Complete mitochondrial genomes and phylogenetic relationships of the genera *Nephila* and *Trichonephila* (Araneae, Araneoidea)

Hoi-Sen Yong¹, Sze-Looi Song²,³*, Kah-Ooi Chua¹, I. Wayan Suana⁴, Praphathip Eamsobhana⁵, Ji Tan⁶, Phaik-Eem Lim³ & Kok-Gan Chan¹,⁷

Spiders of the genera *Nephila* and *Trichonephila* are large orb-weaving spiders. In view of the lack of study on the mitogenome of these genera, and the conflicting systematic status, we sequenced (by next generation sequencing) and annotated the complete mitogenomes of *N. pilipes*, *T. antipodiana* and *T. vitiana* (previously *N. vitiana*) to determine their features and phylogenetic relationship. Most of the tRNAs have aberrant clover-leaf secondary structure. Based on 13 protein-coding genes (PCGs) and 15 mitochondrial genes (13 PCGs and two rRNA genes), *Nephila* and *Trichonephila* form a clade distinctly separated from the other araneid subfamilies/genera. *T. antipodiana* forms a lineage with *T. vitiana* in the subclade containing also *T. clavata*, while *N. pilipes* forms a sister clade to *Trichonephila*. The taxon *vitiana* is therefore a member of the genus *Trichonephila* and not *Nephila* as currently recognized. Studies on the mitogenomes of other *Nephila* and *Trichonephila* species and related taxa are needed to provide a potentially more robust phylogeny and systematics.

Spiders of the genus *Nephila* Leach, 1815 and genus *Trichonephila* Dahl 1911 are members of the family Nephilidae¹ or subfamily Nephilinae of Araneidae². Before the taxonomic treatment by Kuntner et al.¹, *Trichonephila* species were traditionally treated as members of the genus *Nephila*. *Nephila* and *Trichonephila* are large orb-weaving spiders, with *Trichonephila komaci*³ being the largest species ranging from some 33–40 mm in total length⁴. At different times, they have been treated as members of the family Nephilidae¹,⁴,⁵ and members of the subfamily Nephilinae within the family Araneidae²,⁶,⁷. Kuntner et al.¹ listed two species of *Nephila* and 12 species of *Trichonephila*. In contrast, the World Spider Catalog⁸ recorded 10 species of *Nephila* and 12 species of *Trichonephila*. Recently, a new species *Nephila nandiniae* has been described from Bangladesh⁹. Kuntner et al.¹,⁹ did not include the taxon *Nephila vitiana* (Walckenaer, 1847) in their studies. *N. vitiana* was treated as a valid species by Harvey et al.¹,¹² and listed as an accepted species in the World Spider Catalog, version 21.⁵ It is morphologically very similar to *Trichonephila antipodiana* (Walckenaer, 1841). Both taxa exhibit similar abdominal (opisthosomal) colour polymorphism in the adult females¹,¹². Furthermore, the juvenile spiders in both species possess very different colour patterns from the adults. However, adult female *N. vitiana* is easily distinguished from other members of the *T. antipodiana* species-group by the possession of a red-brown sternum¹⁰,¹².

*Nephila pilipes* (Fabricius, 1793) is distributed from India to China, Vietnam, Philippines, and Australia¹. *T. antipodiana* occurs in China, Philippines to New Guinea, Solomon Islands, and Australia (Queensland), whereas *N. vitiana* (*T. vitiana* in the present study) is confined to Indonesia, Fiji, and Tonga¹.

¹Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia. ²Institute for Advanced Studies, University of Malaya, 50603 Kuala Lumpur, Malaysia. ³Institute of Ocean and Earth Sciences, University of Malaya, 50603 Kuala Lumpur, Malaysia. ⁴Faculty of Science and Mathematics, Mataram University, Mataram, Indonesia. ⁵Department of Parasitology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand. ⁶Department of Agricultural and Food Science, Universiti Tunku Abdul Rahman, 31900 Kampar, Perak, Malaysia. ⁷Guangdong Provincial Key Laboratory of Marine Biology, Institute of Marine Sciences, Shantou University, Shantou 515063, China. ⁸email: szelooi@um.edu.my
Of the Nephila and Trichonephila taxa, only the complete mitochondrial genome of T. clavata (previously N. clavata) has been published\(^1\) and is available in the GenBank. There is no report on the phylogenomics of T. vitiana. In view of the lack of study on the mitogenome of the Nephilidae/Nephilinae, and the conflicting systematic status, we sequenced and annotated the complete mitogenomes of N. pilipes, T. antipodiana and T. vitiana to determine their features and phylogenetic relationship. Therefore, this study aims to elucidate the relationship of Nephila and Trichonephila species and support the taxon vitiana as a valid species of the genus Trichonephila.

Table 1. Gene order and features of mitochondrial genome of Nephila pilipes (NP), Trichonephila antipodiana (TA), Trichonephila vitiana (TV, previously N. vitiana) and Trichonephila clavata (TC, NC_008063). CR, control region; size in bp; minus sign indicates overlap.

