THE DNA BARCODES FOR THE SPECIES DELIMITATION OF THE GENUS *Tylopus* Jeekel, 1968 IN VIETNAM (Diplopoda: Polydesmida: Paradoxosomatidae)

Nguyen Duc Anh[*], **Nguyen Thi Thu Anh, Phung Thi Hong Luong,
Dang Thi Hoa, Nguyen Giang Son**

Institute of Ecology and Biological Resources, VAST, Vietnam

ABSTRACT

The 680 bp fragment of the COI gene was employed for DNA barcodes of the millipede genus *Tylopus* in Vietnam. A total of 22 samples representing for 14 morphological *Tylopus* species were analyzed. The K2P genetic divergence between *Tylopus* species ranges from 12.2% to 18.9% with a mean of 15 ± 1%. The intraspecific divergences are slightly different between species, from 3% to 5%. The AGBD analysis and phylogenetic trees also support 14 morphological species. It is also suggested to have more COI sequences of more species for better barcode reference library and better molecular species identification.

Keywords: Molecular taxonomy, biodiversity, COI gene, millipedes.

*Corresponding author email: ducanh410@yahoo.com

INTRODUCTION

Tylopus is one of the largest genera in the family Paradoxosomatidae Daday, 1889, with 64 valid species distributed in mountainous areas in Indochina and the southern part of China (Golovatch & Enghoff, 1993; Likhitrakarn et al., 2010; Nguyen, 2012). Taxonomy of the genus has been revised by Golovatch & Enghoff (1993) and Likhitrakarn et al. (2010), but species occurring in areas other than Thailand have yet been poorly studied.

To date, 20 *Tylopus* species have been recorded in Vietnam (Attems, 1937, 1953; Golovatch, 1984; Enghoff et al., 2004; Nguyen, 2012; Golovatch, 2019). These species have been found only from the type localities in the original descriptions. Since Vietnam is located in the distribution center of the genus *Tylopus*, the number of *Tylopus* species is expected to be higher than in present. Vietnamese *Tylopus* species can be distinguished by morphological characters, mainly gonopod conformation. Some species are differentiated in minor gonopod structure, for example *Tylopus phanluongi* Nguyen, 2012 and *Tylopus hilaris* (Attems, 1937) (Nguyen, 2012). Thus, molecular data are suggested to be used for species description and recognition. Unfortunately, there are no published reports on molecular data of the genus *Tylopus* except 16S rRNA of several *Tylopus* species (Nguyen et al., 2017).

Molecular based taxonomy has been recently used with sequence diversities in the Folmer segment of the mitochondrial gene, Cytochrome c oxidase subunit I. The COI barcodes have also been used to distinguish millipede species. For example, Wesener (2015) used the COI fragment to synonymize *Glomeris malmivaga* Verhoeff, 1912 with *Glomeris ornata* Koch, 1847. Zhao et al. (2020) used a COI fragment for additional data for new species description. The COI sequences are also used in many barcoding projects relating to myriapods, e.g. Barcoding Fauna Bavarica (Spelda et al., 2011). However,

the database or reference library for DNA barcodes are still limited. In other words, for the identification of a wide set of species, reference barcode libraries are needed. Therefore, more COI data for millipede species should be sequenced and published.

This work is devoted to provide the COI barcodes for species delimitation of the genus *Tylopus* in Vietnam, and furthermore to build up the barcoding reference library for the Vietnamese millipede fauna.

METHOD AND METARIALS

Taxon sampling, identification and DNA extraction

A total of 22 specimens was collected from various part of Vietnam and directly preserved in ethanol 85–90%. All were examined under microscope Olympus SZX10 to exactly identify members of the genus *Tylopus*. All materials and voucher specimens are kept in Department of Soil Ecology, Institute of Ecology and Biological Resources (IEBR, Vietnam), and Field Museum of Natural History (Chicago, USA).

Total genomic DNA was extracted from leg tissue using the DNAeasy Blood & Tissue Kit (Qiagen^{TM}).

