Population genetic structure and demographic history of the East Asian wolf spider *Pardosa astrigera*

Dan Fu¹, Lijuan Liu¹, Ying Cheng¹, Haodong Chen¹, Yufa Luo¹

¹ Key Laboratory of Wetland Biodiversity of the Jianhu Basin of Shaoxing, School of Life and Environmental Sciences, Shaoxing University, Shaoxing 312000, China

Abstract

The wolf spider *Pardosa astrigera* L. Koch, 1878, an important biological control agent for pests in agriculture, is widely distributed in various ecosystems across East Asia. This study used mitochondrial DNA and aimed to provide an in-depth understanding of population genetic structure and evolutionary history throughout the species. Mitochondrial gene sequences from 107 samples of *P. astrigera* from 25 East Asian populations were used for genetic analyses. Our data revealed an asymmetric phylogeographic distribution in two sympatric lineages (1–2) of *P. astrigera* in continental East Asia. The spatio-temporal pattern of two mitotypes of *P. astrigera* in this region gives strong support for a Northeast Asian origin during the late Pleistocene (~1.69 million years ago) and the population expansion time of ~74,340 (58,832–104,236) years ago (during the last glacial period) and dual colonization around East Asia from two directions: from North to South and from East to West. Our phylogeographic results suggested that Pleistocene climate oscillations with subsequent fragmentation events and secondary contacts were the major impact factors of the diversification, geographic distribution, and expansion patterns of *P. astrigera*, and human activities and ballooning probably accelerated its recent dispersal.

Key Words

Last glacial period, late Pleistocene diversification, Lycosidae, mtDNA, phylogeography

Introduction

*Pardosa astrigera* (Araneae: Lycosidae), a wandering spider, inhabits a relatively arid-cold environment. It is a common species in various ecosystems of the temperate and subtropical regions of East Asia and possesses strong environmental adaptability, high dispersal potentiality, and rapid population diversification (Lida et al. 2016; Li et al. 2020). This spider can prey on a variety of pests in agriculture and forestry and has strong hunger tolerance and high fertility, as well as long-term population stability and substantially large populations (Li et al. 2020). Therefore, *P. astrigera* plays a vital role in pest control. It is widespread in East Asian regions such as China, Korea, Japan, and Russia (Far East) (Song et al. 1999; World Spider Catalog 2024). This wolf spider is a dominant species in various ecosystems across Korea (Kim and Yoo 1997) and most parts of China, except Fujian, Guangdong, and Hainan provinces (Song et al. 1999). Previous morphological studies suggested that *P. astrigera* was likely a species complex (Schenkel 1963). Recent investigations have considered the morphological variations among the different geographical populations to be intraspecific (Yin et al. 1997; Chang et al. 2007). Studies on the population genetic structure of *P. astrigera* can estimate genetic diversity and gene flow among its different populations and infer the ancestral populations and the phylogeographic pattern. Such information is essential for effectively protecting and utilizing the natural predators of pests in agriculture and forestry.

The phylogeography of *P. astrigera* distributed in local regions of China was investigated using nuclear ITS2 gene sequences, and two major lineages were observed (Chang et al. 2007). However, for the ancient spider
species with an extensive geographic range, the information concerning genetic structure, spatio-temporal evolution patterns, and demographic history across the whole distribution range in East Asia was still unclear, and the following questions concerning this widespread species have not been investigated: 1) genetic relationships among the East Asian populations based on mitochondrial genes; 2) allopatric divergence among the East Asian regions; and 3) origin and dispersal across East Asia. These questions will provide a better understanding of the _P. astrigera_ population history in East Asia.

