Development of novel PCR primer sets for DNA metabarcoding of aquatic insects, and the discovery of some cryptic species

Masaki Takenaka[1], Koki Yano[2], Tomoya Suzuki[3], Koji Tojo[4,5]

Affiliations

[1] Sugadaira Research Station, Mountain Science Center, University of Tsukuba, Sugadairakougen 1278-294, Ueda, Nagano, Japan

[2] Division of Evolutionary Developmental Biology, National Institute for Basic Biology, Nishigonaka 38, Myodaiji, Okazaki, Aichi, 44-8585, Japan

[3] Graduate School of Global Environmental Studies, Kyoto University, Yoshida Nihonmatsu-cho, Sakyo-ku, Kyoto 606-8501, Japan

[4] Department of Biology, Faculty of Science, Shinshu University, Asahi 3-1-1, Matsumoto, Nagano 390-8621, Japan

[5] Institute of Mountain Science, Shinshu University, Asahi 3-1-1, Matsumoto, Nagano 390-8621, Japan

Co-Correspondence: Masaki Takenaka (masakiplayer@gmail.com) and Koji Tojo (ktojo@shinshu-u.ac.jp)

Abstract

DNA barcoding is a powerful tool that provides rapid, accurate, and automatable species identification by using standardized genetic region(s). It can be a powerful tool in various fields of biology such as for revealing the existence of cryptic species and/or rare species and in environmental science such as when monitoring river biota. Biodiversity reduction in recent times has become one of the most serious environmental issues on a worldwide scale. DNA barcoding techniques require the development of sets of universal PCR primers for DNA metabarcoding. We tried to develop universal primer sets for the DNA barcoding of all insect groups. In this study, we succeeded in designing not only universal primer sets for DNA barcoding regions of almost all insects, which were designed to include a hypervariable site between highly conserved sites, but also primer sets for longer fragment sequences for registration in a database. We confirmed successful amplification for 14 orders, 43 families, and 68 species with DNA barcoding in the mtDNA 16S rRNA region, and for 13 orders, 42 families, and 66 species with DNA barcoding in the mtDNA 12S rRNA region. A key feature is that the DNA fragments of the DNA barcoding regions amplified by these primer sets are both short at about 200-bp, and longer fragment sequences will increase the level of data registration in the DNA database. Such resulting database enhancements will serve as a powerful tool for increasingly accurate assessment of biodiversity and genetic diversity.

Keywords: Biodiversity, DNA barcoding, eDNA, mtDNA mtDNA 12S rRNA, 16S rRNA

Introduction

About 8.7 million eukaryotic species are estimated to inhabit the Earth (Mora et al. 2011). Insects are the largest and most diverse group of organisms on Earth, and about one million species have been described (Grimaldi and Engel 2005; Tojo et al. 2017; Stork 2018). It is known that there are still many undescribed insect species, and new species are still described on a daily basis. A more accurate understanding of species diversity and elucidation of the mechanisms of diversity are important issues. On the other hand, as many species have been evaluated to be threatened with extinction, environmental conservation and species conservation efforts are also urgent tasks (Ceballos et al. 2015; Dirzo et al. 2014). For effective conservation measures of a particular species, it is important to appropriately assess and understand the current state of its biodiversity.

DNA barcoding is a system which provides rapid, accurate, and automatable species identification by using a standardized genetic region(s) (Hebert and Gregory 2005). In general, numerous insects are identified based on their morphological characteristics, but this method requires specialist knowledge and it takes a lot of time to acquire enough skills. Under such circumstances, DNA barcoding can rapidly identify a species by sequencing a standardized short DNA fragment, even if the specimens are difficult to identify by morphology (Hebert and Gregory 2005; Miya et al. 2015). In addition, DNA barcoding even allows species identification of specimens that are not suitable for species identification by means of traditional morphological classification, such as larval specimens or parts of specimens (incomplete specimens). Since this method is easy and fast, and its results are highly reproducible, it is possible use it for a wide range of species to undertake long-term monitoring and gain an understanding of their biodiversity (Hänfling et al. 2016; Uchida et al. 2020; Chucholl et al. 2021).

DNA barcoding is also an effective tool for identifying the existence of cryptic species and/or rare species (Hebert et al. 2004). In recent years, it has been reported that many cryptic species or undescribed species have been being discovered by conducting DNA barcoding-based genetic analyses (Vuataz et al. 2013; Saitoh et al. 2015; Struck et al. 2018; Takenaka and Tojo 2019; Yano et al. 2019; Ohnishi et al. 2021; Tojo et al. 2021). Of course, DNA barcoding does not replace traditional taxonomy (Schindel 2005). There is no doubt that highly experienced taxonomists are still required to scrutinize the taxonomic descriptions of such assessed species. As an interesting example of the use of DNA barcoding in recent studies, it was possible to understand past biodiversity by detecting the DNA of a particular fish from hundreds of years ago collected from seafloor sediments (Kuwae et al. 2020). Kudoh et al. (2020) identified a particular herbivorous insect using leaves with external foliage feeding marks and environmental DNA (eDNA) techniques. Moreover, for endangered species, DNA based non-invasive assessment of

biodiversity and corresponding genetic diversity is a breakthrough technique (Sekiya et al. 2017; Ahn et al. 2020; Yamazaki et al. 2020). Such methods are also expected to be applied to various other fields in addition to taxonomy.

For aquatic organisms, eDNA in aquatic environments has also facilitated the detection of an aquatic vertebrate species (Miya et al. 2015). It can be a powerful tool in the various fields of biology and environmental science such as in monitoring river biota. Biodiversity reduction in recent times has become one of the most serious environmental issues on a worldwide scale. In particular, freshwater organisms tend to account for a high proportion of Red List species. Under such circumstances, it is necessary to monitor biological fauna and flora to facilitate the conservation of biodiversity. For this purpose, eDNA analysis offers a powerful molecular tool capable of non-invasively surveying species richness within many ecosystems (Bohmann et al. 2014; Deiner et al. 2017; Uchida et al. 2020).

These techniques require the development of sets of universal PCR primers for DNA metabarcoding. For fish species, Miya et al. (2015) designed a set of universal PCR primers (i.e., “MiFish”) for the metabarcoding of eDNA. Also, these primers have been developed for various other animal taxa (e.g., “MiMammal” for mammals: Ushio et al. 2017; “MiBird” for birds: Ushio et al. 2018; “MiDeca” for crustaceans, especially Decapoda: Komai et al. 2019). As for insects, which have the highest species diversity on the Earth, the mitochondrial DNA (mtDNA) cytochrome c oxidase subunit I (COI) region is frequently targeted using Folmer's universal primer set for DNA barcoding (Folmer et al. 1994). However, it is not a suitable primer set for DNA metabarcoding, as the protein-coding gene for amino acids of the mtDNA COI region is not highly conserved (Deagle et al. 2014); its third codon in particular is detected with a high number of polymorphisms. As such polymorphisms tend to be concentrated on the third base of each codon, we considered that it was not suitable for the development of a highly versatile primer set for amplification of short fragment sequences.

Many previous studies on non-insect groups using metabarcoding of eDNA have used primer sets developed from within their ribosomal RNA region (Miya et al. 2015; Ushio et al. 2017, 2018; Komai et al. 2019). Therefore, we also tried to design a universal primer set suitable for DNA metabarcoding of insects based on the ribosomal RNA region on the mtDNA.

The ideal characteristics of DNA fragments for a DNA barcoding region are listed below (Valentini et al. 2009; Miya et al. 2015). It must be 1) possible to reliably identify the specific insect species from it, so it needs to be completely the same or with only a minimal difference from other individuals of the same species, but with clear differences to the sequences of other species, 2) a homologous standardized region that is also able

to be used for amplification in all insect groups, 3) in a target region which has sufficient phylogenetic information to easily assign undescribed species to a taxon (genus or family), 4) a highly preserved, reliable, robust fragment, 5) suitable for amplification of a short fragment and contain sufficient sequence variations in order to correctly assign the insect species.

Therefore, in this study, we tried to develop universal primer sets for DNA barcoding of all insect groups, by the methods set our below and as in Miya et al. (2015). Firstly, the primer sets developed in this study are applicable to all insect groups (especially aquatic insects), and the region contained between these versatile primers includes polymorphism-rich sites (hypervariable regions). Second, although the target region is a short-length sequence (about 200 bp), it is able to reliably distinguish species, even closely related species, and is also effective at capturing fragmented DNA such as samples recovered from eDNA present in an aquatic environment. However, it takes a lot of effort and is expensive to enrich a database that refers to the region amplified by using our newly designed primer sets. In order to enhance the database more efficiently, we designed versatile universal primer sets that amplify not only short fragments for DNA barcoding, but also longer fragments including a targeted barcoding region that can also be used for phylogenetic analyses. We consider that this primer set that can also amplify longer fragments is an ideal tool for phylogenetic studies; it will therefore be adopted as the optimal method and lead to the enhancement of databases that refer to eDNA.

We also examined whether the DNA region we selected for DNA metabarcoding contained sufficient polymorphisms to be effective in species differentiation even between closely related species, and also whether cryptic species could be readily detected from it. In aquatic insects, it is known that closely related species are niche-differentiated, each adapting to various river microhabitats (Ohgitani and Nakamura 2008; Ohgitani et al. 2021; Okamoto and Tojo 2021; Okamoto et al. 2021). Heptageniid mayflies are a typical group exhibiting niche differentiation between closely related species (Ohgitani and Nakamura 2008; Tojo 2010) and are therefore suitable for testing the newly developed primers in this study. Previous studies using molecular markers have reported cases of discovering undescribed species and/or cryptic species as a result of phylogenetic analysis of species inhabiting a wide range and/or a variety of environments (Ueda et al. 2012; Yano et al. 2019). Also, the study it based on the hypothesis that *Epeorus aesculus* (Heptageniidae) also contains a cryptic species because *Epeorus aesculus* Imanishi, 1934, inhabit in a relatively wide range of river flow (Ohgitani and Nakamura 2008). From these viewpoints, we assessed and verified the detection ability and sensitivity of this newly developed DNA metabarcoding primer set using heptageniid mayflies, including *E. aesculus*.