DNA amplification and sequencing

A fragment of the mitochondrial gene, Cytochrome c Oxidase Subunit I (COI) was amplified using the polymerase chain reaction (PCR). The set of universal primers COI-1F20 (5'-ACTCTAATCATAAGG AT-3') and COI-1R19 (5'-TAAACCTCCGGGTGACCAA-3') (Nguyen et al., 2017) was used to amplify a 680 bp fragment of the COI gene. The PCR conditions were as following: denaturation at 94 °C for 5 min., 38 cycles of 94 °C for 45 seconds, 42 °C for 45 seconds, 72 °C for 1 minute, and final 72 °C for 5 min. After thermal cycling, 2 μl PCR products was screened for potentially successful amplification of a fragment of 16 S or COI by using electrophoresis in 1.2% Agarose- TBE 1X. The electrophoresis was performed in conditions of 100 mA, 120 V and 20 minutes.

About 20 µl of successfully amplified PCR products were purified using the ExosapIT with the company protocol. The purified PCR products were sent to sequence using an Applied Biosystems automatic sequencer (ABI3130 XL) at the Institute of DNA Technology (Vietnam) using the same primer sets as for initial PCR.

Alignment and phylogenetic analysis

Each successful sequence was checked and edited manually using BioEdit ver.7.1 (Hall, 1999), and confirmed using BLASTN 2.6.0+ searches (Zhang et al., 2000). All confirmed sequences were aligned using Cluster X ver.2.0 (Larkin et al., 2007) the ambiguous nucleotide sites and gaps were removed using MEGA X (Kumar et al., 2018).

Table 1. Analyzed species/specimens, deposition voucher, collection data and GenBank accession number

No.	Species	Locality	Voucher	COI	Notes
1	*Tylopus crassipes*	Sapa, Lao Cai	IEBR-Myr 92	KX096920	Nguyen et al. (2017)
2	*Tylopus hilaroides*	Cuc Phuong, Ninh Binh	IEBR-543	MW384914	In this study
3	*Tylopus hilaroides*	Cuc Phuong, Ninh Binh	IEBR-198	MW384918	
4	*Tylopus hilaroides*	Cuc Phuong, Ninh Binh	SVE-149	MW384905	
5	*Tylopus hilaroides*	Cuc Phuong, Ninh Binh	SVE-173	MW384904	
6	*Tylopus hilaroides*	Tam Dao, Vinh Phuc	SVE-55	MW384903	
7	*Tylopus nodulipes*	Huong Son, Ha Tinh	IEBR-105	MW384919	
8	*Tylopus nodulipes*	Minh Hoa, Quang Binh	IEBR-557	MW384912	
9	*Tylopus roseiparaterga*	Ba Vi, Ha Noi	SVE-70	MW384902	
10	*Tylopus roseiparaterga*	Tam Dao, Vinh Phuc	IEBR-185A	KX096923	Nguyen et al. (2017)
11	*Tylopus sapaensis*	Sa Pa, Lao Cai	IEBR-93	MW384908	In this study
12	*Tylopus spinisternus*	Bi Doup - Nui Ba, Lam Dong	IEBR-234	MW384916	
13	*Tylopus* sp.1	Ba Vi, Ha Noi	SVE-73	MW384901	
14	*Tylopus* sp.1	Ba Vi, Ha Noi	SVE-74	MW384900	
15	*Tylopus* sp.2	Phong Nha - Ke Bang, Quang Binh	IEBR-210	MW384917	
16	*Tylopus* sp.2	Phong Nha - Ke Bang, Quang Binh	IEBR-IPE6	MW384907	
17	*Tylopus* sp.3	Sa Pa, Lao Cai	IEBR-556	MW384913	
18	*Tylopus* sp.4	Son Dong, Bac Giang	IEBR-509	MW384915	
19	*Tylopus* sp.5	Hoang Lien, Lao Cai	IEBR-558	MW384911	
20	*Tylopus* sp.6	Tam Dao, Vinh Phuc	IEBR-603	MW384910	
21	*Tylopus* sp.7	Muong Nhe, Dien Bien	IEBR-617	MW384909	
22	*Tylopus* sp.8	Bach Ma, Thua Thien Hue	IEBR-740	MW384906	
23	*Oxidus gigas*	Sapa, Lao Cai	IEBR-Myr 113	KX096921	Nguyen et al. (2017)
24	*Oxidus gigas*	Duc Xuan, Ha Giang	IEBR-Myr 516	KX096928	
25	*Oxidus riukiaria*	Ryukyu, Japan	IEBR-H500	KX096926	
26	*Oxidus riukiaria*	Ryukyu, Japan	IEBR-H500J	KX096927	
27	*Oxidus gracilis*	Taiwan	IEBR-549	KX096931	
28	*Oxidus gracilis*	Ryukyu, Japan	IEBR-466	KX096924	
29	*Oxidus gracilis*	Ryukyu, Japan	IEBR-471	KX096925	

