Molecular Phylogenetics and Evolution 127 (2018) 146–155

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Early Cretaceous greenhouse pumped higher taxa diversification in spiders

Lili Shao[a,b], Shuqiang Li[a,*]

[a] Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China
[b] University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Keywords:
Araneae
Phylogenomics
Fossil spider
Diversification
Angiosperm

ABSTRACT

The Cretaceous experienced one of the most remarkable greenhouse periods in geological history. During this time, ecosystem reorganization significantly impacted the diversification of many groups of organisms. The rise of angiosperms marked a major biome turnover. Notwithstanding, relatively little remains known about how the Cretaceous global ecosystem impacted the evolution of spiders, which constitute one of the most abundant groups of predators. Herein, we evaluate the transcriptomes of 91 taxa representing more than half of the spider families. We add 23 newly sequenced taxa to the existing database to obtain a robust phylogenomic assessment. Phylogenetic reconstructions using different datasets and methods obtain novel placements of some groups, especially in the Synspermiata and the group having a retrolateral tibial apophysis (RTA). Molecular analyses indicate an expansion of the RTA clade at the Early Cretaceous with a hunting predatory strategy shift. Fossil analyses show a 7-fold increase of diversification rate at the same period, but this likely owes to the first occurrence of spiders in amber deposit. Additional analyses of fossil abundance show an accumulation of spider lineages in the Early Cretaceous. We speculate that the establishment of a warm greenhouse climate pumped the diversification of spiders, in particular among webless forms tracking the abundance of insect prey. Our study offers a new pathway for future investigations of spider phylogeny and diversification.

1. Introduction

The Cretaceous period (145–66 Million years ago (Ma)) offers one of the best examples of a greenhouse climate on Earth because high average global temperature lasted until about 70 Ma (Wang et al., 2014). The progressive breakup of Pangaea as a result of intense tectonic activity led to a near total reorganization of global ecosystem including the extinction of dominant groups and subsequent diversification of novel taxa (Coiffard et al., 2012; Lloyd et al., 2008). Of particular interest, the Cretaceous Terrestrial Revolution (KTR; also referred to as the angiosperm revolution) saw an explosive radiation of angiosperms and the replacement of ferns and gymnosperms between 125 and 80 Ma (Lloyd et al., 2008).

Angiosperms were the fundamental players in terrestrial evolution. As primary producers, their domination provided new ecological and evolutionary opportunities for insects such as herbivorous beetles (Zhang et al., 2018), lepidopterans (Wahlberg et al., 2013), and ants (Moreau et al., 2006). Among vertebrates, crown squamates (lizards and snakes) appeared in the Jurassic and underwent a Cretaceous radiation (Bronzati et al., 2015). Mammals and birds experienced a strong increase in diversification, mainly after the end-Cretaceous crisis. Nevertheless, the KTR increased ecospace diversity, possibly precipitated mammalian interordinal diversification (O'Leary et al., 2013), and may have provided opportunities for modern birds to diversify into new adaptive-zones (Prum et al., 2015). Therefore, change in the Cretaceous ecosystem appears to have strongly influenced spatial and temporal population dynamics.

Spiders (Order Araneae) are an important element of the food chain. The group consists of more than 46,000 described species (World Spider Catalog 2017) that occur almost everywhere from Arctic islands to dry desert regions. Their evolution has witnessed mass extinctions, radical alterations in terrestrial floras, continental rearrangements, and changes in key environmental parameters. Throughout this time, spiders have developed a suite of morphological characteristics and remarkably diverse modes of life (Selden and Penney, 2010). In particular, the evolution of the orb web has been long considered a “key innovation” in spiders (Bond and Opell, 1998). The ancient origin of the orb web and web change or loss associates with spider diversification (Blackledge et al., 2009; Dimitrov et al., 2016; Garrison et al., 2016). Nevertheless, it remains unclear how environmental changes, especially during the Cretaceous, influenced spider diversification.

Molecular phylogenetic analyses have mostly driven our understanding of the historical pattern of spider diversification. Multi-family analyses have incrementally covered parts of the spider tree of life and

* Corresponding author at: 1 Beichen West Road, Chaoyang District, Beijing 100101, China.
E-mail address: lisq@ioz.ac.cn (S. Li).

https://doi.org/10.1016/j.ympev.2018.05.026
Received 26 July 2017; Received in revised form 14 May 2018; Accepted 19 May 2018
Available online 24 May 2018
1055-7903/

resolved the evolution of some groups such as Mygalomorphae (Bond et al., 2012), Palpimanoidea (Wood et al., 2012), orb-weavers (Blackledge et al., 2009; Dimitrov et al., 2016), sparassids (Moradmand et al., 2014), and nearly all spider groups (Wheeler et al., 2016). Phylogenomic studies rejected the prevailing paradigm for orb web evolution and corroborated several unexpected results, such as the position of Leptonetidae (Bond et al., 2014; Fernández et al., 2014; Garrison et al., 2016). Still, much of the phylogenetic framework for spiders remains uncertain. Molecular phylogenies cannot directly place extinct lineages, and yet the fossil record is our primary window onto the diversification of ancient life. Fossil spiders supply a handful of calibration points in molecular studies and only a few quantitative analyses of fossils exist (Selden and Penney, 2010). Thus, phylogenies and the fossil record provide two complementary perspectives that can resolve temporal variation in biodiversity.

Herein, we reconstruct the phylogenetic history of spiders using previously acquired and de novo transcriptomes. We perform divergence dating and diversification rate shift analyses to hypothesize the tempo and mode of diversification at the level of family. We re-examine the global fossil record to assess relative abundance of different groups. Finally, we reconstruct the ancestral predatory strategies among different groups and suggest some of the possible drivers of spider diversification during the Early Cretaceous greenhouse.