Because of the haploid maternal inheritance of mitochondrial genes and their high mutation rate and abundance in cells, they are widely used as molecular markers to analyze the population genetic structure and phylogeography of broadly distributed animal species (Avise 2000). To assess the mtDNA population genetic structure and potential dispersal routes of _P. astrigera_, we studied its genetic structure and phylogeography using mitochondrial COI, 16S, and NADH1 genes. We recently collected individuals of the spider from China and augmented our sampling with sequences from GenBank covering nearly the entire geographic range of _P. astrigera_. The current data set covers 25 distinct localities (21 from China, three from South Korea, and one from Japan) around East Asia. It thus offers a more holistic insight into _P. astrigera_ matriline history than have prior studies.

**Materials and methods**

**Sampling, sequencing, and sequence analyses**

We collected 19 samples of _P. astrigera_ from 6 provinces of China (Suppl. material 1: table S1) and sequenced 3 mitochondrial genes (COI, 16S, and NADH1) using three primer pairs (Suppl. material 1: table S2) according to the procedures of Luo and Li (2022). Additional sequences of the species from Korea, Japan and 14 provinces of China were taken from Chang et al. (2007), Li et al. (2011), and GenBank (Suppl. material 1: table S1). The mtDNA sequences of 107 colonists from 25 geographic sites in East Asia were used for genetic analyses (Fig. 1a; Suppl. material 1: table S1). BioEdit was used to check the quality of the DNA sequences (Hall 1999). Sequence alignments were done with MAFFT v7 (Katoh and Standley 2013). Unique haplotypes were identified using DnaSP v5 (Librado and Rozas 2009). The data matrix is available from the authors and will be submitted to the Dryad database (online at http://datadryad.org/).

**Phylogeny and haplotype network reconstruction**

A matrix combining the sequences of the COI, 16S, and NADH1 markers was constructed and used to obtain the phylogenetic relationships of 107 _P. astrigera_ individuals using the Bayesian inference (BI) and maximum likelihood (ML) approaches. _Pardosa laura_ was used as an outgroup. Bayesian analysis was performed in MrBayes V3.2.1 (Ronquist and Huelsenbeck 2003). We used jModelTest to select the best-fit evolutionary model for each codon base or gene partition (Posada 2008) with the Bayesian information criterion (Suppl. material 1: table S3). The Markov chain Monte Carlo (MCMC) simulation was run for 15 million generations and sampled every 1,000 generations, and unlinked parameters among partitions were used in the chain runs. Tracer v1.5 was employed to check effective sample size (ESS) values (>200) for each parameter (Rambaut and Drummond 2009), and a 50% majority rule consensus tree was computed after discarding the first 25% of sampled trees as burn-in. The ML tree was reconstructed using IQ-TREE v1.6.12 (Nguyen et al. 2015) as implemented in the W-IQ-Tree web server (Trifinopoulos et al. 2016). The substitution model for each codon base or gene partition was automatically selected, and an ultrafast bootstrap with 1,000 iterations was conducted to estimate nodal support across the topology (Minh et al. 2013; Kalyaanamoorthy et al. 2017; Hoang et al. 2018). To compare genetic connections between geographic populations, three haplotype networks were inferred in PopART (Leigh and Bryant 2015) using unrooted minimum-spanning trees (Bandelt et al. 1999) based on the separate mtDNA gene sequences. Based on climate, geographical divisions, distributional data for the _P. astrigera_ species from the literature (Song et al. 1999; World Spider Catalog 2024), and information from our samples. We delimited 6 areas, including Northwest China (Qinghai+Gansu, China); Northeast Asia (Neimenggu+Heilongjiang, China+Korea+Japan); East China (Jiangsu+Shandong+Zhejiang, China); Southwest China (Yunnan+Guizhou, China); Central China (Hubei+Hunan+Anhui, China); and North China (Hebei+Henan+Shanxi+Sanxi, China).

**Genetic diversity and distance calculation**

We defined six populations from the sampled regions for the haplotype network reconstruction above. Population genetic diversity was inferred from the COI, 16S, and NADH1 sequences. Estimations of genetic diversity (number of haplotypes, nh; nucleotide diversity, π; haplotype diversity, h) from each population were performed in DnaSP 5.10.01 (Librado and Rozas 2009). Genetic distances among haplotypes and overall mean distances of _P. astrigera_ populations were estimated using MEGA v5 (Tamura et al. 2011).

