Gene 590 (2016) 298–306

Contents lists available at ScienceDirect

Gene

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

Research paper

Characterization of the complete mitogenomes of two *Neoscona* spiders (Araneae: Araneidae) and its phylogenetic implications

Zheng-Liang Wang, Chao Li, Wen-Yuan Fang, Xiao-Ping Yu *

Zhejiang Provincial Key Laboratory of Biometrology and Inspection and Quarantine, College of Life Sciences, China Jiliang University, Hangzhou, Zhejiang 310018, People's Republic of China

ARTICLE INFO

Article history:
Received 18 February 2016
Received in revised form 15 April 2016
Accepted 25 May 2016
Available online 31 May 2016

Keywords:
Mitogenome
Araneae
Araneidae
Neoscona doenitzi
Neoscona nautica
Phylogeny

ABSTRACT

The complete mitogenomes of two orb-weaving spiders *Neoscona doenitzi* and *Neoscona nautica* were determined and a comparative mitogenomic analysis was performed to depict evolutionary trends of spider mitogenomes. The circular mitogenomes are 14,161 bp with A+T content of 74.6% in *N. doenitzi* and 14,049 bp with A+T content of 78.8% in *N. nautica*, respectively. Both mitogenomes contain a standard set of 37 genes typically presented in metazoans. Gene content and orientation are identical to all previously sequenced spider mitogenomes, while gene order is rearranged by tRNAs translocation when compared with the putative ancestral gene arrangement pattern presented by *Limulus polyphemus*. A comparative mitogenomic analysis reveals that the nucleotide composition bias is obviously divergent between spiders in suborder Opisthothelae and Mesothelae. The loss of D-arm in the $trnS^{UCN}$ among all of Opisthothelae spiders highly suggested that this common feature is a synapomorphy for entire suborder Opisthothelae. Moreover, the $trnS^{AGN}$ in araneoids preferred to use TCT as an anticodon rather than the typical anticodon GCT. Phylogenetic analysis based on the 13 protein-coding gene sequences consistently yields trees that nest the two *Neoscona* spiders within Araneidae and recover superfamily Araneoidea as a monophyletic group. The molecular information acquired from the results of this study should be very useful for future research on mitogenomic evolution and genetic diversities in spiders.

1. Introduction

The typical orb weavers, an informal name for the taxa from family Araneidae, are one of the most diverse groups of spiders that have the ability to weave wheel-shaped orb webs composed of a variety of silks (Hormiga and Griswold, 2014). With more than 3000 described species in 169 genera, Araneids represent the third-largest species-rich lineage of spiders next to Salticidae and Linyphiidae (Platnick, 2015). As generalist predators of arthropods, spiders belong to this family have important roles on maintaining most terrestrial ecosystems, and hence served extensively as model organisms for the ecological, behavioral and evolutionary research in arachnology (Nyffeler and Knörnschild, 2013; Eggs and Sanders, 2013). However, current knowledge of the evolutionary relationships among araneids is still limited, as most of the results were based on assessment of morphological similarity or phylogenetic analysis of a small number of taxa. Although some well-supported lineages (subfamily Araneinae and the 'argiopoid' clade) were identified, relationships within the family are still poorly understood (Scharff and Coddington, 1997). In addition, the interfamilial relationships between Araneidae and other orb-weaving spiders in superfamily Araneoidea remain to be satisfactorily resolved (Hormiga and Griswold, 2014).

More recently, molecular sequence data are used to enhance the current understanding of araneid relationships. For example, using five molecular markers (one mitochondrial gene *16S rRNA*, and four nuclear genes *18S rDNA*, *28S rDNA*, *his3* and *wingless*) along with a comprehensive taxon sample, Dimitrov et al. (2012) found that Araneidae was strongly supported as a monophyly. However, none of the molecular analysis of Blackledge et al. (2009) recovered the monophyly of araneids by using the same markers. It is a remarkable fact that all of the molecular markers used above are relatively short in length and the nucleotide information provided is undoubtedly extremely limited. Alternatively, the complete mitogenome sequences, ranging in size from 14 kb to 20 kb, have far been used as genetic marker at genomic level in the phylogenetic study of arthropods. Despite of an ongoing debate concerning their utility in phylogenetics (Cameron et al., 2004; Talavera and Vila, 2011), mitogenomes have proven to be more superior and insightful to reconstruct phylogenetic relationships, as they possess a number of potentially informative features such as length variation, altered transfer RNA (tRNA) anticodons or secondary structures,

