Diversity and Distributions WILEY

The first comprehensive DNA barcode reference library of Chinese *Tanytarsus* (Diptera: Chironomidae) for environmental DNA metabarcoding

Xiao-Long Lin[1,2] 🆔 | Lidong Mo[3] | Wen-Jun Bu[1] | Xin-Hua Wang[1]

[1]College of Life Sciences, Nankai University, Tianjin, China

[2]Department of Natural History, NTNU University Museum, Norwegian University of Science and Technology, Trondheim, Norway

[3]Institute of Integrative Biology, ETH Zurich (Swiss Federal Institute of Technology), Zurich, Switzerland

Correspondence
Lidong Mo, Institute of Integrative Biology, ETH Zurich (Swiss Federal Institute of Technology), Zurich, Switzerland.
Email: lidong.mo@usys.ethz.ch

Editor: Jong Seong Khim

Abstract

Aim: Reliable DNA barcode reference libraries are critical for biodiversity monitoring based on Environmental DNA (eDNA). DNA barcoding has proven successful in species delimitation and association of life stages in many groups of animals, such as non-biting midges (Chironomidae). We use the genus *Tanytarsus* van der Wulp, 1874, second largest genus of the family Chironomidae, to test whether DNA barcodes can be used to identify unknown and cryptic species. A comprehensive DNA barcode library will be built to facilitate further eDNA studies for biodiversity monitoring and conservation in freshwater ecosystems.

Location: Mainland China.

Methods: In this study, the cytochrome *c* oxidase subunit 1 (COI) DNA barcodes from 298 individuals, representing 56 morphospecies from China, were analysed, including 136 sequences from previous studies and 162 newly generated DNA barcodes. 298 COI barcodes were applied into the Automatic Barcode Gap Discovery (ABGD) to compare the operational taxonomic units (OTU) number resulting from the Barcode Index Number (BIN) based "Barcode Gap Analysis" and neighbour joining (NJ) tree.

Results: Total 298 barcode sequences were applied to 66 BINs, including 20 singleton BINs and 46 concordant BINs. ABGD yielded 55 OTUs with a prior intraspecific divergence of $P_{max} = 3.59\%$. By Integrating DNA barcodes with morphology, we revealed 21 putative species new to science and two unidentified species from the 56 well-separated morphospecies. Besides, several cryptic species are uncovered using the NJ tree and TCS haplotype network based on COI DNA barcodes.

Main conclusions: Our results demonstrate that DNA barcodes enable the discovery and identification of *Tanytarsus* non-biting midges. A comprehensive DNA barcode reference library could speed up accurate species delimitation and discovery by providing a reliable resource for taxonomic studies, and improve the application of eDNA metabarcoding in biomonitoring of freshwater ecosystems.

KEYWORDS
bioindicators, China, cryptic diversity, DNA barcoding, species delimitation

1 | INTRODUCTION

The ongoing climate change and human disturbances lead to a continuous decline of global biodiversity (Butchart et al., 2010), which greatly impacts the sustainability of our planet (Díaz et al., 2006). Efforts to conserve biodiversity at regional, national and international scales (Butchart et al., 2010; UNEP, 2011) depend on precise knowledge on taxonomy and species distributions (Kelly et al., 2014; Thomsen & Willerslev, 2015). As such, conservation practices are still limited by the poor understanding of biodiversity in numerous un-described taxa from massive geographical regions (Vié et al., 2009). Biological monitoring in aquatic ecosystems is particularly dependent on taxonomic expertise (Kelly et al., 2014). To overcome these challenges, DNA-based approaches provide effective tools to identify aquatic organisms. DNA barcodes—standardized genetic markers for species identification (Hebert, Cywinska, & Ball, 2003; Hebert, Ratnasingham, & Waard, 2003)—have become the most useful method for estimating species diversity and revealing cryptic species diversity (Hajibabaei et al., 2016; Hebert et al., 2004; Kanturski et al., 2018; Young et al., 2019). Due to the recent advances in sequencing technologies and bioinformatics, DNA metabarcoding (Tab erlet et al., 2012; Yu et al., 2012) based on organismal and environmental DNA (eDNA) (Macher et al., 2018; Majaneva et al., 2018) is increasingly used for biodiversity assessments and biomonitoring of aquatic biota (Elbrecht & Steinke, 2019; Elbrecht et al., 2017; Morinière et al., 2017; Sun et al., 2019; Valentini et al., 2016). Coupled with sensitive, cost-efficient and ever-advancing DNA sequencing technologies, eDNA metabarcoding is thus critical for biodiversity monitoring and environmental policy formulation (Kelly et al., 2014). Within a single standardized sample of freshwater and metabolic wastes, eDNA from entire communities across various taxonomic groups can be analysed simultaneously by comparison with barcode libraries (Deiner et al., 2017; Kelly et al., 2014). Yet, these biodiversity assessments are still highly limited by the availability of reference barcode libraries. More comprehensive and taxa specific barcode libraries are thus integral to monitoring and conserving aquatic biodiversity. In Europe, several freshwater macroinvertebrate DNA barcode libraries on a national level have been built (Leese et al., 2018; Morinière et al., 2017, 2019; Weigand et al., 2019). However, the high diversity of freshwater macroinvertebrate species (e.g. caddisflies and chironomids) occurs in Asia, and only a minor portion of this diversity has been barcoded to date (Lin et al., 2015; Zhou et al., 2007).

Chironomids (Diptera: Chironomidae) are the most abundant species among benthic invertebrates (Armitage et al., 1995; Ferrington, 2008). Chironomidae comprise more than 6,300 accepted species (P. Ashe pers comm.) and occur in all zoogeographical regions, even Antarctica. As a major component of biodiversity, Chironomids are not only important bioindicators for freshwater ecosystems monitoring (Ferrington, 2008), but also valuable sources for phylogenetic and biogeographical studies (Brundin, 1966; Ekrem et al., 2018; Lin et al., 2018c; Silva & Ekrem, 2016). Because it is difficult to identify Chironomidae species by traditional morphological approaches, the actual species number is likely to highly exceed the currently estimated number of 6,300 species. Accurate species identifications are thus urgently needed as a stepping stone to facilitate evolutionary and biodiversity studies.

