ARPHA Preprints

PREPRINT

Author-formatted, not peer-reviewed document posted on 22/01/2025

DOI: https://doi.org/10.3897/arphapreprints.e147415

Mitochondrial 16S DNA data and voucher specimen collection of Japanese aquatic Coleoptera and Hemiptera for environmental DNA metabarcoding

Naoyuki Nakahama, Kei Hirasawa, Masaya Kato, Kohei Watanabe, Seikan Kurata, Masakazu Hayashi

Mitochondrial 16S DNA data and voucher specimen collection of Japanese aquatic Coleoptera and Hemiptera for environmental DNA metabarcoding

Running head: *16S DNA dataset of aquatic insects*

Naoyuki Nakahama[1,2*], Kei Hirasawa[3], Masaya Kato[4], Kohei Watanabe[5], Seikan Kurata[6], Masakazu Hayashi[7]

[1]*Institute of Natural and Environmental Sciences, University of Hyogo, Sanda City, Hyogo, 669–1546, Japan*

[2]*Museum of Nature and Human Activities, Hyogo, 6 Yayoigaoka, Sanda Hyogo, 669–1546, Japan*

[3]*Aquamarine Inawashiro Kingfishers Aquarium, Inawashiro Town, Fukushima, 969–3283, Japan*

[4]*Graduate School of Agriculture, Osaka Metropolitan University, Sakai City, Osaka, 599–8531, Japan*

[5]*Ishikawa Insect Museum, Hakusan City, Ishikawa, 920–2113, Japan*

[6]*Tomakomai Experimental Forest, Field Science Center for Northern Biosphere, Hokkaido University, Tomakomai City, Hokkaido, 053–0035, Japan*

[7]*Hoshizaki Green Foundation, Sono, Izumo, 691–0076, Japan*

ORCID

NN: 0000-0003-3106-8289

KH: 0009-0008-2296-3699

KW: 0000-0002-8761-232X

SK: 0000-0003-3825-3938

MH: 0000-0003-3708-9005

*Correspondence: Naoyuki Nakahama

Institute of Natural and Environmental Sciences, University of Hyogo, 6 chome, Yayoigaoka, Sanda, Hyogo, 669-1546, Japan

E-mail: naoyuki.halobates@gmail.com

ORCID: 0000-0003-3106-8289

Abstract

Aquatic Coleoptera and Hemiptera insects primarily inhabit lentic waters, many of which are at risk of extinction due to development, agriculture, and invasive alien species. Environmental DNA (eDNA) analysis has recently emerged as a powerful tool for conducting comprehensive distribution surveys. The cytochrome c oxidase subunit I (COI) universal primers are conventionally used for DNA barcoding by non-specific amplification and frequent amplification failures. Primers in the mitochondrial 16S region that alleviate these issues have been developed and are considered helpful for eDNA analysis. Therefore, it is necessary to accumulate reference sequences of the 16S region in aquatic Coleoptera and Hemiptera insects. However, molecular identification at the genus or species level remains challenging, as only a few of these insect groups in Japan have registered both COI and 16S reference DNA sequences. Therefore, we constructed a comprehensive dataset of the 16S region of mitochondrial DNA for these insects distributed in Japan. As a result of this study, we were able to obtain partial sequences of the 16S region from 140 taxa (35.5% of Japanese aquatic Coleoptera species or subspecies) and 58 taxa (45.3% of Japanese aquatic Hemiptera species or subspecies). These voucher specimens were deposited in four research institutions. The DNA sequence datasets are expected to significantly contribute as an essential database for eDNA analysis and other DNA metabarcoding studies.

Keywords: aquatic insects, distribution survey, DNA metabarcoding, environmental DNA, high-throughput sequencer, lentic water, mitochondrial DNA

Introduction

Aquatic insects play vital roles in maintaining ecosystems (Macadam and Stockan 2015; Raitif et al. 2019). For example, they drive nutrient cycles in terrestrial and aquatic environments and serve as predators of vectors of infectious diseases (Hupfer et al. 2015; Dambach 2020). Therefore, the conservation of aquatic insect diversity is a critical global issue (Macadam and Stockan 2015). However, wetland ecosystems are declining worldwide, and many aquatic insects are threatened with extinction (Collier et al. 2016; IUCN 2024). As many of these threats stem from human activities, such as land development, invasive alien species, pesticides, and climate change (Collier et al. 2016; IUCN 2024), aquatic insect diversity will continue to decline in the future unless conservation efforts are implemented.

In Japan, as many aquatic insects, such as Coleoptera and Hemiptera, are critically threatened, it is imperative to monitor their distribution and conservation activities (Ministry of the Environment, Government of Japan, 2014; Hayashi et al. 2020; Nakajima et al. 2020). In total, 30.4% of these aquatic Coleoptera insects and 24.6% of aquatic Hemiptera insects are listed in various ranks on Japan's Ministry of Environment's Red List, 2019 (Hayashi et al. 2020). The risk and urgency primarily arise because many of these groups inhabit lentic waters. More than 60% of both aquatic Coleoptera and Hemiptera insect species reside in lentic waters such as rice paddies, ponds, and puddles (Hayashi et al. 2020; Nakajima et al. 2020). Hence, the abandonment or intensification of paddies and ponds, along with the presence of alien species and pesticides, lead to population decline (Ministry of the Environment, Government of Japan 2014; Nakajima et al. 2020). In particular, invasive alien species rapidly degrade the biodiversity in lentic waters, even when the potential habitat area remains unchanged (Gherardi and

Acquistapace 2007; Ohba 2011). Therefore, understanding the distribution of conservation target species is essential for formulating effective conservation measures. Hence, regular and comprehensive surveys are crucial for developing conservation policies.

