Aquatic insect community structure revealed by eDNA metabarcoding derives indices for environmental assessment

Noriko Uchida [Corresp., 1], **Kengo Kubota** [1], **Shunsuke Aita** [2], **So Kazama** [1]

[1] Department of Civil and Environmental Engineering, Tohoku University, Sendai, Miyagi, Japan

[2] School of Engineering, Tohoku University, Sendai, Miyagi, Japan

Corresponding Author: Noriko Uchida
Email address: noriko.uchida.s8@dc.tohoku.ac.jp

Environmental DNA (eDNA) analyses provide an efficient and objective approach for monitoring and assessing ecological status; however, there has been little study of eDNA of aquatic insects, despite its potential as a useful indicator of river health. Here, we investigated community structures of aquatic insects using eDNA and evaluated the applicability of eDNA data for calculating assessment indices. Field surveys were conducted to sample river water for eDNA at six reaches from upstream to downstream of two rivers in Japan in July and November 2016. At the same time, aquatic insects were collected using a traditional Surber net survey method. Communities of aquatic insects were revealed using eDNA by metabarcoding targeting the Cytochrome Oxidase subunit I gene in mitochondrial DNA. The eDNA results revealed 93 families and 75 genera of aquatic insects, which was triple that detected by the Surber net survey (especially for families in Diptera and Hemiptera). Seasonal differences of communities were discriminated by eDNA and Surber net survey data. Furthermore, environmental assessment indices (i.e., EPT index and Chironomidae index) calculated using operational taxonomic unit richness at the genus level showed positive correlations with total nitrogen concentration. Our results demonstrate that eDNA can reveal spatial and temporal differences of aquatic insect communities and that environmental assessment indices can be obtained using eDNA.

Aquatic insect community structure revealed by eDNA metabarcoding derives indices for environmental assessment

Noriko Uchida[1], Kengo Kubota[1], Shunsuke Aita[2], So Kazama[1]

[1]Department of Civil and Environmental Engineering, Tohoku University, Sendai, Miyagi, JAPAN
[2]School of Engineering, Tohoku University, Sendai, Miyagi, JAPAN

Corresponding Author:
Noriko Uchida[1]
6-6-06, Aza-Aoba, Aramaki, Aoba-ku, Sendai, Miyagi, 980-8579, Japan
Email address: noriko.uchida.s8@dc.tohoku.ac.jp

Abstract

Environmental DNA (eDNA) analyses provide an efficient and objective approach for monitoring and assessing ecological status; however, there has been little study of eDNA of aquatic insects, despite its potential as a useful indicator of river health. Here, we investigated community structures of aquatic insects using eDNA and evaluated the applicability of eDNA data for calculating assessment indices. Field surveys were conducted to sample river water for eDNA at six reaches from upstream to downstream of two rivers in Japan in July and November 2016. At the same time, aquatic insects were collected using a traditional Surber net survey method. Communities of aquatic insects were revealed using eDNA by metabarcoding targeting the Cytochrome Oxidase subunit I gene in mitochondrial DNA. The eDNA results revealed 93 families and 75 genera of aquatic insects, which was triple that detected by the Surber net survey (especially for families in Diptera and Hemiptera). Seasonal differences of communities were discriminated by eDNA and Surber net survey data. Furthermore, environmental assessment indices (i.e., EPT index and Chironomidae index) calculated using operational taxonomic unit richness at the genus level showed positive correlations with total nitrogen concentration. Our results demonstrate that eDNA can reveal spatial and temporal differences of aquatic insect communities and that environmental assessment indices can be obtained using eDNA.

Introduction

Stream ecosystems are threatened by global climate change and anthropogenic impacts, including damming, water abstraction, and land-use changes (WWF, 2016). For sustainable development and resource use of freshwater, there is a need to effectively manage stream

environments, which requires effective methods and indicators to measure and assess environmental impacts. Aquatic insects are commonly used as indicators due to their high sensitivity to deterioration of water quality. They are also a core component of the ecological food web in river ecosystems by feeding on producers and being preyed upon by higher consumers. Thus, monitoring aquatic insect fauna is an effective way of assessing environmental and ecological status. However, conventional surveys based on kick net or Surber net sampling are subject to several limitations. First, traditional field sampling processes result in data bias because the success and quality of a survey depend on the investigator's ability and the accessibility of sampling sites. In addition, direct sampling methods inherently involve damaging natural habitats and organisms. Second, the subsequent process of sorting and morphological identification is time-consuming and requires experts in taxonomic identification (Baird and Hajibabaei, 2012; Haase et al., 2006). These difficulties have created a bottleneck for performing high-frequency or long-term biological monitoring.