Tanytarsus van der Wulp, 1874 (Figure 1) is the second most species-rich genus within Chironomidae, with more than 400 accepted named species and 39 of them from China. The genus *Tanytarsus* was created by van der Wulp (1874) based on the type species *Tanytarsus signatus* (van der Wulp, 1859). Various species groups

FIGURE 1 Body photographs of the genus *Tanytarsus* in ethanol. (a) larva of *Tanytarsus simantoteuus* Sasa, Suzuki & Sakai, 1998; (b) pupa of *Tanytarsus simantoteuus* Sasa, Suzuki & Sakai, 1998; (c) adult male of *Tanytarsus formosanus* Kieffer, 1912

and species of *Tanytarsus* have been revised over the past decades (e.g. Cranston, 2000; Ekrem, 2001a; Ekrem, 2001b; Ekrem, 2002; Ekrem et al., 2003; Giłka & Paasivirta, 2009; Glover, 1973; Lin et al., 2018a; Reiss & Fittkau, 1971; Sanseverino, 2006; Sasa, 1980; Sasa, 1983; Sasa & Kawai, 1987), and their evolutionary history has been clarified recently by a molecular phylogenetic study based on multiple loci (Lin et al., 2018c). However, there remain challenges limiting the taxonomic disentangling of *Tanytarsus* in China: (a) there are still many poorly investigated regions, such as eastern Palearctic and Oriental regions; (b) life stages are incomplete as rearing is difficult; and (c) species boundaries remain uncertain in morphologically determined species. Although a few public DNA barcodes of *Tanytarsus* from China were published in previous studies (Lin et al., 2015, 2018a), the reference library still remains incomplete.

In this study, we selected 298 individuals of 56 morphological *Tanytarsus* species collected from China to test the utility of the cytochrome *c* oxidase subunit 1 (COI) barcodes for species delimitation in the species-rich genus *Tanytarsus*.

2 | MATERIALS AND METHODS

2.1 | Taxon sampling and taxonomy

Field sampling was conducted in China from 2008 to 2019 using methods, such as malaise trap, light trap, sweeping net, drift nets and kick-nets. Sampled specimens (Table S1) were preserved in ethanol (85% for adults and 95% for immatures) and stored at 4°C in the dark until further processing. Selected samples were dissected under a stereomicroscope and identified under a compound microscope using available references (e.g. Ekrem, 2002; Kawai, 1991; Lin et al., 2018a; Sasa, 1983; Sasa & Kawai, 1987). Tissue samples (three legs or head-thorax) were used for genomic DNA extraction. Afterwards, the exoskeleton of head-thorax was cleansed and mounted in Euparal® on the same microscope slide as its corresponding body parts (Sæther, 1969).

2.2 | Molecular laboratory protocols

We divided each specimen into two parts for molecular experiments and analysis. **Part 1**. DNA of most specimens was extracted using Qiagen DNA Blood and Tissue Kit at the College of Life Sciences, Nankai University and Department of Natural History, NTNU University Museum. PCR amplifications of COI barcodes with the universal primers LCO1490 and HCO2198 (Folmer et al., 1994) were mimicked following the protocol from Lin et al. (2018a). Sanger sequencing of the purified PCR products was carried on the ABI 3730 at the BGI (Beijing, China) and BigDye 3.1 at the MWG Eurofins (Ebersberg, Germany). **Part 2**. DNA extraction, PCR amplification and high-throughput sequencing of the specimens were executed at the Canadian Centre for DNA Barcoding (CCDB, University of Guelph, Canada) using standard high-throughput protocols (deWaard et al., 2008).

2.3 | Data analysis

Raw sequences were edited and assembled in SeqMan (DNASTAR), and aligned using MUSCLE (Edgar, 2004) implemented in MEGA X (Kumar et al., 2018) to check stop codons.

In addition to the novel data provided here, we searched available public *Tanytarsus* records from China in the Barcode of Life Data System (BOLD, http://www.boldsystems.org/) (Ratnasingham & Hebert, 2007). In total, 298 records (Figure S1; Table S1) with COI barcode of Chinese *Tanytarsus* were assigned to our present database (16 April 2020). A distribution map of the 298 records from China was created using the packages "ggplot2" (Wickham et al., 2018) and "mapdata" (Becker & Wilks, 2019) with R (version 3.5.2).

The "Barcode Gap Analysis" tool on BOLD was used to calculate sequence divergences for the present dataset, including the mean and maximum pairwise distances for intraspecific divergences, the mean and minimum pairwise distances for interspecific divergences, and minimum genetic distances to the nearest neighbour.

In addition, all 298 barcode sequences were applied to the Barcode Index Number (BIN) system on BOLD. The BIN system clusters DNA barcodes to generate OTUs with a threshold of 2.2% for a rough differentiation between interspecific and intraspecific genetic distances (Ratnasingham & Hebert, 2013).

A neighbour joining (NJ) (Saitou & Nei, 1987) tree was constructed based on the aligned dataset using the Kimura 2-Parameter (K2P) model (Kimura, 1980) with 1,000 bootstrap replicates and pairwise deletion in MEGA X.

2.4 | OTU delineation based on DNA barcodes using Automatic Barcode Gap Discovery (ABGD)

Compared to distance-based methods, the single-locus and tree-based delineation methods, such as the Generalized Mixed Yule Coalescent (GMYC) (Pons et al., 2006) and Bayesian Poisson tree processes (PTP) approaches (Zhang et al., 2013), are suspected to cause oversplitting (Lin et al., 2018b; Pentinsaari et al., 2017; Talavera et al., 2013). As a distance-based method, Automatic Barcode Gap Discovery (ABGD) (Puillandre et al., 2012) partitions sequences into groups based on comparisons of pairwise distances with higher reliability. Since ABGD provides a more reliable approach, our data were applied into the ABGD to compare the OTU number resulting from the BIN-based "Barcode Gap Analysis" and NJ-tree. ABGD analysis was carried out on 17 April 2020 using the web interface (https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html). We used the K2P model and applied X = 1.0 as relative barcode gap width, and kept default settings for remaining parameters.
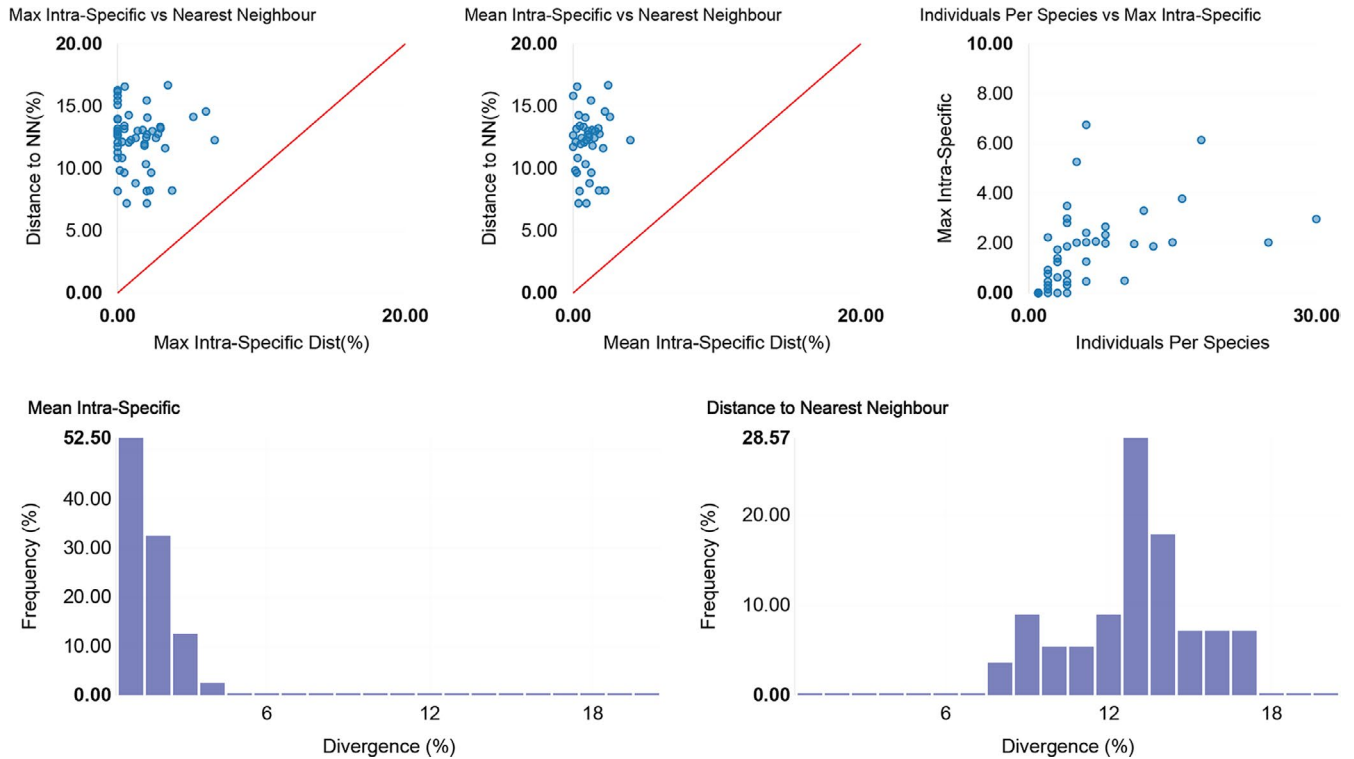
FIGURE 2 Barcode Gap Analysis of 298 COI barcode sequences of 56 *Tanytarsus* species from China. Two distance distribution histograms show the mean intraspecific divergence and distances to nearest neighbour. Three scatter plots are provided to confirm the existence and magnitude of the Barcode Gap. The first two scatter plots show the overlap of the max and mean intraspecific distances vs. the interspecific (nearest neighbour) distances. The third scatter plot plots the number of individuals in each species against their max intraspecific distances, as a test for sampling bias