The reliability of the alignment was estimated using distance estimation and model of Kimura two parameters (K2P) performed in MEGA X (Kimura, 1980). The nucleotide frequencies were statistically calculated using MEGA X. The COI sequences were translated into amino acids for confirmation using transversion code in MEGA X.

The phylogenetic tree was reconstructed using the Maximum Likelihood (ML) analysis with the best model resulted from the ModelFinder (Kalyaanamoorthy et al., 2017) performed in IQTREE ver.1.6.2 for Windows (Minh et al., 2020). Models with the lowest BIC scores (Bayesian Information Criterion) are considered to describe the best substitution pattern. Codon positions included were $1^{st}+2^{nd}+3^{rd}+$Noncoding. Bayesian Inference (BI) analysis was performed using the MrBayes ver 3.2 (Ronquist et al., 2012) with 10 million generations, heating parameter of 0.06 and sampling every 1,000 generations.

The AGBD analysis was performed using the online server (https://bioinfo.mnhn.fr/abi/public/abgd/abgd web.html) to recognize the number of genetic groups. The parameters are default setting except the selected distance = Kimura80 and the relative gap with (X) = 1.0.

All specimens and DNA vouchers were deposited in the Department of Soil Ecology, Institute of Ecology and Biological Resources (IEBR), Vietnam Academy of Science and Technology (VAST), Hanoi, Vietnam.

All nucleotide sequences were deposited in GenBank with accession numbers. Collection localities, specimen vouchers, and GenBank accession numbers are presented in Table 1.

RESEARCH

Dataset statement

The final aligned dataset of the gene COI consists of 570 bp, and has nucleotide frequencies of 20.4, 42.1, 23.4 and 14.1 for A, T, G and C, respectively. The GC content accounts for 37% of total nucleotides. The COI dataset contains 157 (29.9%) parsimony informative and 177 (33.7%) variable sites.

