
Cenozoic


a
Proportion of sites
1
0.5
0
0
1
500
genome-wide dN/dS
Relaxed selection
K = 0.95
P < 0.0000001
water spider
other spiders
b
Holocene
Pleistocene
Effective population size (x 10⁴)
80
70
60
50
40
30
20
10
0
10⁴
10⁵
10⁶
Years (g=1, μ= 0.3 x 10⁻⁸)


a
29%
20.5%
20.5%
10%
8%
4%
2%
2%
2%
2%
Alcaligenaceae
Comamonadaceae
Moraxellaceae
Flavobacteriaceae
Rhodocyclaceae
Chitinophagaceae
Cytophagaceae
Sphingobacteriales
Spirosomataceae
Weeksellaceae
c
LTR
DNA element
LINE
SINE
others
others
SINE
LINE
DNA element
LTR
0
2
4
6
8
b
Weeksellaceae
Spirosomataceae
Sphingobacteriales
Cytophagaceae
Chitinophagaceae
Rhodocyclaceae
Flavobacteriaceae
Moraxellaceae
Comamonadaceae
Alcaligenaceae
water
soil
rhizosphere
animal_enteric
sediment
absent
present


a
GF
RTK
Glucose
Glut
MEK
PI3K
mTOR
PLCy
Pyruvate
E3 ubiquitin ligase complex
VHL
ElonginB
CUL2
Rbx1
ElonginC
Normaxia
HIF-1α
ROS
Mitochondrion
PHD
Continuous hypoxia
Hypoxia
Acetyl-CoA
PDH
+P
PDK-1
TCA cycle
O2
Ubiquitin mediated proteolysis
ATP
HIF-1β
p300CBF
Nucleus
Vascular tone
such as iNOS, eNOS
Angiogenesis
such as ANGPT and Flt-1
Promote anaerobic metabolism
such as HK, ENO1, PGK1, PFKL, GAPDH, and ALDOA.
Regulate proliferation and apoptosis
Bcl2
Promotes glucose uptake
Glut1
Increase oxygen delivery
Reduce oxygen consumption
b
Acetyl-carnitine
CPT1A
PDH
Pyruvate
Acetyl CoA
citrate
succinyl-CoA
GDP
Pi
GTP
CoA-SH
succinate
TCA cycle