3 | RESULTS

3.1 | DNA barcode analysis

The aligned 298 sequences ranged from 607 to 658 base pairs, including 211 sequences with a full barcode length of 658 base pairs. These 298 sequences were assigned to 66 BINs, which included 20 singleton BINs and 46 concordant BINs. 23 new BINs were added to BOLD, most represent new species entries for the system (Figure 2; Table S1). 20 BINs representing 20 species are singleton. The following ten species were represented by two BINs: *Tanytarsus biwatrifurcus* Sasa & Kawai, 1987, *Tanytarsus formosanus* Kieffer, 1912, *Tanytarsus sinuatus* Goetghebuer, 1936, *Tanytarsus* sp. 1XL, *Tanytarsus* sp. 3XL, *Tanytarsus* sp. 4XL, *Tanytarsus* sp. 17XL, *Tanytarsus tamagotoi* Sasa, 1983, *Tanytarsus tongmuensis* Lin et al., 2018, *Tanytarsus yunosecundus* Sasa, 1984. The mean intraspecific divergence of all species was 0.76% with a maximum of 6.75% in *T. biwatrifurcus*. The mean interspecific divergence was 12.4% with the minimum 7.2% between *Tanytarsus oscillans* Johannsen, 1932 and *Tanytarsus unagiseptimus* Sasa, 1985 (Figure 2).

3.2 | Species discrimination

The neighbour joining tree based on 298 COI DNA barcodes comprises 56 well separated clusters representing 33 named and 23

unnamed species. Among the 23 unnamed species, 21 putative species might be new to science, and the remaining two are unidentified. In general, the results showed a preferable matching pattern between DNA barcode clusters and morphospecies in *Tanytarsus*.

3.3 | Species without identification

As the larvae of some *Tanytarsus* species are not described and the females are difficult to differentiate, a larva of *Tanytarsus* sp. 31XL (BOLD:AEC3486) and a larva and a female of *Tanytarsus* sp. 19XL (BOLD:ADC5364) could not be morphologically identified to the species level.

3.4 | Cryptic species diversity

The DNA barcodes we generated in this study have revealed several potential cryptic species in the *Tanytarsus chinyensis*-, *mcmillani*-, *norvegicus*-, *signatus*-, *tamakutibasi*-, *thaicus*- and *triangularis* species groups, and some other species (Figure S2 and Table S1). For instance, the NJ tree and TCS haplotype network based on 31 DNA barcode sequences of the *Tanytarsus tamakutibasi* species group revealed seven OTUs (Figure 3). Four putative cryptic species (*Tanytarsus* sp. 1XL, *Tanytarsus* sp. 14XL, *Tanytarsus* sp. 25XL