Genetic distances

The K2P distance between the taxa ranges (*Tylopus* and *Oxidus* species) from 0% to 22.5% (Fig. 1); overall genetic distance is 16%. The mean interspecific distance between *Tylopus* species is about 15%±1%. The maximum divergence is 18.9% between *Tylopus roseiparaterga* and *Tylopus* sp.5, and the minimum distance is 12.2% between *Tylopus* sp.6 and *Tylopus* sp.7 (Table 2).
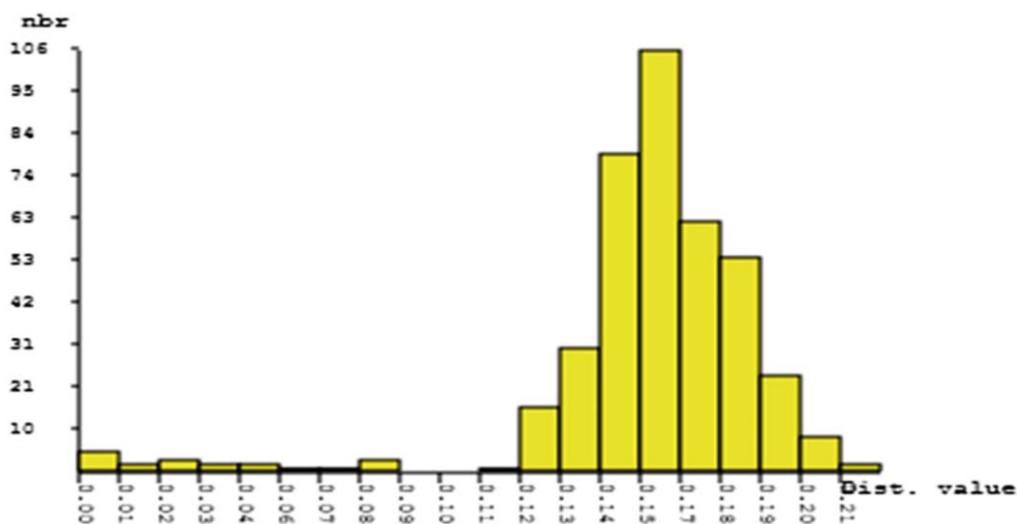
Figure 1. Interspecific/intraspecific COI variability (K2P): distance to nearest neighbor

Table 2. Estimates of Evolutionary Divergence over Sequence Pairs between *Tylopus* species (%)

Species	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Tylopus nodulipes (**1**)													
Tylopus hilaroides (**2**)	15.7												
Tylopus sp.2 (**3**)	17.2	17.4											
Tylopus spinisternus (**4**)	16.5	15.6	14.3										
Tylopus sp.4 (**5**)	13.8	15.7	15.7	13.9									
Tylopus sp.3 (**6**)	16.0	16.9	12.6	12.6	14.9								
Tylopus sp.5 (**7**)	15.8	15.5	14.8	16.0	15.3	14.5							
Tylopus sp.6 (**8**)	15.4	15.5	14.0	13.9	12.6	13.1	14.1						
Tylopus sp.7 (**9**)	15.6	16.6	14.5	15.4	17.0	16.7	15.4	12.2					
Tylopus sp.8 (**10**)	18.4	18.6	15.8	17.7	15.6	15.6	17.3	15.6	15.8				
Tylopus crassipes (**11**)	18.3	17.6	17.1	16.7	16.3	15.6	18.1	15.5	17.5	18.4			
Tylopus sapaensis (**12**)	17.5	14.9	15.1	17.4	16.7	17.9	15.6	14.7	15.6	16.9	17.9		
Tylopus roseiparaterga (**13**)	16.4	18.0	16.6	17.2	16.7	17.1	18.9	15.2	14.8	16.7	13.3	16.7	
Tylopus sp.1 (**14**)	17.9	15.2	14.7	15.8	16.3	15.6	16.2	14.1	14.6	18.5	17.0	13.1	16.2

Within the genus *Tylopus*, intraspecific divergences are slightly different between species, from 3% (for *Tylopus* sp.3) to 5% (for *Tylopus hilaroides* and *Tylopus* sp.2). Within the genus *Oxidus*, three species have significantly genetic distances, 13.2−14.1% between *O. gigas* and *O. gracilis*; 14.0−15.1% between *O. gigas* and *O. riukiaria*; 12.1−13.2% between *O. gracilis* and *O. riukiaria*. The AGBD analysis also indicates 17 groups with prior maximal distance P = 0.022–0.036. This result is more likely to satisfy with 17 morphological species including three *Oxidus* species.

Phylogenetic analysis

The phylogenetic trees were reconstructed for the gene COI using two methods of Maximum Likelihood (ML) and Bayesian Inference (BI). For ML analysis, we consider clades with bootstrap values below 65% to be weekly supported, between 65−85% to be moderately supported, and more than 85% to be strongly supported. For BI analysis, clades will be well supported if BI posterior probability is less than 0.65 bpp; be moderately supported if BI is between 0.65−0.85 bpp; and be well supported if BI is more than 0.85 bpp (Nguyen et al., 2017).