158

159 **Materials and Methods**

160 **Development of primer sets**

161 In order to perform DNA metabarcoding, we focused on the mtDNA 16S rRNA and 12S
162 rRNA regions because these ribosomal RNA regions have been reported to provide
163 almost the same potential to correctly identify different individual species as the mtDNA
164 COI region, which is the standard DNA barcoding region (Collins et al. 2019). We
165 considered that these rRNA regions also have the advantage of having fewer intraspecific
166 polymorphisms than the COI region. Also, for previous studies of other vertebrates and
167 invertebrates, primer sets have been developed for DNA metabarcoding in the 16S rRNA
168 or 12S rRNA regions (Miya et al. 2015; Ushio et al. 2017, 2018; Komai et al. 2019). The
169 general DNA barcoding region for insects is the mtDNA COI region, but this could not
170 be as effective as the versatile primers sets in this study due to the presence of
171 polymorphisms every three bases.

172 In order to select a few suitable regions, whole or partial mitogenome sequences of
173 various insect groups, to increase versatility, were downloaded from GenBank: aquatic
174 insects [Ephemeroptera, Odonata, Plecoptera, each family of Hemiptera (Hemiptera s.
175 lat.: Belostomatidae, Nepidae, Gerridae, Corixidae), Corydalidae, Trichoptera, each
176 family of Coleoptera (Dytiscidae, Gyrinidae, Lampyridae, Dryopoidae), and each family
177 of Diptera (Simuliidae, Culicidae, Tipulidae)], and Apterygota (Diplura, Archaeognatha,
178 Zygentoma). Initially, we referred only to the sequences of aquatic insects because we
179 were focusing on the DNA barcoding of aquatic insects. However, since aquatic insects
180 include a wide range of insect groups, the results were used to search for a genetic region
181 that could be applied to almost all insect groups. All sequences were aligned using
182 MAFFT v7.222 (Katoh & Standley 2013) with the default set of parameters. The highly
183 versatile areas were graphically represented using MEGA 7.0.26 (Kumar et al. 2016) and
184 highly versatile regions were identified by means of visual inspection.

185 For the mtDNA 16S rRNA region, we searched for a highly versatile region using all
186 data sets of all insect groups downloaded. However, it was not possible to design a single
187 set for all insects in the highly versatile mtDNA 12S rRNA region for amplification;
188 therefore, we designed three specialized primer sets to amplify the mtDNA 12S rRNA
189 region for each of the three groups: 1) Hemimetabola, 2) Holometabola excluding
190 Trichoptera, and 3) Trichoptera. In designing these generic primers, we applied the
191 following recommendations made by Miya et al. (2015), paying attention to both ends of
192 each primer so that not only the complementarity on the 3'-end, but also the region with
193 high complementarity on the 5'-end were included.

194

Testing the versatility of the newly developed primers

To evaluate the versatility of the primers designed in this study, PCR amplification was conducted using the total genomic DNA extracted and purified from a variety of insect groups stored in the Tojo laboratory of Shinshu University, Japan (Table 1). Each total genomic DNA sample was used to amplify DNA fragments [the mtDNA 16S rRNA and 12S rRNA regions] by polymerase chain reaction (PCR) with sets of primers designed in this study. Regarding each reaction, 1.0 μL of 10x Ex Taq buffer, 0.8μL dNTP Mixture (included 25 mM $MgCl_2$), 0.05μL of 5U/μL Ex Taq polymerase (TAKARA, Shiga), 0.25μL of each primer, 1.0μL of extracted DNA in total 10μL. The PCR protocol for the DNA barcoding region in the mtDNA 16S rRNA and 12S rRNA regions was: 94 °C for 1 min; 30× (94 °C for 1 min, 50 °C for 30 sec or 1min, 72 °C for 30 sec); 72 °C for 3 min. The PCR protocol for the long fragment in the mtDNA 16S rRNA and 12S rRNA regions was: 94 °C for 1 min; 30× (94 °C for 1 min, 51 °C for 30 sec, 74 °C for 1 min); 74 °C for 3 min. The PCR products were purified using ExoSAP-IT Express (Thermo Fisher Scientific K.K., Tokyo, JP). Sequencing of purified DNA fragments was outsourced to Eurofins Genomics (Tokyo, Japan). A BigDye Terminator Cycle Sequence Kit v3.1 (ABI) was used prior to sequencing with the ABI sequencer. Sequence data have been submitted to the DNA data-bank of Japan (DDBJ database; Accession numbers are given in Table 1). All sequence data were aligned using MAFFT v7.222 (Katoh and Standley 2013). Phylogenetic analyses were performed by the Neighbor-Joining (NJ) method using MEGA 7.0.26 (Kumar et al. 2016).

Evaluation of interspecific variations and phylogenetic analysis

To check for genetic variation between closely related species and within a species (detection of cryptic species/lineages), this study focused on *Epeorus aesculus* (Heptageniidae, Ephemeroptera). We used the total genomic DNA of 14 specimens of *E. aesculus* Imanishi, 1934, from four localities including topotype specimens (i.e., specimens collected from the type locality: Kurobe-goro-zawa, Toyama, Toyama Prefecture) (Table S1).

With respect to the outgroups, we added appropriate DNA sequence data on *Epeorus dayongensis* (MK6422986, MT112895), *Epeorus herklotsi* (MG870104, NC_039612), *Epeorus carinatus* (MT112896), and *Afronurus yixingensis* (MK642297). Each total genomic DNA sample was used to amplify DNA fragments [the mtDNA 16S rRNA and 12S rRNA regions] by polymerase chain reaction (PCR) with sets of primers designed in this study. The PCR products were purified using ExoSAP-IT Express (Thermo Fisher Scientific K.K., Tokyo, JP). Sequencing of purified DNA fragments was outsourced to Eurofins Genomics (Tokyo, Japan). A BigDye Terminator Cycle Sequence Kit v3.1

(ABI) was used prior to sequencing with the ABI sequencer. Sequence data have been submitted to the DNA data-bank of Japan (DDBJ database; Accession numbers are given in Table S1).

Sequence alignment and editing were performed using the same methods for each gene separately using ATGC bundled with GENETYX ver. 15.2 (GENETYX Corporation). All sequence data were aligned using MAFFT v7.222 (Katoh and Standley, 2013). Phylogenetic analyses were performed by Bayesian analysis using MrBayes v3.2.6 (Ronquist et al. 2012). The program Kakusan4 (Tanabe 2007) was used to select appropriate models based on Schwarz's Bayesian Information Criterion (BIC; Schwarz, 1978). Best-fit substitution models were chosen as follows: HKY + G for the mtDNA 16S rRNA; HKY + G for the mtDNA 12S rRNA regions. Bayesian MCMC simulations were run for 10 million generations, sampling every 1000 generations. The output files were checked for convergence after removing a 10% burn-in by examining Effective Sampling Size (ESS > 200) using Tracer v1.6 (Rambaut et al. 2014). Then, this was visualized in the resulting tree created by FigTree v1.3.1 (Rambaut 2009).

Results and Discussion

Design of versatile primer sets for DNA metabarcoding

We designed a primer set, "MtInsects-16S", for amplification of the DNA barcoding regions in the mtDNA 16S rRNA region that is applicable to almost all insect groups. We also designed a primer "AQdb-16S" that can be used to PCR-amplify longer DNA fragments containing the DNA barcoding region and this is highly effective for phylogenetic analyses. This primer set will be very effective and useful when registering reference sequences. (Table 2; Fig. 1). The positional relationship of each primer is shown in Figure S1.

On the other hand, in the mtDNA 12S rRNA region, we designed primer sets that amplified the DNA barcoding region for each of three groups: 1) Hemimetabola, 2) Holometabola excluding Trichoptera, and 3) Trichoptera. For Hemimetabola, "MtInsects-12S_HemHol_F" and "MtInsects-12S_Hem_R" were used as the primer set to amplify the DNA barcoding region. For Holometabola excluding Trichoptera, "MtInsects-12S_HemHol_F" (the same as Hemimetabola) and "MtInsects-12S_Hol_R" were used as the primer sets to amplify the DNA barcoding region. For Trichoptera, "MtInsects-12S_cad_F" and "MtInsects-12S_cad_R" were used as the primer sets to amplify the DNA barcoding region (Table 2; Fig. 1).

We also designed three primer sets to amplify longer fragments including each of the three DNA barcoding regions of each group: 1) Hemimetabola, 2) Holometabola

excluding Trichoptera, and 3) Trichoptera. In Hemimetabola, "MtInsects-12S_HemHol_F" (the same as DNA barcoding) and "AQdb-12S_Hem_R" were used as a primer set; in Holometabola excluding Trichoptera, "MtInsects-12S_HemHol_F" (the same as DNA barcoding) and "AQdb-12S_Hol_R"; were used as a primer set, and in Trichoptera, "MtInsects-12S_cad_F" and "AQdb-12S_cad_R" were used as a primer set (Table 2; Fig. 1). The positional relationship of each primer is shown in Figure S2.

The above primer information is shown in Table 2. All primers were designed to put a hypervariable region between highly conserved regions (Fig. 2). It shows a comparison of the polymorphisms of primer sites in all insects used in the search for primer sites in this study (Table S2-13, Fig. 1). The concordance rate of each locus (the graph above) and rate of each nucleic acid sequence (include indel) in the full length mtDNA 16S rRNA, 12S rRNA regions, and the COI region, which is the standard DNA barcoding region for insects, are shown in Fig S3. These results show that the primer sets for DNA metabarcoding developed in this study have been optimized to select the best sites because the primer regions have a higher nucleic sequence concordance rate than other loci, except for the regions with a lot of indels. Also, we suggested that the mtDNA COI region, which is the standard DNA barcoding region for insects, is an unsuitable DNA barcoding region to use for all insects because there was no region with a continuous high concordance rate. For the mtDNA COI region, it was not possible to find regions with high concordance rates, even at the mayfly order level.

Versatility and nucleotide polymorphisms of the DNA metabarcoding regions

In this study, using each of the applicable primer sets designed for PCR and Sanger sequencing, we succeeded in the PCR amplification and sequencing of DNA barcoding regions and longer fragments, including the respective DNA barcoding regions, for all species used. We confirmed successful amplification for 14 orders, 43 families, and 68 species using DNA barcoding in the mtDNA 16S rRNA region, and for 13 orders, 42 families, and 66 species for DNA barcoding in the mtDNA 12S rRNA (Table 1). It is necessary to confirm the methodology developed in this study in practice in future DNA barcoding and/or DNA metabarcoding research (e.g., eDNA), but there is no doubt that all the primers designed in this study have high versatility. Although PCR amplification was successful in all examined species of the various insect groups tested, care should be taken when using primer sets for the mtDNA 12S rRNA region in which Tm values differ by about 10℃ between the forward and reverse primers. In the mtDNA 12S rRNA region, the GC nucleotides contained at the sites where the primers were designed was biased, so the Tm values of these primers could not be made uniform (Fig. S3).