In recent years, environmental DNA (eDNA) analysis has emerged as a valuable tool to estimate the distribution of wildlife (Ruppert et al. 2019; Duarte et al. 2023). It has helped monitor many aquatic insects as well (Doi et al. 2017; Wang et al. 2024). In addition to the species-specific eDNA analysis primer designing (Doi et al. 2017; Ogata et al. 2023), metabarcoding has also helped in analyzing aquatic insect fauna, contributing significantly to understanding their distribution (Takenaka et al. 2024; Wang et al. 2024).

In aquatic insects, metabarcoding using eDNA is essential for improving identification accuracy, which relies on the availability of reference sequences of each species. However, in Japan, the development of aquatic insect DNA libraries is slow (Kishimoto-Yamada et al. 2022). For example, only 8% of Coleoptera and 6% of Hemiptera taxa sequences have been registered in the Barcode of Life Data System (BOLD). Although the DNA barcode regions of Cytochrome c oxidase subunit I (COI) have been determined for several regions and taxa of aquatic insects (Inai et al. 2022; Hayashi et al. 2024a, b), they are not comprehensive enough to cover Japanese taxa. Additionally, the COI region poses several challenges as a marker for eDNA metabarcoding analysis, including the requirement of longer sequence length (*i.e.*, ca. 650 bp) and frequent mutation of primer, which may result in non-specific amplification and amplification failure (Takenaka et al. 2023). Thus, there were significant issues with the marker's data availability and accumulation for eDNA metabarcoding analysis.

Here, we focused on the 16S region of mitochondrial DNA. Compared to the COI region, it has a slower mutation rate (Papadopoulou et al. 2010) and poses a lower risk of PCR

failure or pseudogene amplification in some taxonomic groups (Yano et al. 2020; Takenaka et al. 2023). Recently, a primer set targeting the 16S region ('MtInsects-16S') was developed. Because of the short PCR product lengths (*i.e.*, ca. 200 bp), this primer set is considered suitable for eDNA analysis (Takenaka et al. 2023). Moreover, Takenaka et al. (2024) also elucidated its applicability for metabarcoding with eDNA in stream insects, particularly in Ephemeroptera, Plecoptera, and Trichoptera (Takenaka et al. 2024). Based on these studies, we concluded that the 16S region is well-suited for eDNA metabarcoding analysis, and the database enrichment in aquatic Coleoptera and Hemiptera insects is essential.

In this study, we constructed the sequence datasets of the 16S region for aquatic Coleoptera and Hemiptera insects that spend a significant portion of their life cycle in or on water. These groups are predominantly found in lentic waters in Japan and include many endangered species (Hayashi et al. 2020). The dataset is anticipated to significantly contribute to future distribution surveys using eDNA metabarcoding and support the construction of conservation policies based on these surveys.

Materials and Methods

Target taxa and DNA sampling

This study includes the taxa listed in the 'List of Aquatic Coleoptera and Hemiptera of Japan' (Nakajima 2024, accessed in October 2024) as the target taxa. The list comprises Coleoptera taxa (Chrysomelidae, Dytiscidae, Gyrinidae, Haliplidae, Hydraenidae, Hydrophiloidea, Microsporioidea, and Noteridae) and Hemiptera taxa (Gerromorpha and Nepomorpha), which majorly spend their life cycle in water or on the surface, as documented in Japan. The scientific name of each taxon follows Nakajima (2024).

Additionally, the taxa Gelastocoridae and Heteroceridae were included, as listed in Nakajima et al. (2020), and the nomenclature was based on the same source (Nakajima et al. 2020). Although *Micronecta* spp. is categorized under Corixidae in Nakajima (2024), prior studies have classified it under Micronectidae, a practice we adopt here (Meyin A Ebong et al. 2016; Ye et al. 2023; Xie et al. 2024).

In total, we collected 199 Coleoptera taxa and 68 Hemiptera taxa. All sampled specimens were preserved in 100% ethanol for further analysis.

DNA extraction and library preparation

For individuals with a body length of approximately 5 mm or more, one middle or hind leg was cut off and used for DNA extraction, while the entire body was used for those with a body length of less than 5 mm. Genomic DNA was extracted using the Chelex method according to the protocol by Miura et al. (2017). DNA was extracted from all samples without destroying the body, and they were deposited as voucher specimens in four institutions: Aquamarine Inawashiro Kingfishers Aquarium, Hoshizaki Institute for Wildlife Protection, Ishikawa Insect Museum, Museum of Nature and Human Activities, Hyogo (Table S1).

A single pair of primers was used to amplify the partial sequence of the mitochondrial 16S DNA; MtInsects-16S_F: 5′-GGACGAGAAGACCCTWTAGA-3′ and MtInsects-16S_R: 5′-GGACGAGAAGACCCTWTAGA-3′ (Takenaka et al. 2023). The first PCR was performed as a 10 µl reaction volume containing 1 ng template DNA, 5 µl of 2×PCR Buffer for KOD -Multi & Epi-, 0.2 µl KOD -Multi & Epi- (TOYOBO), and 0.1µl of 20 µM for each primer. The PCR cycle for the COII gene was as follows: template denaturation at 94°C for 12 min, followed by 30 cycles of denaturation at 94°C for 1 min, annealing at

50°C for 1 min, and extension at 72°C for 0.5 min; followed by a final extension at 72°C for 3 min. Library preparation after the first PCR followed the MPM-seq protocol by Suyama et al. (2022). Subsequent paired-end sequencing was conducted using 2 × 250 bp cycle run on an Illumina MiSeq Sequencer (Illumina, San Diego, USA) and with the MiSeq Reagent Nano Kit v.2 (500 cycles).

Assembling sequences

Data pre-processing, quality control, and identification of representative sequences for the mitochondrial 16S gene were conducted using Claident v.0.2.2019.05.10 (Tanabe and Toju 2013), as described by Suyama et al. (2022) and Kurata et al. (2024).