The use of environmental DNA (eDNA) is a novel biological monitoring method to overcome these difficulties. Owing to its simple sampling method (i.e., grab sampling of water, soil, etc.), eDNA can reduce the sampling bias caused by an investigator's ability (Rees et al., 2014; Smart et al., 2015). It can also minimize the bias related to accessibility of the site, especially for lotic environments, because water and suspended materials are mixed in these environments, but eDNA can detect not only lotic animals but also lentic ones (Fernández et al., 2018; Macher et al., 2018, Deiner et al. 2016). In addition, owing to it requiring only non-biotic samples, it can overcome ethical issues such as habitat disturbance and the sacrifice of animals associated with field sampling. Furthermore, DNA-based identification immediately provides higher taxonomic resolution than morphological identification (Carew et al., 2013; Elbrecht and Leese, 2015; Hajibabaei et al., 2011; Serrana et al., 2018). Skills in analyzing DNA can be acquired in a shorter time than those for morphological identification. Although the characteristics of eDNA remain obscure (e.g., production and degradation rates, and transportation dynamics), it can provide ecological information that is unobtainable via traditional surveys. For example, it can be used to evaluate diversity, including whole fauna (fish: Ushio et al., 2018) and multiple phyla (eukaryotes: Deiner et al., 2016).

Aquatic insects including Ephemeroptera, Plecoptera, Trichoptera (abbreviated name of these groups in combination: EPT), and Diptera have been gradually targeted in eDNA studies in the field of river management (Bista et al., 2017; Fernandes et al., 2018; Fernández et al., 2018; Hajibabaei et al., 2019a; Macher et al., 2018; Mächler et al., 2019). However, little is known about the spatial and temporal differences of community structures of aquatic insects (Bush et al., 2019; Roussel et al., 2015). Furthermore, there is a need to examine the possibility of using eDNA data for environmental assessment indices focusing on aquatic insects in stream ecosystems in a Japanese context.

Against this background, the present study was established to evaluate the applicability of eDNA for environmental assessment. First, we revealed the community structures of aquatic insects using eDNA and the traditional Surber net survey method. Subsequently, we investigated

whether eDNA data and the Surber net survey data can discern the spatial and temporal differences of aquatic insect communities. Furthermore, we evaluated the relationships between water quality parameters and assessment indices derived from each method.

Materials & Methods

2.1 eDNA sampling, filtration, and DNA extraction

Field samplings were conducted at Hirose River and Natori River, which are located in the Natori River basin, Miyagi Prefecture, northeast Japan. The length of the channel of Hirose River is 45.2 km and its catchment area is 315.9 km^2. Natori River is 55.0 km long and its catchment area is 623.0 km^2 (not including the Basin of Hirose River). Sampling was conducted in July and November 2016 at six reaches from upland to lowland regions along the two rivers (sites H1–H3 and N1–N3; see Figure 1 and Table S1). These are temperate rivers that originate in the mountains and flow through the hills at their middle reaches and through urbanized flatlands at their lower reaches, and finally flow into the Pacific Ocean.

Water samples for eDNA analysis were collected at the same sites and on the same days. The plastic bottles for eDNA sampling were sterilized with 10% chlorine bleach (Kao Corporation, Tokyo, Japan) and rinsed with tap water in the laboratory, and subsequently washed three times with river water at the collection site before sampling. At each site, flowing surface water was collected and transported to the laboratory on ice in a cool box. Water samples were filtered on the same day using vacuum filtration with 47-mm diameter glass-fiber filters with 0.7-μm pore size (GF/F, Whatman, 1 L/filter (Mächler et al., 2016)). These filtered samples were stored at −20°C until DNA extraction. DNA was extracted from the filters through lysis using proteinase K at 56°C for 30 min. After incubation, the supernatant was subjected to phenol-chloroform-isoamyl alcohol extraction and ethanol precipitation. Eventually, the elution was purified using the OneStep PCR Inhibitor Removal Kit (Zymo Research, Irvine, CA, USA) with a final volume of 100 μl.