A phylogenetic cladogram was constructed to investigate whether species could be identified using the DNA barcoding regions amplified using each primer set designed in this study (Fig. 3). As a result, the species analyzed could be identified and did not in any case share the same genotype among different species. In particular, as the DNA barcoding regions selected using our newly designed primer sets retain sufficient polymorphisms to identify each species, they can even differentiate between species of the same genus (Fig. 3).

Our study detected only two exceptional cases of inadequate DNA barcoding, both of which were due to species characteristics or for taxonomically problematic species. The marker (primer set) of the mtDNA 16S rRNA region we designed could not distinguish between two sub-species, *Hesperocorixa distanti hokkensis* and *Hesperocorixa distanti distanti*. However, these subspecies cannot be differentiated even using the sequence data of the mtDNA COI region (Yano et al. 2020). Therefore, these species cannot be differentiated by genetic markers or are possibly not in fact differentiated at the subspecies level. Regarding the second case, in the mtDNA 12S rRNA region, no genetic differentiation was observable between *Ephemera japonica* and *Ephemera strigata*. These species are sister species to each other, and are widely distributed sympatrically in the Japanese Islands (Okamoto and Tojo 2021). We confirmed interspecies introgression in areas where these two species had re-contacted after speciation (Takenaka et al. unpublished data). Therefore, we do not think such cases nullify the versatility of the primers we have developed.

DNA barcoding method sufficiently sensitive to detect cryptic species

It is known that there are cryptic species within the mayfly species, *Epeorus aesculus* (Ogitani and Nakamura 2008; Tojo 2010). Therefore, this is a suitable species or species group to assess the potential of the genetic region we proposed for phylogenetic analysis. *Epeorus aesculus* were collected from multiple streams in the Japan Alps, including the type locality, and genetic analyses were conducted using the primer sets developed in this study to amplify longer fragments of their mtDNA 16S rRNA and 12S rRNA regions. As a result, a clade (i.e., Clade B) was detected that was largely genetically differentiated from another clade (i.e., Clade A: including the topotype specimens) collected at Kurobe-goro-zawa, which is the type locality of *E. aesculus* (Fig. 4). This means that *E. aesculus*, which has been treated as one species, has intraspecial large scale genetic differentiation, i.e., a cryptic species. Although it is necessary to investigate more samples and their morphology in detail, the results of this study adequately identified the existence of a known cryptic species or an undescribed species.

Today, DNA barcoding is a powerful tool for assessing biodiversity (Struck et al. 2018; Mosa et al. 2019). Most importantly, the genetic region of longer fragments amplified by the primer sets newly designed in this study contain a high number of polymorphisms by which detection of even cryptic species is possible. In addition, our development of the DNA barcoding methodology for short DNA fragments of about 200-bp will undoubtedly bring about further significant innovation in DNA metabarcoding methodology. Especially for aquatic insects, further development in the utilization of eDNA is anticipated. Under such circumstances, we believe that this study constitutes a methodological breakthrough that will underpin significant research advances.

Conclusion

In this study, we succeeded in developing universal primer sets and newly selected optimal DNA barcoding regions. A key feature is that the DNA fragments amplified by these primer sets are both short at about 200-bp. By developing these primer sets for DNA barcoding or DNA metabarcoding that can be applied to almost all insect groups, it will enable relatively easy long-term monitoring of insect species composition and species diversity. In addition, it is also expected to facilitate the discovery of many undescribed species and/or cryptic species that have been overlooked or are difficult to identify morphologically.

We have established a barcoding region in the mtDNA ribosomal region; however, it is possible that closely related species that could not be identified using ribosomal regions will become identifiable in the future. For fish, it has been reported that the universal primer “MiFish” for the metabarcoding primers cannot differentiate species or amplify the barcoding region of some groups, so new primer sets focusing on specific groups have been designed (genus Anguilla: Takeuchi et al. 2019; Salmonidae: Morita et al. 2019; Cichlidae: Doble et al. 2019).

In the case of insects, if it is not possible to identify species using the ribosome DNA regions developed in this study, it may be able to identify them by using the mtDNA COI region, which is the traditional standard DNA barcoding region. In such cases, it would be best to make effective use of the existing database. Also, we recommend using multiple genetic regions to improve the accuracy and reliability of species identification.

Finally, we were able to design not only DNA barcoding region markers, but also primer sets for longer fragment sequences for registration in the database. These primer sets that can amplify longer fragments can also be used for phylogenetic analyses, which will increase the level of data registration in the DNA database. In addition, such resulting

database enhancements will serve as a powerful tool for increasingly accurate assessment of biodiversity and genetic diversity.

Supplementary Information

Table S1. List of specimens of *Epeorus* mayflies examined in this study, sequence types, and the GenBank accession numbers

Table S2. Nucleotide sequences of universal primer (MtInsects-16S_F) and homologous region of 557 sequences of insects used for development in this study.

Table S3. Nucleotide sequences of universal primer (MtInsects-16S_R) and homologous region of 557 sequences of insects used for development in this study.

Table S4. Nucleotide sequences of universal primer (AQdb-16S_F) and homologous region of 557 sequences of insects used for development in this study.

Table S5. Nucleotide sequences of universal primer (AQdb-16S_R) and homologous region of 558 sequences of insects used for development in this study.

Table S6. Nucleotide sequences of universal primer (MtInsects-12S_HemHol_F) and homologous region of 491 sequences of insects used for development in this study.

Table S7. Nucleotide sequences of universal primer (MtInsects-12S-Hem_R) and homologous region of 265 sequences of insects used for development in this study.

Table S8. Nucleotide sequences of universal primer (MtInsects-12S_Hol_R) and homologous region of 209 sequences of insects used for development in this study.

Table S9. Nucleotide sequences of universal primer (MtInsects-12S_cad_F) and homologous region of 31 sequences of trichopteran insects used for development in this study.

Table S10. Nucleotide sequences of universal primer (MtInsects-12S_cad_R) and homologous region of 31 sequences of trichopteran insects used for development in this study.

Table S11. Nucleotide sequences of universal primer (AQdb-12S_Hem_R) and homologous region of 265 sequences of insects used for development in this study.

Table S12. Nucleotide sequences of universal primer (AQdb-12S_Hol_R) and homologous region of 214 sequences of insects used for development in this study.

Table S13. Nucleotide sequences of universal primer (AQdb-12S_cad_R) and homologous region of 31 sequences of trichopteran insects used for development in this study.

Figure S1. Position of primers sets in the mtDNA 16S rRNA region. Reference sequence data used the mtDNA 16S rRNA region in the complete mitochondrion genome of *Ephemera orientalis* (NC_012645).

Figure S2. Position of primers sets in the mtDNA 12S rRNA region. Reference sequence data used the mtDNA 12S rRNA region in the complete mitochondrion genome of *Ephemera orientalis* (NC_012645).

Figure S3. Graph of the nucleotide concordance rate (0.0-1.0; above) and the proportion of adenine (A: green), thymine (T: red), guanine (G: purple), cytosine (C: blue) and indel (white) (0.0-1.0; below) in each nucleotide position in the full length of each of the mtDNA 16S rRNA, 12S rRNA, and the COI regions. The nucleotide concordance rate of each locus is shown by gray bars, and the region of each newly designed primer is shown black bars.

Acknowledgements

We express our thanks to Dr. Y. Iwasaki (National Institute of Advanced Industrial Science and Technology), Dr. N. Uchida (Tohoku University), Dr. N. Kondo (National Institute of Environmental studies), Dr. Y. Hasebe (Kanagawa Prefectural Government) for valuable advice and encouragement. We are indebted to Tojo Lab. members (Shinshu University) and Niimi Lab. members (NIBB) for their cooperation with the field research and collection of specimens. This study was supported by grants from the River Environment Fund (22-1215-021 to KT, 2017-5211-025 to KT, 2021-5311-005 to MT) of River and Watershed Environment Management.

Contributions

M.T. designed, managed the study, and performed sample collection; M.T., K.Y. mainly performed laboratory work and phylogenetic analyses; M.T., K.Y., T.S. designed Figures and Tables; M.T., K.T. gathered funds; M.T., K.Y., and M.T., K.Y., T.S., K.T. wrote and reviewed the manuscript.

Ethics declarations

The authors declare that they have no competing interests. The experiments comply with the current laws of the country in which they were performed.

References

Ahn H, Kume M, Terashima Y, Ye F, Kameyama S, Miya M, Yamashita Y, Kasai A (2020) Evaluation of fish biodiversity in estuaries using environmental DNA metabarcoding. PloS ONE 15: e0231127. https://doi.org/10.1371/journal.pone.0231127

Bohmann K, Evans A, Gilbert MTP, Carvalho GR, Creer S, Knapp M, Douglas WY, De Bruyn M (2014) Environmental DNA for wildlife biology and biodiversity monitoring. Trends Ecol. Evol. 29: 358-367. https://doi.org/10.1016/j.tree.2014.04.003

Ceballos G, Ehrlich PR, Barnosky AD, García A, Pringle RM, Palmer TM (2015) Accelerated modern human–induced species losses: Entering the sixth mass extinction. Sci. Adv. 1: e1400253. https://doi.org/10.1126/sciadv.1400253

Chucholl F, Fiolka F, Segelbacher G, Epp LS (2021) eDNA detection of native and invasive crayfish species allows for year-round monitoring and large-scale screening of lotic systems. Front. Environ. Sci. 9: 23. https://doi.org/10.3389/fenvs.2021.639380

Collins RA, Bakker J, Wangensteen OS, Soto AZ, Corrigan L, Sims DW, Genner MJ, Mariani S (2019) Non-specific amplification compromises environmental DNA metabarcoding with COI. Methods Ecol. Evol. 10: 1985-2001. https://doi.org/10.1111/2041-210X.13276

Deagle BE, Jarman SN, Coissac E, Pompanon F, Taberlet P (2014) DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. Biol. Lett. 10: 20140562. https://doi.org/10.1098/rsbl.2014.0562