Non-demultiplexed fast files were required for quality control and data analysis using Claident. The non-demultiplexed fast files (261 bp) were generated from the BCL files using bcl2fastq v.1.8.4 (Illumina). During this step, non-demultiplexed fast reads were sorted based on the index reads (index1: 9 bp, index2: 5 bp), and the last position of the raw reads were trimmed (--use-bases-mask Y260n,I9,I5,Y260n), as per the settings in Kurata et al. (2024). The Claident command of *clsplitseq* was then used to demultiplex the nonmultiplexed fast reads, specifying the indices and primer sequences with a quality threshold of the index sequence set to 30 (--minqualtag = 30). The option of "--truncateN = enable" was included in *clsplitseq* because 0–3 Ns were added to the beginning of the primer sequences.

To identify the overlaps between the forward and reverse reads of the three genes, the *clconcatpair* command was used with the same settings mentioned in Kurata et al. (2024): *i.e.*, the --mode = OVL argument was used to generate concatenated reads from the forward and reverse sequences. In addition, any low-quality reads were filtered out using the

clfilterseq command settings from Kurata et al. (2024): *i.e.*, --maxplowequal = 0.1 --minqual = 27, to remove positions with a quality score lower than Q27. Finally, the *clcleanseqv* parameters were used to remove noisy and chimeric sequences with the following settings: --derepmode = FULLLENGTH --primarymaxnmismatch=0 --secondarymaxnmismatch = 1 --pnoisycluster = 0.5. The *clclasseqv* command was used to identify representative sequences for each sample and each gene, with a 99% identity threshold (--minident = 0.99). The mitochondrial 16S sequences were directly detected via the *clclasseqv* step.

Phylogenetic analysis

The phylogenetic relationships were constructed based on the 16S sequences to check the correctness of each obtained sequence of aquatic Coleoptera and Hemiptera insects. Mitochondrial sequence alignment was performed using ClustalW (Thompson et al., 1994) and an algorithm implemented in MEGA X (Kumar et al. 2018). The maximum likelihood (ML) tree was generated using IQ-TREE 2.2.0 (Minh et al. 2020) and ultrafast bootstrapping (UB) with 1000 replicates. The substitution models were estimated using ModelFinder (Kalyaanamoorthy et al. 2017) within IQ-TREE version 2.2.0 based on the BIC value. The best models chosen for each locus were GTR+F+I+G4 for Coleoptera insects and TVM+F+I+G4 for Hemiptera insects.

The resulting phylogenetic trees were visualized and edited using FigTree version 1.4.4. For taxa where phylogenetic analysis did not indicate monophyly (*Anacaena okinawana*, *Berosus japonicus*, *Copelatus zimmermanni*, *Elmomorphus brevicornis*, *Mesovelia miyamotoi*, *Neohydrocoptus* sp., *Peltodytes intermedius*, *Peltodytes sinensis*, and *Regimbartia attenuata*), BLAST searches were performed to identify matching sequences

for related taxa within the same genus.

Results and Discussion

After clustering in Claident, we obtained 16–5422 reads for aquatic Coleoptera (median:1,379; average:1,442) and 6–7093 reads for aquatic Hemiptera insects (median, 1867; average, 2277). All species analyzed via BLAST searches were identified as closely related, indicating that accurate sequences were obtained.

This study successfully determined the sequences for 70.3% and 85.0% of the collected Coleoptera and Hemiptera taxa, respectively. The failures in sequence identification were likely due to experimental errors or primer mismatches. Finally, we obtained 140 Coleoptera taxa (including species and subspecies), representing 68 genera (Fig. 1, Table 1, Tables S1, 2). These covered 35.5% and 72.0% of total taxa and genera of aquatic Coleoptera insects in Japan. Similarly, 58% of taxa were determined for aquatic Hemiptera, representing 31 genera (Fig. 2, Table 1, Tables S1, 3). These accounted for 45.3% and 70.5% of total taxa and genera of aquatic Hemiptera insects in Japan.

Although barcode sequences for Japanese aquatic insects have been previously published (Inai et al. 2022, Takenaka et al. 2024), a few databases in Japan comprehensively cover these insects, with exceptions like Chironomidae (National Institute for Environmental Studies 2013) and Odonata insects (Kishimoto-Yamada et al. 2022). The addition of 16S sequences determined in this study and deposited in Genbank significantly increases the coverage of aquatic Coleoptera and Hemiptera insects. The expanded dataset is expected to contribute significantly to eDNA analyses in the future.

In addition to determining the 16S sequences of aquatic Coleoptera and Hemiptera insect taxa, this study prepared voucher specimens deposited in four institutions. Many

227 undescribed Coleoptera and Hemiptera insects in Japan are still being discovered and
228 described, making the phylogenetic analysis critical for taxonomic revision (Hayashi et al.
229 2020; Nakajima et al. 2020). On the other hand, since many of these taxa are classified as
230 endangered (Ministry of the Environment, Government of Japan 2014), the voucher
231 specimen collections, which include both the morphological and DNA sequence
232 information, housed in Japanese research institutions and museums are of utmost
233 importance for future conservation and taxonomic efforts.
234 The DNA sequences obtained in our study are intended primarily as reference sequences
235 for DNA meta-barcoding, such as eDNA analysis. These sequences, being relatively short
236 (~200bp), are not suitable for phylogenetic analysis; therefore, they do not provide accurate
237 estimation of phylogenetic relationships. For example, Dytiscidae and Hydrophilidae were
238 estimated as polyphyletic (Fig. 1), which contradicts the phylogenetic relationships of
239 Coleoptera estimated through genomic-level analysis (McKenna et al. 2019; Cai et al.
240 2022). In addition, *Heterlimnius masakazui* and *H. yoshitomii* (Elmidae) could not be
241 distinguished as two different species using the obtained sequences. These limitations
242 clearly stem from the short sequence length. Therefore, for robust phylogenetic analysis, it
243 is necessary to integrate these sequences with data from other loci.
244 In addition to eDNA analysis, the sequence data obtained would be highly valuable for
245 other DNA metabarcoding research. For instance, it could facilitate the analysis of feeding
246 habits from fecal samples of insectivores, ensuring molecular identification from small
247 tissue samples (Rytkönen et al. 2019; Ingala et al. 2021). Our comprehensive DNA
248 sequence dataset would enhance the detailed identification of food organism taxa.
249 Furthermore, it could aid in the molecular identification of samples with degraded and
250 fragmented DNA, such as museum specimens or parts of a deceased organism found in the