FIGURE 3 Genetic analyses of 31 COI barcodes of the *Tanytarsus tamakutibasi* species group from China. (a) Neighbour joining tree of *T. tamakutibasi* group based on K2P distance; numbers on branches represent bootstrap based on 1,000 replicates; scale represents K2P genetic distances. (b). TCS haplotype network based the COI barcodes of the *T. tamakutibasi* group. Mutations are shown as lines on the branches
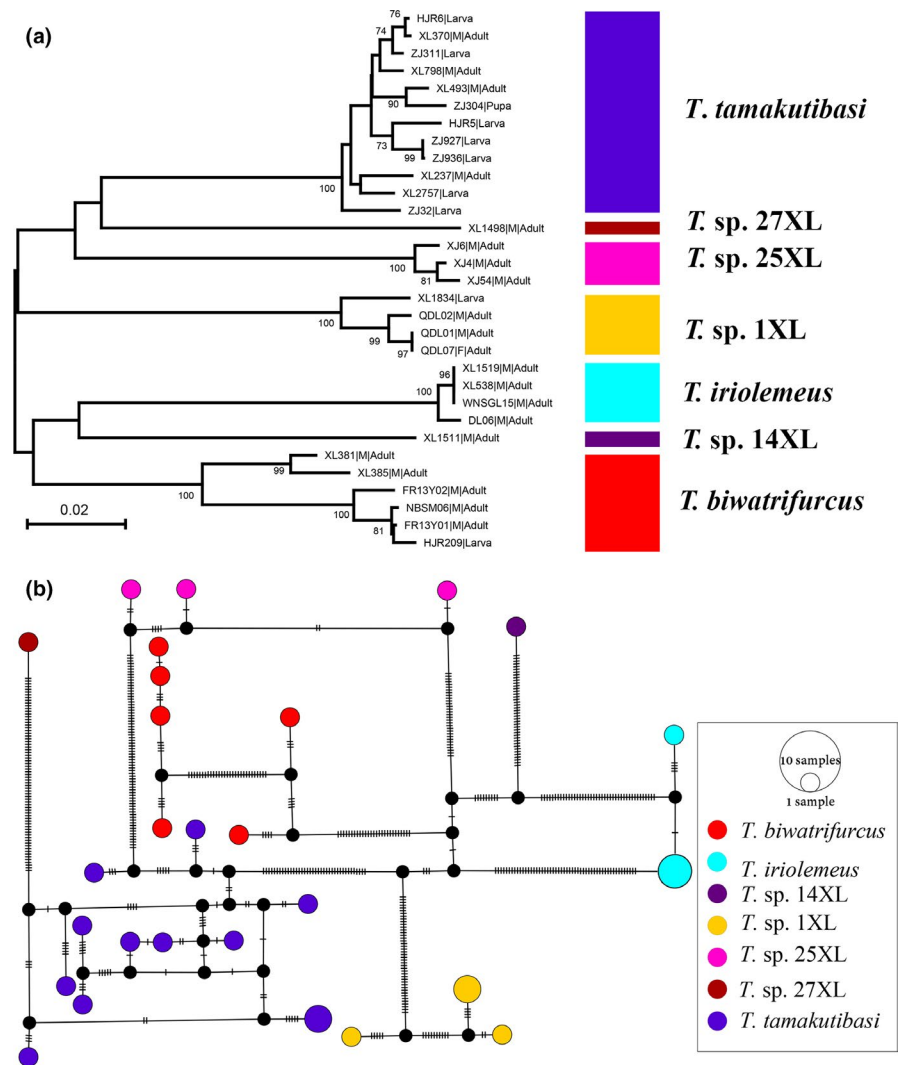
and *Tanytarsus* sp. 27XL) are closely related to *T. biwatrifurcus*, *T. iriolemeus* Sasa & Suzuki, 2000 and *Tanytarsus tamakutibasi* Sasa, 1983 with highly similar hypopygia. Among these seven species, five were well differentiated by multiple nuclear markers in the previous molecular phylogeny of *Tanytarsus* (Lin et al., 2018c, Figure 3). Since immatures were associated with adults by DNA barcodes, a few morphological diagnoses of species within the *T. tamakutibasi* species group were detected.

Additionally, similar cases were observed in the *Tanytarsus thaicus* and *Tanytarsus triangularis* species groups. Within the *T. thaicus* species group, three genetic lineages are well supported by the NJ tree and TCS haplotype network based on COI barcodes (Figure 4). These two potential cryptic species, *Tanytarsus* sp. 6XL and *Tanytarsus* sp. 17XL, are closely related to *Tanytarsus thaicus* Moubayed, 1990 based on morphology. *T.* sp. 17XL could be separated from *T. thaicus* by additional loci (Lin et al., 2018c, Figure 1). Within the *T. triangularis* species group, *Tanytarsus* sp. 20XL, *Tanytarsus simantoteuus* Sasa, Suzuki & Sakai, 1998 and *Tanytarsus tobaoctadecimus* Kikuchi & Sasa, 1990 are difficult to separate morphologically. DNA barcodes of the above three species were clustered into three OTUs in both the NJ tree and the haplotype network (Figure 5), indicating

T. sp. 20XL as a potential cryptic species. After examination of these vouchers, some diagnostic characters were found in immatures and adult males associated with DNA barcodes. Moreover, the above three species have also been confirmed by additional multiple nuclear markers (Lin et al., 2018c, Figure 1).

3.5 | Life stages association

In this study, the larvae of 24 species, pupae of three species and adult females of five species were associated with their adult males by DNA barcodes (Figure S2; Table S1). Among above cases, the larvae of 22 species, pupae of two species and adult females of four species have not been described yet.

3.6 | OTU delineation based on DNA barcodes using Automatic Barcode Gap Discovery (ABGD)

As some morphospecies showed comparably high intraspecific divergence, no definite "barcode gap" was observed based on pairwise
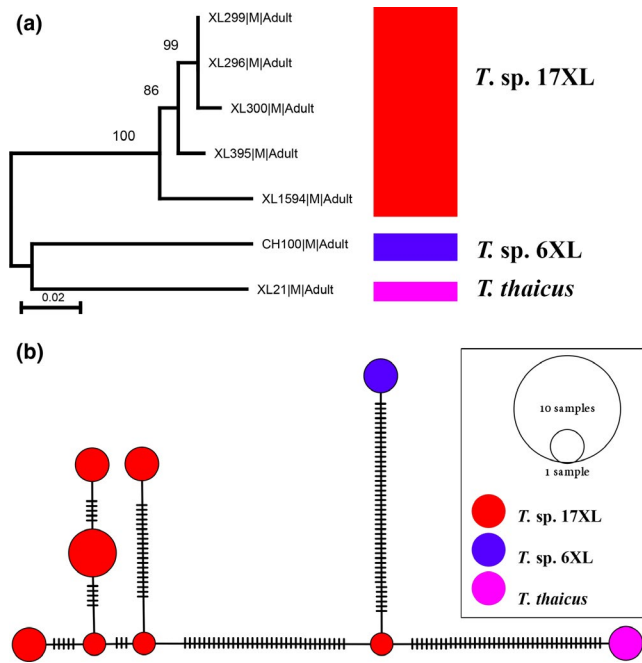
FIGURE 4 Genetic analyses of seven COI barcodes of the *Tanytarsus thaicus* species group from China. (a) Neighbour joining tree of *T. thaicus* group based on K2P distance; numbers on branches represent bootstrap based on 1,000 replicates; scale represents K2P genetic distances. (b). TCS haplotype network based the COI barcodes of the *T. thaicus* group. Mutations are shown as lines on the branches

distance (Figure 6). ABGD analysis of the present dataset recognized 55 OTUs with a prior intraspecific divergence of P_{max} = 3.59%. Minor discrepancies were observed in *Tanytarsus okuboi* Sasa & Kikuchi, 1986 and *T.* sp. 4XL, which was clustered in one group in ABGD, but recognized as two OTUs in both the BIN system and the NJ tree.

4 | DISCUSSION

4.1 | *Tanytarsus* DNA barcode reference library

In this study, we investigated the taxonomy of Chinese *Tanytarsus* non-biting midges and generated a DNA barcode reference library. This library includes 298 records for 66 BINs representing 56 putative species. Until now, 33 species with a Linnaean name, and the remaining 23 species lack of a Linnaean name including 21 new species to science and two unidentified species. Overall, DNA barcode coverage is nearly complete for known *Tanytarsus* from China. Our results also suggest a high incidence of cryptic species diversity, indicating that the number of species in *Tanytarsus* is likely to be much higher than previously recognized. To date, 99 species within *Tanytarsus* have been recorded in Japan (Yamamoto & Yamamoto, 2014). Although 56 morphological species of *Tanytarsus* have been barcoded here, wider sampling of *Tanytarsus* is needed for building a more comprehensive DNA barcode reference library.

