



OPEN

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

Hoi-Sen Yong<sup>1</sup>, Sze-Looi Song<sup>2,3</sup>✉, Kah-Ooi Chua<sup>1</sup>, I. Wayan Suana<sup>4</sup>,  
Prapathip Eamsobhana<sup>5</sup>, Ji Tan<sup>6</sup>, Phaik-Eem Lim<sup>3</sup> & Kok-Gan Chan<sup>1,7</sup>

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 Nephilidae<sup>1</sup> or subfamily Nephilinae of Araneidae<sup>2</sup>. Before the taxonomic treatment by Kuntner et al.<sup>1</sup>, *Trichonephila* species were traditionally treated as members of the genus *Nephila*. *Nephila* and *Trichonephila* are large orb-weaving spiders, with *Trichonephila komaci*<sup>3</sup> being the largest species ranging from some 33–40 mm in total length<sup>3</sup>. At different times, they have been treated as members of the family Nephilidae<sup>1,4,5</sup>, and members of the subfamily Nephilinae within the family Araneidae<sup>2,6,7</sup>.

Kuntner et al.<sup>1</sup> listed two species of *Nephila* and 12 species of *Trichonephila*. In contrast, the World Spider Catalog<sup>2</sup> recorded 10 species of *Nephila* and 12 species of *Trichonephila*. Recently, a new species *Nephila nandinia* has been described from Bangladesh<sup>8</sup>. Kuntner et al.<sup>1,9</sup> did not include the taxon *Nephila vitiana* (Walckenaer, 1847) in their studies. *N. vitiana* was treated as a valid species by Harvey et al.<sup>4</sup> and listed as an accepted species in the World Spider Catalog, version 21.5<sup>2</sup>. It is morphologically very similar to *Trichonephila antipodiana* (Walckenaer, 1841). Both taxa exhibit similar abdominal (opisthosomal) colour polymorphism in the adult females<sup>10,11</sup>. 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 sternum<sup>4,12</sup>.

*Nephila pilipes* (Fabricius, 1793) is distributed from India to China, Vietnam, Philippines, and Australia<sup>2</sup>. *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 Tonga<sup>2</sup>.

<sup>1</sup>Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia. <sup>2</sup>Institute for Advanced Studies, University of Malaya, 50603 Kuala Lumpur, Malaysia. <sup>3</sup>Institute of Ocean and Earth Sciences, University of Malaya, 50603 Kuala Lumpur, Malaysia. <sup>4</sup>Faculty of Science and Mathematics, Mataram University, Mataram, Indonesia. <sup>5</sup>Department of Parasitology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand. <sup>6</sup>Department of Agricultural and Food Science, Universiti Tunku Abdul Rahman, 31900 Kampar, Perak, Malaysia. <sup>7</sup>Guangdong Provincial Key Laboratory of Marine Biology, Institute of Marine Sciences, Shantou University, Shantou 515063, China. ✉email: szelooi@um.edu.my

Gene	Size				Intergenic sequence	Start/stop codons
	NP	TA	TV	TC		
<i>trnM</i> (cat)	65	58	49	49	-16: -8: -8: -8	
<i>nad2</i>	957	948	948	948	-10: -2: -2: -4	ATT/TAA: ATA/TAG: ATA/TAA: ATA/TAG
<i>trnW</i> (tca)	56	58	56	52	-10: -21: -29: -7	
<i>trnY</i> (gta)	46	52	68	51	-16: 2: -4: -9	
<i>trnC</i> (gca)	59	48	40	42	-5: -5: 2: -12	
<i>cox1</i>	1536	1536	1536	1554	3: 3: 3: 48	TTA/TAA: TTA/TAA: TTA/TAA: ATG/TAA
<i>cox2</i>	669	663	663	618	0: 6: 2: 0	TTG/TAA: TTG/TAA: TTG/TAA: ATT/TAA
<i>trnK</i> (ctt)	60	54	57	55	-18: -15: -14: 1	
<i>trnD</i> (gtc)	57	63	63	40	-8: -15: -15: -2	
<i>atp8</i>	156	156	156	159	-4: -4: -4: -4	ATT/TAA: ATT/TAA: ATT/TAA: ATT/TAA
<i>atp6</i>	663	663	663	663	3: 3: 3: 3	ATA/TAA: ATA/TAA: ATA/TAA: ATA/TAA
<i>cox3</i>	786	786	786	784	-1: 3: 3: 0	TTG/TAA: TTG/TAA: TTG/TAA: TTG/T
<i>trnG</i> (tcc)	65	53	65	65	-4: 8: -16: -16	
<i>nad3</i>	324	321	333	333	-5: -5: -5: -19	ATT/TAA: ATT/TAA: ATT/TAA: ATT/TAA
<i>trnL2</i> (taa)	53	52	52	63	-12: 2: -10: 0	
<i>trnN</i> (ggt)	66	54	59	46	-11: -9: -2: -3	
<i>trnA</i> (tgc)	50	48	48	51	-14: -5: -5: 3	
<i>trnS1</i> (tct)	67	55	55	47	3: -4: -3: 9	
<i>trnR</i> (tcg)	55	70	52	53	-19: -26: -9: -9	
<i>trnE</i> (ttc)	60	51	50	41	-28: -14: -12: -9	
<i>trnF</i> (gaa)	60	55	55	50	-7: -1: -2: 0	
<i>nad5</i>	1644	1635	1635	1635	-6: -8: -5: -2	ATA/TAG: ATT/TAA: ATA/TAA: ATA/TAA
<i>trnH</i> (gtg)	66	66	63	61	0: 0: 1: 2	
<i>nad4</i>	1275	1275	1275	1275	12: 6: 6: 6	ATT/TAA: ATA/TAA: ATA/TAA: ATA/TAA
<i>nad4L</i>	270	264	270	270	-4: 6: -1: -3	ATT/TAA: ATA/TAA: ATT/TAA: ATT/TAA
<i>trnP</i> (tgg)	53	43	49	52	7: 14: 9: 8	
<i>nad6</i>	426	426	429	426	0: 2: -1: 3	ATT/TAA: ATT/TAA: ATT/TAA: ATT/TAG
<i>trnI</i> (gat)	69	64	68	54	-17: -15: -19: -7	
<i>cob</i>	1131	1131	1131	1131	-1: -1: -1: -15	ATT/TAG: ATA/TAG: ATA/TAG: ATA/TAG
<i>trnS2</i> (tga)	58	59	59	73	0: -2: 0: 6	
<i>trnT</i> (tgt)	55	60	59	53	-11: -14: 0: -15	
<i>nad1</i>	921	921	921	906	-17: -12: -13: -10	ATT/TAG: ATA/TAG: ATT/TAG: ATA/TAG
<i>trnL1</i> (tag)	55	53	54	66	19: 0: -1: 0	
<i>rrnL</i>	1048	1042	1050	1046	19: 24: 32: 0	
<i>trnV</i> (tac)	63	57	43	82	-1: 1: 7: 0	
<i>rrnS</i>	693	702	699	695	0: 0: 0: 0	
<i>trnQ</i> (ttg)	61	65	65	64	0: 0: 0: 0	
CR	498	428	511	848		