wild. This is because PCR is more likely to succeed when primers targeting short PCR products are used (Nakahama and Isagi 2017). Thus, our sequence dataset is expected to be applicable not only for eDNA analysis but also for a variety of molecular research studies.

There are two main limitations for the future DNA sequence dataset of Japanese aquatic Coleoptera and Hemiptera insects. The first limitation is the need for comprehensive enrichment of the sequence dataset. This study provided the data for just under 40% of Japanese aquatic Coleoptera and Hemiptera insects. However, a significant number of endangered species, such as *Kirkaldyia deyrolli* (Belostomatidae) and *Dytiscus sharpi* (Dytiscidae) (Ministry of the Environment, Government of Japan 2014), remain unsequenced. Therefore, further studies to enrich the dataset are essential. The second limitation concerns the determination of longer sequences of the 16S region. Although 16S region primers targeting a longer PCR product were previously designed (AQdb-16S_F and AQdb-16S_R) (Takenaka et al. 2023), this study used primers targeting a shorter PCR product (MtInsects-16S_F and MtInsects-16S_R) to accommodate the MiSeq Reagent Nano Kit v.2 (500 cycles) used for sequencing. Recent advancements in high-throughput sequencing technologies, such as Nanopore, which allow the determination of longer sequences, have facilitated the construction of various DNA barcoding databases (Menegon et al. 2017; Hebert et al. 2023). Moving forward, the construction of the dataset by determining longer DNA sequences would enhance the utility, not only for eDNA analysis but also for phylogenetic analysis.

Acknowledgements

We thank Asato Noishiki, Jun Nakajima, Kazutaka Togashi, Shigeyuki Yoshii, Shintarou Yasui, Shun Yamasaki, Toshimasa Mitamura, Yoshiaki Sakai, Leo Takechi and Yusuke Tomimori for providing their collection materials. We also thank Minatsu Mukai, Shiho Araki, Shunya Tamaru and Yuko Ichimachi, Kazutaka Yamada and Takeshi Yamasaki for helping of curation our sequence data and voucher specimens.

Funding Program

JSPS KAKENHI (19K15856 and 24K01781)

Author Contributions

Naoyuki Nakahama conceived the ideas. Kei Hirasawa, Masaya Kato, Kohei Watanabe and Masakazu Hayashi collected aquatic insect samples. Naoyuki Nakahama and Seikan Kurata performed sequence analysis and interpret the analysis results. Naoyuki Nakahama Kei Hirasawa, Masaya Kato, Kohei Watanabe and Masakazu Hayashi curated the samples and made the voucher specimens. All authors contributed to write the manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

Cai C, Tihelka E, Giacomelli M, Lawrence JF, Ślipiński A, Kundrata R, Yamamoto S, Thayer MK, Newton AF, Leschen RAB, Gimme; ML, Lü L, Engel MA, Bouchard P, Huang D, Pisani D, Donoghue, PC. (2022) Integrated phylogenomics and fossil data illuminate the evolution of beetles. Royal Society Open Science 9(3): 211771. https://doi.org/10.1098/rsos.211771

Collier KJ, Probert, PK, Jeffries, M (2016) Conservation of aquatic invertebrates: concerns, challenges and conundrums. Aquatic Conservation: Marine and Freshwater Ecosystems 26(5): 817–837. https://doi.org/10.1002/aqc.2710

Dambach, P (2020) The use of aquatic predators for larval control of mosquito disease vectors: opportunities and limitations. Biological Control 150: 104357. https://doi.org/10.1016/j.biocontrol.2020.104357

Doi H, Katano I, Sakata Y, Souma R, Kosuge T, Nagano M, Ikeda K, Yano K, Tojo K (2017) Detection of an endangered aquatic heteropteran using environmental DNA in a wetland ecosystem. Royal Society Open Science 4(7): 170568. https://doi.org/10.1098/rsos.170568

Duarte S, Simões, L, Costa, FO (2023) Current status and topical issues on the use of eDNA-based targeted detection of rare animal species. Science of The Total Environment 904: 166675. https://doi.org/10.1016/j.scitotenv.2023.166675

Gherardi F, Acquistapace P (2007) Invasive crayfish in Europe: the impact of *Procambarus clarkii* on the littoral community of a Mediterranean lake. Freshwater Biology 52(7): 1249–1259. https://doi.org/10.1111/j.1365-2427.2007.01760.x

Hayashi M, Nakajima J, Ishida K, Kitano T, Yoshitomi H (2020) Species diversity of aquatic Hemiptera and Coleoptera in Japan. Japanese Journal of Systematic

Entomology 26(2): 191–200.