Previous studies have shown the great potential of macroinvertebrate metabarcoding for biodiversity assessments of freshwater ecosystems (Elbrecht & Steinke, 2019; Elbrecht et al., 2017). Compared with classical morphological approaches, biodiversity assessments based on metabarcoding will reduce costs and time by generating millions of DNA fragments from thousands of specimens at a time. However, few barcode libraries for macroinvertebrates have been built in China, which are crucial for biomonitoring. The present study is an important contribution to build a more comprehensive DNA barcode library for Chinese chironomids.

4.2 | Taxonomy

DNA barcodes have proven useful in differentiating chironomid species and providing novel insight into their taxonomy (Giłka et al., 2018; Lin et al., 2015; Lin et al., 2018a; Song et al., 2018; Stur et al., 2019). We show that DNA barcoding is an efficient tool for species delimitation and recognition of cryptic species in *Tanytarsus*. A number of immatures of some species associated with DNA barcodes are discovered for the first time as putative new species. In general, the NJ tree and ABGD based on DNA barcodes yielded concordant OTUs corresponding to morphospecies. With a lower threshold on BOLD, BIN yields a bit higher OTUs. However, using a single locus such as COI to differentiate species can cause an overestimation of the species diversity due to deep intraspecific divergence (Lin et al., 2015, 2018b) or an underestimation of species diversity due to recent speciation. Incomplete lineage sorting (Willyard et al., 2009), horizontal gene flow (Naciri et al., 2012), recent speciations and insufficient taxon sampling (Luo et al., 2015; Zhang et al., 2010), can lead to inaccurate species delimitation. Therefore, an integrative taxonomy with molecular, morphological and ecological data is crucial to explore species boundaries (Galimberti et al., 2012; Hebert et al., 2004; Schutze et al., 2017; Zheng et al., 2020). For instance, species boundaries of closely related species within two *Tanytarsus* species groups were explored based on both morphology and DNA, and eight cryptic species were uncovered and described (Lin et al., 2018a, 2018b). Although we reveal many putative new species by DNA barcodes here, a further taxonomic study together with nuclear markers and morphology is still needed to diagnose and describe these putative new taxa.

5 | CONCLUSION

Our results demonstrate that DNA barcodes enable the discovery and identification of cryptic new species of *Tanytarsus* non-biting midges. Our DNA barcode reference library of Chinese *Tanytarsus* included 298 records for 66 BINs representing 56 putative species, indicating 21 putative new species. A comprehensive DNA barcode reference library will speed up accurate species delimitation and discovery by providing a reliable reference resource for taxonomic studies. Furthermore, eDNA metabarcoding will be improved in biomonitoring of freshwater ecosystem with more integrative DNA barcode reference libraries.

FIGURE 5 Genetic analyses of 20 COI barcodes of the *Tanytarsus triangularis* species group from China. (a) Neighbour joining tree of *T. triangularis* group based on K2P distance; numbers on branches represent bootstrap based on 1,000 replicates; scale represents K2P genetic distances. (b). TCS haplotype network based the COI barcodes of the *T. triangularis* group. Mutations are shown as lines on the branches
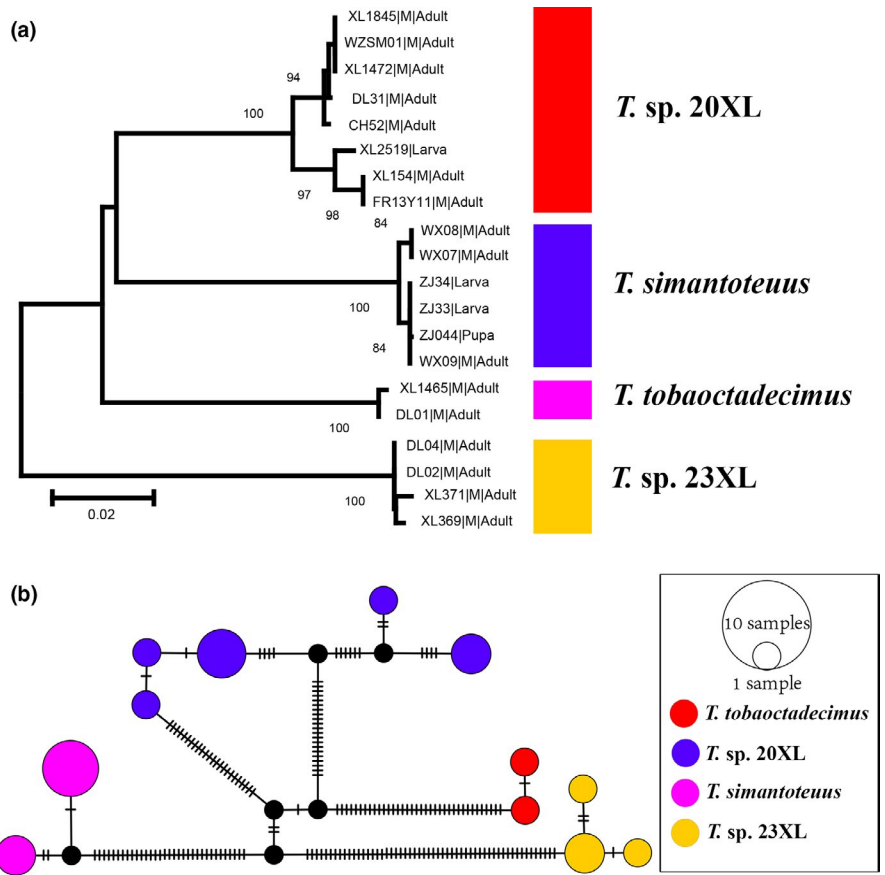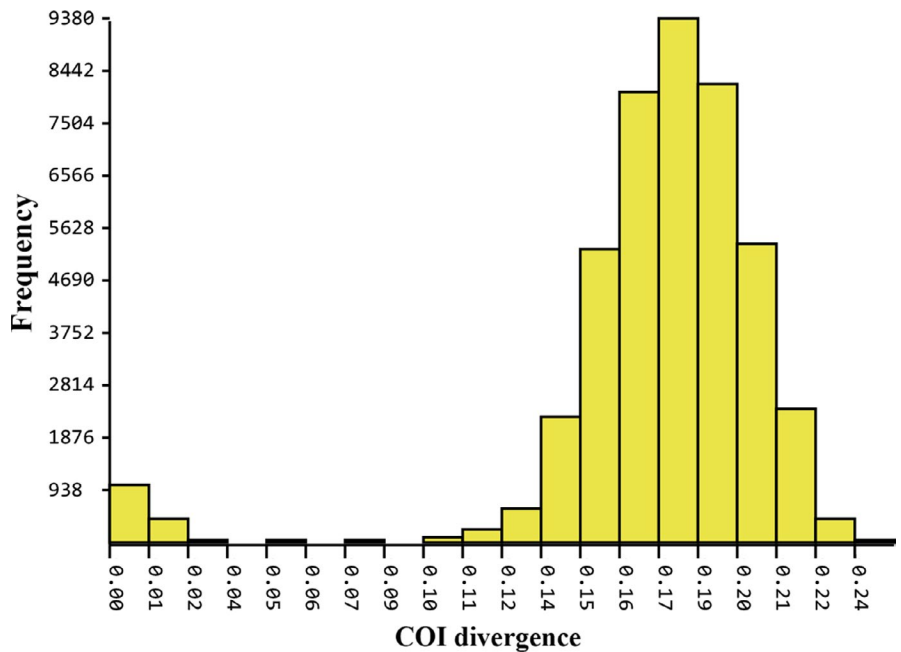

FIGURE 6 Histogram of pairwise K2P distances of 298 COI barcodes of Chinese *Tanytarsus*

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ddi.13209.