**Estimations of divergence time and mtDNA substitution rate**

Divergence times and mtDNA substitution rates were inferred based on the combined COI, 16S, and NADH1 sequences. They were implemented in BEAST v1.6.1 (Drummond and Rambaut 2007).
Figure 1. a. Sampling regions of the East Asian wolf spider *Pardosa astrigera* and the haplotype distribution of the 16S gene. Detailed sampling information is presented in Suppl. material 1: table S1 in the Supplemental materials; b. Time-tree of the species *Pardosa astrigera* based on the combined mitochondrial COI, 16S, and NADH1 sequences. Blue dots mark the four fossil calibration nodes. Numbers at nodes indicate the main divergence times of the haplotypes of the species. **“** indicates stable branches with Bayesian support >0.90.
First, a test was run to determine if a molecular clock model was appropriate based on the harmonic mean (HM) of the likelihood values of the MCMC samples (Newton and Raftery 1994) and a more accurate stepping-stone (SS) sampling method (Xie et al. 2011) in MrBayes v3.2.1. We found that the total data sets did not conform to a strict molecular clock (P<0.001) and that a relaxed molecular clock method was more suitable for estimating divergence time. Therefore, we estimated divergence times using the uncorrelated lognormal relaxed molecular clock model. The tree prior was set to the birth-death and Yule specification process models. We repeated the analysis with birth-death and Yule prior to assess the sensitivity of our results to tree-prior specification. According to the marginal likelihood estimated using stepping-stone sampling, the birth-death model outperformed the Yule model. Hence, subsequent analyses focused on the chronogram resulting from the birth-death prior. JModelTest (Posada 2008) was used to select the best-fitting substitution model for each gene or codon base partition under the Bayesian information criterion (Suppl. material 1: table S4). To obtain reliable results, we ran five independent MCMC tree searches for 300 million generations and sampled every 1,000 generations. The convergence of MCMC chains was assessed in TRACER v1.5 (Rambaut and Drummond 2009). TreeAnnotator v1.6.1 (in the BEAST package) was used to produce a maximum clade credibility (MCC) tree with median heights after discarding 25% of the trees obtained from the BEAST analysis were input into RASP. The 1,000 random trees were used for the biogeographical reconstruction; Sanmartín et al. 2001). BBM was implemented in RASP v3.0 (Yu et al. 2014) using a maximum of two areas per node. The 10,000 trees obtained from the BEAST analysis were input into RASP. The 1,000 random trees were used for the biogeographical reconstruction (BBM). The MCC tree produced in the BEAST analysis was used as the input tree. Based on the dated chronogram for P. astrigera, we pruned the tree to remove outgroups. We divided the range into the same six biogeographical areas as above.

Results

Sequences

The 96, 81, and 76 sequences were obtained for the COI, 16S, and NADH1 genes from all samples of P. astrigera, respectively. The aligned data of the sequences from P. astrigera had lengths of 932 base pairs (bp) for COI, 574 bp for 16S, and 584 bp for NADH1. In total, 62, 17, and 38 haplotypes were found for the COI, 16S, and NADH1 gene sequences, respectively. Among the 62 unique COI haplotypes, Hap20 dominated in 8.3% of the samples, Hap39 in 6.3%, and each of Hap3, Hap15, and Hap52 in 4.2%; among the 17 unique 16S haplotypes, Hap4 dominated in 55.6% of the samples, Hap15 in 14.8%, and Hap9 in 7.4%; and among the 38 unique NADH1 haplotypes, Hap13 dominated in 36.8% of the samples, Hap22 in 6.6%, and Hap1 in 5.3%. The concatenated mitochondrial genes comprised 2,090 bp. All sequences are deposited in GenBank (for accession numbers, see Suppl. material 1: table S1).
Molecular phylogeny