Hayashi M, Souma R, Iwata Y (2024a) DNA barcodes of Hydraenidae (Coleoptera) of Japan. Bulletin of the Hoshizaki Green Foundation 27: 111–117 [In Japanese]

Hayashi M, Souma R, Iwata Y (2024b) DNA Barcodes of Aquatic Adephaga (Coleoptera) of Shimane Prefecture, Japan. Special Bulletin of the Hoshizaki Green Foundation 34: 1–20. [In Japanese]

Hebert, PD, Floyd, R, Jafarpour, S, Prosser, SW (2023) Barcode 100k specimens: in a single nanopore run. Molecular Ecology Resources 25: e14028. https://doi.org/10.1111/1755-0998.14028

Hupfer M, Jordan S, Herzog C, Ebeling C, Ladwig R, Rothe M, Lewandowski, J (2019) Chironomid larvae enhance phosphorus burial in lake sediments: Insights from long-term and short-term experiments. Science of the Total Environment 663: 254–264. https://doi.org/10.1016/j.scitotenv.2019.01.274

Inai K, Wakimura K, Kato M (2022) Operational taxonomic units of cox1 amplified from bulk community DNA samples in Takamigawa River, Nara, Japan, Mendeley Data V1 https://doi.org/10.17632/ pfhf4kt37s.1

Ingala MR, Simmons NB, Wultsch C, Krampis K, Provost KL, Perkins SL (2021) Molecular diet analysis of neotropical bats based on fecal DNA metabarcoding. Ecology and Evolution 11(12): 7474–7491. https://doi.org/10.1002/ece3.7579

IUCN (2024) The IUCN Red List of Threatened Species. Version 2024-2. [Accessed 27 Nov. 2024] https://www.iucnredlist.org

Kalyaanamoorthy S, Minh BQ, Wong TK, Von Haeseler A, Jermiin, LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nature Methods 14(6): 587–589. https://doi.org/10.1038/nmeth.4285

347 Kishimoto-Yamada, K, Jinbo U, Nakahama N (2022) Review of current status of DNA
348 barcoding of Japanese insects. The Nature and Insects 57: 2–6 [In Japanese]
349 Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary
350 genetics analysis across computing platforms. Molecular Biology and Evolution 35:
351 1547–1549. https://doi.org/10.1093/molbev/msy096
352 Kurata S, Mano S, Nakahama N, Hirota SK, Suyama Y, Ito M (2024) Development of
353 mitochondrial DNA cytochrome c oxidase subunit I primer sets to construct DNA
354 barcoding library using next-generation sequencing. Biodiversity Data Journal 12:
355 e117014. https://doi.org/10.3897/BDJ.12.e117014
356 Macadam CR, Stockan, JA (2015) More than just fish food: ecosystem services provided
357 by freshwater insects. Ecological Entomology 40: 113–123.
358 https://doi.org/10.1111/een.12245
359 McKenna DD, Shin S, Ahrens D, Balke M, Beza-Beza C, Clarke, DJ, Donath A, Escalona
360 HE, Friedrich F, Letsch H, Liu S, Maddison D, Mayer C, Misof B, Murin PJ, Niehuis
361 O, Peters RS, Podsiadlowski L, Pohl H, Scully ED, Yan EV, Zhou X, Ślipiński A,
362 Beutel RG (2019) The evolution and genomic basis of beetle diversity. Proceedings of
363 the National Academy of Sciences 116(49): 24729–24737.
364 https://doi.org/10.1073/pnas.1909655116
365 Menegon M, Cantaloni C, Rodriguez-Prieto A, Centomo C, Abdelfattah A, Rossato M,
366 Bernardi M, Xumerle L, Loader S, Delledonne M (2017) On site DNA barcoding by
367 nanopore sequencing. PLoS One 12(10): e0184741.
368 https://doi.org/10.1371/journal.pone.0184741
369 Meyin A Ebong S, Petit E, Le Gall P, Chen PP, Nieser N, Guilbert E, Njiokou F, Marsollier
370 L, Guégan GF, Pluot-Sigwalt D, Eyangoh S, Harry, M (2016) Molecular species

371 delimitation and morphology of aquatic and sub-aquatic bugs (Heteroptera) in
372 Cameroon. PLoS One 11(5): e0154905. https://doi.org/10.1371/journal.pone.0154905
373 Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A,
374 Lanfear R (2020) IQ-TREE 2: new models and efficient methods for phylogenetic
375 inference in the genomic era. Molecular Biology and Evolution 37(5): 1530–1534.
376 https://doi.org/10.1093/molbev/msaa015
377 Ministry of the Environment, Government of Japan (2014) Threatened wildlife of Japan-
378 red data book, Insects. Japan Wildlife Research Center, Tokyo, Japan, 509 pp. [In
379 Japanese]
380 Miura K, Higashiura Y, Maeto K (2017) Evaluation of easy, non-destructive methods of
381 DNA extraction from minute insects. Applied Entomology and Zoology 52: 349–352.
382 https://doi.org/10.1007/s13355-017-0481-4
383 Nakahama N, Isagi Y (2017) Availability of short microsatellite markers from butterfly
384 museums and private specimens. Entomological Science 20(1): 3–6.
385 https://doi.org/10.1111/ens.12215
386 Nakajima J (2024) List of Aquatic Coleoptera and Hemiptera of Japan. [Accessed 23. Oct
387 2024] http://kuromushiya.com/mlist/mlist.html
388 Nakajima J, Hayashi M, Ishida K, Kitano T, Yoshitomi H (2020) Aquatic Coleoptera and
389 Hemiptera of Japan. Bun-ichi Sogo Shuppan, Tokyo, Japan, 352 pp. [In Japanese]
390 National Institute for Environmental Studies (2013) Chironomid DNA barcode database.
391 Biodiversity Division, National Institute for Environmental Studies, Tsukuba, Japan,
392 under CC BY 4.0. [Accessed 10 Dec. 2024]
393 https://www.nies.go.jp/yusurika/en/index.html
394 Ogata S, Nishiwaki A, Yamazoe K, Sugai K, Takahara T (2023) Discovery of unknown

new ponds occupied by the endangered giant water bug Kirkaldyia deyrolli (Hemiptera: Heteroptera: Belostomatidae) by combining environmental DNA and capture surveys. Entomological Science, 26(1): e12540. https://doi.org/10.1111/ens.12540