Both phylogenic trees also show the distinct separation of all *Tylopus* species and *Oxidus* species. All *Tylopus* species are clearly distinct from each other (Fig. 2). Two large linages are obviously separated with strong supports of bootstrap and BI values (99% and 1.00 bpp, respectively). The first clade includes samples of three species: *Tylopus hilaroides*, *T. sapaensis* and *Tylopus* sp.1. The separation of three species is well to moderately supported by bootstrap and BI values (80−90% and 0.74–0.9 bpp, respectively). The second clade consists of samples from 11 species. However, the relationship between species in the second clade is not well resolved (Fig. 2b) or very poorly supported by

bootstrap values (Fig. 2a). It seems not to have same common ancestor for 11 *Tylopus* species in the second clade. This problem is mainly due to lack of more species for deeper analysis. Therefore, it is suggested to have more samples from different species for further analysis.

DISCUSSION

The genus *Tylopus* is placed into the Sulciferini with typical characters of spiral solenophore completely sheathing solenomere, presence of postfemoral demarcation and processes (Jeekel 1968). Until now, no relationship analysis among sulciferine genera has been made, except only Golovatch & Enghoff (1993) reported the closer relationship among *Tylopus* species and *Oxidus gracilis*. They recommended that *Oxidus gracilis* or the genus *Oxidus* can be a sister of the genus *Tylopus* because both genera have similar gonopod characters, such as presences of lamina **l**, process **h**, spine **z**. The close relationship was also strongly supported by molecular evidences (Nguyen et al., 2017).