**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.

Of the *Nephila* and *Trichonephila* taxa, only the complete mitochondrial genome of *T. clavata* (previously *N. clavata*) has been published<sup>13</sup> 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*<sup>13</sup>. The lengths of *Nephila* and *Trichonephila* mitogenomes are similar to those reported for araneoid taxa ranging from 14,032 bp in *Argiope perforata*<sup>14</sup> to 14,687 bp in *Cyclosa japonica*<sup>15</sup> (NC\_044696). The complete mitogenome of *T. antipodiana* has the smallest size compared to those of other araneoid 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



Region	A + T %				AT skew				GC skew			
	NP	TA	TV	TC	NP	TA	TV	TC	NP	TA	TV	TC
Whole genome	75.8	76.4	76.3	76.0	-0.045	-0.054	-0.056	-0.053	0.240	0.271	0.266	0.242
Protein coding genes	75.1	75.8	76.2	74.9	-0.161	-0.175	-0.067	-0.181	0.078	0.079	0.280	0.091
1st codon position	69.2	71.0	74.9	69.6	-0.055	-0.070	-0.090	-0.037	0.235	0.258	0.400	0.252
2nd codon position	68.8	67.7	70.7	67.8	-0.415	-0.443	-0.103	-0.467	-0.091	-0.105	0.242	-0.105
3rd Codon position	87.5	88.8	82.4	87.4	-0.043	-0.053	-0.033	-0.074	0.111	0.143	0.326	0.205
tRNA genes	77.7	77.5	77.9	78.8	-0.018	0.006	-0.004	-0.008	0.117	0.144	0.131	0.125
rRNA genes	78.6	78.3	77.8	78.8	0.001	0.015	0.022	0.020	-0.005	-0.018	-0.041	-0.003
Control region	74.1	79.7	74.5	81.7	0.107	-0.004	0.013	0.007	0.166	0.379	0.153	0.049
J strand	74.3	75.2	75.2	73.9	-0.171	-0.180	-0.176	-0.185	0.268	0.306	0.306	0.295
N strand	77.3	77.6	77.7	77.8	0.079	0.082	0.072	0.082	0.189	0.214	0.229	0.189

**Table 2.** A + T content (%), AT and GC skewness of the *Nephila* and *Trichonephila* mitogenomes. NP, *Nephila pilipes*; TA, *Trichonephila antipodiana*; TV, *Trichonephila vitiana* (previously *N. vitiana*); TC, *Trichonephila clavata* NC\_008063.

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)<sup>16</sup>; *Epeus alboguttatus* (Salticidae)<sup>17</sup>; *Wadicosa fidelis* (Lycosidae)<sup>18</sup>; *Ebrechtella tricuspadata* (Thomisidae)<sup>19</sup>; *Lyrognaethus crotalus* (Theraphosidae)<sup>20</sup>; and *Cheiracanthium trivale* (Cheiracanthidae), and *Dystera silvatica* (Dysteridae)<sup>21</sup>.

**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 *cox3* to 81.4% for *atp8* in *T. vitiana*, 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 (TAA 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. vitiana* 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/TAG in *N. pilipes* and ATT/TAA in *T. antipodiana*). The *nad2* PCG in *N. pilipes* and the three other *Trichonephila* mitogenomes have different start and/or stop codons (Table 1).