2.2 Library preparation and sequencing

Target regions (cytochrome oxidase subunit I gene (COI) in mitochondrial DNA) of extracted DNA were amplified using the universal primer for invertebrates developed by Folmer et al (1994). The primer set of LCO1490 (5′-GGT CAA CAA ATC ATA AAG ATA TTG G-3′) as forward primer and HC02198 (5′-TAA ACT TCA GGG TGA CCA AAA AAT CA-3′) as a reverse primer, resulting in an amplification of a 658-bp fragment. For MiSeq library preparation, a three-step PCR was conducted. The first PCR was performed in a total volume 20 μl PCR mixture containing 10 μl of Taq™ HS Low DNA (TaKaRa, Kyoto, Japan), 0.4 μl each of 10 μM forward and reverse primers, 17.2 μl ultra-pure water, and 2.0 μl of template DNA. The PCR conditions were as follows: 35 cycles at 94°C for 5 s, 50°C for 30 s, 68°C for 10 s; and a final extension at 68°C for 7 min. The fragment size of amplicons and concentrations were

verified by electrophoresis using the Agilent 2100 Bioanalyzer DNA7500 kit (Agilent, Santa Clara, CA, USA). PCR products were purified using the Agencourt AMPure XP (Beckman Coulter, Brea, CA, USA) and the purified products were used as templates for the following. The second PCR was performed using Ex Taq Hot Start Version (TaKaRa, Kyoto, Japan) to add the overhang sequences that required amplification with the Nextera XT Index Kit for Illumina MiSeq analysis. The PCR conditions were as follows: 94°C for 2 min; followed by 5 cycles of 94°C for 30 s, 50°C for 30 s, 72°C for 30 s; and a final extension at 72°C for 5 min. The amplicons were purified in the same way as those from the first PCR, and the purified products were used as templates for the following. The third PCR was performed using Ex Taq Hot Start Version and Nextera XT Indice Kit v2 Set A (Illumina, San Diego, CA, USA). The PCR conditions were followed: 94°C for 2 min; followed by 8 cycles of 94°C for 30 s, 50°C for 30 s, 72°C for 30 s; and a final extension at 72°C for 5 min. After purification, the final PCR amplicons were quantified using the Qubit dsDNA High Sensitivity Kit. The sequencing of prepared libraries was performed following the manufacturer's instructions for MiSeq.

2.3 Bioinformatics

The sequence lengths of amplicons were 658 bp, so the forward and reverse reads in our study could not be merged when using the MiSeq Reagent Kit v3 (600 cycles). Elbrecht and Leese (2017) demonstrated that invertebrate species could be identified at the reverse side of the COI region through an *in silico* PCR approach. Therefore, we conducted a subsequent analysis using the reverse side sequence. At first, raw sequence reads were subjected to the Trimmomatic v0.36 software to discard low-quality sequences and read sequence lengths of <150 bp. Filtered reads were clustered into operational taxonomic units (OTUs) using QIIME (Caporaso et al., 2010) with an identity cut-off value of 97% which is a common approach for invertebrate metabarcoding analyses (Macher et al., 2018); subsequently, OTUs with singleton sequences were removed. The most frequently occurring sequences in each OTU were extracted as representative sequences. The assignment was performed against 3,433,026 sequences retrieved from the NCBI database by the following search criteria: cytochrome [all fields] AND oxidase [all fields] AND mitochondrion [filter]. After assignment, eight orders, namely Ephemeroptera, Plecoptera, Trichoptera, Diptera, Coleoptera, Odonata, Megaloptera, and Hemiptera, which mostly include aquatic insect species, were extracted using the QIIME script "filter_taxonomy_from_table.py." Subsequently, representative sequences of extracted OTUs were subjected to chimera check. Taxonomic identification was performed by BLAST search using the QIIME script "assign_taxonomy.py". Since the conventional environmental assessment indices (%EPT, %Diptera, %Chironomidae) require at least family-level resolution for taxonomic identification, we employed two thresholds for taxonomic identification: 97% identity for genus-level assignation and 85% identity for family-level assignation. A threshold of 97% is commonly used for species/genus-level assignment for macroinvertebrates (Hebert et al., 2003), but may cause loss of sequence depth. According to our investigation using a subset of sequence data of EPT registered in NCBI, sequence identity for the COI region of EPT was 99% at intra-