Ohba SY (2011) Impact of the crayfish *Procambarus clarkii* on the giant water bug *Kirkaldyia deyrolli* (Hemiptera) in rice ecosystems. Japanese Journal of Environmental Entomology and Zoology 22(2): 93–98. https://doi.org/10.11257/jjeez.22.93

Papadopoulou A, Anastasiou I, Vogler AP (2010) Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. Molecular Biology and Evolution 27(7): 1659–1672. https://doi.org/10.1093/molbev/msq051

Raitif J, Plantegenest M, Roussel JM (2019) From stream to land: Ecosystem services provided by stream insects to agriculture. Agriculture, Ecosystems & Environment 270: 32–40. https://doi.org/10.1016/j.agee.2018.10.013

Ruppert KM, Kline RJ, Rahman MS (2019) Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. Global Ecology and Conservation 17: e00547. https://doi.org/10.1016/j.gecco.2019.e00547

Rytkönen S, Vesterinen EJ, Westerduin C, Leviäkangas T, Vatka E, Mutanen M, Välimäki P, Hukkanen M, Suokas M, Orell M (2019) From feces to data: A metabarcoding method for analyzing consumed and available prey in a bird-insect food web. Ecology and Evolution 9(1): 631–639. https://doi.org/10.1002/ece3.4787

Suyama Y, Hirota SK, Matsuo A, Tsunamoto Y, Mitsuyuki C, Shimura A, Okano K (2022) Complementary combination of multiplex high-throughput DNA sequencing for molecular phylogeny Ecological Research 37: 171–181. https://doi.org/10.1111/1440-1703.12270

419 Takenaka M, Hasebe Y, Yano K, Okamoto S, Tojo K, Seki M, Sekiguchi S, Jitsumasa T,
420 Morohashi N, Handa Y, Sakaba T (2024) Environmental DNA metabarcoding on
421 aquatic insects: Comparing the primer sets of MtInsects-16S based on the mtDNA 16S
422 and general marker based on the mtDNA COI region. Environmental DNA 6(4): e588.
423 https://doi.org/10.1002/edn3.588
424 Takenaka M, Yano K, Suzuki T, Tojo K (2023) Development of novel PCR primer sets for
425 DNA barcoding of aquatic insects, and the discovery of some cryptic species.
426 Limnology 24(2): 121–136. https://doi.org/10.1007/s10201-022-00710-5
427 Tanabe AS, Toju H (2013) Two new computational methods for universal DNA barcoding:
428 A benchmark using barcode sequences of bacteria, archaea, animals, fungi, and land
429 plants. PLoS One 11(3): e0152242. https://doi.org/10.1371/journal.pone.0152242
430 Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of
431 progressive multiple sequence alignment through sequence weighting, position-
432 specific gap penalties and weight matrix choice. Nucleic Acids Research 22: 4673–
433 4680. https://doi.org/10.1093/nar/22.22.4673
434 Wang Z, Li F, Wu F, Guo F, Gao W, Zhang Y, Yang Z (2024) Environmental DNA and
435 remote sensing datasets reveal the spatial distribution of aquatic insects in a disturbed
436 subtropical river system. Journal of Environmental Management 351: 119972.
437 https://doi.org/10.1016/j.jenvman.2023.119972
438 Xie BJ, Chen PP, Damgaard J, Xie JY, Xie Q, Wang YH (2024) Paraphyly of the subgenus
439 *Micronecta* (Micronecta) Kirkaldy, 1897 (Hemiptera: Heteroptera: Micronectidae)
440 based on mitochondrial genomes and nuclear rDNAs. Arthropod Systematics &
441 Phylogeny 82: 77–87. https://doi.org/10.3897/asp.82.e108906
442 Yano K, Takenaka M, Mitamura T, Tojo K (2020) Identifying a “pseudogene” for the

mitochondrial DNA COI region of the corixid aquatic insect, *Hesperocorixa distanti* (Heteroptera, Corixidae). Limnology 21: 319–325. https://doi.org/10.1007/s10201-020-00618-y

Ye Z, Damgaard J, Hädicke CW, Zhu X, Mazzucconi SA, Hebsgaard MB, Xie T, Yang H, Bu W (2023) Phylogeny and historical biogeography of the water boatmen (Insecta: Hemiptera: Heteroptera: Nepomorpha: Corixoidea). Molecular Phylogenetics and Evolution 180: 107698. https://doi.org/10.1016/j.ympev.2022.107698

Table 1. Number of Japanese aquatic Coleopteta and Hemiptera taxa used in this study, of which sequences were obtained.