488 Deiner K, Bik HM, Mächler E, Seymour M, Lacoursière-Roussel A, Altermatt F, Creer
489 S, Bista I, Lodge DM, de Vere N, Pfrender ME, Bernatchez L (2017) Environmental
490 DNA metabarcoding: Transforming how we survey animal and plant communities.
491 Mol. Ecol. 26: 5872-5895. https://doi.org/10.1111/mec.14350
492 Dirzo R, Young HS, Galetti M, Ceballos G, Isaac NJ, Collen B (2014) Defaunation in the
493 Anthropocene. Science 345: 401-406. https://doi.org/10.1126/science.1251817
494 Doble CJ, Hipperson H, Salzburger W, Horsburgh GJ, Mwita C, Murrell DJ, Day JJ
495 (2020) Testing the performance of environmental DNA metabarcoding for surveying
496 highly diverse tropical fish communities: A case study from Lake Tanganyika.
497 Environmental DNA 2: 24-41. https://doi.org/10.1002/edn3.43
498 Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for
499 amplification of mitochondrial cytochrome c oxidase subunit I from diverse
500 metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3: 294-9.
501 Grimaldi D, Engel MS (2005) Evolution of the Insects. Cambridge University Press,
502 Cambridge.
503 Hänfling B, Lawson Handley L, Read DS, Hahn C, Li J, Nichols P, Blackman RC, Oliver
504 A, Winfield IJ (2016) Environmental DNA metabarcoding of lake fish communities
505 reflects long-term data from established survey methods. Mol. Ecol. 25: 3101-3119.
506 https://doi.org/10.1111/mec.13660
507 Hebert PD, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one:
508 DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes*
509 *fulgerator*. Proc. Natl. Acad. Sci. 101: 14812-14817.
510 https://doi.org/10.1073/pnas.0406166101
511 Hebert PD, Gregory TR (2005) The promise of DNA barcoding for taxonomy. Syst. Biol.
512 54: 852-859. https://doi.org/10.1080/10635150500354886
513 Hebert PD, Gregory TR (2005) The promise of DNA barcoding for taxonomy. Syst. Biol.
514 54: 852-859. https://doi.org/10.1080/10635150500354886
515 Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7:
516 improvements in performance and usability. Mol. Biol. Evol. 30: 772-780.
517 https://doi.org/10.1093/molbev/mst010
518 Komai T, Gotoh RO, Sado T, Miya M (2019) Development of a new set of PCR primers
519 for eDNA metabarcoding decapod crustaceans. Metabarcoding Metagenom. 3:
520 e33835. https://doi.org/10.3897/mbmg.3.33835
521 Kudoh A, Minamoto T, Yamamoto S (2020) Detection of herbivory: eDNA detection
522 from feeding marks on leaves. Environmental DNA 2: 627-634.

523 Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis
524 version 7.0 for bigger datasets. Mol. Biol. Evol. 33: 1870-1874.
525 https://doi.org/10.1093/molbev/msw054
526 Kuwae M, Tamai H, Doi H, Sakata MK, Minamoto T, Suzuki Y (2020) Sedimentary
527 DNA tracks decadal-centennial changes in fish abundance. Commun. Biol. 3: 1-12.
528 https://doi.org/10.1038/s42003-020-01282-9
529 Miya M, Sato Y, Fukunaga T, Sado T, Poulsen JY, Sato K, Minamoto T, Yamamoto S,
530 Yamanaka H, Araki H, Iwasaki W (2015) MiFish, a set of universal PCR primers
531 for metabarcoding environmental DNA from fishes: detection of more than 230
532 subtropical marine species. Royal Soc. Open Sci. 2: 150088.
533 https://doi.org/10.1098/rsos.150088
534 Mora C, Tittensor DP, Adl S, Simpson AG, Worm B (2011) How many species are there
535 on Earth and in the ocean?. PLoS Biol. 9: e1001127.
536 https://doi.org/10.1371/journal.pbio.1001127
537 Morita K, Sahashi G, Miya M, Kamada S, Kanbe T, Araki H (2019) Ongoing localized
538 extinctions of stream-dwelling white-spotted charr populations in small dammed-off
539 habitats of Hokkaido Island, Japan. Hydrobiologia 840: 207-213.
540 https://doi.org/10.1007/s10750-019-3891-1
541 Mosa KA, Gairola S, Jamdade R, El-Keblawy A, Al Shaer KI, Al Harthi EK, ShabanaHA,
542 Mahmoud T (2019) The promise of molecular and genomic techniques for
543 biodiversity research and DNA barcoding of the Arabian Peninsula flora. Front.
544 Plant Sci. 9: 1929. https://doi.org/10.3389/fpls.2018.01929
545 Okamoto S, Saito T, Tojo K (2021) Geographical fine-scaled distributional
546 differentiation caused by niche differentiation in three closely related mayflies.
547 Limnol. https://doi.org/10.1007/s10201-021-00673-z.
548 Okamoto, S., & Tojo, K. (2021). Distribution patterns and niche segregation of three
549 closely related Japanese ephemerid mayflies: a re-examination of each species'
550 habitat from "megadata" held in the "National Census on River Environments".
551 Limnol. 22: 227-287. https://doi.org/10.1007/s10201-021-00654-2
552 Ohgitani M, Nakamura H (2008) Distribution and Seasonal Population Change of
553 Heptageniidae Nymph in the Oguro River (The Branch of Tenryu River). The
554 Annals of environmental science, Shinshu University 30: 57-66. [Japanese]
555 Ohnishi O, Takenaka M, Okano R, Yoshitomi H, Tojo K (2021) Wide-scale gene flow,
556 even in Insects that have lost their flight ability: presence of dispersion due to a
557 unique parasitic ecological strategy of piggybacking hosts. Zool. Sci. 38: 122-139.
558 https://doi.org/10.2108/zs200088

559 Rambaut A (2009) FigTree, Version 1.3.1. Available at: http://tree.bio.ed.ac.uk/
560 software/figtree/ (cited 25 March 2019).
561 Rambaut A, Suchard MA, Xie D, Drummond AJ (2014) Tracer, version 1.6, MCMC trace
562 analysis package. Available at: http://tree.bio.ed.ac.uk/software/tracer/ (cited 25
563 March 2019).
564 Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu
565 L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian
566 phylogenetic inference and model choice across a large model space. Syst. Biol. 61:
567 539-542. https://doi.org/10.1093/sysbio/sys029
568 Saitoh T, Sugita N, Someya S, Iwami Y, Kobayashi S, Kamigaichi H, Higuchi A, Asai S,
569 Yamamoto Y, Nishiumi I (2015) DNA barcoding reveals 24 distinct lineages as
570 cryptic bird species candidates in and around the Japanese Archipelago. Mol. Ecol.
571 Resour. 15: 177-186. https://doi.org/10.1111/1755-0998.12282
572 Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis
573 version 7.0 for bigger datasets. Mol. Biol. Evol. 33: 1870-1874.
574 https://doi.org/10.1093/molbev/msw054
575 Schindel DE, Miller SE (2005) DNA barcoding a useful tool for taxonomists. Nature 435:
576 17.
577 Schwarz G (1978) Estimating the dimension of a model. Ann. Stat. 6: 461-464.
578 Sekiya T, Ichiyanagi H, Tojo K (2017) Establishing of genetic analyses methods of feces
579 from the water shrew, *Chimarrogale platycephalus* (Erinaceidae, Eulipotyphla).
580 JSM Biol. 2: 1010.
581 Stork NE (2018) How many species of insects and other terrestrial arthropods are there
582 on Earth?. Annu. Rev. Entomol. 63: 31-45. https://doi.org/10.1146/annurev-ento-
583 020117-043348
584 Struck TH, Feder JL, Bendiksby M, Birkeland S, Cerca J, Gusarov VI, Kistenich S,
585 Larsson KH, Liow LH, Nowak MD, Stedje B, Bachmann L, Dimitrov D (2018)
586 Finding evolutionary processes hidden in cryptic species. Trends Ecol. Evol. 33:
587 153-163. https://doi.org/10.1016/j.tree.2017.11.007
588 Takenaka M, Tojo K (2019) Ancient origin of a dipteromimid mayfly family endemic to
589 the Japanese Islands and its genetic differentiation across tectonic faults. Biol. J.
590 Linn. Soc. 126: 555-573. https://doi.org/10.1093/biolinnean/bly192
591 Takeuchi A, Iijima T, Kakuzen W, Watanabe S, Yamada Y, Okamura A, Horie N,
592 Mikawa N, Miller MJ, Kojima T, Tsukamoto K (2019) Release of eDNA by different
593 life history stages and during spawning activities of laboratory-reared Japanese eels
594 for interpretation of oceanic survey data. Sci. Rep. 9: 1-9.
595 https://doi.org/10.1038/s41598-019-42641-9

596 Tanabe AS (2007) Kakusan: a computer program to automate the selection of a nucleotide
597 substitution model and the configuration of a mixed model on multilocus data. Mol.
598 Ecol. Notes 7: 962-964. https://doi.org/10.1111/j.1471-8286.2007.01807.x
599 Tojo K (2010) The current distribution of aquatic insects in habiting river systems, with
600 respect to their population and genetic structure. In Harris EL, Davies NE (eds).
601 Insect habitats: characteristics, diversity and management. Nova Science, New
602 York: 157-161.
603 Tojo K, Sekiné K, Takenaka M, Isaka Y, Komaki S, Suzuki T, Schoville SD (2017)
604 Species diversity of insects in Japan: their origins and diversification processes.
605 Entomol. Sci. 20: 357-381. https://doi.org/10.1111/ens.12261
606 Ueda S, Nozawa T, Matsuzuki T, Seki RI, Shimamoto S, Itino T (2012) Phylogeny and
607 phylogeography of *Myrmica rubra* complex (Myrmicinae) in the Japanese Alps.
608 Psyche 2012: 1-7. https://doi.org/10.1155/2012/319097
609 Uchida N, Kubota K, Aita S, Kazama S (2020) Aquatic insect community structure
610 revealed by eDNA metabarcoding derives indices for environmental assessment.
611 PeerJ 8: e9176. https://doi.org/10.7717/peerj.9176
612 Ushio M, Fukuda H, Inoue T, Makoto K, Kishida O, Sato K, Murata K, Nikaido M, Sado
613 Y, Takeshita M, Iwasaki W, Yamanaka H, Kondoh M, MiyaM. (2017)
614 Environmental DNA enables detection of terrestrial mammals from forest pond
615 water. Mol. Ecol. Resour. 17: e63-e75. https://doi.org/10.1111/1755-0998.12690
616 Ushio M, Murata K, Sado T, Nishiumi I, Takeshita M, Iwasaki W, Miya M (2018)
617 Demonstration of the potential of environmental DNA as a tool for the detection of
618 avian species. Sci. Rep. 8: 1-10. https://doi.org/10.1038/s41598-018-22817-5
619 Yamazaki H, Sekiya T, Nagayama S, Hirasawa K, Tokura K, Sasaki A, Ichiyanagi H,
620 Tojo K (2020) Development of microsatellite markers for a soricid water shrew,
621 *Chimarrogale platycephalus*, and their successful use for individual identification.
622 Genes Genet. Syst. 95: 201-210. https://doi.org/10.1266/ggs.20-00017
623 Yano K, Takenaka M, TojoK (2019) Genealogical position of Japanese populations of
624 the globally distributed mayfly *Cloeon dipterum* and related species (Ephemeroptera,
625 Baetidae): A molecular phylogeographic analysis. Zool. Sci. 36: 479-489.
626 https://doi.org/10.2108/zs190049
627 Yano K, Takenaka M, Mitamura T, Tojo K (2020) Identifying a “pseudogene” for the
628 mitochondrial DNA COI region of the corixid aquatic insect, *Hesperocorixa distanti*
629 (Heteroptera, Corixidae). Limnology 21: 319-325. https://doi.org/10.1007/s10201-
630 020-00618-y
631 Valentini A, Pompanon F, Taberlet P (2009. DNA barcoding for ecologists. Trends Ecol.
632 Evol. 24: 110-117. https://doi.org/10.1016/j.tree.2008.09.011