specific, 85% at intra-genus, 83% at intra-family, and 80% at intra-order levels (see Text S1). Therefore, BLAST assignment was conducted with minimum identity of 97% (assigned at the genus level) and 85% (assigned at least at the family level) and maximum e-value of 10^{-50} (Fernández et al., 2018). To compare the communities between samples, we subsampled the number of sequences in each sample to a uniform number. Either 250 reads for the family-level analysis or 150 reads for genus-level analysis were picked at random according to the smallest numbers of sequence reads (see Table S2).

2.4 Aquatic insect sampling using a Surber net and measuring environmental parameters

After eDNA sampling but on the same day, conventional aquatic insect collection was conducted by the Surber net method. We used a Surber net of 250-μm mesh size, in a 30 × 30 cm quadrat at randomly selected riffle and pool habitats at each site (total collection area: 0.18 m^2/reach). Collected invertebrates were placed in 99.5% ethanol solution and morphologically identified using a stereomicroscope (Leica MZ APO; Leica, Germany) by referring to the identification key for the aquatic insects of Japan (Kawai and Tanida, 2018). Because morphological identification was difficult for some aquatic insects, particularly Chironomidae and some Baetidae individuals, population abundance and richness were summarized at the family level.

At the same time as aquatic insect sampling, environmental parameters were measured. Water temperature, electrical conductivity (EC; TOA DKK CM-21P; Japan), and pH (TOA-DKK HM-20P; Japan) were measured in the field. Fifty milliliters of river water was sampled from each site to obtain the concentrations of total phosphorus (TP) and total nitrogen (TN). TP and TN were measured in our laboratory using a QuAAtro-2HR (BLTEC Corporation, Japan).

2.5 Community structure analysis

We assessed the dissimilarity of the community structure of aquatic insects using eDNA or Surber net survey data. Community dissimilarities were calculated based on the Sørensen index (binary Bray-Curtis index) using presence/absence of OTU data for eDNA and detected taxa for Surber net survey using package “vegan” (Oksanen et al., 2019) in R ver. 3.4.0 (R core team, 2018). Using the ordination of dissimilarity, non-metric multidimensional scaling (nMDS) was performed to visualize the similarity in community structures using the “metaMDS” function in package “vegan.” Further, the correlations between community structures and environmental parameters (i.e., water temperature, EC, TN, and TP) were tested using the function “envfit” in package “vegan”.

2.6 Environmental assessment indices

The applicability of biological information obtained from eDNA to existing environmental assessment indicators, namely, EPT index, Diptera index, and Chironomidae index (Reynoldson and Metcalfe-Smith, 1992), was evaluated. These indices are the ratios of the number of individuals/richness of EPT taxa to the total number of individuals/richness of the eight orders observed in the samples (Ephemeroptera, Plecoptera, Trichoptera, Diptera, Coleoptera, Odonata,

Megaloptera, and Hemiptera). For the Surber net survey data, the EPT index was calculated using the abundance or the richness at the family/genus level of EPT. The Diptera/Chironomidae index uses the same method as the EPT index, but using Diptera/Chironomidae instead of EPT. For eDNA-analyzed samples, each index was calculated using the richness of OTUs or groups by taxonomic name at the assigned family/genus level. Here, the OTU richness refers to the number of OTUs included in the sample, and the assigned family/genus richness refers to the number of families/genera included in the sample (see Text S2 for formulas).