Abbreviations: PCGs, protein-coding genes; COI–III, cytochrome *c* oxidase subunits I–III; ND1–6 and 4 L, NADH dehydrogenase subunits 1–6 and 4 L; ATP6 and ATP8, ATPase subunits 6 and 8; Cytb, cytochrome *b*; tRNA, transfer RNA; rRNA, ribosomal RNA; 16S rRNA and 12S rRNA, large and small subunits ribosomal RNA; L1, $trnL^{UUR}$; L2, $trnL^{CUN}$; S1, $trnS^{AGN}$; S2, $trnS^{UCN}$; DHU arm, dihydrouridine arm; ML, Maximum likelihood; BI, Bayesian Inference.

* Corresponding author.
E-mail address: yuxiaoping19630306@163.com (X.-P. Yu).

http://dx.doi.org/10.1016/j.gene.2016.05.037
0378-1119/

nucleotide composition bias, gene rearrangement and models of control of replication and transcription (Boore et al., 1998; Dowton et al., 2002; Castresana et al., 1998; Masta, 2000; Macey et al., 2000).

Arthropod mitogenome usually contains a standard set of 13 protein-coding genes (PCGs), 2 ribosomal RNAs (the large and small ribosomal subunits, rRNAs), 22 tRNAs and a control region (D-loop or A + T-rich region) with highly variable length (Boore, 1999; Taanman, 1999). Typically, tRNAs in metazoan mitogenomes fold into a highly conserved cloverleaf secondary structure, with an aminoacyl stem (AA-arm), a dihydrouracil arm (DHU or D-arm), an anticodon stem, a TΨC arm (T-arm) and a smaller variable extra loop. Most metazoan mitochondrial genes for $trnS^{AGN}$ lack the sequence to form the *D*-arm, which significantly suggested that this feature is occurred before the divergence of metazoans (Garey and Wolstenholme, 1989). Additionally, many severely truncated or atypical tRNA structures, such as missing either D-arm or T-arm, or unpaired AA-arm, are observed in many arthropod mitogenomes, especially found in those of nematodes and Opisthothelae spiders (Montiel et al., 2006; Masta and Boore, 2004, 2008). As the evolutionary loss one of the paired stem regions is a rare event, organisms that possessed the similar non-canonial tRNA structures might be share the evolutionary history and have a relative closer lineage. For example, a comparative analysis of the sea spider mitogenomes revealed that rare structural changes in three tRNAs (*trnA*, *trnK* and $trnS^{AGN}$) unite pycnogonids as a clade (Masta et al., 2010).

The complete mitogenome sequences have been determined for more than 1100 species of arthropods, including 118 species of arachnids. Within Araneae, only 24 complete mitogenomes (7 belong to araneids) are available in GenBank, which is impeding our ability to understand of phylogenetic relationships among spiders. Therefore, the complete mitogenomes from additional araneids are required to improve the accuracy of phylogenetic inference of Araneidae in Araneae. In the present study, we sequenced the mitogenomes of two typical orb-weaving spiders *Neoscona doenitzi* and *N. nautica* (Araneae: Araneidae), which often found in fields and gardens. New characteristics of truncated tRNA genes and novel gene order rearrangement of the two mitogenomes were analyzed. Additionally, phylogenetic studies and comparative mitogenomic analysis with other published spider mitogenomes were also performed to provide the molecular basis for understanding diversification and phylogenetic relationship of Araneae.

2. Materials and methods

2.1. Sample collection and DNA extraction

Specimens of *N. doenitzi* and *N. nautica* were collected from the road side vegetation by sweeping in Yuyao county (E121°09′, N30°03′), Zhejiang Province, China. The place is one of field stations for our studying biodiversity and pest biocontrol, where there are no endangered or protected species, and hence no specific permits were required for our collecting. All collections were preserved in 95% ethanol and maintained at 4 °C. Total genomic DNA was extracted from the legs of one adult for each specimen using the DNeasy Tissue Kit (QIAGEN Co., Germany) according to the manufactures protocol.

2.2. PCR amplification, cloning and sequencing

The complete mitogenomes of *N. doenitzi* and *N. nautica* were generated by amplification of overlapping fragments using six pairs of long PCR primers (Table S1), which were designed based on the consensus nucleotides sequences of the complete mitogenome of five spider species (*Argiope amoena*: NC_024282, *Nephila clavata*: NC_008063, *Telamonia vlijmi*: NC_024287, *Habronattus oregnonesis*: NC_005942 and *Araneus ventricosus*: NC_025634).