633 Vuataz L, Sartori M, Gattolliat JL, Monaghan MT (2013) Endemism and diversification
634 in freshwater insects of Madagascar revealed by coalescent and phylogenetic
635 analysis of museum and field collections. Mol. Phylogenet. Evol. 66: 979-991.
636 https://doi.org/10.1016/j.ympev.2012.12.003
637

638 **Figure Legends**

639 **Figure 1.** Concordance rate of primers site sequences. Gray bands show the sequences of
640 each primer, while under gray bands show stability of the sequence within compared taxa.
641
642 **Figure 2.** The barcoding region and the primer region were arranged using some of the
643 sequences used in the development of each primer. The DNA barcoding regions
644 suggested in this study in both the mtDNA 16S rRNA and 12S rRNA regions have high
645 polymorphism and versatility by placing a hypervariable region between regions with
646 highly conserved regions. Legends, green: "A" of nucleic acid sequence, red: "T" of
647 nucleic acid sequence, "G" of nucleic acid sequence, "C" of nucleic acid sequence.
648
649 **Figure 3.** A phylogenetic cladogram based on concatenated data of the mtDNA 16S
650 rRNA and the 12S rRNA regions to examine species identification sensitivity. The same
651 species are shown by solid lines, and the same genus is shown by dashed lines.
652
653 **Figure 4.** The estimated phylogenetic relationships (ML methods) based on the mtDNA
654 16S rRNA and 12S rRNA regions.
655
656
657

Table 1. List of specimens examined in this study, sequence types, and the GenBank accession numbers

				GPS*		The GenBank accession numbers			
			Locality name	Latitude	Longitude	mtDNA 16S rRNA		mtDNA 12S rRNA	
Family	Species			(N)	(E)	Barcoding	Database	Barcoding	Database
Insecta									
Zygentoma									
Lepismatidae	*Thermobia domestica*	Zyg_1	Niimi laboratory strain	-	-	LC656038	-	-	-
Ephemeroptera									
Baetidae	*Cloeon dipterum*	Eph_1	Japan:Nara,Sango,Tatsunokita	34.59496	135.68168	LC656039	LC656119	LC656806	LC656864
Heptageniidae	*Bleptus fasciatus*	Eph_2	Japan:Nagano,Azumino,Horigane	36.30581	137.78880	LC656040	LC656120	LC656807	LC656865
	Bleptus michinokuensis	Eph_3	Japan: Fukushima,Fukushima,Tsuchiyuonsen	37.67972	140.24361	LC656041	LC656121	LC656808	LC656866
Isonychiidae	*Isonychia japonica*	Eph_4	Japan:Nagano,Ina,Nishiharuchika	35.82555	137.95057	LC656042	-	LC656809	LC656867
		Eph_5	Japan:Okayama,Maniwa,Kojiro	35.10613	133.64750	LC656043	LC656122	LC656810	LC656868
Isonychiidae	*Isonychia shima*	Eph_6	Japan:Miyagi,Osaki,Naruko-onsen	38.74012	140.74065	LC656044	LC656123	LC656811	LC656869
Ephemerellidae	*Drunella basalis*	Eph_7	Japan:Nagano,Matsumoto,Azumi	36.24880	137.63759	LC656045	LC656124	LC656812	LC656870
Leptophlebiidae	*Paraleptophlebia japonica*	Eph_8	Japan:Kochi,Nakatosa,Kure	33.30871	133.19434	LC656046	LC656125	LC656813	LC656871
		Eph_9	Japan:Kochi,Muroto,Sakihama	33.46041	134.19602	LC656047	-	-	-
Polymitarcyidae	*Ephoron shigae*	Eph_10	Japan: Niigata, Ojiya,Iwasawa	37.24570	138.80470	LC656048	LC656126	LC656814	LC656872
Potamanthidae	*Potamanthus formosus*	Eph_11	Japan:Oita,Oita,Shimohanda	33.16303	131.64708	LC656049	LC656127	LC656815	LC656873
Ephemeridae	*Ephemera japonica*	Eph_12	Japan:Kochi,Muroto,Sakihama	33.44480	134.21730	LC656050	LC656128	LC656816	LC656874
	Ephemera strigata	Eph_13	Japan:Aichi,Okazaki,Yomogyu	34.94153	137.26815	-	-	LC656817	LC656875

Siphlonuridae	*Siphlonurus* sp.	Eph_14	Japan:Nagano,Chikuma,Wakamiya	36.49517	138.13783	LC656051	LC656129	LC656818	LC656876
Ameletidae	*Ameletus* sp.	Eph_15	Japan:Nagano,Matsumoto,Azumi	36.24880	137.63759	LC656052	LC656130	LC656819	LC656877
Dipteromimidae	*Dipteromimus tipuliformis*	Eph_16	Japan:Tokyo,Hinode,Oguno	35.77498	139.19145	LC656053	-	-	-
		Eph_17	Japan:Kochi,Kitagawa,Kashiwagi	33.47795	134.07485	LC656054	LC656131	LC656820	-
Odonata									
Aeshnidae	*Aeshna crenata*	Odo_1	Japan:Nagano,Matsumoto	-	-	LC656055	-	-	-
		Odo_2	Japan:Nagano,Azumino	-	-	LC656056	-	-	-
		Odo_3	Japan:Nagano,Shiojiri	-	-	LC656057	-	-	-
	Aeshna juncea	Odo_4	Japan:Nagano,Matsumoto	-	-	LC656058	LC656132	LC656821	LC656878
Libellulidae	Deielia phaon	Odo_5	Japan:Okayama,Okayama	-	-	LC656059	-	-	-
		Odo_6	Japan:Ibaraki,Kasama	-	-	LC656060	LC656133	LC656822	-
Libellulidae	*Sympetrum depressiusculum*	Odo_7	Japan:Nagano	-	-	LC656061	LC656134	LC656823	LC656879
		Odo_8	Japan:Nagano,Matsumoto	-	-	LC656062	-	-	-
	Libellula quadrimaculata	Odo_9	Japan:Nagan,Matsumoto	-	-	LC656063	LC656135	LC656824	LC656880
	Sympetrum croceolum	Odo_10	Japan:Nagan,Matsumoto	-	-	LC656064	LC656136	LC656825	LC656881
	Sympetrum parvulum	Odo_11	Japan:Nagan,Matsumoto	-	-	LC656065	LC656137	LC656826	LC656882
	Pseudothemis zonata	Odo_12	Japan:Nagano,Matsumoto	-	-	LC656066	-	-	-
Corduliidae	*Somatochlora* sp.	Odo_13	Japan:Nagan,Matsumoto	-	-	LC656067	LC656138	LC656827	LC656883
	Cordulia sp.	Odo_14	Japan:Nagan,Matsumoto	-	-	LC656068	LC656139	LC656828	LC656884
Megapodagrionidae	*Rhipidolestes amamiensis amamiensis*	Odo_15	Japan:Kagoshima,Amami	-	-	LC656069	LC656140	LC656829	LC656885
	Rhipidolestes amamiensis tokunoshimensis	Odo_16	Japan:Kagoshima,Tokunoshima	-	-	LC656070	-	-	-
	Rhipidolestes shozoi	Odo_17	Japan:Okinawa,Nago	-	-	LC656071	-	-	-

Euphaeidae	*Euphaea yayeyamana*	Odo_18	Japan:Okinawa,Ishigaki	-	-	LC656072	-	-	-
	Euphaea formosa	Odo_19	Taiwan	-	-	LC656073	LC656141	LC656830	LC656886
Coenagrionidae	*Ischnura asiatica*	Odo_20	Japan:Nagano,Azumino	-	-	LC656074	LC656142	LC656831	LC656887
Plecoptera									
Scopuridae	*Scopura montana*	Ple_1	Japan:Fukushima,Tadami,Tadami	37.38463	139.27148	LC656075	LC656143	LC656832	LC656888
	Scopura longa	Ple_2	Japan:Niigata,Sado,Koda	38.20202	138.43325	LC656076	LC656144	LC656833	LC656889
	Scopura quattuorhamulata	Ple_3	Japan:Hokkaido,Hidaka,Mitsuiwa	42.79407	142.45286	LC656077	LC656145	-	-
Peltoperlidae	*Cryptoperla japonica*	Ple_4	Japan:Kochi,Muroto,Murotsu	33.34711	134.18271	LC656078	LC656146	LC656834	LC656890
		Ple_5	Japan:Niigata,Sado,Hayoshi	38.11197	138.39410	LC656079	LC656147	LC656835	LC656891
	Yoraperla uenoi	Ple_6	Japan:Niigata,Uonuma,Oshirakawa	-	-	LC656080	-	-	-
Pteronarcyidae	*Pteronarcys californica*	Ple_7	USA:Idaho,Twin Falls	42.33860	-114.28771	LC656081	LC656148	LC656836	LC656892
Chloroperlidae	*Sweltsa* sp.	Ple_8	Japan:Nagano,Matsumoto,Azumi	36.24880	137.63759	-	LC656149	LC656837	LC656893
	Chloroperlidae Gen. sp.	Ple_9	Japan:Nagano	-	-	LC656082	-	-	-
Orthoptera									
Acrididae	*Chorthippus fallax*	Ort_1	Japan:Shizuoka,Fujinomiya,Awakura	-	-	LC656083	-	LC656838	LC656894
Dermaptera									
Labiduridae	*Labidura riparia*	Der_1	Niimi laboratory strain	-	-	LC656084	LC656150	LC656839	LC656895
Phasmatodea									
Phasmatidae	*Ramulus mikado*	Pha_1	Japan:Nagano,Azumino,Akashina	36.36432	137.94301	LC656085	LC656151	LC656840	LC656896
Mantodea									
Mantidae	*Tenodera sinensis*	Man_1	Japan:Nagano,Ueda,Sanada	36.45611	138.35000	LC656086	LC656152	LC656841	LC656897
Hemiptera									
Belostomatidae	*Appasus japonicus*	Hem_1	Japan:Nagano,Chino	-	-	LC656087	LC656153	LC656842	LC656898