Results

3.1 Community structure of aquatic insects revealed by eDNA analysis

Overall, 1,235,176 sequences (50,728–168,413 sequences/sample) passed the sequence quality filter (Table S2 for detail of metabarcoding outputs). These sequences were used to create OTUs based on 97% sequence identity. As a result, 90,948 OTUs were formed. Out of these, 66,176 OTUs included just one sequence (singletons), which were excluded from the analysis. Therefore, a total of 1,169,000 sequences (47,443–161,461 sequences/sample), generating 24,773 OTUs, were analyzed. After a BLAST search, we found that 8.0% of the total sequences were assigned to aquatic insects at the family level (sequence identity ≥ 85%, see Materials & Methods **2.3**), and only 4.1% of them were assigned at the genus level (sequence identity ≥ 97%) (Table S2).

eDNA metabarcoding detected 93 families and 104 genera (before subsampling), and 63 families and 75 genera (after subsampling at sequence depths of 250 reads and 150 reads, respectively; see Table S2 and Table S3 for OTU table before subsampling). Among these, 26 families were common with the Surber net survey results. Even after subsampling, the total number of taxa detected by eDNA was almost double compared with that in the Surber net survey (Table S4). Specifically, eDNA detected 27 genera of Chironomidae (Diptera), while we could not distinguish any of them by morphological identification. In addition, eDNA detected taxa that were mostly distributed in riparian/terrestrial habitats (e.g., Culicidae; Diptera, Cicadidae; Hemiptera) and lentic habitats (Aeshnidae and Epiophlebiidae; Odonata).

According to the subsampling results of eDNA, the taxa detected in both months were 26 families/25 genera, those detected only in July were 26 families/41 genera, and those detected only in November were 11 families/9 genera. Among these, we found that the three families (Ephemerellidae, Chironomidae, Simuliidae) were commonly detected among all sites and both months (see Table S5 for details of assignment results at family-level). On the other hand, a number of unique taxa were detected in communities of site H1 (11 genera in July and 9 genera in November; see Table S6 for details of assignment results at genus-level).

Earlier studies reported that the use of eDNA in lotic systems tends to enable the detection of more taxa than a traditional survey (Macher et al., 2018); this contrasts with the case in pond systems, where transportation is very low (Hajibabaei et al., 2019a). This is because DNA is transported downstream, resulting in the additional detection of upstream communities, which

are overlooked by conventional methods. In the case of fish eDNA, it is decomposed and transported after release from organisms, with a decrease of eDNA concentration of 73% within 900 m downstream of the source (Nukazawa et al., 2018). Even 50–250 m downstream of the source, eDNA is not detected when target organisms' abundance or biomass is small (Jane et al., 2015; Pilliod et al., 2014). Thus, DNA sampled in rivers probably includes some DNA originating from abundant organisms inhabiting up to around 1 km upstream. While source materials of eDNA differ depending on the organism (e.g., mucus for fish eDNA (Takeuchi et al., 2019), saliva for terrestrial mammals (Rodgers and Mock, 2015; Ushio et al., 2017), and exuvia for aquatic arthropod (Deiner and Altermatt, 2014)), the nature of eDNA in lotic systems may be in a similar manner. The interval between sampling sites in our study is about 3–5 km, so the eDNA contamination between samples is assumed to be negligible. In addition, the length of the amplified sequence in this study is relatively long (658 bp). The longer sequences are degraded faster than shorter sequences, resulting in shorter maintenance time and transporting distance (Jo et al., 2017). Hence, our eDNA samples describe the spatial differences of communities. In addition, eDNA can also be used to detect taxa that are usually difficult to capture from Surber net surveys in lotic locations, such as terrestrial organisms (Deiner et al., 2016b; Mächler et al., 2014). As in this previous study, these taxa were detected in our samples. These results indicate that eDNA sampled from river ecosystems provides a diverse taxonomic list that differs from that from traditional Surber net sampling.

From the results of family-level assignment, nine families were not detected in any eDNA samples, but were found with Surber net sampling, namely, two ephemeropterans (Isonychiidae and Siphlonuridae), one plecopteran (Chloroperilidae), trichopteran (Apataniidae), one dipteran (Blephariceridae), and four coleopteran (Gyrinidae, Hydrophilidae, Psephenidae, and Ptilodactylidae) taxa. Their sequences have mismatches to the primers used in this study, resulting in the failure of PCR primer amplification. Therefore, the primers should be modified or new primers should be developed to analyze these five families. Some refined primer sets for the metabarcoding of aquatic invertebrates have been developed (Elbrecht and Leese, 2017; Hajibabaei et al., 2012). (Hajibabaei et al., 2019b) also suggested that the use of multiple universal primers enables a broader range of taxa to be covered.