| Gene | Size NP | Size TA | Size TV | Size TC | Intergenic sequence | Start/stop codons |
|------|---------|---------|---------|---------|---------------------|-------------------|
| trnM(cat) | 65 | 58 | 49 | 49 | −16: −8: −8: −8 | ATT/TAA: ATA/TAG: ATA/TAA: ATA/TAG |
| nad2 | 957 | 948 | 948 | 948 | −10: −2: −2: −4 | ATT/TAA: ATA/TAA: ATA/TAA: ATA/TAA |
| trnW(tca) | 56 | 58 | 56 | 52 | −10: −21: −29: −7 | |
| trnY(gta) | 46 | 52 | 68 | 51 | −16: 2: −4: −9 | |
| trnC(gca) | 59 | 48 | 40 | 42 | −5: −5: −2: −12 | |
| cox1 | 1536 | 1536 | 1536 | 1554 | 3: 3: 3: 3 | TTA/TAA: TTA/TAA: TTA/TAA: TTA/TAA |
| cox2 | 669 | 663 | 663 | 618 | 0: 6: 2: 0 | TGG/TAA: TGG/TAA: TGG/TAA: TGG/TAA |
| trnK(ctt) | 60 | 54 | 57 | 55 | −18: −15: −14: 1 | |
| trnD(gtc) | 57 | 63 | 63 | 40 | −8: −15: −15: −2 | |
| nad3 | 324 | 321 | 333 | 333 | −5: −5: −5: −19 | ATT/TAA: ATT/TAA: ATT/TAA: ATT/TAA |
| trnL2(taa) | 53 | 52 | 52 | 63 | −12: 2: −10: 0 | |
| trnN(gtt) | 66 | 54 | 59 | 46 | −11: 9: −2: −3 | |
| trnA(tgc) | 50 | 48 | 48 | 51 | −14: −5: −5: 3 | |
| trnS1(tct) | 67 | 55 | 55 | 47 | 3: −4: −3 | |
| trnK(ctt) | 55 | 70 | 52 | 53 | −19: −26: −9: −9 | |
| trnF(gaa) | 60 | 55 | 55 | 50 | −7: −1: −2: 0 | |
| nad5 | 1644 | 1635 | 1635 | 1635 | −6: −8: −5: −2 | ATA/TAG: ATT/TAA: ATT/TAA: ATT/TAA |
| trnH(gtg) | 66 | 66 | 63 | 61 | 0: 0: 1: 2 | |
| nad4 | 1275 | 1275 | 1275 | 1275 | 12: 6: 6: 6 | ATT/TAA: ATT/TAA: ATT/TAA: ATT/TAA |
| nad4L | 270 | 264 | 270 | 270 | −6: 6: −1: −3 | ATT/TAA: ATT/TAA: ATT/TAA: ATT/TAA |
| trnF(ggg) | 53 | 43 | 49 | 52 | 7: 14: 9: 8 | |
| nad6 | 426 | 426 | 429 | 426 | 0: 2: −1: 3 | ATT/TAA: ATT/TAA: ATT/TAA: ATT/TAG |
| trnG(gat) | 69 | 64 | 68 | 54 | −17: −15: −19: −7 | |
| cob | 1131 | 1131 | 1131 | 1131 | −1: −1: −1: −15 | ATT/TAG: ATT/TAG: ATT/TAG: ATT/TAG |
| trnS2(cca) | 58 | 59 | 59 | 73 | 0: −2: 0: 6 | |
| trnT(tgt) | 55 | 60 | 59 | 53 | −11: 14: 0: −15 | |
| nad1 | 921 | 921 | 921 | 906 | −17: −12: −13: −10 | ATT/TAG: ATT/TAG: ATT/TAG: ATT/TAG |
| trnL1(tag) | 55 | 53 | 54 | 66 | 19: 0: −1: 0 | |
| rrnL | 1048 | 1042 | 1050 | 1046 | 19: 24: 32: 0 | |
| trnI(gat) | 63 | 57 | 43 | 82 | −1: 1: 7: 0 | |
| trnS | 693 | 702 | 699 | 695 | 0: 0: 0: 0 | |
| trnQ(tgg) | 61 | 65 | 65 | 64 | 0: 0: 0: 0 | |
| CR | 498 | 428 | 511 | 848 | |

Of the Nephila and Trichonephila taxa, only the complete mitochondrial genome of T. clavata (previously N. clavata) has been published\(^1\) and is available in the GenBank. There is no report on the phylogenomics of T. vitiana. In view of the lack of study on the mitogenome of the Nephilidae/Nephilinae, and the conflicting systematic status, we sequenced and annotated the complete mitogenomes of N. pilipes, T. antipodiana and T. vitiana to determine their features and phylogenetic relationship. Therefore, this study aims to elucidate the relationship of Nephila and Trichonephila species and support the taxon vitiana as a valid species of the genus Trichonephila.

Results and discussion

Mitogenome features. The total lengths of the complete mitogenomes of N. pilipes, T. antipodiana and T. vitiana (previously N. vitiana) are 14,117 bp, 14,029 bp and 14,108 bp, respectively (Table 1; Table S2; Fig. 1). These three mitogenomes are shorter than those reported for T. clavata\(^1\). The lengths of Nephila and Trichonephila mitogenomes are similar to those reported for araneoid taxa ranging from 14,032 bp in Argope perforata\(^4\) to 14,687 bp in Cyclosa japonica\(^15\) (NC_044696). The complete mitogenome of T. antipodiana has the smallest size compared to those of other araneid taxa; the shortest so far reported is 14,032 bp in A. perforata. The gene arrangement in Nephila and Trichonephila mitogenomes is identical to those of other araneid spiders included in...
this study (Table S2; Fig. S1). All the present three mitogenomes (N. pilipes, T. antipodiana and T. vitiana) have 13 PCGs, two rRNA genes, 22 tRNAs, a non-coding A + T rich control region, and a large number of intergenic sequences (spacers and overlaps) (Table 1; Table S2; Fig. 1).

Besides, all three mitogenomes of N. pilipes, T. antipodiana, and T. vitiana are AT-rich (Table 2). These mitogenomes have negative values for AT skewness and positive values for GC skewness indicating the bias toward the use of Gs over Cs. Although an overall negative AT skewness value and positive GC skewness value are observed for the whole mitogenomes, they are variable for individual genes in different mitogenomes (Table 2).

The A + T content for the N strand in the Nephila and Trichonephila mitogenomes is slightly higher than that for the J strand: with negative skewness value for the J strand and positive skewness value for the N strand (Table 2). The GC skewness value is positive for both the J and N strands, with the respective values for the J strand higher than those of the N strand.