	Appasus major	Hem_2	Japan:Aomori,Aomori	-	-	LC656088	LC656154	LC656843	LC656899
Nepidae	*Laccotrephes japonensis*	Hem_3	South Korea:Gyeongsangbuk	-	-	LC656089	LC656155	LC656844	LC656900
	Nepa hoffmanni	Hem_4	Japan:Wakayama,Hashimoto	-	-	LC656090	LC656156	LC656845	LC656901
		Hem_5	Japan:Aichi,Toyota	-	-	LC656091	-	-	-
		Hem_6	Japan:Wakayama,Hashimoto	-	-	LC656092	-	-	-
	Ranatra chinensis	Hem_7	Japan:Okinawa,Kume	-	-	LC656093	LC656157	LC656846	LC656902
Corixidae	*Hesperocorixa distanti hokkensis*	Hem_8	Japan:Fukushima,Shirakawa	-	-	LC656094	LC656158	LC656847	LC656903
	Hesperocorixa distanti distanti	Hem_9	Japan:Hokkaido,Ishikari	-	-	LC656095	LC656159	-	-
	Hesperocorixa kolthoffi	Hem_10	Japan:Fukuoka,Kitakyushu	-	-	LC656096	LC656160	-	-
Neuroptera									
Myrmeleontidae	*Hagenomyia micans*	Neu_1	Japan:Nagano,Matsumoto,Shinagura	36.29241	137.99015	LC656097	LC656161	LC656848	LC656904
Mantispidae	*Mantispa japonica*	Neu_2	Japan:Nagano,Matsumoto,Satoyamabe	36.24843	138.01241	LC656098	-	LC656849	LC656905
		Neu_3	Japan:Ishikawa,Nanao,Notojimamagari	37.14830	136.98442	LC656099	-	-	-
Diptera									
Chironomidae	Chironomidae Gen. sp.	Dip_1	Japan:Gifu,Takayama,Okuhidaonsengonakao	36.23879	137.59551	LC656100	LC656162	LC656850	LC656906
	Chironomidae Gen. sp.	Dip_2	Japan:Gifu,Takayama,Okuhidaonsengonakao	36.23879	137.59551	LC656101	-	-	-
Coleoptera									
Dytiscidae	*Agabus japonicus*	Col_1	Japan:Niigata,Gosen	-	-	LC656102	LC656163	LC656851	LC656907
	Cybister chinensis	Col_2	Japan:Nagano,Matsumoto	-	-	LC656103	LC656164	LC656852	LC656908
Lucanidae	*Dorcus montivagus*	Col_3	Japan:Niigata,Yuzawa,Mitsumata	-	-	LC656104	LC656165	LC656853	LC656909

Meloidae	*Meloe coarctatus*	Col_4	Japan:Nagano,Matsumoto,Iriyamabe	-	-	LC656105	LC656166	LC656854	LC656910
	Meloe sp.	Col_5	Japan:Tochigi,Nikko,Chugushi	-	-	LC656106	-	-	-
		Col_6	Japan:Tochigi,Nikko,Chugushi	-	-	LC656107	-	-	-
Lepidoptera									
Pieridae	*Eurema hecabe*	Lep_1	Niimi laboratory strain	-	-	LC656108	LC656167	LC656855	LC656911
Trichoptera									
Stenopsychidae	*Stenopsyche marmorata*	Tri_1	Japan:Nagano,Matsumoto,Mizukuma	36.26313	137.98338	LC656109	LC656168	LC656856	LC656912
	Stenopsyche schmidi	Tri_2	Japan:Kagoshima,Amami	-	-	LC656110	LC656169	LC656857	LC656913
Rhyacophilidae	*Rhyacophila* sp.	Tri_3	Japan:Nara,Kamikitayama	-	-	LC656111	LC656170	LC656858	LC656914
	Rhyacophila verecunda	Tri_4	Japan:Hyogo,Yabu,Oyachoyokoiki	-	-	LC656112	-	LC656859	LC656915
Limnephilidae	*Pseudostenophylax sp.*	Tri_5	Japan:Nagano,Takayama,Maki	36.64366	138.44631	LC656113	LC656171	LC656860	LC656916
Limnocentropodidae	*Limnocentropus insolitus*	Tri_6	Japan:Nagano,Matsumoto,Misayama	36.27342	138.02556	LC656114	LC656172	LC656861	LC656917
Helicopsychidae	*Helicopsyche* sp.	Tri_7	Japan:Aichi,Okazaki,Ishihara	34.95250	137.39753	LC656115	LC656173	LC656862	LC656918
Hydroptilidae	*Palaeagapetus ovatus*	Tri_8	Japan:Nagano,Matsumoto,Azumi	36.25544	137.63853	LC656116	LC656174	-	-
		Tri_9	Japan:Nagano,Matsumoto,Azumi	36.25545	137.63853	LC656117	-	-	-
Calamoceratidae	*Anisocentropus kawamurai*	Tri_10	Japan:Tokushima,Tokushima,Hata	33.97891	134.50219	LC656118	-	LC656863	LC656919

*For rare species, there were samples for which detailed points could not be shown.

659

Table 2. Information on the newly designed primer sets in this study

Gene	Primer name	Primer direction	Primer sequence (5'-3')	Target groups	purpose
mtDNA 16S rRNA	MtInsects-16S_F	Forward	GGACGAGAAGACCCTWTAGA	All	DNA barcoding
	MtInsects-16S_R	Reverse	ATCCAACATCGAGGTCGCAA	All	DNA barcoding
	AQdb-16S_F	Forward	TRACYGTRCAAAGGTAGC	All	Phylogeny and database
	AQdb-16S_R	Reverse	CCGGTYTRAACTCARATCATGT	All	Phylogeny and database
mtDNA 12S rRNA	MtInsects-12S_HemHol_F	Forward	GTGCCAGCHDYYGCGGTTA	Hemimetabola & Holometabola	DNA barcoding &Phylogeny and database
	MtInsects-12S_Hem_R	Reverse	HATARDRGGGTMTCTAATCC	Hemimetabola	DNA barcoding
	MtInsects-12S_Hol_R	Reverse	TARTAGGGTATCTAATCCTAG	Holometabola	DNA barcoding
	MtInsects-12S_cad_F	Forward	TTGKGCCAGCARTYGCGGTWA	Trichoptera	DNA barcoding
	MtInsects-12S_cad_R	Reverse	WATARTRGRGTATCTAATYC	Trichoptera	DNA barcoding
	AQdb-12S_Hem_R	Reverse	CTACTWTGTTACGACTTRT	Hemimetabola	Phylogeny and database
	AQdb-12S_Hol_R	Reverse	TAMWYCTACTWTGTTACGACTT	Holometabola	Phylogeny and database
	AQdb-12S_cad_R	Reverse	ARYGACGGGCAATATGTRC	Trichoptera	Phylogeny and database

660

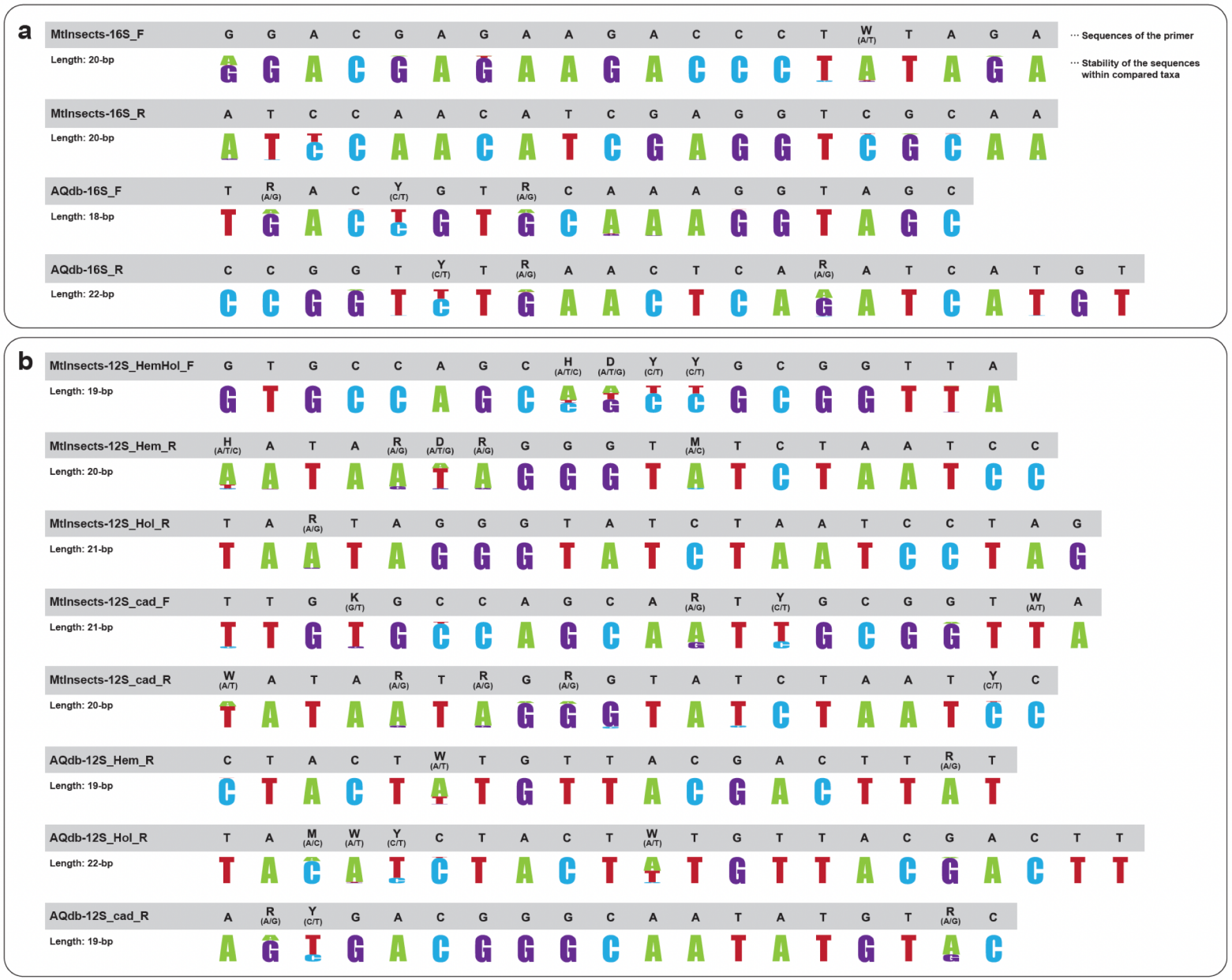
a
MtInsects-16S_F
Length: 20-bp
MtInsects-16S_R
Length: 20-bp
AQdb-16S_F
Length: 18-bp
AQdb-16S_R
Length: 22-bp
··· Sequences of the primer
··· Stability of the sequences within compared taxa
b
MtInsects-12S_HemHol_F
Length: 19-bp
MtInsects-12S_Hem_R
Length: 20-bp
MtInsects-12S_Hol_R
Length: 21-bp
MtInsects-12S_cad_F
Length: 21-bp
MtInsects-12S_cad_R
Length: 20-bp
AQdb-12S_Hem_R
Length: 19-bp
AQdb-12S_Hol_R
Length: 22-bp
AQdb-12S_cad_R
Length: 19-bp