In addition to the primer issue, the sequence identity threshold used for taxonomic identification can be another source of problems for eDNA analysis. To evaluate the discrepancy between the reference library and the query sequence, we investigated interspecific, intergeneric, and interfamilial genetic identity (Text S1). As a result, an 85% identity threshold for the family level and a 97% identity threshold for the genus level were employed for taxonomic assignment in this study. However, some species might not even have reached this threshold and went undetected. To overcome this issue, the accumulation of reference sequence data is essential. It is known that geographically separated intra-species have low sequence identity of the COI gene (Takenaka and Tojo, 2019). Therefore, the accumulation of genetic information of local aquatic insects and the construction of a database are necessary to improve the taxonomic assignment, in order to avoid failure. Because failing to detect some specific taxonomic groups could directly affect the

assessment results based on the richness of taxa, there is an urgent need to overcome these issues.

3.2 Relationships between communities and environmental parameters

Community dissimilarities among all samples using the Sorensen index for eDNA and Chao index for Surber net data were plotted on nMDS coordinate axes (Figure 2). Visually, communities were divided into two clusters, each of which contained samples taken from the same month. Further, the uppermost site of Hirose River (H1) was plotted in isolation from the other sites in both months for three datasets. The ordination was significantly correlated with the water temperature in all three datasets ('envfit' function; eDNA at genus-level: $R^2 = 0.56$, $p = 0.016$, eDNA at family-level: $R^2 = 0.81$, $p = 0.001$, Surber net at the family level: $R^2 = 0.57$, $p = 0.021$, see details for Table S7). Furthermore, the ordination based on eDNA data at genus-level resolution was significantly correlated with TN concentration ('envfit' function; eDNA at genus-level: $R^2 = 0.50$, $p = 0.043$, eDNA at family-level: $R^2 = 0.51$, $p = 0.046$), but Surber net data was not correlated.

Bista et al. (2017) demonstrated that eDNA targeting the community of Chironomidae (Diptera) showed different clusters associated with seasonal differences in a lake system. Similarly, the present research revealed that eDNA targeting the whole aquatic insect community also revealed seasonal differences in the river system. The isolated plotting of site H1 was understandable because the landscape of site H1 differs from the others and it is a mountain stream, so the taxa there should also differ. Indeed, according to the taxonomic assignment results by metabarcoding, many unique taxa were found at site H1 (Table S5 and S6). In terms of the community differences in the Surber net survey data, these were likely due to the absence of taxa detected at other sites, rather than the presence of unique taxa at site H1 (Table S4).

In our dataset, the higher the TN concentration, the larger the stream order (Table S1). Therefore, the ordination of community dissimilarity using genus-level resolution data from eDNA may be related to river order. Considering that other studies illustrated that eDNA can reveal the spatial differences of the fauna community in river systems (fungi: Matsuoka et al., 2019; macroinvertebrates:(Hajibabaei et al., 2019b); Fernández et al., 2019), spatial differences of communities could only be discriminated by analysis using genus-level resolution in the present study.

3.3 Environmental assessment indices derived from eDNA

Since TN could be assumed to be a chemical indicator of water pollution, the relationship between gradations of TN and biological environmental assessment indices (%EPT, %Diptera, and %Chironomidae) was examined (Figure 3). The results showed that %EPT and %Chironomidae derived from eDNA at genus-level resolution showed sufficient effect sizes with significant rank correlation with TN (Spearman's rank correlation; %EPT: $r = -0.59$, $p = 0.049$, %Chironomidae; $r = 0.69$, $p = 0.014$. See all results in Table S7). However, eDNA and the Surber net results at family-level resolution did not show significant correlations. These

results indicate that, when eDNA data are obtained at higher taxonomic resolution, the sensitivity of biological indices to environmental factors can be improved. In contrast, the sensitivity could be impaired if biological indices are obtained using coarse taxonomic resolution.
Emilson et al. (2017) reported that assessment indices (EPT index and Chironomidae index) derived from DNA metabarcoding using macroinvertebrate tissue samples highly correlated with indices derived from the morphological survey. The present study demonstrated that the indices obtained from eDNA can also be used as a new assessment method. Although the biological indices obtained from eDNA were compared only by TN in this study, it is possible to develop new environmental indicators using eDNA data by comparisons with more chemical pollution indicators such as BOD and DOC.
While our manuscript was under review, a study that detected macroinvertebrate eDNA and applied this to assessment of the environmental status of a river was reported (Fernández et al., 2019). The environmental assessment score was calculated based on the presence/absence of indicator macroinvertebrates at family-level identification and demonstrated that eDNA data could be used for the monitoring program that they used (IBMWP). On the other hand, the present study showed that EPT and Chinoromidae indices calculated using OTU richness required genus-level resolution and showed a clearer response to organic pollution than when using family-level resolution.