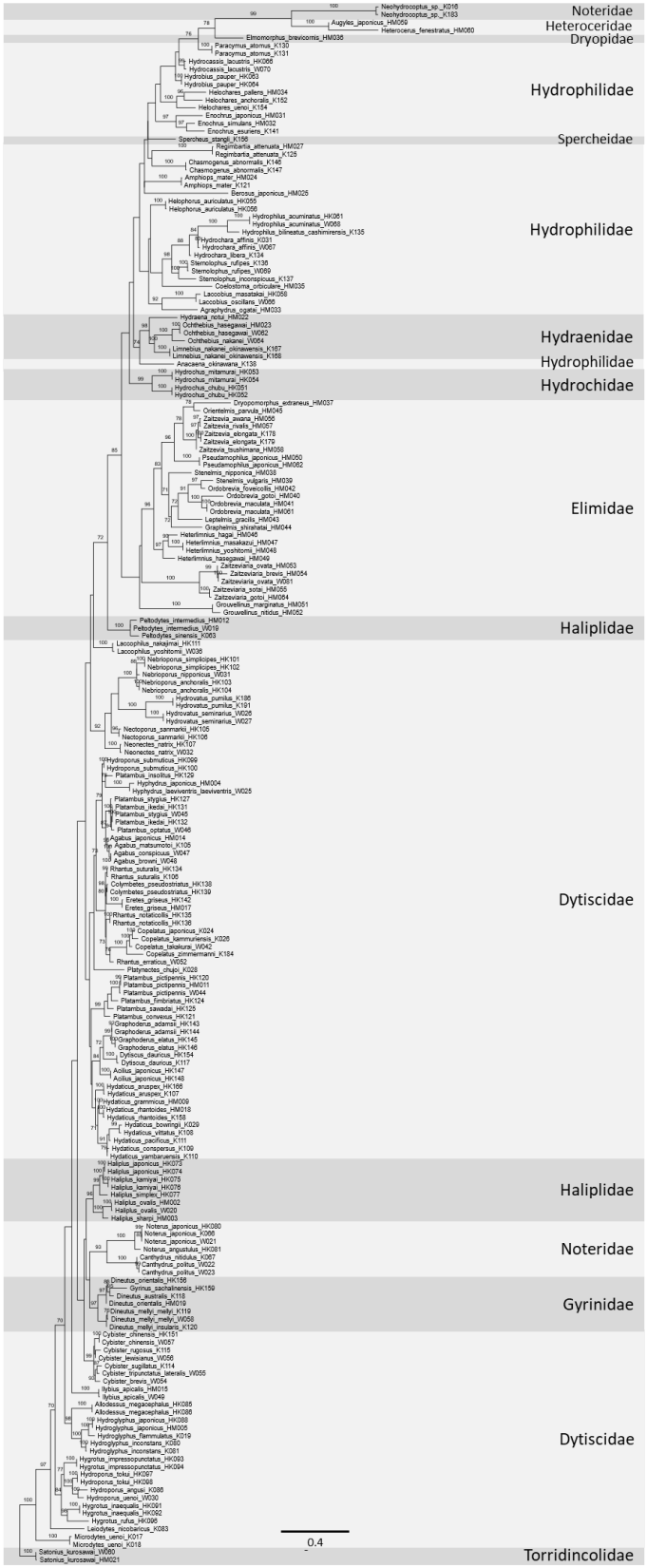
Order name	Family name	No. of taxa recorded in Japan	No. of taxa used for sequencing	No. of taxa for which sequences were obtained	Percentage of all taxa that have been sequenced
Coleoptera	Haliplidae	13	10	7	53.8%
	Noteridae	16	6	5	31.3%
	Dytiscidae	139	94	63	45.3%
	Gyrinidae	18	7	5	27.8%
	Torridincolidae	1	1	1	100.0%
	Hydraenidae	44	4	4	9.1%
	Hydrochidae	5	2	2	40.0%
	Helophoridae	5	1	1	20.0%
	Hydrophilidae (only aquatic species)	81	40	24	29.6%
	Spercheidae	1	1	1	100.0%
	Dryopidae	4	2	1	25.0%
	Elmidae	62	29	24	38.7%
	Chrysomelidae (only aquatic species)	2	0	0	0.0%
	Heteroceridae	3	2	2	66.7%
	Total	394	199	140	35.5%
Hemiptera	Nepidae	7	5	4	57.1%
	Belostomatidae	5	2	2	40.0%
	Micronectidae	9	3	2	22.2%
	Corixidae	22	8	7	31.8%
	Naucoridae	1	1	1	100.0%
	Aphelocheiridae	3	1	1	33.3%
	Notonectidae	12	3	3	25.0%
	Pleidae	3	3	2	66.7%
	Helotrephidae	1	0	0	0.0%
	Mesoveliidae	7	3	2	28.6%
	Hydrometridae	5	5	3	60.0%
	Veliidae	24	15	14	58.3%
	Gerridae	27	17	15	55.6%
	Hermatobatidae	1	1	1	100.0%
	Gelastocoridae	1	1	1	100.0%
	Total	128	68	58	45.3%
Total		522	267	198	37.9%

Figure Legends

Figure 1. The phylogenetic tree of aquatic Coleoptera insects in this study. Maximum likelihood (ML) tree for mitochondrial 16S sequences. The node numbers represent ML ultrabootstrap values. In each branch, only values with ultrabootstrap > 70 are denoted in the tree.

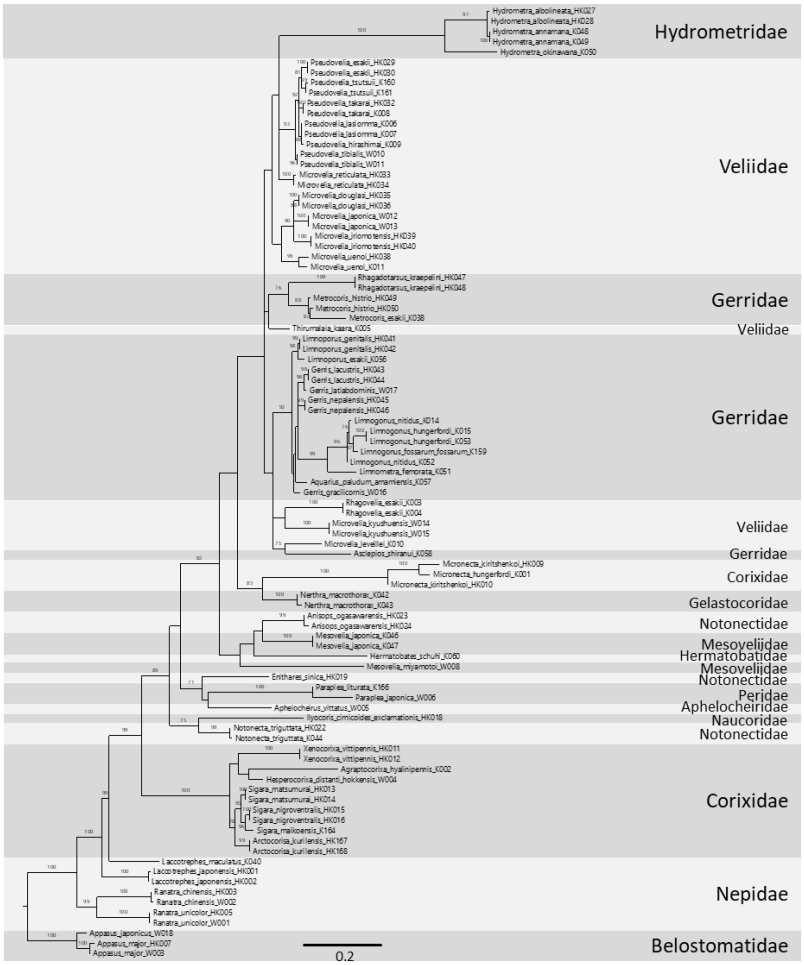
Figure 2. The phylogenetic tree of aquatic Hemiptera insects in this study. Maximum likelihood (ML) tree for mitochondrial 16S sequences. The node numbers represent ML ultrabootstrap values. In each branch, only values with ultrabootstrap > 70 are denoted in the tree.