The mitogenomes of both Nephila and Trichonephila are characterized by many more intergenic overlaps than spacers (Table 1; Table S2). The longest spacer in N. pilipes (19 bp) is between trnL1 and rrnL as well as between rrnL and trnV; that in T. antipodiana (24 bp) is between rrnL and trnV; that in T. vitiana (32 bp) between rrnL and trnV; and that in T. clavata (48 bp) between cox1 and cox2. The respective largest overlaps were: – 29 bp

**Figure 1.** Complete mitogenomes of Nephila pilipes, Trichonephila antipodiana, T. vitiana (previously N. vitiana) and T. clavata with BRIG visualization showing the protein-coding genes, rRNAs and tRNAs. GC skew is shown on the outer surface of the ring whereas GC content is shown on the inner surface. The anticodon of each tRNAs is shown in parentheses. Figure generated by BRIG Development version (0.95-dev:0004) (http://brig.sourceforge.net/).
between trnW and trnY in T. vitiana; – 28 bp between trnE and trnF in N. pilipes; – 26 bp between trnR and trnE in T. antipodiana; and – 19 bp between nad3 and trnL2 in T. clavata.

A larger number of intergenic overlaps than spacers is also evident in the mitogenomes of other spiders: Tetragnatha maxillosa, and Tet. nitens (Tetragnathidae)22; Epeus alboguttatus (Salticidae)23; Nadinus falsis (Lycosidae)24, Ebrechetta tricuspidata (Thomisidae)25, Lygnaithus crotalus (Theraphosidae)26, and Cheiracanthium trivale (Cheiracanthidae), and Dystera silvatica (Dysteridae)27.

### Protein-coding genes and codon usage.

The A + T content for PCGs ranges from 69.7% for cox3 to 82.0% for atp8 in N. pilipes, 71.3% for cox1 to 83.4% for atp8 in T. antipodiana, 71.7% for cox1 to 81.4% for atp8 in N. pilipes, and 71.3% for cox3 to 83.4% for atp8 in T. clavata (Table S3). Interestingly, the AT skewness values are negative for the 13 PCGs in N. pilipes, T. antipodiana, and T. clavata; the AT skewness has both positive (nad4, nad4L, and nad5 PCGs) and negative values (the other PCGs) in T. vitiana. All the 13 PCGs in T. vitiana mitogenome have positive GC skewness value (Table S3). The mitogenomes of N. pilipes, T. antipodiana and T. clavata have negative GC values for nad1, nad4, nad4L and nad5 PCGs.

The PCGs of Nephila and Trichonephila mitogenomes are characterized by four start codons: ATA, ATT, TTG and TTA in N. pilipes, T. antipodiana and T. vitiana; ATA, ATT, ATG and TTG in T. clavata (Table 1; Table S2). Two complete stop codons (TA and TAG) are present in the Nephila and Trichonephila mitogenomes. In addition, T. clavata has a truncated incomplete T stop codon. ATT is the commonest start codon in N. pilipes (8 PCGS), while ATA is the commonest in T. antipodiana, T. vitiana and T. clavata (each with 6 PCGS).

Nephila pilipes has identical start/stop codons with the other three Trichonephila mitogenomes for atp8 (ATT/TAA), atp6 (ATA/TAA) and nad3 (ATT/TAA); T. antipodiana and T. vitiana for cox1 (TTA/TAA), cox2 (TTG/TAA), cox3 (TTG/TAA) and nad6 (ATT/TAA); and T. antipodiana and T. clavata for nad4L (ATT/TAA). The mitogenomes of T. antipodiana, T. vitiana and T. clavata have identical start/stop codons for nad4 (ATA/TAA), T. vitiana and T. clavata have identical ATA/TAA codons for nad5 (ATA/TA in N. pilipes and AT/TA in T. antipodiana). The nad2 PCG in N. pilipes and the three other Trichonephila mitogenomes have different start and/or stop codons (Table 1).

The most common start codon with ATA in other spiders includes Tet. maxillosa (5 PCGS) and Tet. nitens (5 PCGS)22; D. silvatica (6 PCGS)22; E. alboguttatus (5 PCGS)23; W. fidelis (5 PCGS)23, and E. tricuspidata (7 PCGS)24. Spiders with ATT as the most common start codon include: C. trivale (5 PCGS)22; L. crotalus (6 PCGS)25; Araneus ventricosus (Araneidae) (7 PCGS)22; Argiope ochiai (Araneidae) (4 PCGS)22; Habronattus orsogonensis (Salticidae) (6 PCGS)25, and Argyroneta aquatica (Cybaeidae) (6 PCGS)25. In six species of Dysteridae spiders, ATA is the commonest start codon in only one species (Parachthes teruelis); the other species have AT as the commonest start codon28.

TAA is the commonest stop codon in N. pilipes (9 PCGs), T. antipodiana (10 PCGs), T. vitiana (11 PCGS), and T. clavata (9 PCGS), excepting: TAG for cob, nad1 and nad5 in N. pilipes; nad1, nad2 and cob in T. antipodiana; cob and nad1 in T. vitiana; and nad2, nad6, cob and nad1 in T. clavata (Table 1; Table S2).

There is a reported AGT as the most common stop codon in A. ventricosus (9 PCGS)22, Neoscona scylla (Araneidae) (12 PCGS)27, Tet. maxillosa (8 PCGS) and Tet. nitens (10 PCGS)22, Evarcha coreana (Salticidae) (9 PCGS)22, W. fidelis (7 PCGS)23, E. tricuspidata (5 PCGS)23, Uroea compactilis (Ocobediidae) (6 PCGS)25, C. trivales (7 PCGS) and D. silvatica (7 PCGS)22, L. crotalus (8 PCGS)25, H. orsogonensis (5 PCGS)24, A. aquatica (4 PCGS and 6 truncated T)22, Mesabolivar sp. 1 (Phocidae) (8 PCGS) and Mesabolivar sp. 2 (11 PCGS)25, and E. alboguttatus (8 PCGS)23.