Conclusions

The present study demonstrated that eDNA can be used to describe differences between community structures of aquatic insects in two seasons as well as in different locations in river systems. In addition, the community as derived from eDNA analysis was shown to be correlated with water pollution with higher sensitivity than under the traditional community surveying. The EPT index and Chironomidae index at the genus level derived from eDNA data showed significant correlations with TN concentration. In addition to providing the ecological and environmental information described above, eDNA has remarkable advantages, such as requiring minimal sampling effort, having high taxonomic resolution, and being applicable to a broad range of species. We believe that eDNA can be useful for monitoring the long-term trends of community structure changes associated with environmental changes such as climate change and other human activities.

Acknowledgements

The authors would like to thank Mr. Kengo Watanabe (Tohoku University, Japan) and Dr. Yasuyuki Takemura (National Institute for Environmental Studies, Japan) for their assistance on field survey and experiments. We also greatly appreciate Dr. Kei Nukazawa (University of Miyazaki, Japan), Dr. Yasuhiro Takemon (Kyoto University, Japan) and Dr. Masahiro Ryo (Free University of Berlin, Germany) for providing meaningful comments on our paper.

References

Baird, D.J., Hajibabaei, M., 2012. Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. Mol. Ecol. 21, 2039–2044. https://doi.org/10.1111/j.1365-294X.2012.05519.x

Bista, I., Carvalho, G.R., Walsh, K., Seymour, M., Hajibabaei, M., Lallias, D., Christmas, M., Creer, S., 2017. Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. Nat. Commun. 8, 14087. https://doi.org/10.1038/ncomms14087

Bush, A., Compson, Z.G., Monk, W.A., Porter, T.M., Steeves, R., Emilson, E., Gagne, N., Hajibabaei, M., Roy, M., Baird, D.J., 2019. Studying Ecosystems With DNA Metabarcoding: Lessons From Biomonitoring of Aquatic Macroinvertebrates. Front. Ecol. Evol. 7, 434. https://doi.org/10.3389/fevo.2019.00434

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Peña, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R., 2010. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7, 335–336. https://doi.org/10.1038/nmeth.f.303

Carew, M.E., Pettigrove, V.J., Metzeling, L., Hoffmann, A.A., 2013. Environmental monitoring using next generation sequencing: rapid identification of macroinvertebrate bioindicator species. Front. Zool. 10, 45. https://doi.org/10.1186/1742-9994-10-45

Chao, A., Chiu, C.-H., 2016. Species Richness: Estimation and Comparison, in: Balakrishnan, N., Colton, T., Everitt, B., Piegorsch, W., Ruggeri, F., Teugels, J.L. (Eds.), Wiley StatsRef: Statistics Reference Online. John Wiley & Sons, Ltd, Chichester, UK, pp. 1–26. https://doi.org/10.1002/9781118445112.stat03432.pub2

Deiner, K., Altermatt, F., 2014. Transport Distance of Invertebrate Environmental DNA in a Natural River. PLoS ONE 9, e88786. https://doi.org/10.1371/journal.pone.0088786

Deiner, K., Fronhofer, E.A., Mächler, E., Walser, J.-C., Altermatt, F., 2016a. Environmental DNA reveals that rivers are conveyer belts of biodiversity information. Nat. Commun. 7, 12544. https://doi.org/10.1038/ncomms12544

Deiner, K., Fronhofer, E.A., Mächler, E., Walser, J.-C., Altermatt, F., 2016b. Environmental DNA reveals that rivers are conveyer belts of biodiversity information. Nat. Commun. 7, 12544. https://doi.org/10.1038/ncomms12544

Elbrecht, V., Leese, F., 2017. Validation and Development of COI Metabarcoding Primers for Freshwater Macroinvertebrate Bioassessment. Front. Environ. Sci. 5. https://doi.org/10.3389/fenvs.2017.00011