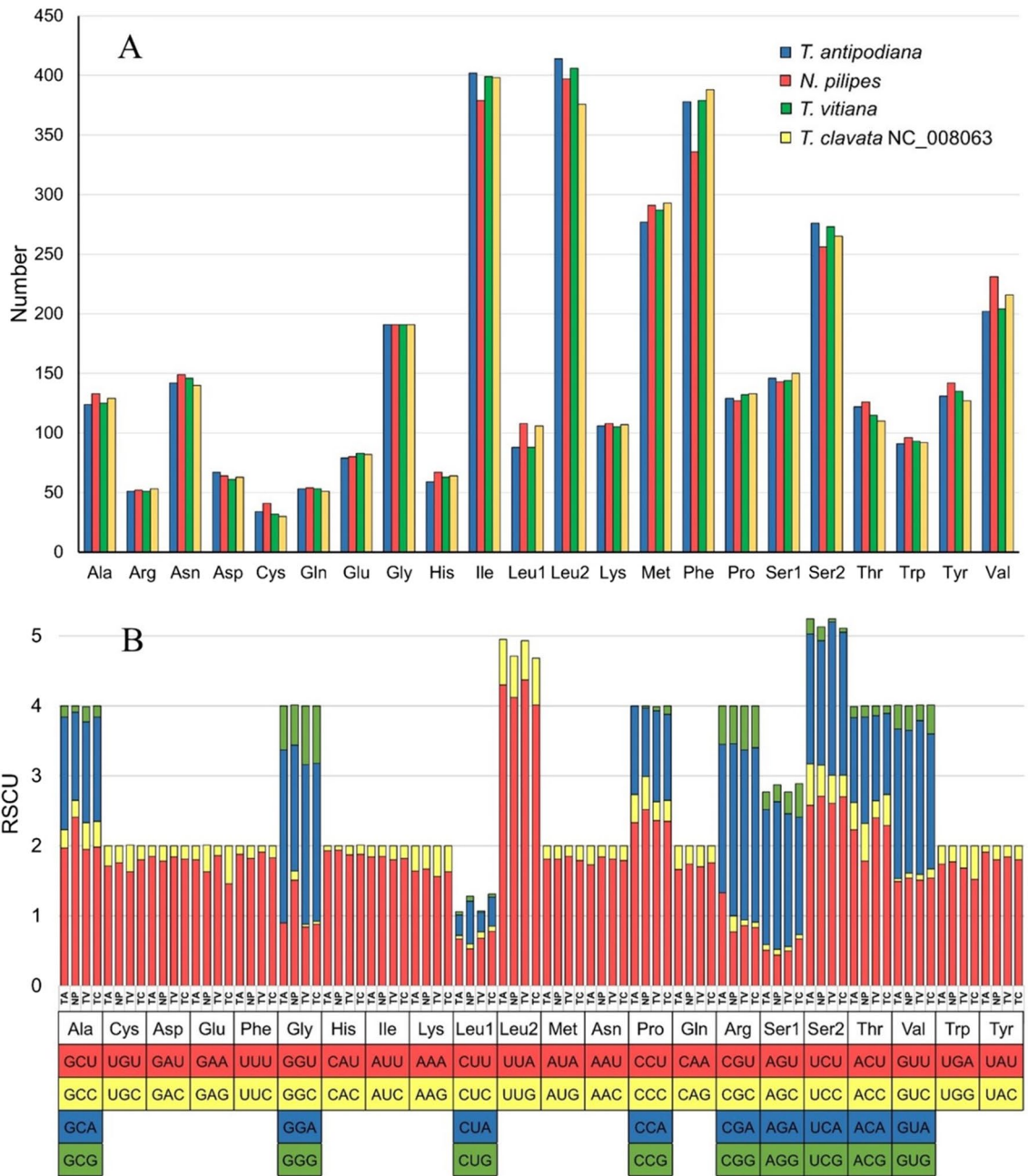
The most common start codon with ATA in other spiders includes *Tet. maxillosa* (5 PCGs) and *Tet. nitens* (5 PCGs)<sup>16</sup>; *D. silvatica* (6 PCGs)<sup>21</sup>; *E. alboguttatus* (5 PCGs)<sup>17</sup>; *W. fidelis* (5 PCGs)<sup>18</sup>; and *E. tricuspadata* (7 PCGs)<sup>19</sup>. Spiders with ATT as the most common start codon include: *C. trivale* (5 PCGs)<sup>21</sup>; *L. crotalus* (6 PCGs)<sup>20</sup>; *Araneus ventricosus* (Araneidae) (7 PCGs)<sup>22</sup>; *Argiope ocula* (Araneidae) (4 PCGs)<sup>23</sup>; *Habronattus oregonensis* (Salticidae) (6 PCGs)<sup>24</sup>; and *Argyroneta aquatica* (Cybaeidae) (6 PCGs)<sup>25</sup>. In six species of Dysteridae spiders, ATA is the commonest start codon in only one species (*Parachtes teruelis*); the other species have ATT as the commonest start codon<sup>26</sup>.

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).

TAA has been reported to be the most common stop codon in *A. ventricosus* (9 PCGs)<sup>24</sup>, *Neoscona scylla* (Araneidae) (12 PCGs)<sup>27</sup>, *Tet. maxillosa* (8 PCGs) and *Tet. nitens* (10 PCGs)<sup>16</sup>, *E. alboguttatus* (8 PCGs)<sup>17</sup>, *Evarcha coreana* (Salticidae) (9 PCGs)<sup>28</sup>, *W. fidelis* (7 PCGs)<sup>18</sup>, *E. tricuspadata* (5 PCGs)<sup>19</sup>, *Uroctea compactilis* (Oecobiidae) (6 PCGs)<sup>29</sup>, *C. trivales* (7 PCGs) and *D. silvatica* (7 PCGs)<sup>21</sup>, *L. crotalus* (8 PCGs)<sup>20</sup>, *H. oregonensis* (5 PCGs)<sup>24</sup>, *A. aquatica* (4 PCGs and 6 truncated T)<sup>25</sup>, *Mesabolivar* sp. 1 (Phocidae) (8 PCGs) and *Mesabolivar* sp. 2 (11 PCGs)<sup>30</sup>, and *E. alboguttatus* (8 PCGs)<sup>17</sup>.