In the present study, truncated incomplete stop codon (T) is detected only for cox3 in T. clavata (Table 1; Table S2). The incomplete stop codon has been reported for L. crotalus20. Truncated stop codons are however not uncommon in the animal world. Examples of spider mitogenomes with incomplete T stop codons are: E. tricuspidata22; Tet. maxillosa and Tet. nitens22; A. perforata22; A. ocula22; A. ventricosus22; E. alboguttatus22; E.
In general, the incomplete T stop codon in spiders involve the nad genes. Other incomplete stop codons may also be present in spider mitogenomes. Both T and TA stop codons are present in *Mesabolivar* sp. 130 and two species of *Neoscona*. *H. appenica* and five species of *Parachtes* have TA stop codon for two to four PCGs, while only *H. appenica* and three species of *Parachtes* have T stop codon in one or two PCGs. Incomplete TT stop codon has been reported for *nad4L* in *C. triviale*21. Incomplete stop codons are presumed to be completed by post-translational polyadenylation32.

The frequency of individual amino acid varies among the congeners of *Trichonephila* as well as the genera *Nephila* and *Trichonephila* (Fig. 2). However, the most frequently utilized codons are highly similar in these mitogenomes. The predominant amino acids (with frequency above 200) in all the four mitogenomes are isoleucine (Ile), leucine2 (Leu2), methionine (Met), phenylalanine (Phe), serine2 (Ser2), and valine (Val) (Table S4).

Analysis of the relative synonymous codon usage (RSCU) reveals the biased usage of A/T than G/C at the third codon position (Fig. 2). The frequency of each codon is very similar across all the four spider mitogenomes.
Transfer RNA genes. All the tRNAs of the whole Nephila and Trichonephila mitogenomes are AT-rich (Table 2), with positive AT skewness value in N. pilipes, T. vitiana and T. clavata; the GC skewness value is positive for all the four mitogenomes. Most of the tRNAs in Nephila and Trichonephila mitogenomes have aberrant clover-leaf secondary structure, including truncated aminoacyl acceptor stem and mismatched (lacking well-paired) aminoacyl acceptor stem (Fig. 4).

Sixteen tRNAs in the Nephila and Trichonephila mitogenomes do not possess a TΨC arm: seven in N. pilipes and 10 each in T. antipodiana, T. vitiana and T. clavata (Fig. 4). There are also tRNAs with complete loss of TΨC stem (trnD in N. pilipes; trnV in T. antipodiana; and trnK in T. clavata) and complete loss of TΨC loop (trnR and trnQ in N. pilipes and trnK in T. vitiana).

Two tRNAs (trnA, trnS2) do not have DHU arm in all the Nephila and Trichonephila mitogenomes. Other tRNAs without DHU arm are: trnR in N. pilipes; and trnS1 and trnT in T. clavata. The complete loss of DHU loop involves trnQ in N. pilipes, trnN and trnV in T. antipodiana and T. clavata, and trnV in T. vitiana (Fig. 4).

Many tRNAs in spider mitogenomes have been reported to lack a well-paired aminoacyl acceptor stem, a TΨC arm, and a DHU arm. None of the 22 tRNA sequences in H. oregonensis mitogenome have the potential to form a fully paired, seven-member aminoacyl acceptor stem. Mismatched aminoacyl acceptor stem has been reported to be a shared characteristic among spider mitogenomes. It has been postulated that the missing 3' acceptor stem sequence is post-translationally modified by the RNA-editing mechanism. In A. aquatica mitogenome, the tRNAs are characterized by mismatched aminoacyl acceptor stem, and excepting trnS1 and trnS2 (both with only TΨC loop), the remaining tRNAs lack a TΨC arm. The armless tRNA secondary structures are conserved across the family Dysderidae.
Control region. The length of the non-coding control region in *N. pilipes* (498 bp), *T. antipodiana* (428 bp) and *T. vitiana* (511 bp) is much shorter than that of *T. clavata* (848 bp) (Table 1; Table S2). Spider mitogenomes with less than 800 bp for the control region include: *N. nautica* (455 bp) and *N. doenitzi* (566 bp); *E. coreana*.
Figure 5. Fifteen tandem repeats of ATAGA motif with TATATCATAT stretch (except one each with TAT, TATGTACATAT, and TATATACATA) in the control region of *Trichonephila clavata* checked using Tandem Repeats Finder (http://tandem.bu.edu/trf/trf.html).

(697 bp)²⁰; *T. nitens* (690 bp)²⁰; *H. oregonsis* (716 bp)²⁰; *U. compactilis* (688 bp)²⁹; and *L. crotalus* (356 bp)²⁰. Examples of spider mitogenomes with greater than 800 bp are: *Tet. maxillosa* (864 bp)²⁰; *E. tricuspisdapida* (859 bp)²⁰; *C. triviale* (985 bp), *D. sylvatica* (954 bp)²⁰; *E. alboguttatus* (968 bp)²⁶; and *A. aquatica* (2047 bp)²⁵.

The A+T content of the control region of *Nephila* and *Trichonephila* mitogenomes is AT-rich (Table 2), with negative AT skewness value in *T. antipodiana* and positive values in *N. pilipes*, *T. vitiana* and *T. clavata* (Table S3). The GC skewness value is positive for all four mitogenomes.

The close affinity of *T. antipodiana* and *Trichonephila pilipes* supports the *Nephila*–*Trichonephila* clade being basal to other araneid subfamilies (Fig. 6; Fig. S2). The genera *Clitaetra* (containing *N. tetragnathoides* and *N. pilipes*) and *Trichonephila* (now *Trichonephila clavata*) form a clade distinct from other genera of Araneidae. 