493 **Figure 1**



494

Figure 2


Hydrometridae
Veliidae
Gerridae
Veliidae
Gerridae
Veliidae
Gerridae
Corixidae
Gelastocoridae
Notonectidae
Mesoveliidae
Hermatobatidae
Mesoveliidae
Notonectidae
Peridae
Aphelocheiridae
Naucoridae
Notonectidae
Corixidae
Nepidae
Belostomatidae
0.2

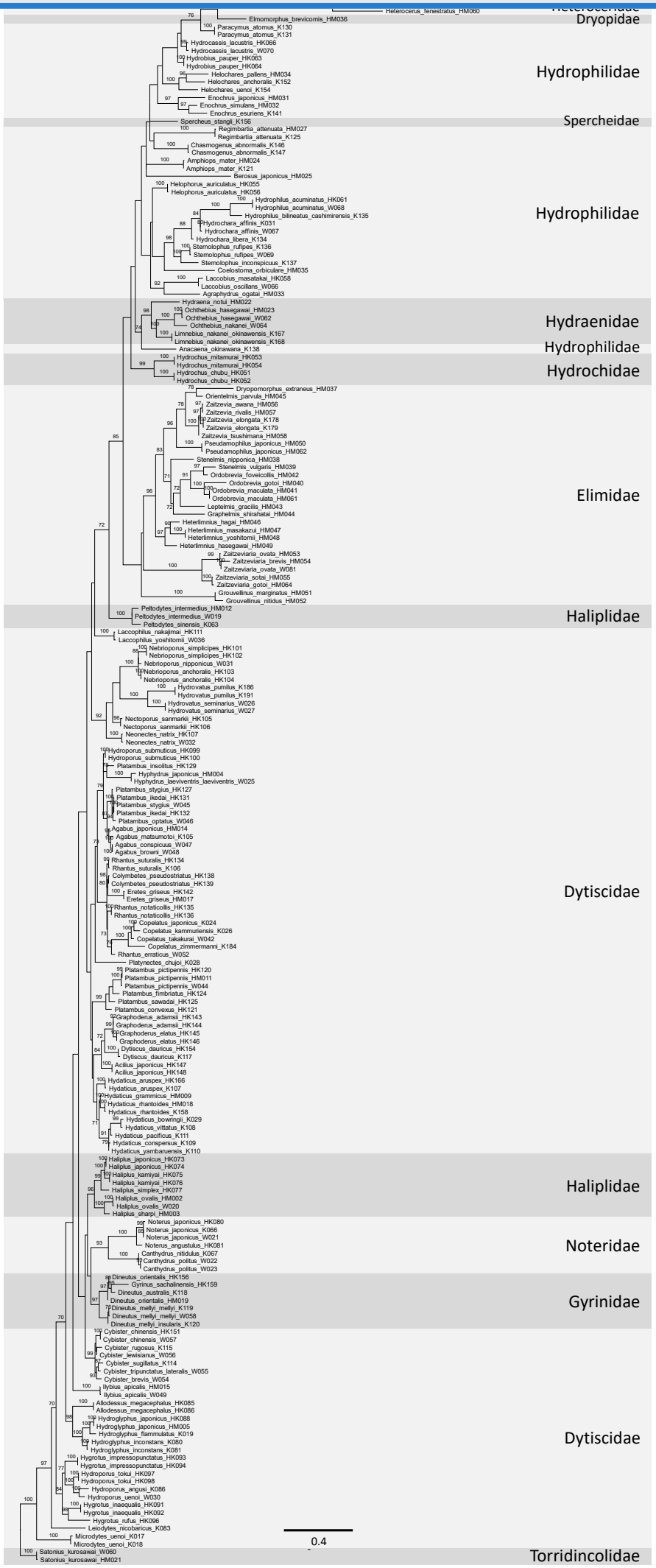
Supplementary Information

Table S1. Sample list used in this study (including both scientific and Japanese names). In each sample, the collection site and date, and collector are indicated. For the collection sites, only the prefecture is listed for conservation perspectives.

Table S2. A list of each of the aquatic Coleoptera insects treated in this study for each family (including both scientific and Japanese names). Sample IDs are given for taxa for which sequences were obtained.

Table S3. A list of each of the aquatic Hemiptera insects treated in this study, for each family (including both scientific and Japanese names). Sample IDs are given for taxa for which sequences were obtained.

Author-formatted, not peer-reviewed document posted on 22/01/2025. DOI: https://doi.org/10.3897/arphapreprints.e147415



ARPHA Preprints *Author-formatted, not peer-reviewed document posted on 22/01/2025.* DOI: https://doi.org/10.3897/arphapreprints.e147415