In the dated phylogenetic tree, *P. astrigera* is composed of the two lineages (1–2; Fig. 1b). Lineage 1 consists of the sample (CF5) from Neimonggu, China, the four samples (LY1–3 and LY5) from Henan, China, and the sample (MS3) from Heilongjiang province of China, and the lineage spiders occur on a clade of the ML tree (Suppl. material 1: fig. S1b); and Lineage 2 consists of all other individuals of *P. astrigera* across continental East Asia, and the lineage samples cluster together in a clade of the BI tree (Suppl. material 1: fig. S1a). The BI and ML trees (Suppl. material 1: fig. S1) indicated five sister groups (2–a). The samples in Group 2a are from Neimenggu, Shanxi, and Zhejiang provinces of China and South Korea; Group 2b from Heilongjiang and Zhejiang provinces of China and Japan; Group 2c from Shandong, Gansu, and Hubei provinces of China; Group 2d from Shanxi and Zhejiang provinces of China; and Group 2e from Neimenggu, Hebei, Shandong, and Shanxi provinces of China and South Korea. However, the relevance of these groups is not clear in the molecular phylogeny.

Haplotype network

The haplotype network analyses based on both 16S and *NADH1* genes revealed clear genetic structuring among the East Asian populations of *P. astrigera* (Fig. 2a, b); in contrast, those based on the *COI* gene showed a complex and intricate haplotype connection pattern (Fig. 2c). In 16S and *NADH1* networks, all studied individuals clustered into two lineages (1–2; Fig. 2a, b). Lineage 1 comprises three haplotypes from Neimenggu and Henan regions in China; Lineage 2 is composed of the 35 and 14 haplotypes from all analyzed regions, with one (Hap13) and two (Hap4 and Hap15) universal haplotypes occurring at the center of a star-like cluster of rare haplotypes in the *NADH1* and 16S networks, respectively (Fig. 2a, b). The *COI* network structure consists of 62 haplotypes with Hap20 at the center, indicating that the wolf spider has complex population genetic relationships and a high flow among populations (Fig. 2c).

Genetic diversity

Northeast Asia and North China populations exhibited higher genetic diversity than the other four regional populations based on *COI* and *NADH1* genes (Table 1). The genetic diversity of the East China population was higher than that of the Central China population based on the two protein-coding genes (Table 1). Among all analyzed populations, the population of Southwest China had the lowest genetic diversity based on *COI* and 16S genes, whereas Northwest China had the lowest genetic diversity based on the *NADH1* gene (Table 1).

Haplotype divergence times and population expansion time

Based on fossil calibrations (Suppl. material 1: table S5), the estimated initial haplotype divergence of *P. astrigera* started at ~1.69 Ma (95% credibility interval, CI: 0.95–2.66 Ma; Fig. 1b). Diversification of the five haplotypes (Lineage 1) from Heilongjiang, Neimenggu, and Henan provinces of China occurred ~0.98 Ma (95% CI: 0.50–1.44 Ma; Fig. 1b). The two haplotypes from Jiangsu and Zhejiang provinces of China diverged from the other haplotypes of Lineage 2 at ~1.21 Ma (95% CI: 0.62–1.95 Ma; Fig. 1b). The mean substitution rate (*V*) of the concatenated mtDNA sequences for the species was estimated to be 0.0129 (95% CI: 0.0092–0.0163) per million years, equivalent to a divergence rate of 2.58% (1.84%–3.26%) per million years. The τ value was inferred from the mitochondrial allele frequencies and linkage disequilibrium to be 2.244. Using the values of *V* and τ, the estimated generation time since expansion for the *P. astrigera* populations in East Asia was approximately 74,340 (58,832–104,236) years ago.