In the present study, truncated incomplete stop codon (T) is detected only for *cox3* in *T. clavata* (Table 1; Table S2). No incomplete stop codon has been reported for *L. crotalus*<sup>20</sup>. Truncated stop codons are however not uncommon in the animal world. Examples of spider mitogenomes with incomplete T stop codons are: *E. tricuspadata*<sup>19</sup>; *Tet. maxillosa* and *Tet. nitens*<sup>16</sup>; *A. perforata*<sup>14</sup>; *A. ocula*<sup>23</sup>; *A. ventricosus*<sup>22</sup>; *E. alboguttatus*<sup>17</sup>; *E.*





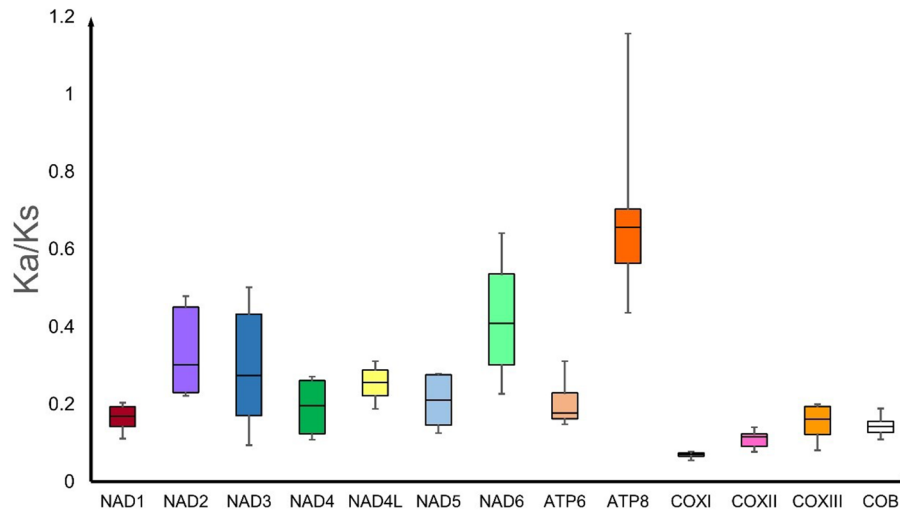
**Figure 2.** Amino acid frequency (A) and relative synonymous codon usage (B) of PCGs in the *Nephila* and *Trichonephila* mitogenomes generated using MEGAX (<https://www.megasoftware.net/>). NP, *Nephila pilipes*; TA, *Trichonephila antipodiana*; TV, *Trichonephila vitiana* (previously *N. vitiana*); TC, *Trichonephila clavata*.

*coreana*<sup>28</sup>; *Neoscona nautica*<sup>31</sup>; *N. scylla*<sup>27</sup>; *H. oregonensis*<sup>24</sup>; *Mesabolivar* sp. 1<sup>30</sup>; *C. triviale*<sup>21</sup>; *D. silvatica*<sup>21</sup>; *U. compactilis*<sup>29</sup>; *A. aquatica*<sup>25</sup>; *W. fidelis*<sup>18</sup>.

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. 1<sup>30</sup> and two species of *Neoscona*<sup>31</sup>. *H. appenicola* and five species of *Parachtes* have TA stop codon for two to four PCGs, while only *H. appenicola* and three species of *Parachtes* have T stop codon in one or two PCGs<sup>26</sup>. Incomplete TT stop codon has been reported for *nad4L* in *C. triviale*<sup>21</sup>. Incomplete stop codons are presumed to be completed by post-translational polyadenylation<sup>32</sup>.

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.



**Figure 3.** Box plot for pairwise divergence of Ka/Ks ratio (mean  $\pm$  SD, and range) for 13 PCGs of *Nephila* ( $n = 1$ ) and *Trichonephila* ( $n = 3$ ) mitogenomes generated using DnaSP6.0. (<http://www.ub.edu/dnasp/>).

The Ka/Ks ratio (an indicator of selective pressure on a PCG) is less than 1 for all the 13 PCGs in *Nephila* and *Trichonephila* mitogenomes, indicating purifying selection (Fig. 3; Table S5). Similar finding has been reported for 17 spider mitogenomes<sup>20</sup>. The sequence of the Ka/Ks ratio ( $cox1 < cox2 < cob < cox3 < nad1 < nad4 < atp6 < nad5 < nad4L < nad3 < nad2 < nad6 < atp8$ ) in *Nephila* and *Trichonephila* species differs from that of ( $cox1 < nad1 < cox2 < nad5 < cob < cox3 < nad4 < atp6 < nad4L < nad3 < nad2 < nad6 < atp8$ ) reported for 17 spider mitogenomes<sup>20</sup>. The *cox1* gene with the lowest Ka/Ks ratio in spider mitogenomes, representing fewer changes in amino acids, supports its use as a molecular marker for species differentiation and DNA barcoding<sup>33,34</sup>.

**Ribosomal RNA genes.** Of the two rRNA genes in *Nephila* and *Trichonephila* mitogenomes, *rrnS* is much shorter, ranging from 693 bp in *N. pilipes* to 702 bp in *T. antipodiana*, while *rrnL* ranges from 1042 bp in *T. antipodiana* to 1050 bp in *T. vitiana* (Table 1, Table S2). As in other araneid spiders, *rrnL* is located between *trnL1* and *trnV* and *rrnS* between *trnV* and *trnQ* (Fig. 1; Fig. S1).

Both the rRNA genes of the complete mitogenome are AT-rich (Table 2). The AT skewness value is variable among the mitogenomes: positive for both *rrnL* and *rrnS* in *T. antipodiana* and *T. clavata*; negative for both genes in *T. vitiana*; and negative for *rrnL* but positive for *rrnS* in *N. pilipes*. The GC skewness value is negative for *rrnL* and positive for *rrnS* in *N. pilipes*, *T. antipodiana* and *T. clavata* mitogenomes; it is positive for *rrnL* and negative for *rrnS* in *T. vitiana*.

Most spiders have longer *rrnL* than *rrnS* gene: *Tet. maxillosa* and *Tet. nitens*<sup>16</sup>; *C. triviale* and *D. silvatica*<sup>21</sup>; *E. coreana*<sup>28</sup>; *W. fidelis*<sup>18</sup>; *A. perforata*<sup>14</sup>; *L. crotalus*<sup>20</sup>; *E. tricuspidata*<sup>19</sup>, and *A. aquatica*<sup>25</sup>. Some spiders have similar length for *rrnL* and *rrnS*: for example, the length of *rrnL* and *rrnS* is the same (1722 bp) in *N. nautica* and *N. doenitzii*<sup>31</sup>.