This is the first report on using the COI gene for species identification in the genus *Tylopus* and the family Paradoxosomatidae. Prior to this study, the COI distances were calculated using uncorrected p-distance for *Sphaerobelum* species (from 20.2% to 24.4%) (Zhao et al., 2020), *Glomeris* species (from 11.5% to 17.1%) (Wesener, 2015). In comparison with these researches, the genetic distance between *Tylopus* species (from 12.2% to 18.9%, with means of 15%) is narrower than that of Zhao et al. (2020) but relatively similar to that of Wesener (2015). Regarding phylogenetic analysis, all researches also indicated the COI has limitation due to rapidly evolved mitochondrial gene, and recommend to have more genes, especially of nuclear origin, for evolutionary analysis in addition to COI for a better understanding of evolution (Spelda et al., 2011).
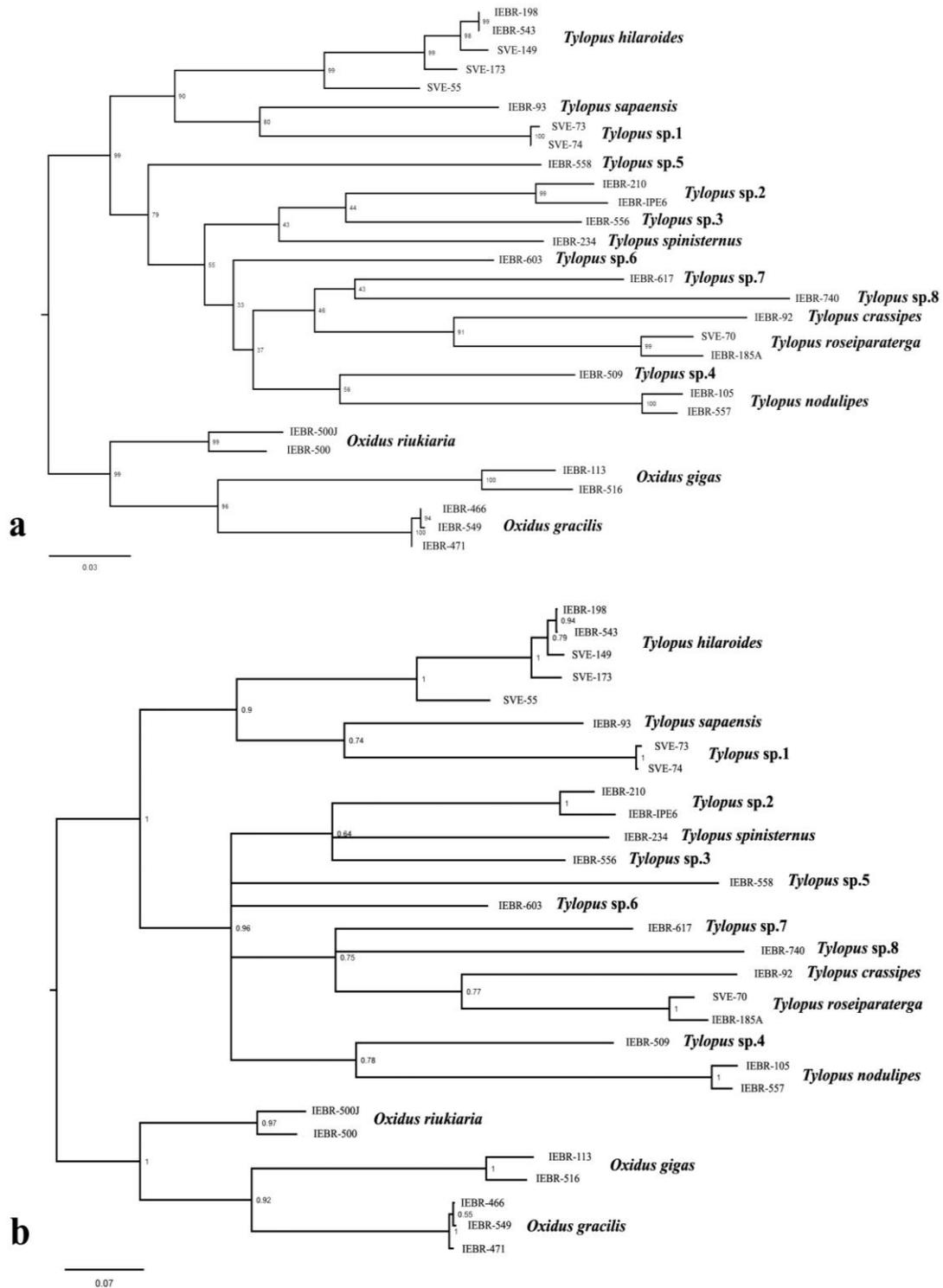
Figure 2. Phylogenetic diagram of the *Tylopus* species inferring from 570 bp COI fragment
a: Maximum Likelihood analysis; **b**: Bayesian Inference analysis.
Values at nodes show the bootstrap and BI

CONCLUSION

The COI barcodes reveals the genetic divergence of 14 *Tylopus* species. The mean interspecific distance is about 15 ± 1% (ranging from 12.2% to 18.9%). The intraspecific divergences are slightly different between species, from 3% to 5%.

A COI fragment is an efficient tool in helping to recognize the *Tylopus* species in particular and millipede species in general. However, a reference barcode library is needed and the genetic variation must be better known for the comprehensive molecular identification.

REFERENCES

Attems C., 1937. Myriapoda 3. Polydesmoidea I. Fam. Strongylosomidae. *Das Tierreich*, 68: 1–300.

Attems C., 1953. Myriopoden von Indochina. Expedition von Dr. C. Dawydoff (1938-1939). *Mémoires du Muséum National d'Histoire Naturelle*, {N. S., Sér. A, Zool.}5(3): 133–230.

Enghoff H., Golovatch S. I. and Nguyen D. A., 2004. A review of the millipede fauna of Vietnam (Diplopoda). *Arthropoda Selecta*, 13(1/2): 29–43.

Golovatch S. I., Enghoff H., 1993. Review of the genus *Tylopus*, with descriptions of new species from Thailand (Diplopoda, Polydesmida, Paradoxosomatidae). *Steenstrupia*, 19(3): 85–125.