Figure 1

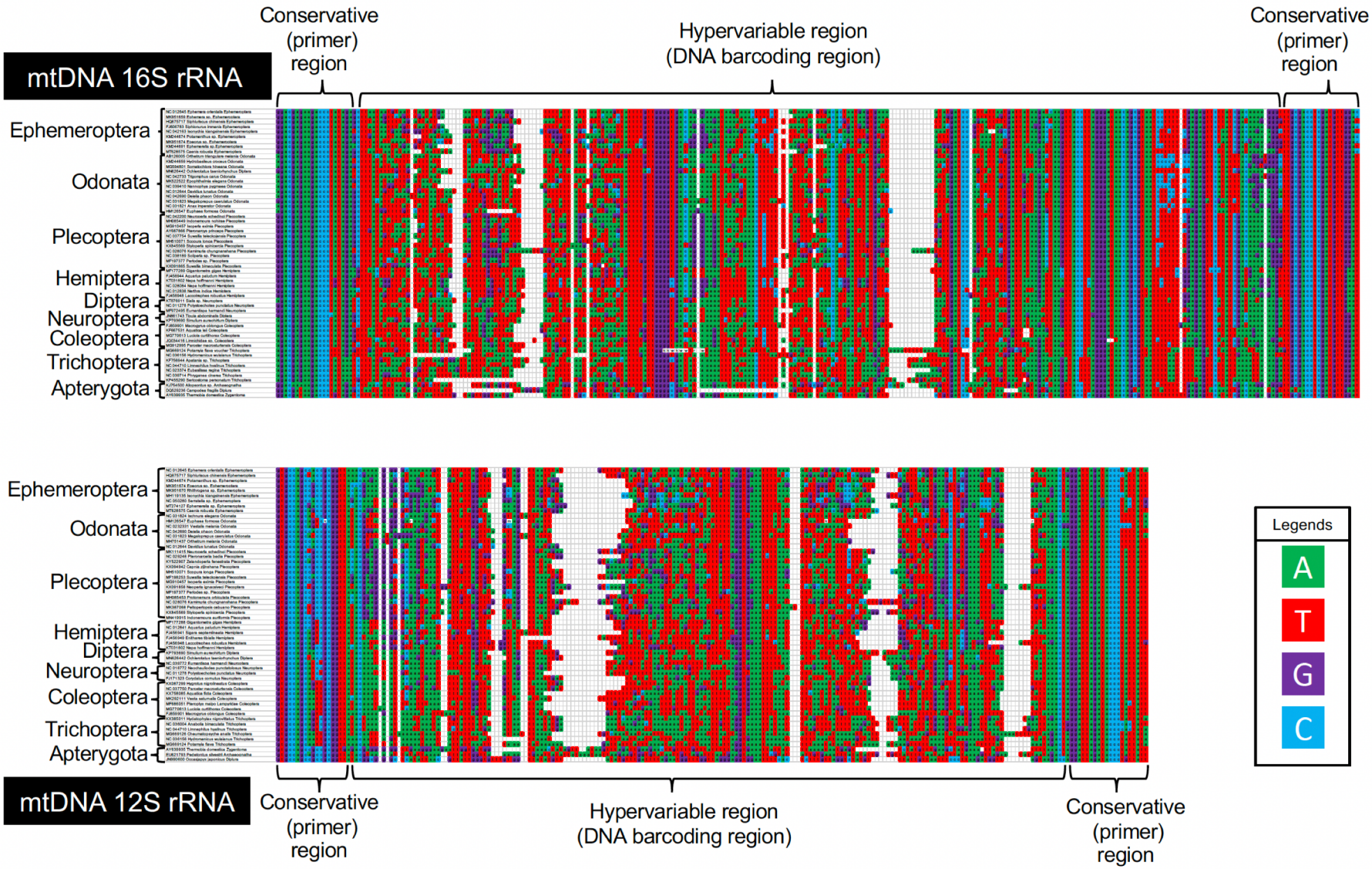
mtDNA 16S rRNA
Conservative (primer) region
Hypervariable region (DNA barcoding region)
Conservative (primer) region
Ephemeroptera
Odonata
Plecoptera
Hemiptera
Diptera
Neuroptera
Coleoptera
Trichoptera
Apterygota
mtDNA 12S rRNA
Conservative (primer) region
Hypervariable region (DNA barcoding region)
Conservative (primer) region
Legends
A
T
G
C