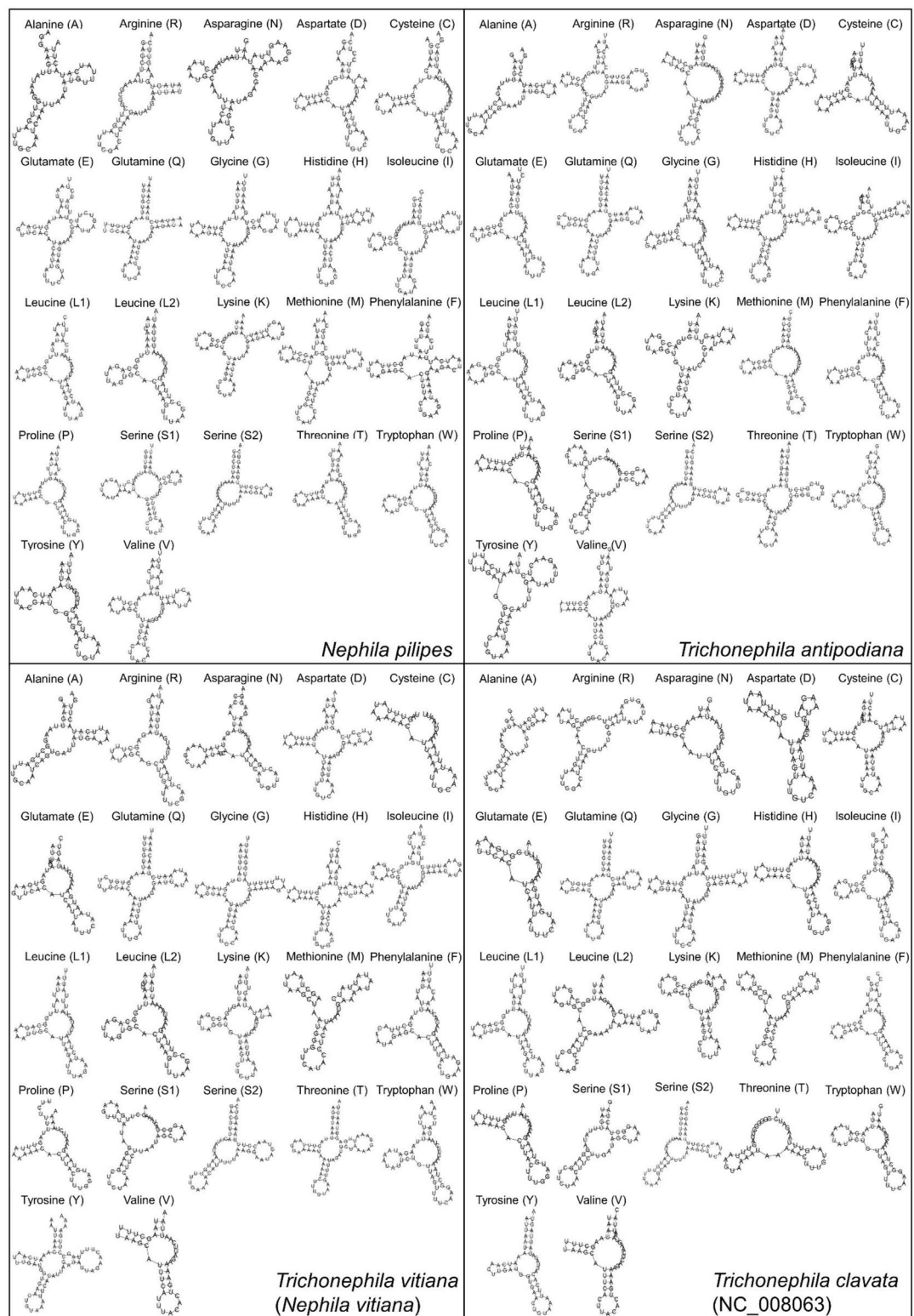
**Transfer RNA genes.** The tRNAs of the whole *Nephila* and *Trichonephila* mitogenomes are AT-rich (Table 2), with positive AT skewness value in *T. antipodiana* and negative value in *N. pilipes*, *T. vitiana* and *T. clavata*; the GC skewness value is positive for all the four mitogenomes.

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<sup>35</sup>. None of the 22 tRNA sequences in *H. oregonensis* mitogenome have the potential to form a fully paired, seven-member aminoacyl acceptor stem<sup>24</sup>. Mismatched aminoacyl acceptor stem has been reported to be a shared characteristic among spider mitogenomes<sup>35</sup>. It has been postulated that the missing 3' acceptor stem sequence is post-translationally modified by the RNA-editing mechanism<sup>24</sup>. 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<sup>25</sup>. The armless tRNA secondary structures are conserved across the family Dysderidae<sup>36</sup>.



**Figure 4.** Cloverleaf structure of the 22 inferred tRNAs in the mitogenomes of *Nephila* and *Trichonephila* mitogenomes obtained from MITOS web-server (<http://mitos.bioinf.uni-leipzig.de/index.py>).

**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)<sup>31</sup>; *E. coreana*



```

9961 atatagatat atacatat agatatatac atatagatat atatacat atagatatat
10021 acatatatag atatatacat atatagatat atacatat agatatatac atatagatat
10081 atatacat atagatat acatatatag atatatacat atatagatat atagatatat
10141 acatatatag atatatacat atatagatat gtacatat agatatatac ataaataata
10201 tcagctaatt aagctaattg gttcatacct cataaatgaa aat

```

**Figure 5.** Fifteen tandem repeats of ATAGA motif with TATATACATAT stretch (except one each with TAT, TATGTACATAT, and TATATACATAA) in the control region of *Trichonephila clavata* checked using Tandem Repeats Finder (<http://tandem.bu.edu/trf/trf.html>).