Golovatch S. I., 1984. Contribution to the millipede fauna of Vietnam (Diplopoda), II. *Acta Zoologica Hungarica*, 30(1–2): 53–77.

Golovatch S. I., 2019. On several new or poorly-known Oriental Paradoxosomatidae (Diplopoda: Polydesmida), XXVII. *Arthropoda Selecta*, 28(4): 459–478.

Hall T. A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41: 95–98.

Jeekel C. A. W., 1968. On the classification and geographical distribution of the family Paradoxosomatidae (Diplopoda, Polydesmida). Bronder Offset Rotterdam, private.

Kalyaanamoorthy S., Minh B. Q., Wong T. K. F., von Haeseler A., Jermiin L. S., 2017. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14: 587–589. https://doi.org/10.1038/nmeth.4285

Kimura M., 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16: 111–120.

Kumar S., Stecher G., Li M., Knyaz C., Tamura K., 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*, 35(6): 1547–1549.

Larkin M. A., Blackshields G., Brown N. P., Chenna R., McGettigan P. A., McWilliam H., Valentin F., Wallace I. M., Wilm A., Lopez R., Thompson J. D., Gibson T. J., Higgins D. G., 2007. Clustal W and Clustal X version 2.0. *Bioinformatics*, 23: 2947−2948.

Likhitrakarn N., Golovatch S. I., Prateepasen R. & Panha S., 2010. Review of the genus *Tylopus* Jeekel, 1968, with descriptions of five new species from Thailand (Diplopoda, Polydesmida, Paradoxosomatidae). *ZooKeys*, 72: 23–68.

Minh B. Q., Schmidt H. A., Chernomor O., Schrempf D., Woodhams M. D., von Haeseler A., Lanfear R., 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era.

Molecular and Biological Evolution, 37: 1530–1534. https://doi.org/10.1093/molbev/msaa015

Nguyen A. D., 2012. *Tylopus* millipedes in Vietnam (Diplopoda: Polydesmida: Paradoxosomatidae: Sulciferini), with descriptions of five new species. *The Raffles Bulletin of Zoology*, 60(2): 289–311.

Nguyen A. D., Korsós Z., Jang K. H., Hwang U. W., 2017. A revision and phylogenetic analysis of the millipede genus *Oxidus* Cook, 1911 (Polydesmida: Paradoxosomatidae). *European Journal of Taxonomy*, 293: 1–22. https://doi.org/10.5852/ejt.2017.293

Ronquist F., Teslenko M., van der Mark P., Ayres D. L., Darling A., Höhna S., Larget B., Liu L., Suchard M. A. & Huelsenbeck J. P., 2012. MRBAYES 3.2: Efficient Bayesian phylogenetic inference and model selection across a large model space. *Systematic Biology*, 61: 539–542. https://doi.org/10.1093/sysbio/sys029

Spelda J., Reip H., Oliveira Biener U., Melzer R., 2011. Barcoding Fauna Bavarica: Myriapoda - a contribution to DNA sequence-based identifications of centipedes and millipedes (Chilopoda, Diplopoda). *ZooKeys*, 156: 123–139. https://doi.org/10.3897/zookeys.156.2176

Wesener T., 2015. No millipede endemics north of the Alps? DNA-Barcoding reveals *Glomeris malmivaga* Verhoeff, 1912 as a synonym of *G. ornata* Koch, 1847 (Diplopoda, Glomerida, Glomeridae). *Zootaxa*, 3999(4): 571–580. https://doi.org/10.11646/zootaxa.3999.4.7

Zhang Z., Schwartz S., Wagner L., Miller W., 2000. A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology*, 7(1–2): 203–214.

Zhao Y., Yu J. & Liu W., 2020. A molecular-based phylogeny of the millipede genus *Sphaerobelum* Verhoeff, 1924, with the first record of the genus from mainland China (Diplopoda: Sphaerotheriida: Zephroniidae). *Annales de la Société entomologique de France* (N.S.), 56(4): 341–348. https://doi.org/10.1080/00379271.2020.1811153