Figure 2

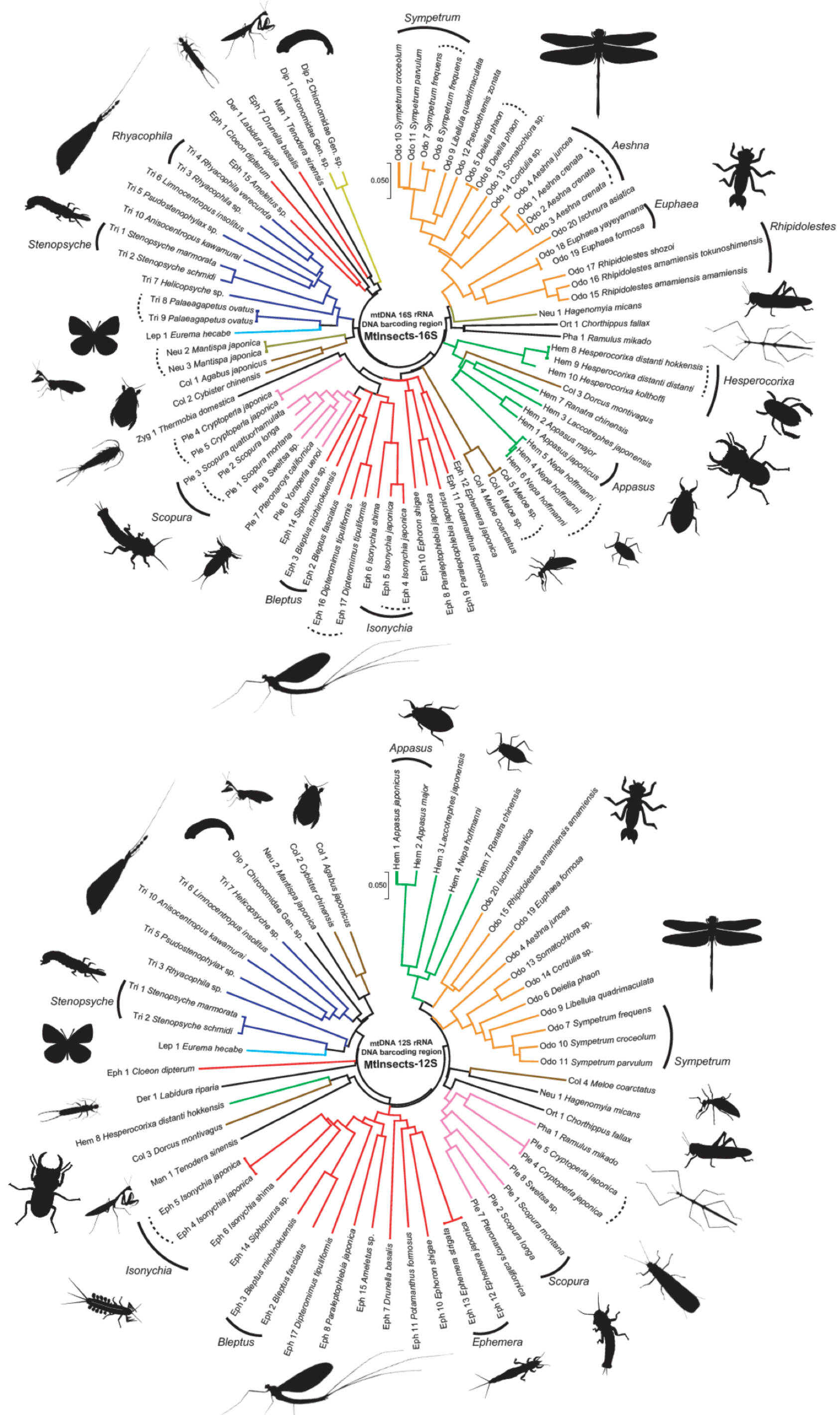
mtDNA 16S rRNA
DNA barcoding region
MtInsects-16S
0.050
Sympetrum
Aeshna
Euphaea
Rhipidolestes
Hesperocorixa
Appasus
Isonychia
Bleptus
Scopura
Stenopsyche
Rhyacophila
Odo 10 Sympetrum croceolum
Odo 11 Sympetrum parvulum
Odo 7 Sympetrum frequens
Odo 8 Sympetrum frequens
Odo 9 Libellula quadrimaculata
Odo 12 Pseudothemis zonata
Odo 5 Deielia phaon
Odo 6 Deielia phaon
Odo 13 Somatochlora sp.
Odo 14 Cordulia sp.
Odo 4 Aeshna juncea
Odo 1 Aeshna crenata
Odo 2 Aeshna crenata
Odo 3 Aeshna crenata
Odo 20 Ischnura asiatica
Odo 18 Euphaea yayeyamana
Odo 19 Euphaea formosa
Odo 17 Rhipidolestes shozoi
Odo 16 Rhipidolestes amamiensis tokunoshimensis
Odo 15 Rhipidolestes amamiensis amamiensis
Neu 1 Hagenomyia micans
Ort 1 Chorthippus fallax
Pha 1 Ramulus mikado
Hem 8 Hesperocorixa distanti hokkensis
Hem 9 Hesperocorixa distanti distanti
Hem 10 Hesperocorixa kolthoffi
Col 3 Dorcus montivagus
Hem 7 Ranatra chinensis
Hem 3 Laccotrephes japonensis
Hem 2 Appasus major
Hem 1 Appasus japonicus
Hem 5 Nepa hoffmanni
Hem 4 Nepa hoffmanni
Hem 6 Nepa hoffmanni
Col 5 Meloe sp.
Col 6 Meloe sp.
Col 4 Meloe coarctatus
Eph 12 Ephemera japonica
Eph 11 Potamanthus formosus
Eph 9 Paraleptophlebia japonica
Eph 8 Paraleptophlebia japonica
Eph 10 Ephoron shigae
Eph 4 Isonychia japonica
Eph 5 Isonychia japonica
Eph 6 Isonychia shima
Eph 17 Dipteromimus tipuliformis
Eph 16 Dipteromimus tipuliformis
Eph 2 Bleptus fasciatus
Eph 3 Bleptus michinokuensis
Eph 14 Siphlonurus sp.
Ple 6 Yoraperla uenoi
Ple 7 Pteronarcys californica
Ple 9 Sweltsa sp.
Ple 1 Scopura montana
Ple 2 Scopura longa
Ple 3 Scopura quattuorhamulata
Ple 5 Cryptoperla japonica
Ple 4 Cryptoperla japonica
Zyg 1 Thermobia domestica
Col 2 Cybister chinensis
Col 1 Agabus japonicus
Neu 3 Mantispa japonica
Neu 2 Mantispa japonica
Lep 1 Eurema hecabe
Tri 9 Palaeagapetus ovatus
Tri 8 Palaeagapetus ovatus
Tri 7 Helicopsyche sp.
Tri 2 Stenopsyche schmidi
Tri 1 Stenopsyche marmorata
Tri 10 Anisocentropus kawamurai
Tri 5 Psudostenophylax sp.
Tri 6 Limnocentropus insolitus
Tri 3 Rhyacophila sp.
Tri 4 Rhyacophila verecunda
Eph 15 Ameletus sp.
Eph 1 Cloeon dipterum
Der 1 Labidura riparia
Eph 7 Drunella basalis
Man 1 Tenodera sinensis
Dip 1 Chironomidae Gen. sp.
Dip 2 Chironomidae Gen. sp.
mtDNA 12S rRNA
DNA barcoding region
MtInsects-12S
0.050
Appasus
Sympetrum
Scopura
Ephemera
Bleptus
Isonychia
Stenopsyche
Hem 1 Appasus japonicus
Hem 2 Appasus major
Hem 3 Laccotrephes japonensis
Hem 4 Nepa hoffmanni
Hem 7 Ranatra chinensis
Odo 20 Ischnura asiatica
Odo 15 Rhipidolestes amamiensis amamiensis
Odo 19 Euphaea formosa
Odo 4 Aeshna juncea
Odo 13 Somatochlora sp.
Odo 14 Cordulia sp.
Odo 6 Deielia phaon
Odo 9 Libellula quadrimaculata
Odo 7 Sympetrum frequens
Odo 10 Sympetrum croceolum
Odo 11 Sympetrum parvulum
Col 4 Meloe coarctatus
Neu 1 Hagenomyia micans
Ort 1 Chorthippus fallax
Pha 1 Ramulus mikado
Ple 5 Cryptoperla japonica
Ple 4 Cryptoperla japonica
Ple 8 Sweltsa sp.
Ple 1 Scopura montana
Ple 2 Scopura longa
Ple 7 Pteronarcys californica
Eph 12 Ephemera japonica
Eph 13 Ephemera strigata
Eph 10 Ephoron shigae
Eph 11 Potamanthus formosus
Eph 7 Drunella basalis
Eph 15 Ameletus sp.
Eph 8 Paraleptophlebia japonica
Eph 17 Dipteromimus tipuliformis
Eph 2 Bleptus fasciatus
Eph 3 Bleptus michinokuensis
Eph 14 Siphlonurus sp.
Eph 6 Isonychia shima
Eph 4 Isonychia japonica
Eph 5 Isonychia japonica
Man 1 Tenodera sinensis
Col 3 Dorcus montivagus
Hem 8 Hesperocorixa distanti hokkensis
Der 1 Labidura riparia
Eph 1 Cloeon dipterum
Lep 1 Eurema hecabe
Tri 2 Stenopsyche schmidi
Tri 1 Stenopsyche marmorata
Tri 3 Rhyacophila sp.
Tri 5 Psudostenophylax sp.
Tri 10 Anisocentropus kawamurai
Tri 6 Limnocentropus insolitus
Tri 7 Helicopsyche sp.
Dip 1 Chironomidae Gen. sp.
Neu 2 Mantispa japonica
Col 2 Cybister chinensis
Col 1 Agabus japonicus

Figure 3

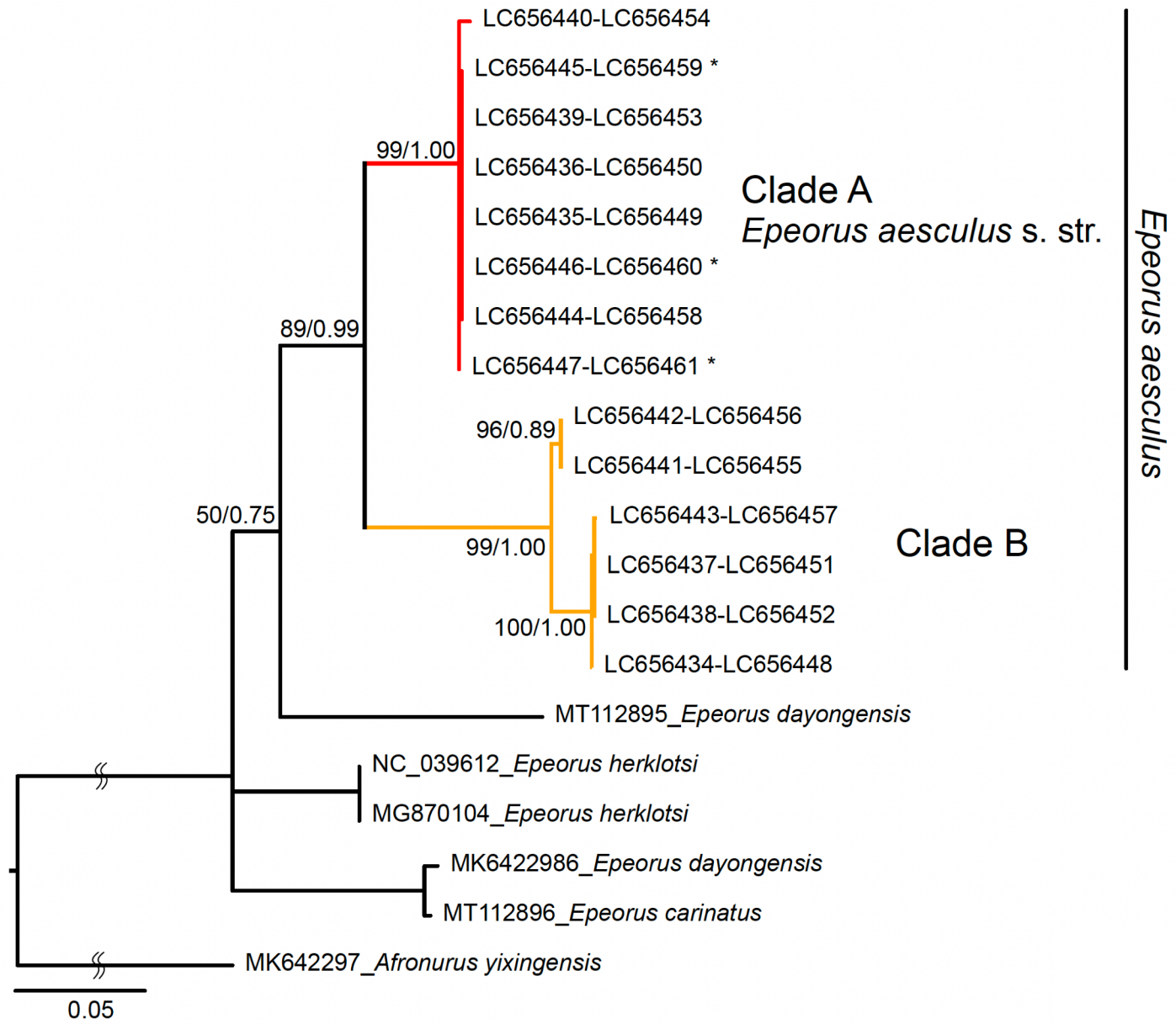
LC656440-LC656454
LC656445-LC656459 *
LC656439-LC656453
99/1.00
LC656436-LC656450
Clade A
Epeorus aesculus s. str.
LC656435-LC656449
LC656446-LC656460 *
89/0.99
LC656444-LC656458
LC656447-LC656461 *
Epeorus aesculus
96/0.89
LC656442-LC656456
LC656441-LC656455
50/0.75
LC656443-LC656457
99/1.00
Clade B
LC656437-LC656451
100/1.00
LC656438-LC656452
LC656434-LC656448
MT112895_Epeorus dayongensis
NC_039612_Epeorus herklotsi
MG870104_Epeorus herklotsi
MK6422986_Epeorus dayongensis
MT112896_Epeorus carinatus
MK642297_Afronurus yixingensis
0.05

Figure 4