All long PCRs were performed in a total volume of 25 μL containing 2 μL DNA template, 2.5 μL 10 × LA PCR buffer (Mg^{2+}-plus), 2 μL dNTPs (10 mM), 1 μL each primer (10 mM), 0.25 μL LA Taq polymerase (Takara Co., Dalian, China) and 16.25 μL sterile distilled H_2O. Amplification conditions consist of an initial denaturation for 5 min at 95 °C, followed by 40 cycles of denaturation for 10s at 92 °C, annealing for 30s at 50 °C, elongation for 4 min at 68 °C in the initial 20 cycles and increase by 20s per cycle in the final 20 cycles and a final extension step for 10 min at 72 °C. All PCR products were isolated by electrophoresis in 1.0% agarose gel, purified with QIAquick Gel Extraction Kit (Qiagen, Germany), ligated to the pUC19 vector (Promega, USA) and subjected to pair end sequencing at Invitrogen (Shanghai, China).

2.3. Sequencing assembling, annotation and analysis

The overlapping PCR product sequences were checked and assembled using BioEdit 7.0.5.3 (Hall, 1999). Protein-coding genes (PCGs) and ribosomal RNA genes (rRNAs) were identified by BLAST searches in GenBank and subsequently by alignment with genes of other spider species. The putative tRNAs were either detected in combined approach using tRNAscan-SE 1.2.1 (Lowe and Eddy, 1997) and ARWEN 1.2.3.c (Laslett and Canback, 2008), or by extensive inspection of intergenic regions by eye. The annotated complete mitogenome sequences of *N. doenitzi* and *N. nautica* are deposited in the GenBank database under accession number KR259805 and KR259804, respectively. MEGA 4.0 was used to calculate the nucleotide composition and codon usage (Kumar et al., 2008). Potential inverted repeats or palindromes in the control region were determined using Tandem Repeats Finder (http://tandem.bu.edu/trf/trf.html). AT-skew and GC-skew were calculated according to the following formulae by Perna and Kocher (1995): AT-skew = (A − T)/(A + T) and GC-skew = (G − C)/(G + C).

2.4. Phylogenetic analysis

Two analytical methods, Maximum likelihood (ML) and Bayesian Inference (BI), were used to reconstruct a mitogenome phylogeny of Araneae. A total of 24 spider mitogenomes available in GenBank were included in our phylogenetic analysis, with the horseshoe crab *Limulus polyphemus* (NC_003057) used as an outgroup. GenBank accession numbers are shown in Table 1.

All the 13 PCGs of each selected species were manually extracted from the mitogenome. Nucleotide sequences were translated to amino acid sequences prior to the multiple alignment and translated back afterwards by MEGA 4.0. Ambiguously aligned regions were omitted by Gblocks version 0.91b using the default settings (Talavera and Castresana, 2007), except for changing 'allowed gap position' to 'with half'. The resultant alignments concatenated as a combined matrix with 10,821 sites in length. As the data partitioning schemes can influence the accuracy of phylogenetic reconstruction in mitochondrial phylogenomics, PartitionFinder v1.1.1 was used to select the optimal partition strategy and models (Lanfear et al., 2012). We created an input configuration file that contained a total of 39 partitions, corresponding to individual codon position of each PCG (13 1st codon positions, 13 2nd codon positions and 13 3rd codon positions). We used the "greedy" algorithm (heuristic search) with branch lengths estimated as "unlinked" and Bayesian information criterion (BIC) to search for the best-fit scheme (Table S2). Model selection was done with ModelTest 3.7 for nucleotide sequences (Posada and Crandall, 1998). ML analysis was performed by RAxML 8.0.0 (Stamatakis, 2006) through CIPRES Science Gateway (Miller et al., 2011) and the note support values were evaluated using 1000 replications of rapid bootstrapping implemented in RAxML. BI analysis was conducted by MrBayes version 3.1.2 (Ronquist and Huelsenbeck, 2003) with the following conditions: 1,000,000 generations, four chains (one cold chain and 3 hot chains) and a burn-in step for the first 1000 generations. The confidence values of BI tree were expressed as the Bayesian posterior probabilities in percentages.