(697 bp)<sup>20</sup>; *T. nitens* (690 bp)<sup>17</sup>; *H. oregonensis* (716 bp)<sup>24</sup>; *U. compactilis* (688 bp)<sup>29</sup>; and *L. crotalus* (356 bp)<sup>20</sup>. Examples of spider mitogenomes with greater than 800 bp are: *Tet. maxillosa* (864 bp)<sup>17</sup>; *E. tricuspidata* (859 bp)<sup>19</sup>; *C. triviale* (985 bp), *D. sylvatica* (954 bp)<sup>21</sup>; *E. alboguttatus* (968 bp)<sup>16</sup>; and *A. aquatica* (2047 bp)<sup>25</sup>.

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 control region of *Nephila* and *Trichonephila* mitogenomes is characterized by: (i) many simple tandem repeats and palindrome; (ii) long poly-nucleotide; and (iii) several stem-loop structures in these spider mitogenomes. The presence of 15 tandem repeats of ATAGA motif with TATATACATAT stretch (except one each with TAT, TATGTACATAT, and TATATACATAA) in *T. clavata* (Fig. 5) is a unique feature for this orb-weaving spider. Five 135-bp tandem repeats and two 363-bp tandem repeats have been identified in the putative control region of *A. aquatica*<sup>25</sup>. A long tandem repeat region comprising three full 215 bp and a partial 87 bp is present in the control region of *W. fidelis* mitogenome<sup>18</sup>.

**Phylogenetic analysis.** An early study based on one nuclear (18S) and two mitochondrial (COXI and 16S) markers revealed that *N. pilipes* and *N. constricta* Karsch, 1879 formed a clade that was sister to all other *Nephila* species<sup>37</sup>. This finding was supported by molecular phylogenetic study based on three nuclear and five mitochondrial genes which indicated that the genus *Nephila* was diphyetic, with true *Nephila* (containing *N. pilipes* and *N. constricta*) and the other species (now genus *Trichonephila* according to Kuntner et al.<sup>1</sup>) being sister to the genus *Clitaetra* Simon, 1889<sup>38</sup>. Large genetic difference (Fixed Differences, FD = 80%) between *N. pilipes* and other *Nephila* (now *Trichonephila*) species [*N. edulis* (Labillardière), *N. plumipes* (Latreille, 1804) and *N. tetragathoides* (Walckenaer, 1841)] in Australasia had also been reported based on allozyme data<sup>4</sup>.

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). The genera *Nephila* and *Trichonephila* form a clade distinct from other genera of Araneidae. *T. antipodiana* and *T. vitiana* are closer related in the lineage containing also *T. clavata*, while *N. pilipes* is distinctly separated from these *Trichonephila* species. The araneid subfamilies Araneinae (genera *Araneus*, *Cyclosa*, *Hypsosinga* and *Neoscona*), Argiopinae (genus *Argiope*), Cyrtarachninae (genus *Cyrtarachna*) and Cyrtophorinae (genus *Cyrtophora*) form a clade distinct from the *Nephila-Trichonephila* clade.

Araneinae does not form a monophyletic group, with the genus *Cyclosa* being basal to the other Araneinae genera (*Araneus*, *Hypsosinga* and *Neoscona*), as well as the monophyletic subfamilies Argiopinae and Cyrtophorinae (Fig. 6; Fig. S2). Argiopinae and Cyrtophorinae form a lineage distinct from the Araneinae lineages comprising *Neoscona* and (*Araneus-Hypsosinga*), Cyrtarachninae is basal to the above araneid subfamilies. A large, representative taxonomic sampling is needed to reconstruct a robust phylogeny.

Both the BI and ML trees based on two rRNA (*rrnL* and *rrnS*) sequences reveal identical clades as 15 mt-genes and 13 PCGs (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 rRNA genes alone are not suitable for reconstructing phylogeny at the higher taxonomic level.

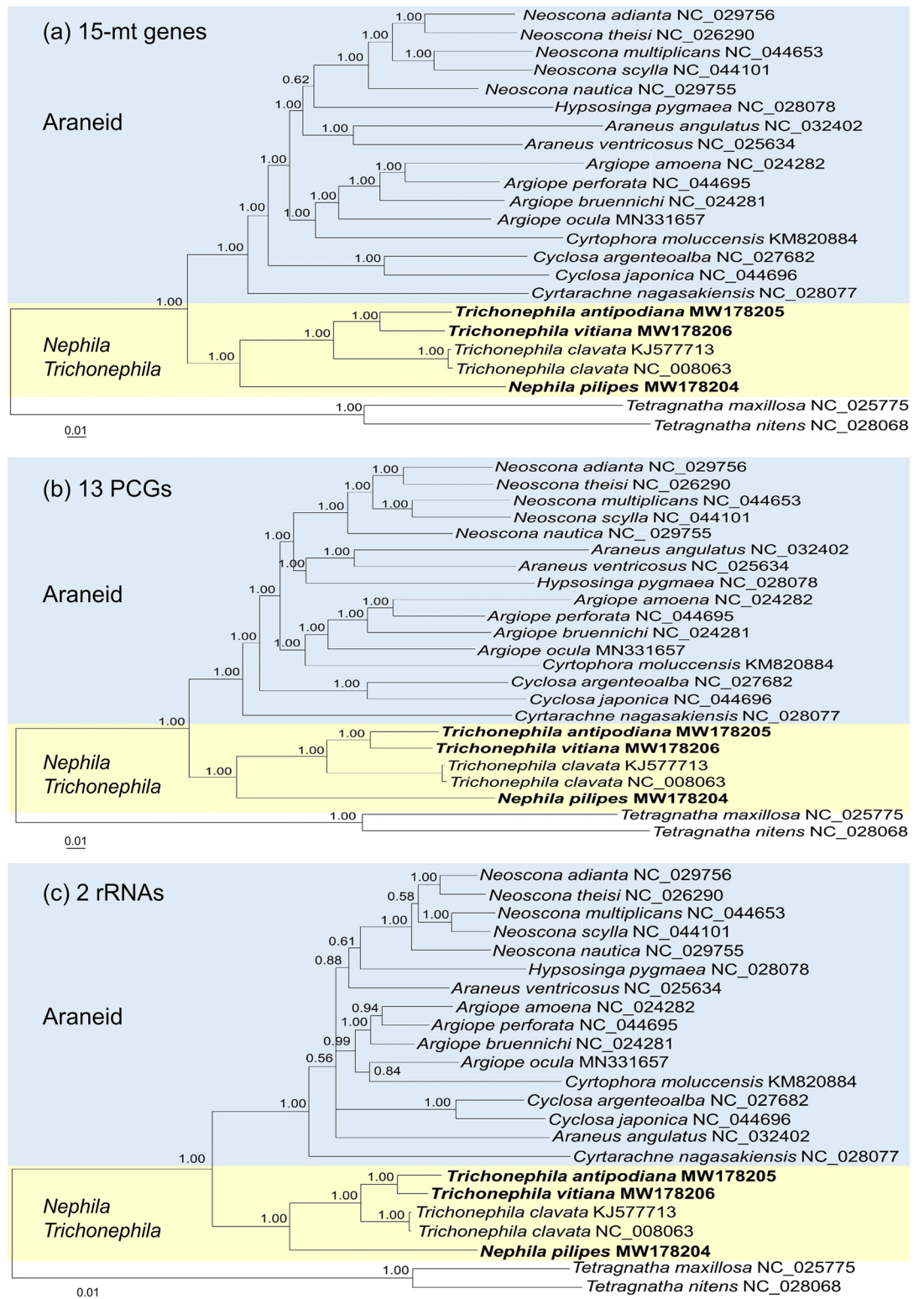
In a recent study based on 13 protein-coding genes of the complete mitogenome, Nephilidae (represented by *T. clavata*) is basal to the family Araneidae<sup>19</sup>. Our present study, with the inclusion of *N. pilipes*, *T. antipodiana* and *T. vitiana* (previously *N. vitiana*) as well as *T. clavata* 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 recognized<sup>2</sup>.