2. Materials and methods

2.1. Sampling, extraction, and assembly

New transcriptomic data were generated for 23 specimens representing major groups of spiders. Previously available transcriptomic and genomic data were collected for 68 spider samples plus 14 outgroup taxa (Tables S1 and S2). All specimens used in this study were legally collected. Samples were flash frozen in liquid nitrogen and mRNA was extracted using the TRIzol total RNA extraction method (Life Technologies). Purification of mRNA, library preparation, sequencing, and quality control were done by Novogene Bioinformatics Technology Co. Ltd. Sequencing libraries were generated using NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) following manufacturer's recommendations and index codes were added to attribute sequences to each sample. Briefly, mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. Fragmentation was carried out using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer (5X). Then first and second strand cDNA were successively synthesized. For detailed methods of library preparation and sequencing refer to the supplementary methods. Raw data (raw reads) in the fastq format were processed by removing reads containing adapter, reads containing ploy-N, and low quality. Next, clean reads were assembled using default parameters in Trinity v2.0.5 (Grabherr et al., 2011). Previously available sequence data were acquired from the NCBI database, quality-checked, and trimmed using the FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and FASTXToolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html) before assembly with Trinity.

2.2. Orthology determination and data filtering

Putative orthologs were determined for each species with HaMStR v13.2.3 (Ebersberger et al., 2009) using the Arthropoda core ortholog set. The resulting orthologous groups (OGs) were further processed by filtering them with fewer than half of all sample species while discarding amino acid sequences with length shorter than 75 sites. The remaining OGs were then aligned with MAFFT v7.222 (Katoh and Standley, 2016) and processed with Aliscore (Kück et al., 2010) and ALICUT (http://zfmk.de/web/ZFMK_Mitarbeiter/KckPatrick/Software/AliCUT/Download/index.en.html). Missing data, evolutionary rate, and compositional heterogeneity have been shown to compromise phylogenomic analysis. Therefore, we regrouped the initial 979 OGs (dataset A). Each OG was ranked by evolutionary rate following Telford et al. (2014) as follows: using an alignment for each gene, a maximum-likelihood (ML) tree was calculated using RAxML v8.0.0 (Stamatakis, 2014) and the total length of that tree (in estimated substitutions per position across all branches) was divided by the number of taxa on the tree to give an estimate of the rate of evolution for each gene. Half of the slowest-evolving genes were concatenated to form dataset B. Each OG was also ranked by missing data ratio: half of the OGs with lowest missing data ratio (lower than 0.1535) were then concatenated to form dataset C. BaCoCa (Kück and Struck, 2014) and two metrics (the χ^2-test of heterogeneity and relative composition frequency variability; RCFV) (Zhong et al., 2011) were used to identify the OGs with amino acid compositional heterogeneity. OGs with p-values below 0.01 were removed because the composition significantly deviated from homogeneity. Then, 490 OGs with the lower RCFV value (lower than 0.1053) formed dataset D. The full dataset was also optimized using MARE (Meyer et al., 2011) to find highly informative subsets of OGs (dataset E). Detailed information for all the five datasets was summarized in Table S3.

2.3. Phylogenetics analyses

OGs for each dataset were concatenated using FASconCAT v1.0 after alignment (Kück and Meusemann, 2010). A meta-partition analysis using the PartitionFinderProtein v1.1.1 (Lanfear et al., 2012) was conducted for all five datasets to select the optimal meta-partitions and the best-fit amino acid substitution models. Maximum-likelihood topologies were inferred with RAxML v8.0.0 (Stamatakis, 2014) using partitions as indicated by ParitionFinderProtein, associated best-fit substitution models, and the GAMMA parameter to model rate-heterogeneity. Nodal support was measured with 1,000 fast bootstrap pseudoreplicates. Relatively small datasets B, C and D were also analyzed under the site-heterogeneous mixture model, CAT-GTR + GAMMA, in PhyloBayesMPI v1.5a (Lartillot et al., 2013) using the resources available from the CIPRES Science Gateway (Miller et al., 2010). A coalescent-based method as implemented in ASTRAL (Accurate Species Tree Reconstruction ALgorithm; astral v4.10.2; Mirarab et al., 2014) was used to infer a species tree from a series of unrooted gene networks.

2.4. Divergence time estimation

Selected meta-partitions from dataset B were used to estimate the time tree for spiders with each partition having at least partial non-ambiguous sequences for all 105 taxa and a minimum amino acid alignment length of 500 sites. Fourteen calibration points (5 for outgroup calibration nodes and 9 for ingroup calibration nodes) were chosen conservatively to calibrate the inferred phylogeny.

Although discussion persists on the nature of the Ediacaran fauna (including putative arthropods; Erwin and Valentine, 2010), 580 Ma constituted a justifiable maximum hard bound for the origin of the major arthropod clades, which we used as the maximum rootHeight age. To constrain the node of Copepoda + Branchiopoda, we used a secondary calibration: a normal distribution with mean 390 Ma and 95% confidence interval of 513–267 Ma (Misof et al., 2014). The oldest known fossil insect, *Rhyniella praecursor*, was used to constrain the minimum age of 411.5 Ma for the node consisting of Protura + Collembola (Scourfield, 1940). The fossil *Triassolestodes asiaticus* represented the order crown Odonata, and we set the crown Odonata node with a minimum age of 237 Ma (Kohli et al., 2016). The earliest true insect fossil, *Rhyniognatha hirsti*, was used to constrain the minimum age of Dicondylia at 411.5 Ma (Engel and Grimaldi, 2004). Although the unequivocal land plant megafossils date back into the mid-Silurian (about 425 Ma) (e.g., Hart et al., 2013), the earliest