Table 1. Summary of genetic diversity in *P. astrigera* from East Asia based on mtDNA sequences. The number of haplotypes (nh), nucleotide diversity (ρ), haplotype diversity (h), Fu’s FS, and Tajima’s D are shown. Northeast Asia (Neimenggu+Heilongjiang, China+Korea+Japan); Northwest China (Qinghai+Gansu, China); North China (Hebei+Henan+Shanxi+Sanxi, China); Central China (Hubei+Hunan+Anhui, China); East China (Jiangsu+Shandong+Zhejiang, China); and Southwest China (Yunnan+Guizhou, China).

<table>
<thead>
<tr>
<th>Population</th>
<th>COI</th>
<th>16S</th>
<th>NADH1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ni nh</td>
<td>h</td>
<td>Fu’s FS</td>
</tr>
<tr>
<td>Northeast Asia</td>
<td>10 9</td>
<td>0.014 ± 0.978 ±</td>
<td>-3.323 -0.90274**</td>
</tr>
<tr>
<td>Asia</td>
<td>0.003</td>
<td>0.054</td>
<td></td>
</tr>
<tr>
<td>Northwest China</td>
<td>11 10</td>
<td>0.007 ±</td>
<td>0.962 ±</td>
</tr>
<tr>
<td>China</td>
<td>0.002</td>
<td>0.046</td>
<td></td>
</tr>
<tr>
<td>North China</td>
<td>27 23</td>
<td>0.01 ±</td>
<td>0.986 ±</td>
</tr>
<tr>
<td>Central China</td>
<td>13 10</td>
<td>0.009 ±</td>
<td>0.923 ±</td>
</tr>
<tr>
<td>China</td>
<td>0.002</td>
<td>0.069</td>
<td></td>
</tr>
<tr>
<td>East China</td>
<td>23 21</td>
<td>0.01 ±</td>
<td>0.992 ±</td>
</tr>
<tr>
<td>China</td>
<td>0.001</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>Southwest China</td>
<td>12 7</td>
<td>0.006 ±</td>
<td>0.879 ±</td>
</tr>
<tr>
<td>China</td>
<td>0.002</td>
<td>0.075</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>96 58</td>
<td>0.01 ±</td>
<td>0.979 ±</td>
</tr>
<tr>
<td>(East Asia)</td>
<td>0.001</td>
<td>0.006</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2. a. 16S haplotype network of the Pardosa astrigera spiders; b. Network from NADH; c. Network from COI: NHJ, Northeast Asia (Neimenggu+Heilongjiang, China+Korea+Japan); SGS, North China (Hebei+Henan+Shanxi+Sanxi, China); QHS, Northwest China (Qinghai+Gansu, China); YUG, Southwest China (Yunnan+Guizhou, China); HUA, Central China (Hubei+Hunan+Anhui, China); JSZ, East China (Jiangsu+Shandong+Zhejiang, China).
Demographic history

The neutrality test of mitochondrial genes showed all six separate populations, and the total population had negative values of Fu’s FS/Tajima’s D or significantly negative values of Tajima’s D (Table 1). These values indicate historical population size expansions, genetic hitchhiking, and/or selection (Tajima 1989; Fu 1997). The 16S and NADH1 network structures in which one or two universal haplotypes occurred at the center of a star-like cluster of rare haplotypes indicated rapid demographic expansion of the *P. astrigera* population in East Asia (Fig. 2a, b). The unimodal mismatch distribution and its strong bias toward low divergence values, as distinguished by 0- and 1-nucleotide changes, illustrated relatively recent expansion from a small number of ancestors (Rogers and Harpending 1992) or a range expansion with high levels of migration between neighboring demes (Ray et al. 2003; Excoffier 2004; Fig. 3a–c). The BSP analyses revealed that the East Asian *P. astrigera* adequate population size underwent four stages: (1) a gradual growth since ~50,000 years ago; (2) a slight fall between 8,000 and 1,800 years ago; (3) a sharp decrease from 1,800 to 600 years ago; and (4) a rapid rise after ~600 years ago (Fig. 3d).