DATA AVAILABILITY STATEMENT

The list of all specimen records, images, trace files and collateral information is publicly accessible on BOLD through the dataset "DS-2020TACH", DOI: dx.doi.org/10.5883/DS-2020TACH.

ORCID

Xiao-Long Lin [iD] https://orcid.org/0000-0001-6544-6204

REFERENCES

Armitage, P. D., Cranston, P. S., & Pinder, L. C. V. (1995). *The chironomidae: Biology and ecology of non-biting midges* (XII, p. 572).Chapman & Hall.

Becker, R., & Wilks, A. (2019). R. Version by Ray Brownrigg. Mapdata: Extra Map Databases. R package version 2.3.0: Updated at: https://CRAN. R-project. org/package= mapdata, accessed.

Brundin, L. (1966). Transantarctic relationships and their significance, as evidenced by chironomid midges with a monograph of the subfamilies Podonominae and Aphroteniinae and the austral Heptagynae. *Kongliga Svenska Vetenskaps Academiens Nya Handlingar*, 11, 1–472.

Butchart, S. H. M., Walpole, M., Collen, B., van Strien, A., Scharlemann, J. P. W., Almond, R. E. A., Baillie, J. E. M., Bomhard, B., Brown, C., Bruno, J., Carpenter, K. E., Carr, G. M., Chanson, J., Chenery, A. M., Csirke, J., Davidson, N. C., Dentener, F., Foster, M., Galli, A., ... Watson, R. (2010). Global biodiversity: Indicators of recent declines. *Science*, 328(5982), 1164–1168. https://doi.org/10.1126/science.1187512

Cranston, P. S. (2000). Monsoonal tropical *Tanytarsus* van der Wulp (Diptera: Chironomidae) reviewed: New species, life histories and significance as aquatic environmental indicators. *Australian Journal of Entomology*, 39(3), 138–159. https://doi.org/10.1046/j.1440-6055.2000.00170.x

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., & De Vere, N. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, 26(21), 5872–5895. https://doi.org/10.1111/mec.14350

deWaard, J. R., Ivanova, N. V., Hajibabaei, M., & Hebert, P. D. N. (2008). Assembling DNA barcodes: Analytical protocols. In M. Cristofre (Ed.), *Methods in molecular biology: environmental genetics* (pp. 275–293). Humana Press Inc.

Díaz, S., Fargione, J., Chapin, F. S. III, & Tilman, D. (2006). Biodiversity loss threatens human well-being. *PLoS Biology*, 4(8), e277. https://doi.org/10.1371/journal.pbio.0040277

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Ekrem, T. (2001a). Diagnoses and immature stages of some Australian *Tanytarsus* van der Wulp (Diptera : Chironomidae). *Australian Journal of Entomology*, 40, 312–325. https://doi.org/10.1046/j.1440-6055.2001.00246.x

Ekrem, T. (2001b). A Review of Afrotropical *Tanytarsus* van der Wulp (Diptera: Chironomidae). *Tijdschrift Voor Entomologie*, 144(1), 5–40. https://doi.org/10.1163/22119434-99900054

Ekrem, T. (2002). A review of selected South- and East Asian *Tanytarsus* v.d. Wulp (Diptera: Chironomidae). *Hydrobiologia*, 474(1–3), 1–39. https://doi.org/10.1023/A:1016527603086

Ekrem, T., Stur, E., Orton, M. G., & Adamowicz, S. J. (2018). DNA barcode data reveal biogeographic trends in Arctic non-biting midges. *Genome*, 61(11), 787–796. https://doi.org/10.1139/gen-2018-0100

Ekrem, T., Sublette, M. F., & Sublette, J. E. (2003). North American *Tanytarsus* I. Descriptions and Keys to Species in the *eminulus*, *gregarius*, *lugens* and *mendax* Species Groups (Diptera : Chironomidae). *Annals of the Entomological Society of America*, 96(3), 265–328. https://doi.org/10.1603/0013-8746(2003)096[0265:Natida]2.0.Co;2

Elbrecht, V., & Steinke, D. (2019). Scaling up DNA metabarcoding for freshwater macrozoobenthos monitoring. *Freshwater Biology*, 64(2), 380–387. https://doi.org/10.1111/fwb.13220

Elbrecht, V., Vamos, E. E., Meissner, K., Aroviita, J., & Leese, F. (2017). Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution*, 8(10), 1265–1275. https://doi.org/10.1111/2041-210X.12789

Ferrington, L. C. (2008). Global diversity of non-biting midges (Chironomidae; Insecta-Diptera) in freshwater. *Hydrobiologia*, 595, 447–455. https://doi.org/10.1007/s10750-007-9130-1

Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3(5), 294–299.

Galimberti, A., Romano, D. F., Genchi, M., Paoloni, D., Vercillo, F., Bizzarri, L., Sassera, D., Bandi, C., Genchi, C., Ragni, B., & Casiraghi, M. (2012). Integrative taxonomy at work: DNA barcoding of taeniids harboured by wild and domestic cats. *Molecular Ecology Resources*, 12(3), 403–413. https://doi.org/10.1111/j.1755-0998.2011.03110.x

Giłka, W., & Paasivirta, L. (2009). Evaluation of diagnostic characters of the *Tanytarsus chinyensis* group (Diptera: Chironomidae), with description of a new species from Lapland. *Zootaxa*, 2197, 31–42. https://doi.org/10.5281/zenodo.189527

Giłka, W., Paasivirta, L., Gadawski, P., & Grabowski, M. (2018). Morphology and molecules say: *Tanytarsus latens*, sp. nov. from Finland (Diptera: Chironomidae). *Zootaxa*, 4471(3), 569–579. https://doi.org/10.11646/zootaxa.4471.3.8.

Glover, B. (1973). The Tanytarsini (Diptera: Chironomidae) of Australia. *Australian Journal of Zoology Supplementary Series*, 21(23), 403–478. https://doi.org/10.1071/AJZS023

Hajibabaei, M., Baird, D. J., Fahner, N. A., Beiko, R., & Golding, G. B. (2016). A new way to contemplate Darwin's tangled bank: How DNA barcodes are reconnecting biodiversity science and biomonitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1702), 20150330. https://doi.org/10.1098/rstb.2015.0330

Hebert, P. D. N., Cywinska, A., & Ball, S. L. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(1512), 313–321. doi: https://doi.org/10.1098/rspb.2002.2218.