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*<sup>37–39</sup>. 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, *Nephilingis* Kuntner, 2013, *Trichonephila* and *Clitaetra*<sup>1</sup>. The genus *Clitaetra* is basal to the genera *Herennia*, *Nephilengys*, *Nephilingis*, and *Trichonephila*.

Mitochondrial genomes have been applied particularly to studies regarding phylogeny and evolution of insects<sup>40</sup>. 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*<sup>21</sup>. 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. komaci* 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 related 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  $-20^{\circ}\text{C}$  freezer until use for DNA extraction. *N. pilipes* and *T. antipodiana* were collected in Kelantan, Peninsular Malaysia ( $6.1254^{\circ}\text{N}$ ,  $102.4253^{\circ}\text{E}$ ), and *T. vitiana* from Lombok, Indonesia ( $8.6510^{\circ}\text{S}$ ,  $116.3249^{\circ}\text{E}$ ). 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<sup>41</sup>. 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 ( $2 \times 150$  bp paired-end reads) (Illumina, USA)<sup>42</sup>.

**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<sup>43</sup>. Ambiguous nucleotides and raw sequence reads with lower than Q20 Phred score were trimmed and removed using CLC genomic workbench v.7.0.4 (Qiagen, Germany). Quality-filtered DNA sequences were mapped against the reference mitogenome *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<sup>41</sup>. On the other hand, demultiplexed raw sequence reads that were free of sequencing adapter were subjected for de novo assembly using NOVOplasty with different lengths of k-mer<sup>44</sup>. The assembled genomes from both softwares were aligned and examined for terminal repeats to evaluate their circularity and completeness. The mitogenome sequences of *N. pilipes*, *T. antipodiana* and *T. vitiana* (previously *N. vitiana*) have been deposited in GenBank under the accession numbers MW178204, MW178205 and MW178206, respectively.

**Gene annotation, visualization and comparative analysis.** The assembled mitogenomes were submitted to MITOS web-server (<http://mitos.bioinf.uni-leipzig.de/index.py>) for an initial gene annotation<sup>45</sup>. 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)<sup>46</sup> 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<sup>26</sup>. The gene boundaries as well as the start and stop codons of PCGs were determined following multiple sequence alignment using ClustalW<sup>47</sup>. The overlapping and intergenic spacer regions were curated manually<sup>21</sup>. The nucleotide composition, amino acid frequency and relative synonymous codon usage (RSCU) in the complete mitogenomes were calculated in MEGA X<sup>48</sup>. The ratios of non-synonymous substitutions (Ka) and synonymous (Ks) substitutions for all PCGs were estimated in DnaSP6.0<sup>49</sup>. The skewness of the mitogenomes was determined from formulae:  $\text{AT skew} = (A - T)/(A + T)$  and  $\text{GC skew} = (G - C)/(G + C)$ <sup>50</sup>. Inverted repeats or palindromes in the control region were checked using Tandem Repeats Finder (<http://tandem.bu.edu/trf/trf.html>)<sup>51</sup>. The circular mitogenomes of the spiders were visualized using Blast Ring Image Generator (BRIG)<sup>52</sup>.

**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<sup>16</sup> 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. MAFFT was used for alignment of the individual nucleotide and amino acid sequences of PCG and rRNA gene sequences<sup>53</sup>. The poorly aligned and highly divergent regions were trimmed.

Alignments of individual genes 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<sup>54</sup> for phylogenetic analysis. The best-fit nucleotide substitution models for maximum likelihood (ML) analysis were determined using ModelFinder<sup>55</sup> based on the Bayesian information criterion<sup>56</sup>. A ML analysis was performed using IQ-tree<sup>57</sup> incorporated in PhyloSuite under ultrafast bootstrap algorithm with 10,000 replicates. The phylogenetic trees constructed were visualized in MEGA X<sup>48</sup>.