The present phylogenetic trees based on 13 PCGs and 15 mt-genes (13 PCGs and 2 rRNA genes) reveal identical topology with very good nodal support based on ML and BI methods (Fig. 6, Fig. S2). However, the genera *Araneus* and *Argiope* do not form monophyletic lineages, and the genus *Cyclosa* is the most basal genus to the other araneid genera. This result indicates that the rDNA genes alone are not suitable for reconstructing phylogeny at the higher taxonomic level.

In a recent study based on 13 protein-coding genes of the complete mitogenome, Nephilidae (represented by *T. clavata*) is basal to the family Araneidae²⁶. Our present study, with the inclusion of *N. pilipes*, *T. antipodiana* and *T. vitiana* (previously *N. vitiana*) as well as *T. clavata* and additional recently published mitogenomes of Araneidae supports the *Nephila–Trichonephila* clade being basal to other araneid subfamilies (Fig. 6; Fig. S2). The close affinity of *T. vitiana* with *T. antipodiana* and *T. clavata* indicates that it is a member of the genus *Trichonephila* and not *Nephila* as currently recognized²⁶.

The close affinity between *T. antipodiana* and *T. vitiana* is also reflected by their genetic distance: 8.65% based on 13 PCGs and 8.62% based on 15 mt-genes. On the other hand, the genetic distance between *T. vitiana* and *N. pilipes* is 21.68% based on 13 PCGs and 21.56% based on 15 mt-genes. Based on 15 mt-genes, the genetic distance between *Trichonephila* species ranges from 8.62 to 13.41% (Table S6).

Studies based on morphological data and mitochondrial and nuclear gene sequences have indicated closer relationship of *T. antipodiana* with *T. clavata* than with *N. pilipes*³⁷–³⁹. Based on anchored hybrid enrichment (AHE) targeted-sequencing approach with 585 single copy orthologous loci, the genus *Nephila* is basal to the genera *Herennia* Thorell, 1877, *Nephilengys* L. Koch, 1872, *Nephilengys* Kuntner, 2013, *Trichonephila* and *Clitaetra*. The genus *Clitaetra* is basal to the genera *Herennia*, *Nephilengys*, *Nephilengys*, and *Trichonephila*.

Mitochondrial genomes have been applied particularly to studies regarding phylogeny and evolution of insects³⁶. A recent study on spider mitogenomes covered only 12 species of Araneidae: 1 species of *Trichonephila*,
Figure 6. Bayesian inference phylogenetic tree based on (a) 13 PCGs and 2 rRNA genes, (b) 13 protein-coding genes, and (c) 2 rRNA genes of the whole mitogenomes of Nephila, Trichonephila and other araneid taxa with Tetragnatha taxa as outgroup. Trichonephila vitiana (previously Nephila vitiana); Numeric values at the nodes are Bayesian posterior probabilities. Figures generated by Mr Bayes v.3.1.2 (https://nbisweden.github.io/MrBayes/download.html).
2 species of Araneus, 2 species of Argiope, 1 species of Cyclosa, 1 species of Cyrtarachne, 1 species of Hypsosinga, and 4 species of Neoscona\(^2\). Our present study has added 1 species of Nephila, 2 species of Trichonephila, 2 species of Argiope, 1 species of Cyrtophora, and 1 species of Neoscona. The taxon sampling is however still very limited compared to the large number of Araneid species. Studies on the mitogenomes of *T. konunci* and *T. plumipes* as well as other *Nephila* and *Trichonephila* species and related taxa will provide a potentially more robust phylogeny and systematics.

**Conclusion**

The whole mitogenomes of *N. pilipes*, *T. antipodiana* and *T. vitiana* (previously *N. vitiana*) possess 37 genes (13 protein-coding genes, two rRNA and 22 tRNA genes), a non-coding control region and intergenic spacer and overlap sequences. Most of the tRNAs have aberrant clover-leaf secondary structure, including loss of TΨC stem and DHU arm as well as truncated and mismatched (lacking well-paired) aminoacyl acceptor stem. The gene arrangement is identical to those of other araneid mitogenomes. Based on 13 protein-coding genes (PCGs) and 15 mitochondrial genes (13 PCGs and two rRNA genes), *Nephila* and *Trichonephila* form a clade distinctly separated from the other araneid subfamilies/genera. *T. antipodiana* and *T. vitiana* are closer to each other than to another member *T. clavata* of the same lineage, and this lineage is separated distinctly from *N. pilipes*, supporting the placement of *vitiana* as a member and valid species of *Trichonephila*. The present study on the mitogenomes of limited taxonomic sampling reveals similar genetic distance between *Nephila*, *Trichonephila* and six other araneid genera, lending support for consideration of *Nephila* and *Trichonephila* as members of the family Araneidae. A more extensive taxonomic sampling of *Nephila* and *Trichonephila* species and related taxa is needed to reconstruct a robust phylogeny based on complete mitogenomes.

**Materials and methods**

**Sample collection.** Adult female spiders were collected from their webs with an insect sweep net. They were preserved in absolute ethanol and stored in −20 °C freezer until use for DNA extraction. *N. pilipes* and *T. antipodiana* were collected in Kelantan, Peninsular Malaysia (6.1254° N, 102.4253° E), and *T. vitiana* from Lombok, Indonesia (8.6510° S, 116.3249° E). The *Nephila* and *Trichonephila* spiders are not endangered or protected by law. No permits are needed to study these spiders.

**Mitochondrial DNA extraction, sample preparation and genome sequencing.** The extraction of mitochondrial DNA was performed as previously described\(^2\). The purified mitochondrial DNA was quantified using Qubit dsDNA High Sensitivity Assay Kit (Life Technologies, USA) and normalized to a final concentration of 50 ng for sample and library preparation using Nextera DNA Sample Preparation Kit. Size estimation of the library was performed on a 2100 Bioanalyzer using High Sensitivity DNA analysis kit (Agilent Technologies) and a real-time quantification of the library was carried out in an Eco Real-Time PCR System using KAPA Library Quantification Kit. The library was sequenced using the Illumina MiSeq Desktop Sequencer (2 × 150 bp paired-end reads) (Illumina, USA)\(^2\).