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H., & Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. *Proceedings of the National Academy of Sciences of the United States of America*, 101(41), 14812–14817. https://doi.org/10.1073/pnas.0406166101

Hebert, P. D. N., Ratnasingham, S., & de Waard, J. R. (2003). Barcoding animal life: Cytochrome *c* oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(Suppl 1), S96–S99. https://doi.org/10.1098/rsbl.2003.0025

Kanturski, M., Lee, Y., Choi, J., & Lee, S. (2018). DNA barcoding and a precise morphological comparison revealed a cryptic species in

the *Nippolachnus piri* complex (Hemiptera: Aphididae: Lachninae). *Scientific Reports*, 8(1), 1–16. https://doi.org/10.1038/s41598-018-27218-2

Kawai, K. (1991). Seven new chironomid species (Diptera, Chironomidae) from Japan. *Japanese Journal of Limnology (Rikusuigaku Zasshi)*, 52(3), 161–171. https://doi.org/10.3739/rikusui.52.161

Kelly, R. P., Port, J. A., Yamahara, K. M., Martone, R. G., Lowell, N., Thomsen, P. F., Mach, M. E., Bennett, M., Prahler, E., Caldwell, M. R., & Crowder, L. B. (2014). Harnessing DNA to improve environmental management. *Science*, 344(6191), 1455–1456. https://doi.org/10.1126/science.1251156

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16(2), 111–120. https://doi.org/10.1007/BF01731581

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. https://doi.org/10.1093/molbev/msy096

Leese, F., Bouchez, A., Abarenkov, K., Altermatt, F., Borja, Á., Bruce, K., & Costa, F. O. (2018). Why we need sustainable networks bridging countries, disciplines, cultures and generations for aquatic biomonitoring 2.0: A perspective derived from the DNAqua-Net COST action. *Advances in Ecological Research*, 58, 63–99. https://doi.org/10.1016/bs.aecr.2018.01.001

Lin, X.-L., Stur, E., & Ekrem, T. (2015). Exploring genetic divergence in a species-rich insect genus using 2790 DNA Barcodes. *PLoS One*, 10(9), e0138993. https://doi.org/10.1371/journal.pone.0138993

Lin, X.-L., Stur, E., & Ekrem, T. (2018a). DNA barcodes and morphology reveal unrecognized species of Chironomidae (Diptera). *Insect Systematics & Evolution*, 49(4), 329–398. https://doi.org/10.1163/1876312X-00002172

Lin, X.-L., Stur, E., & Ekrem, T. (2018b). Exploring species boundaries with multiple genetic loci using empirical data from non-biting midges. *Zoologica Scripta*, 47(3), 325–341. https://doi.org/10.1111/zsc.12280

Lin, X.-L., Stur, E., & Ekrem, T. (2018c). Molecular phylogeny and temporal diversification of *Tanytarsus* van der Wulp (Diptera: Chironomidae) support generic synonymies, a new classification and centre of origin. *Systematic Entomology*, 43, 659–677. https://doi.org/10.1111/syen.12292

Luo, A. R., Lan, H. Q., Ling, C., Zhang, A. B., Shi, L., Ho, S. Y. W., & Zhu, C. D. (2015). A simulation study of sample size for DNA barcoding. *Ecology and Evolution*, 5(24), S869–S879. https://doi.org/10.1002/ece3.1846

Macher, J. N., Vivancos, A., Piggott, J. J., Centeno, F. C., Matthaei, C. D., & Leese, F. (2018). Comparison of environmental DNA and bulk-sample metabarcoding using highly degenerate cytochrome *c* oxidase I primers. *Molecular Ecology Resources*, 18(6), 1456–1468. https://doi.org/10.1111/1755-0998.12940

Majaneva, M., Diserud, O. H., Hajibabaei, M., Eagle, S., & Ekrem, T. (2018). Choice of DNA extraction method affects DNA metabarcoding of unsorted invertebrate bulk samples. *Metabarcoding and Metagenomics*, 2, e26664. https://doi.org/10.3897/mbmg.2.26664

Morinière, J., Balke, M., Doczkal, D., Geiger, M. F., Hardulak, L. A., Haszprunar, G., Hausmann, A., Hendrich, L., Regalado, L., Rulik, B., Schmidt, S., Wägele, J.-W., & Hebert, P. D. N. (2019). A DNA barcode library for 5,200 German flies and midges (Insecta: Diptera) and its implications for metabarcoding-based biomonitoring. *Molecular Ecology Resources*, 19(4), 900–928. https://doi.org/10.1111/1755-0998.13022

Morinière, J., Hendrich, L., Balke, M., Beermann, A. J., König, T., Hess, M., Koch, S., Müller, R., Leese, F., Hebert, P. D. N., Hausmann, A., Schubart, C. D., & Haszprunar, G. (2017). A DNA barcode library for Germany′s mayflies, stoneflies and caddisflies (Ephemeroptera, Plecoptera and Trichoptera). *Molecular Ecology Resources*, 17(6), 1293–1307. https://doi.org/10.1111/1755-0998.12683

Naciri, Y., Caetano, S., & Salamin, N. (2012). Plant DNA barcodes and the influence of gene flow. *Molecular Ecology Resources*, 12(4), 575–580. https://doi.org/10.1111/j.1755-0998.2012.03130.x

Pentinsaari, M., Vos, R., & Mutanen, M. (2017). Algorithmic single-locus species delimitation: Effects of sampling effort, variation and nonmonophyly in four methods and 1870 species of beetles. *Molecular Ecology Resources*, 17(3), 393–404. https://doi.org/10.1111/1755-0998.12557

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., Kamoun, S., Sumlin, W. D., & Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, 55(4), 595–609. https://doi.org/10.1080/10635150600852011

Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2012). ABGD, automatic barcode gap discovery for primary species delimitation. *Molecular Ecology Notes*, 21(8), 1864–1877. https://doi.org/10.1111/j.1365-294X.2011.05239.x

Ratnasingham, S., & Hebert, P. D. N. (2007). BOLD: The barcode of life data system (www.barcodinglife.org). *Molecular Ecology Notes*, 7(3), 355–364. https://doi.org/10.1111/j.1471-8286.2006.01678.x

Ratnasingham, S., & Hebert, P. D. N. (2013). A DNA-based registry for all animal species: The barcode index number (BIN) system. *PLoS One*, 8(7), e66213. https://doi.org/10.1371/journal.pone.0066213

Reiss, F., & Fittkau, E. J. (1971). Taxonomie und Ökologie europäisch verbreiteter *Tanytarsus*-Arten (Chironomidae, Diptera). *Archiv Für Hydrobiologie, Supplement*, 40, 75–200.

Sæther, O. A. (1969). Some nearctic podonominae, diamesinae, and orthocladiinae (Diptera: Chironomidae) *Bulletin of the Fisheries Research Board of Canada*, 170, 1–154.