The possible ancestral ranges and dispersal pathways of *P. astrigera* in East Asia were inferred by the BBM analyses (Fig. 4a). The most likely ancestral area of this spider was Northeast Asia, and it possibly dispersed from North China and East China to Northwest China or Southwest China via Central China (Fig. 4b). These results were further supported by the genetic diversity estimations (Table 1) because ancestral populations possess higher genetic diversity than derived populations (Savolainen et al. 2002).

Discussion

Genetic variations of *P. astrigera*

Previous studies suggested that *P. astrigera* was a species complex that showed high intraspecific and interspecific morphological variations (Schenkel 1963; Yin et al. 1997; World Spider Catalog, 2024). Our analyses of the mitochondrial COI, 16S, and NADH1 data also found the largest genetic distance (p-distance) among the *P. astrigera* haplotypes was 0.033 for COI, 0.013 for 16S, and 0.030 for NADH1, and the largest overall mean distance (p-distance) among the six populations was 0.014 for COI, 0.005 for 16S, and 0.010 for NADH1 (Suppl. material 1: table S6). Our DNA data indicated that the wolf spider has a rich genetic diversity, and the six divided geographical populations showed obvious genetic variations (Table 1). Moreover, in the study of Chang et al. (2007),

![Figure 3. Mismatch distributions of *Pardosa astrigera* from the whole sample based on the COI sequences (a) and based on the 16S sequences (b) and the NADH1 sequences (c) independently. Bayesian skyline plots from the whole sample (d); middle lines represent median estimates of the effective population size, and shaded areas represent 95% of the highest posterior densities (95% HPD). The effective population size is presented on a logarithmic scale. The LGM represents the last glacial maximum.](image-url)
two phenotypes (types A and B distinguished by the shape and length of the median apophysis in palpus) of the *P. astrigera* males were recognized in China, and they differed genetically in the presence (type A corresponding to Lineage 1 of this study) or absence (type B corresponding to Lineage 2) of common insertions and deletions in the nuclear ITS2 gene. Therefore, our mtDNA data also supported the two differentiated phenotypes.

The evolutionary potential of species under environmental stress depends on levels of genetic diversity (Black et al. 1992; Gurdebeke et al. 2003). High genetic diversity is an important factor in adaptation to environmental changes. The wolf spider, *P. astrigera*, is widespread in East Asia. As a major natural predator of pests in various semi-arid land ecosystems such as farmland, forests, vegetable fields, tea gardens, and fruit orchards,
it has a wide distribution, a large population size, strong adaptability, and a complex mechanism for protecting offspring (spiderlings clustering on the abdomen and back of mothers after hatching and dispersing after one molt). Therefore, this wolf spider seems to be highly resistant to environmental changes and has the genetic basis to become a dominant and widespread species.

Haplotype relationships and mitochondrial genes

The complex haplotype relationships among populations of *P. astrigera* and the high frequency of gene flow within regions were revealed by the *COI* haplotype network analyses. However, the inferred 16S and NADH1 networks show clear haplotype relationships and a lower degree of divergence than that obtained using *COI*. These findings can be explained by the importance of *COI* in cells and strong natural selection (Arnold 2012). The haplotype network structure revealed by the mitochondrial 16S or NADH1 analyses is consistent with that by nuclear ITS2 analyses (Chang et al. 2007), and we found a similar mutation rate of 16S to ITS2 in *P. astrigera*.