**Analysis of mitogenome.** Raw sequence reads were obtained from the MiSeq system in FASTQ format. The overall quality of the sequences was assessed from their Phred scores using FastQC software\(^3\). Ambiguous nucleotides and raw sequence reads with lower than Q20 Phred score were trimmed and removed using CLC genomics workbench v.7.0.4 (Qiagen, Germany). Quality-filtered DNA sequences were mapped against the reference mitogenome of *T. clavata* (NC_008063), before a de novo assembly was performed on the mapped DNA sequences. Contigs larger than 13 kbp were extracted for a BLAST search against NCBI nucleotide database to identify the mitochondrial genome of the spider species\(^5\). On the other hand, demultiplexed raw sequence reads that were free of sequencing adapter were subjected for de novo assembly using NOVOplasty with different parameters. The assembled mitogenomes were submitted to MITOS web-server (http://mitos.bioinf.uni-leipzig.de/index.py) for an initial gene annotation\(^6\). The coding regions of protein coding genes (PCGs), transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs) were further validated using nucleotide-nucleotide BLAST (BLASTn) and protein–protein BLAST (BLASTp)\(^6\) against the reference mitogenome of *T. clavata* (NC_008063). For tRNA genes that were not identified, we extracted the DNA sequences of their putative coding regions for an additional Infernal prediction with maximum overlap increased to 50\(^\circ\). The gene boundaries as well as the start and stop codons of PCGs were determined following multiple sequence alignment using ClustalW\(^6\). The overlapping and intergenic spacer regions were curated manually\(^9\). The nucleotide composition, amino acid frequency and relative synonymous codon usage (RSCU) in the complete mitogenomes were calculated in MEGA X\(^8\). The ratios of non-synonymous substitutions (Ka) and synonymous (Ks) substitutions for all PCGs were estimated in DnaSP6\(^9\). The skewness of the mitogenomes was determined from formulae: AT skew = (A − T)/(A + T) and GC skew = (G − C)/(G + C)\(^9\). Inverted repeats or palindromes in the control region were checked using Tandem Repeats Finder (http://tandem.bu.edu/trf/trf.html)\(^5\). The circular mitogenomes of the spiders were visualized using Blast Ring Image Generator (BRIG)\(^5\).
Phylogenetic analysis. The complete mitogenomes of *T. clavata* and Araneidae available from GenBank (Table S1) were used for phylogenetic comparison. *Tetragnatha maxillosa* NC_025775 and *Tetragnatha nitens* NC_028068 were used as outgroup taxa. In addition to *T. clavata* (NC_008063), 16 araneid mitogenomes available in the GenBank were retrieved for phylogenetic analysis (Supplementary Table S1). The mitogenomes of *Tetragnatha maxillosa* (NC_025775) and *T. nitens* (NC_028068) were selected as outgroups. The nucleotide and amino acid sequences of 13 PCGs and the nucleotide sequences of 2 rRNA genes of all mitogenomes were extracted for analysis. MAFIT was used for alignment of the individual nucleotide and amino acid sequences of PCG and rRNA gene sequences. The poorly aligned and highly divergent regions were trimmed.

Alignments of individual gene sequences were concatenated into five datasets: (1) nucleotide sequences of 13 PCGs; (2) nucleotide sequences of two rRNA genes; (3) nucleotide sequences of 15 mt-genes (13 PCGs, 2 rRNA genes); (4) amino acid sequences of 13 PCGs; (5) 13 PCGs with the third codon position excluded. The datasets were imported into PhyloSuite for phylogenetic analysis. The best-fit nucleotide substitution models for maximum likelihood (ML) analysis were determined using ModelFinder based on the Bayesian information criterion.

A ML analysis was performed using IQ-tree incorporated in PhyloSuite under ultrafast bootstrap algorithm with 10,000 replicates. The phylogenetic trees constructed were visualized in MEGA X.

Kakusan v.3 was used to determine the best-fit nucleotide substitution models for Bayesian Inference (BI) analyses using the Bayesian Information Criterion. Bayesian analyses were conducted using the Markov chain Monte Carlo (MCMC) method via MrBayes v.3.1.2, with two independent runs of 2 × 10^6 generations with four chains, and with trees sampled every 200th generation. Likelihood values for all post-analysis trees and parameters were evaluated for convergence and burn-in using the "sump" command in MrBayes and the computer program Tracer v.1.5 (http://tree.bio.ed.ac.uk/software/tracer/). The first 200 trees from each run were discarded as burn-in (where the likelihood values were stabilized prior to the burn-in), and the remaining trees were used for the construction of a 50% majority-rule consensus tree. Phylogenetic trees were viewed and edited by FigTree v.1.4.