Kakusan v.3<sup>58</sup> was used to determine the best-fit nucleotide substitution models for Bayesian Inference (BI) analyses using the Bayesian Information Criterion<sup>56</sup>. Bayesian analyses were conducted using the Markov chain Monte Carlo (MCMC) method via MrBayes v.3.1.2<sup>59</sup>, with two independent runs of  $2 \times 10^6$  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<sup>60</sup>.

Received: 19 January 2021; Accepted: 29 April 2021

Published online: 21 May 2021

## References

- 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).
- World Spider Catalog (2020) Natural History Museum Bern. <http://wsc.nmbe.ch>. Accessed 25 July 2020
- Kuntner, M. & Coddington, J. A. Discovery of the largest orbweaving spider species: the evolution of gigantism in *Nephila*. *PLoS ONE* **4**, e7516 (2009).
- 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).
- Kuntner, M. Phylogenetic systematics of the Gondwanan nephilid spider lineage Clitaetrinae (Araneae, Nephilidae). *Zoolog. Scr.* **35**, 19–62 (2006).
- 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).
- Kallal, R. J., Dimitrov, D., Arnedo, M. A., Giribet, G. & Hormiga, G. Monophyly, taxon sampling, and the nature of ranks in the classification of orb-weaving spiders (Araneae: Araneoidea). *Syst. Biol.* **69**, 401–411 (2020).
- Biswas, V. & Raychaudhuri, D. Taxonomic account of the genus *Nephila* (Araneae: Nephilidae) of Bangladesh. *World Sci. News* **123**, 66–75 (2019).
- 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).
- 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).
- Yong, H.-S., Hashim, R., Belabut, D., & Lim, P. E. (2010). Abdominal colour polymorphism in female Asian golden web spider *Nephila antipodiana* (Araneae: Nephilidae).
- Dahl, F. Seidenspinne und Spinnenseide. *Mitteilungen aus dem Museum für Naturkunde in Berlin. Zoologisches Museum und Institut für Spezielle Zoologie (Berlin)* **6**, 2–89 (1912).
- Pan, W.-J., Fang, H.-Y., Zhang, P. & Pan, H.-C. The complete mitochondrial genome of *Nephila clavata* (Araneae: Nephilidae) Chinese population. *Mitochondrial DNA A* **27**, 994–995 (2016).
- 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 (2019).
- Yang, W.-J., Liu, Y., Xu, K.-K., Yang, D.-X. & Li, C. Characterization of the complete mitochondrial genome of *Cyclosajaponica* (Araneae: Araneidae). *Mitochondrial DNA B* **4**, 1877–1878 (2019).
- 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).
- Yang, D., Yan, X., Xu, K., Yang, W. & Li, C. The complete mitochondrial genome of *Epeus alboguttatus* (Araneae: Salticidae). *Mitochondrial DNA B* **4**, 316–317 (2019).
- 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).
- Zhu, H.-F., Wang, Z.-Y., Wang, Z.-L. & Yu, X.-P. Complete mitochondrial genome of the crab spider *Ebrechtella tricuspidata* (Araneae: Thomisidae): a novel tRNA rearrangement and phylogenetic implications for Araneae. *Genomics* **111**, 1266–1273 (2019).
- Kumar, V. *et al.* The Complete Mitochondrial Genome of endemic giant tarantula, *Lyrognaethus crotalus* (Araneae: Theraphosidae) and comparative analysis. *Sci. Rep.* **10**, 1–11 (2020).
- Tyagi, K. *et al.* The gene arrangement and phylogeny using mitochondrial genomes in spiders (Arachnida: Araneae). *Int. J. Biol. Macromol.* **146**, 488–496 (2020).
- Wang, Z.-L., Li, C., Fang, W.-Y. & Yu, X.-P. The complete mitochondrial genome of orb-weaving spider *Araneus ventricosus* (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**, 3318–3319 (2019).
24. Masta, S. E. & Boore, J. L. The complete mitochondrial genome sequence of the spider *Habronattus oregonensis* 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 *Uroctea compactilis* (Arachnida: Oecobiidae). *Mitochondrial DNA B* **4**, 348–349 (2019).
30. Oliveira, R. R. M. *et al.* Complete mitochondrial genomes of three troglophile 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. Ojala, 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. Astrin, J. J. *et al.* Towards a DNA barcode reference database for spiders and harvestmen of Germany. *PLoS ONE* **11**, 00162624 (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. Kuntner, M., Arnedo, M. A., Trontelj, P., Lokovšek, 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 *Clitaetra* Simon, 1889 (Araneae, Araneioidea, Nephilidae) with the first description of the male of the Sri Lankan species *Clitaetra 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**, 95–117 (2020).
41. Yong, H.-S., Song, S.-L., Lim, P.-E., Eamsobhana, P. & Suana, I. W. Complete mitochondrial genome of four *Bactrocera* fruit flies of subgenus *Bactrocera* (Diptera: Tephritidae) and their phylogenetic implications. *PLoS ONE* **11**, e0148201 (2016).
42. Yong, H.-S. *et al.* Mitochondrial genome supports sibling species of *Angiostrongylus costaricensis* (Nematoda: Angiostrongylidae). *PLoS ONE* **10**, e0134581 (2015).
43. Andrews, S. 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. *Cold Spring Harbor Protocols*, (2016), pdb.prot093088.
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. Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* **27**, 573–580 (1999).
52. Alikhan, N.-F., Petty, N. K., Zakour, N. L. B. & Beatson, S. A. BLAST ring image generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* **12**, 402 (2011).
53. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
54. 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).
55. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermini, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
56. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **6**, 461–464 (1978).
57. Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
58. Tanabe, A. S. Kakusan: a computer program to automate the selection of a nucleotide substitution model and the configuration of a mixed model on multilocus data. *Mol. Ecol. Notes* **7**, 962–964 (2007).
59. Huelsenbeck, J. P. & Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755 (2001).
60. 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-5620009) 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](http://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