Both haplotype networks inferred from 16S and NADH1 are composed of two lineages (1–2). Combined with the results of the biogeographical reconstruction (BBM), we speculated that the Lineage 1 spiders were the ancestors of *P. astrigera*. The main haplotypes (Hap 4 and Hap 15 for 16S; Hap 13 for NADH1) at the center of Lineage 2 of each haplotype network were identified in and Hap 15 for 16S; Hap 13 for *P. astrigera*. The main haplotypes (Hap 4 and Hap 15 of 16S) of the ancestors of *P. astrigera* in continental East Asia. Therefore, human movements and trades can change the geographic distribution range of *P. astrigera* and accelerate its dispersal across the East Asian region. Additionally, ballooning is likely another driver for the frequent and/or long-distant dispersal of *P. astrigera* across East Asia. Dispersal by ballooning on silken threads is a well-known behavior in wolf spiders (Matthew 1982; Postiglioni et al. 2017).

Analyses of the mtDNA population genetics revealed that *P. astrigera* may have undergone a recent demographic expansion in East Asia. Our estimation of the population expansion time (~74,340 (58,832–104,236) years ago) of the spider based on fossil calibrations is during the last glaciation. The time is earlier than that (28,000 years ago) inferred by Chang et al. (2007) using a calibration assuming a rate of 2.3% sequence divergence per million years for arthropod mtDNA. Our estimated divergence rate (2.58%) was comparable to those estimated by Papadopoulos et al. (2010), who, through an extensive survey of tenebrionid beetles, obtained a divergence rate of 2.69% per million years for their concatenated *COI* and 16S sequences.

Lycosidae spiders occur globally, with many species reported to have widespread distributions and inhabit various ecosystems (Song et al. 1999; World Spider Catalog, 2024). The wolf spider *P. astrigera* lives mainly in the East Asian regions (Song et al. 1999; World Spider Catalog 2024; Fig. 1a; Suppl. material 1: table S1). Our analyses provide novel insights into the diversification of *P. astrigera* in continental East Asia during the Cenozoic. This spider is thought to originate most likely in Northeast Asia around the Pliocene and possibly dispersed in the major directions from North to South and from East to West. Our findings indicated that the pattern of genetic diversity within populations seems predominantly determined by historical factors but is modified by contemporary aspects.

Phylogeographical pattern and historical demography

Our mtDNA data support that *P. astrigera* comprises two sympatric lineages (1–2). Lineage 1 has a limited geographic distribution north of the Qinling Mountains (Mts) of China, including the sample (CF5) from Neimonggu province of China, the four samples (LY1-3 and LYS) from Henan of China, and the sample (MS3) from Heilongjiang of China; in contrast, Lineage 2 is present across the entire ranges of the species in continental East Asia. This population structure pattern supports the hypothesis that intraspecific phylogeography involves common lineages that are widespread, plus related lineages that are confined to one or a few nearby locales (Avisé 2000). The two lineages of *P. astrigera* early split ~1.69 (0.95–2.66) Ma, diversified since the late Pleistocene in East Asia, and exist sympatrically at multiple locations. The results suggest that the Pleistocene glaciation is probably a major driving factor in its genetic differentiation. One haplotype (Hap 13) of NADH1 and two (Hap 4 and Hap 15 of 16S) occurred in all six sampled populations across continental East Asia, indicating recent frequent dispersal of the wolf spider, probably with human activities. *Pardosa astrigera* is a dominant species and has strong environmental adaptability in various open ecosystems such as farmland, grassland, urban regions, and hills with human activity. Moreover, continental East Asia has a dense population. Therefore, human movements and trades can change the geographic distribution range of *P. astrigera* and accelerate its dispersal across the East Asian region. Additionally, ballooning is likely another driver for the frequent and/or long-distant dispersal of *P. astrigera* across East Asia.