Received: 19 January 2021; Accepted: 29 April 2021
Published online: 21 May 2021

References

1. Kuntner, M. et al. Golden orbweavers ignore biological rules: phylogenomic and comparative analyses unravel a complex evolution of sexual size dimorphism. *Syst. Biol.* 68, 555–572 (2019).
2. World Spider Catalog (2020) Natural History Museum Bern. [http://wsc.nmbe.ch](http://wsc.nmbe.ch). Accessed 25 July 2020
3. Kuntner, M. & Coddington, J. A. Discovery of the largest orbweaving spider species: the evolution of gigantism in *Nephila*. *PLOS ONE* 4, e7516 (2009).
4. Harvey, M. S., Austin, A. D. & Adams, M. The systematics and biology of the spider genus *Nephila* (*Araneae: Nephilidae*) in the Australasian region. *Invertebr. Syst.* 21, 407–451 (2007).
5. Kuntner, M. Phylogenetic systematics of the Gondwanan nephilid spider lineage Clitaetrinae (*Araneae, Nephilidae*). *Zool. Scr.* 35, 19–62 (2006).
6. Dimitrov, D. et al. Rounding up the usual suspects: a standard target-gene approach for resolving the interfamilial phylogenetic relationships of ecribellate orb-weaving spiders with a new family-rank classification (*Araneae, Araneoidea*). *Cladistics* 33, 221–250 (2017).
7. Kallal, R. J., Dimitrov, D., Arnedo, M. A., Giribet, G. & Hormiga, G. Monophyly, taxon sampling, and the nature of ranks in the systematics of ecribellate orb-weaving spiders (*Araneae*). *Cladistics* 35, 19–62 (2009).
8. Busvas, V. & Raychaudhuri, D. Taxonomic account of the genus *Nephila* (*Araneae: Nephilidae*) of Bangladesh. *World Sci. News* 123, 66–75 (2019).
9. Kuntner, M., Coddington, J. A. & Hormiga, G. Phylogeny of extant nephilid orb-weaving spiders (*Araneae: Nephilidae*): testing morphological and ethological homologies. *Cladistics* 24, 147–217 (2008).
10. Suana, I. W. et al. Colour variation and polymorphism in the Giant orb-weaving spider *Nephila vitiana* (*Araneae: Nephilidae*) from Lombok, Indonesia. *J. Sci. Technol. Trop.* 7, 5–8 (2011).
11. Yong, H.-S., Hashim, R., Belabut, D., Lim, P. E. (2010). Abdominal colour polymorphism in female Asian golden web spider *Nephila antennata* (*Araneae: Nephilidae*). *Mitochondrial DNA A* 21, 245–251 (2010).
12. Barth, F. Seidenspinne und Spinnenseide. *Mitteilungen aus dem Museum für Naturkunde in Berlin. Zoologisches Museum und Institut für Spezielle Zoologie (Berlin)* 6, 2–9 (1912).
13. Pan, W.-J., Fang, H.-Y., Zhang, P. & Pan, H.-C. The complete mitochondrial genome of *Nephila clavata* (*Araneae: Nephilidae*) from the Chinese population. *Mitochondrial DNA A* 27, 994–995 (2016).
14. Yang, W.-J., Xu, K.-K., Liu, Y., Yang, D.-X. & Li, C. Complete mitochondrial genome and phylogenetic analysis of *Argiope perforata* (*Araneae: Araneidae*). *Mitochondrial DNA B* 4, 1963–1964 (2015).
15. Yang, W.-J., Liu, Y., Xu, K.-K., Yang, D.-X. & Li, C. Characterization of the complete mitochondrial genome of *Cyclosa japonica* (*Araneae: Araneidae*). *Mitochondrial DNA B* 4, 1877–1878 (2019).
16. Wang, Z.-L., Li, C., Fang, W.-Y. & Yu, X.-P. The complete mitochondrial genome of two *Tetragnatha* spiders (*Araneae: Tetragnathidae*): severe truncation of tRNAs and novel gene rearrangements in *Araneae*. *Int. J. Biol. Sci.* 12, 109 (2016).
17. Yang, D., Yan, X., Xu, K., Yang, W. & Li, C. The complete mitochondrial genome of *Epeus albiguttatus* (*Araneae: Salticidae*). *Mitochondrial DNA B* 4, 316–317 (2019).
18. Wang, Z.-L., Li, C., Fang, W.-Y. & Yu, X.-P. The complete mitochondrial genome of the wolf spider *Wadicosa fidelis* (*Araneae: Lycosidae*). *Mitochondrial DNA A* 27, 3909–3910 (2016).
19. Zhu, H.-F., Wang, Z.-Y., Wang, Z.-L. & Yu, X.-P. Complete mitochondrial genome of the crab spider *Erebichthys tricuspedita* (*Araneae: Thomisidae*): a novel rRNA rearrangement and phylogenetic implications for *Araneae*. *Genomics* 111, 1266–1273 (2019).
20. Kumar, V. et al. The Complete Mitochondrial Genome of endemic giant tarantula, *Lyrognathus crotalus* (*Araneae: Theraphosidae*) and comparative analysis. *Sci. Rep.* 10, 1–11 (2020).
21. Tyagi, K. et al. The gene arrangement and phylogeny using mitochondrial genomes in spiders (*Arachnida: Araneae*). *Mitochondrial DNA A* 27, 1026–1027 (2016).
22. Wang, Z.-L., Li, C., Fang, W.-Y. & Yu, X.-P. The complete mitochondrial genome of orb-weaving spider *Aranea ventricosa* (*Araneae: Araneidae*). *Mitochondrial DNA A* 27, 1926–1927 (2016).
23. Yan, Y., Xu, K.-K., Yang, D.-X., Li, C. & Yang, W.-J. The complete mitochondrial genome of Argiope ocula (Araneae: Araneidae) and its phylogeny. *Mitochondrial DNA B* 4, 3138–3139 (2019).

24. Masta, S. E. & Boore, J. L. The complete mitochondrial genome sequence of the spider *Habronattus orsoginesis* reveals rearranged and extremely truncated tRNAs. *Mol. Biol. Evol.* 21, 893–902 (2004).

25. Liu, M., Zhang, Z. & Peng, Z. The mitochondrial genome of the water spider *Argyroneta aquatica* (Araneae: Cybaeidae). *Zool. Scr.* 44, 179–190 (2015).

26. Pons, J., Bover, P., Bidegaray-Batista, L. & Arnedo, M. A. Arm-less mitochondrial tRNAs conserved for over 30 millions of years in spiders. *BMC Genomics* 20, 665 (2019).

27. Xu, K., Lin, X., Yang, D., Yang, W. & Li, C. Characterization of the complete mitochondrial genome sequence of *Neoscona scylla* and phylogenetic analysis. *Mitochondrial DNA B* 4, 416–417 (2019).