Saitou, N., & Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4(4), 406–425.

Sanseverino, A. M. (2006). A review of the genus Tanytarsus van der Wulp, 1874 (Insecta, Diptera, Chironomidae) from the Neotropical region. Dissertation zur Erlangung des Doktorgrades der Fakultät für Biologie der Ludwig-Maximilians-Universität, München, pp. 306.

Sasa, M. (1980). Studies on chironomid midges of the Tama River. Part 2. Description of 20 species of Chironominae recovered from a tributary. *Research Report from the National Institute for Environmental Studies, Japan*, 13, 9–107.

Sasa, M. (1983). Studies on chironomid midges of the Tama River. Part 5. An observation on the distribution of Chironominae along the main stream in June, with description of 15 new species. *Research Report from the National Institute for Environmental Studies, Japan*, 43, 1–67.

Sasa, M., & Kawai, K. (1987). Studies on chironomid midges of Lake Biwa (Diptera, Chironomidae). *Lake Biwa Research Institute, Otsu, Japan*, 3, 1–119.

Schutze, M. K., Virgilio, M., Norrbom, A., & Clarke, A. R. (2017). Tephritid integrative taxonomy: where we are now, with a focus on the resolution of three tropical fruit fly species complexes. *Annual Review of Entomology*, 62, 147–164. https://doi.org/10.1146/annurev-ento-031616-035518

Silva, F. L., & Ekrem, T. (2016). Phylogenetic relationships of nonbiting midges in the subfamily Tanypodinae (Diptera: Chironomidae) inferred from morphology. *Systematic Entomology*, 41(1), 73–92. https://doi.org/10.1111/syen.12141

Song, C., Lin, X.-L., Wang, Q., & Wang, X.-H. (2018). DNA barcodes successfully delimit morphospecies in a superdiverse insect genus. *Zoologica Scripta*, 47(3), 311–324. https://doi.org/10.1111/zsc.12284

Stur, E., da Silva, F. L., & Ekrem, T. (2019). Back from the Past: DNA barcodes and morphology support *Ablabesmyia americana* Fittkau as a valid species (Diptera: Chironomidae). *Diversity*, 11(9), 173. https://doi.org/10.3390/d11090173

Sun, Z., Majaneva, M., Sokolova, E., Rauch, S., Meland, S., & Ekrem, T. (2019). DNA metabarcoding adds valuable information for

management of biodiversity in roadside stormwater ponds. *Ecology and Evolution*, 9(17), 9712–9722. https://doi.org/10.1002/ece3.5503

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, 21(8), 2045–2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x

Talavera, G., Dincă, V., & Vila, R. (2013). Factors affecting species delimitations with the GMYC model: Insights from a butterfly survey. *Methods in Ecology and Evolution*, 4(12), 1101–1110. https://doi.org/10.1111/2041-210X.12107

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA–An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, 183, 4–18. https://doi.org/10.1016/j.biocon.2014.11.019

UNEP (2011). Report of the tenth meeting of the conference of the parties to the convention on biological diversity. Nagoya, Japan.

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., Bellemain, E., Besnard, A., Coissac, E., Boyer, F., Gaboriaud, C., Jean, P., Poulet, N., Roset, N., Copp, G. H., Geniez, P., Pont, D., Argillier, C., Baudoin, J.-M., … Dejean, T. (2016). Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology*, 25(4), 929–942. https://doi.org/10.1111/mec.13428

Vié, J.-C., Hilton-Taylor, C., & Stuart, S. N. (2009). Wildlife in a changing world: an analysis of the 2008 IUCN Red List of threatened species: IUCN.

Weigand, H., Beermann, A. J., Čiampor, F., Costa, F. O., Csabai, Z., Duarte, S., Geiger, M. F., Grabowski, M., Rimet, F., Rulik, B., Strand, M., Szucsich, N., Weigand, A. M., Willassen, E., Wyler, S. A., Bouchez, A., Borja, A., Čiamporová-Zaťovičová, Z., Ferreira, S., … Ekrem, T. (2019). DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. *Science of the Total Environment*, 678, 499–524. https://doi.org/10.1016/j.scitotenv.2019.04.247

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., &Yutani, H. (2018). Package 'ggplot2': Create elegant data visualisations using the grammar of graphics. R Package.

Willyard, A., Cronn, R., & Liston, A. (2009). Reticulate evolution and incomplete lineage sorting among the ponderosa pines. *Molecular Phylogenetics and Evolution*, 52(2), 498–511. https://doi.org/10.1016/j.ympev.2009.02.011

Wulp, F. M. (1874). Dipterologische Aanteekeningen. *Tijdschrift Voor Entomologie*, 17, 109–148.

Yamamoto, M., & Yamamoto, N. (2014). Family chironomidae. In T. Saigusa (Ed.), *Catalogue of the insects of Japan, Part 1. Diptera (Nematocera - Brachycera Aschiza)* (Vol. 8, pp. 237–362). : The Entomological Society of Japan, Touka Shobo Publisher.

Young, M. R., Proctor, H. C., Dewaard, J. R., & Hebert, P. D. (2019). DNA barcodes expose unexpected diversity in Canadian mites. *Molecular Ecology*, 28(24), 5347–5359. https://doi.org/10.1111/mec.15292

Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., & Ding, Z. (2012). Biodiversity soup: Metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, 3(4), 613–623. https://doi.org/10.1111/j.2041-210X.2012.00198.x

Zhang, A. B., He, L. J., Crozier, R. H., Muster, C., & Zhu, C. D. (2010). Estimating sample sizes for DNA barcoding. *Molecular Phylogenetics and Evolution*, 54(3), 1035–1039. https://doi.org/10.1016/j.ympev.2009.09.014

Zhang, J., Kapli, P., Pavlidis, P., & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, 29(22), 2869–2876. https://doi.org/10.1093/bioinformatics/btt499

Zheng, C., Ye, Z., Zhu, X., Zhang, H., Dong, X., Chen, P., & Bu, W. (2020). Integrative taxonomy uncovers hidden species diversity in the rheophilic genus *Potamometra* (Hemiptera: Gerridae). *Zoologica Scripta*, 49(2), 174–186. https://doi.org/10.1111/zsc.12401

Zhou, X., Kjer, K. M., & Morse, J. C. (2007). Associating larvae and adults of Chinese Hydropsychidae caddisflies (Insecta: Trichoptera) using DNA sequences. *Journal of the North American Benthological Society*, 26(4), 719–742. https://doi.org/10.1899/06-089.1

BIOSKETCH

Xiao-Long Lin is a postdoc at Nankai University, and is mainly interested in evolution, systematics, DNA barcoding and biogeography of the family Chironomidae (Diptera). His work focuses on building the DNA barcode reference library of Chinese chironomids to provide a reliable resource for taxonomic studies and the application of environmental DNA metabarcoding in biomonitoring of freshwater ecosystems.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.