Impacts of the Pleistocene glaciation on the evolution of *P. astrigera*

Many species experience retreats, expansions, and diversification in response to cold-warm climatic oscillations during the Pleistocene glacial periods (Hewitt 2000, 2004; Ding et al. 2011). They often retreated to their southern range or refugia during the cold glaciation periods and expanded their distribution ranges during the warm interglacial periods, and thus their glacial refugia are often located at the south and postglacial expansions are often northward (Fu and Wen 2023). These
evolutionary patterns resulted in a genetic diversity pattern of “southern richness to northern purity” (Hewitt 2000). That was, however, not supported by the present study, which found the Northeast Asian \textit{P. astrigera} population had the highest genetic diversity among the six divided geographic populations from continental East Asia. Moreover, the ancestral area reconstruction of this species suggests its ancestors inhabited Northeast Asia and subsequently dispersed southwestward. Therefore, the Northeast Asian regions were probably the glacial refugia of \textit{P. astrigera}. The refugia affected the phylogeography and genetic structure of the species. That is similar to the results of González-Trujillo et al. (2016), who studied the Pleistocene refugia of \textit{Pardosa sierra} on the Baja California Peninsula. On continental East Asia, the Pleistocene glacial-interglacial climatic cycle plays a crucial role in some species (Fu and Wen 2023). Pleistocene climatic fluctuations have also been shown to influence the distribution patterns of European and North American lycosids (Muster and Berendonk 2006; Ivanov et al. 2023). These results and our findings suggested that similar interglacial dynamics shaped the modern boreal fauna of the wolf spiders in the Holarctic region, and secondary contacts were likely to be common among the allopatric cold-adapted species.

The spiders of \textit{P. astrigera} started during the mid-Pleistocene, resulting in Lineage 1 and Lineage 2, and subsequently both the two lineages rapidly diversified. Therefore, the mid-Pleistocene glaciation is an important driver for the diversification of this spider. The population expansion of \textit{P. astrigera} began during the last glaciation. We speculated that the co-existence of Lineage 1 and Lineage 2 occurred before the last glacial epoch. The last glaciation triggered the expansion of the Lineage 2 populations into the south of the Qinling Ms, whereas the Lineage 1 populations were likely blocked by the Qinling Ms. Thus, multiple glaciations during the Pleistocene affected the arid-cold distribution and diversification of \textit{P. astrigera} in East Asia.

Future directions

This study has allowed understanding of the matrilineage structure of the wolf spider \textit{P. astrigera} and generated hypotheses regarding its origin and dispersal using the mitochondrial \textit{COI}, 16S, and \textit{NADH1} loci. However, while mtDNA data have proven highly useful in phylogeographic analyses, there are several shortcomings to this approach (e.g., Avise 2000). There is a potential mismatch between the dispersal histories of males and females and a likely mismatch between a species tree and any single gene tree. In this paper, we summarize available \textit{COI}, 16S, and \textit{NADH1} data, add new data analyses of patterns across East Asia, and discuss hypotheses on the dispersal history of \textit{P. astrigera}. Robustly testing the historical expansion hypotheses and gaining a deeper understanding of the population structure of this spider will require DNA data from the whole genome, using methods such as RAD-seq. Such data will tease apart male and female dispersal, testing the “female-based” dispersal hypotheses proposed here. They will also allow testing of the “independent lineages” hypothesis through precise measures of genetic isolation and gene flow unavailable to several marker studies. Sampling is the process of choosing a subset of a target population that will serve as its representative. In this paper, the total sampling specimens for some regions are on the low side. However, we hope our study can aid in strategic resampling, reflecting known lineage divergences.

Acknowledgements

The manuscript benefited greatly from comments by Danilo Harms (Hamburg, Germany), Yanfeng Tong (Shen-yang, China) and one anonymous referees. We thank Li Tan, Xuanxuan Chi, Jun Wang, and Zhaojun Bao (Shaoxing, China) for their help in collecting the material and laboratory work. This study was supported by the National Natural Sciences Foundation of China (NSFC-32170463, 31860602, 31660611) and the Zhejiang Provincial Natural Science Foundation of China (LTGN24C140006).

References


### Supplementary material 1

### Supplementary materials and datasets

Authors: Dan Fu, Liujuan Liu, Ying Cheng, Haodong Chen, Yufa Luo

Data type: zip

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zse.100.125246.suppl1