28. Yang, W.-J., Xu, K.-K., Yang, D.-X. & Li, C. Characterization of complete mitochondrial genome of *Evarcha coreana* (Araneae: Salticidae). *Mitochondrial DNA B* 4, 1321–1322 (2019).

29. Kim, J. Y. & Park, Y. C. Complete mitogenome and phylogenetic position of *Uroteca compactalis* (Arachnida: Oecobiidae). *Mitochondrial DNA B* 4, 348–349 (2019).

30. Oliveira, R. R. M. et al. Complete mitochondrial genomes of three troglobiphile cave spiders (Mesabolivar, pholcidae). *Mitochondrial DNA B* 4, 251–252 (2019).

31. Wang, Z.-L., Li, C., Fang, W.-Y. & Yu, X.-P. Characterization of the complete mitogenomes of two *Neoscona spiders* (Araneae: Araneidae) and its phylogenetic implications. *Gene* 590, 298–306 (2016).

32. Ogata, D., Montoya, J. & Attardi, G. tRNA punctuation model of RNA processing in human mitochondria. *Nature* 290, 470–474 (1981).

33. Ashfaq, M. et al. Assembling a DNA barcode reference library for the spiders (Arachnida: Araneae) of Pakistan. *PLoS ONE* 14, e0217086 (2019).

34. Aastrup, J. J. et al. Towards a DNA barcode reference database for spiders and harvestmen of Germany. *PLoS ONE* 11, 0016264 (2016).

35. Masta, S. E. & Boore, J. L. Parallel evolution of truncated transfer RNA genes in arachnid mitochondrial genomes. *Mol. Biol. Evol.* 25, 949–959 (2008).

36. Adrián-Serrano, S., Lozano-Fernandez, J., Pons, J., Rozas, J., & Arnedo, M. A. On the shoulder of giants: mitogenome recovery from non-targeted genome projects for phylogenetic inference and molecular evolution studies. *J. Zool. Syst. Evolut. Res.* (2020).

37. Su, Y.-C. et al. Biogeography and speciation patterns of the golden orb spider genus *Nephila* (Araneae: Nephilidae) in Asia. *Zool. Sci.* 28, 47–55 (2011).

38. Kutterer, M., Arnedo, M. A., Trontelj, P., Lokvsek, T. & Agnarsson, I. A molecular phylogeny of nephilid spiders: evolutionary history of a model lineage. *Mol. Phylogenet. Evol.* 69, 961–979 (2013).

39. Dimitrov, D., Benjamin, S. P. & Hormiga, G. A revised phylogenetic analysis for the spider genus *Clitiaeta* Simon, 1889 (Araneae, Araneoidea, Nephilidae) with the first description of the male of the Sri Lankan species *Clitiaeta thisbe* Simon, 1903. *Bull. Museum Comp. Zool.* 159, 301–323 (2009).

40. Cameron, S. L. Insect mitochondrial genomics: implications for evolution and phylogeny. *Annu. Rev. Entomol.* 59, 95–110 (2014).

41. Young, H.-S., Song, S.-L., Lim, P.-E., Eamsobhana, P. & Suana, I. W. Complete mitochondrial genome of three troglobile cave spiders (Mesabolivar, pholcidae). *BMC Genomics* 20, 665 (2019).

42. Yin, H., Chen, C., Ge, Y., Li, J. & Liu, H. Fruit flies (Diptera: Tephritidae) and their phylogenetic implications. *J. Mol. Biol.* 215, 403–410 (1990).

43. Zhang, D. et al. FastQC: A quality control tool for high throughput sequence data. Cambridge, United Kingdom: Babraham Bioinformatics, Babraham Institute (2010).

44. Dierckxsens, N., Mardulyn, P. & Smits, G. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* 45, e18–e18 (2017).

45. Bernt, M. et al. MITOS: improved de novo metazoan mitochondrial genome annotation. *Mol. Phylogenet. Evol.* 69, 313–319 (2013).

46. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410 (1990).

47. Hung, J.-H. & Weng, Z. Sequence alignment and homology search with BLAST and ClustalW. *J. Mol. Evol.* 41, 353–358 (1995).

48. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549 (2018).

49. Rozas, J. et al. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34, 3299–3302 (2017).

50. Perna, N. T. & Kocher, T. D. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *J. Mol. Evol.* 41, 353–358 (1995).

51.加快实现全体的健康。*Ann. Stat.* 69, 961–979 (2013).

52. Phillips, D. M. A. Mafft: an automated sequence alignment program that combines fast Fourier transform distance matrix andенносии и генерации*. *J. Mol. Evol.* 69, 961–979 (2013).

53. Zhang, D. et al. PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* 20, 348–355 (2020).

54. Huelsenbeck, J. P. & Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17, 754–755 (2001).

55. Rambaut, A. Tree Figure Drawing Tool, version 1.4.0 [computer program]. http://tree.bio.ed.ac.uk/software/figtree/. Accessed 9 October (2012).

Acknowledgements
We thank our respective institutions for their support and encouragement of this research. H.-S. Yong thanks University Malaya Deputy Vice Chancellor Kamila Ghazali for permission to carry out research during the Covid-19 pandemic.

Author contributions
H.-S.Y.: Conceptualization, Supervision, Project administration, Methodology, Analysis, Writing—original draft. S.-L.S.: Supervision, Sequencing, Methodology, Bioinformatics, Analysis, Writing—draft, Resources, Validation.
K.-O.C.: Bioinformatics Analysis, Writing—draft, Validation. I.W.S.: Resources, Writing—review. P.E.: Analysis, Writing—review and editing. J.T.: Writing—review and editing. P.-E.L.: Resources, Writing—review and editing. K.-G.C.: Validation, Resources, Writing—review and editing.

**Funding**
This work was supported by grants from University of Malaya (H-562009) to H.-S. Yong.

**Competing interests**
The authors declare no competing interests.

**Additional information**

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-90162-1.

**Correspondence** and requests for materials should be addressed to S.-L.S.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021