Elbrecht, V., Leese, F., 2015. Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass—Sequence Relationships with an Innovative Metabarcoding Protocol. PLOS ONE 10, e0130324. https://doi.org/10.1371/journal.pone.0130324

Emilson, C.E., Thompson, D.G., Venier, L.A., Porter, T.M., Swystun, T., Chartrand, D., Capell, S., Hajibabaei, M., 2017. DNA metabarcoding and morphological

macroinvertebrate metrics reveal the same changes in boreal watersheds across an environmental gradient. Sci. Rep. 7. https://doi.org/10.1038/s41598-017-13157-x

Fernandes, K., Heyde, M. van der, Bunce, M., Dixon, K., Harris, R.J., Wardell‐Johnson, G., Nevill, P.G., 2018. DNA metabarcoding—a new approach to fauna monitoring in mine site restoration. Restor. Ecol. 26, 1098–1107. https://doi.org/10.1111/rec.12868

Fernández, S., Rodríguez, S., Martínez, J.L., Borrell, Y.J., Ardura, A., García-Vázquez, E., 2018. Evaluating freshwater macroinvertebrates from eDNA metabarcoding: A river Nalón case study. PLoS ONE 13. https://doi.org/10.1371/journal.pone.0201741

Fernández, S., Rodríguez‐Martínez, S., Martínez, J.L., Garcia‐Vazquez, E., Ardura, A., 2019. How can eDNA contribute in riverine macroinvertebrate assessment? A metabarcoding approach in the Nalón River (Asturias, Northern Spain). Environ. DNA 1, 385–401. https://doi.org/10.1002/edn3.40

Folmer, O., Black, M.B., Hoeh, W., Lutz, R., Vrijenhoek, R.C., 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3, 294–299.

Haase, P., Murray-Bligh, J., Lohse, S., Pauls, S., Sundermann, A., Gunn, R., Clarke, R., 2006. Assessing the impact of errors in sorting and identifying macroinvertebrate samples. Hydrobiologia 566, 505–521. https://doi.org/10.1007/s10750-006-0075-6

Hajibabaei, M., Porter, T.M., Robinson, C.V., Baird, D.J., Shokralla, S., Wright, M.T.G., 2019a. Watered-down biodiversity? A comparison of metabarcoding results from DNA extracted from matched water and bulk tissue biomonitoring samples. PLOS ONE 14, e0225409. https://doi.org/10.1371/journal.pone.0225409

Hajibabaei, M., Porter, T.M., Wright, M., Rudar, J., 2019b. COI metabarcoding primer choice affects richness and recovery of indicator taxa in freshwater systems. PLOS ONE 14, e0220953. https://doi.org/10.1371/journal.pone.0220953

Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G.A.C., Baird, D.J., 2011. Environmental Barcoding: A Next-Generation Sequencing Approach for Biomonitoring Applications Using River Benthos. PLOS ONE 6, e17497. https://doi.org/10.1371/journal.pone.0017497

Hajibabaei, M., Spall, J.L., Shokralla, S., van Konynenburg, S., 2012. Assessing biodiversity of a freshwater benthic macroinvertebrate community through non-destructive environmental barcoding of DNA from preservative ethanol. BMC Ecol. 12, 28. https://doi.org/10.1186/1472-6785-12-28

Hebert, P.D.N., Ratnasingham, S., deWaard, J.R., 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc. R. Soc. B Biol. Sci. 270, S96–S99. https://doi.org/10.1098/rsbl.2003.0025

Jane, S.F., Wilcox, T.M., McKelvey, K.S., Young, M.K., Schwartz, M.K., Lowe, W.H., Letcher, B.H., Whiteley, A.R., 2015. Distance, flow and PCR inhibition: eDNA dynamics in two headwater streams. Mol. Ecol. Resour. 15, 216–227. https://doi.org/10.1111/1755-0998.12285

Jo, T., Murakami, H., Masuda, R., Sakata, M.K., Yamamoto, S., Minamoto, T., 2017. Rapid degradation of longer DNA fragments enables the improved estimation of

452 distribution and biomass using environmental DNA. Mol. Ecol. Resour. 17, e25–
453 e33. https://doi.org/10.1111/1755-0998.12685
454 Kawai, T., Tanida, K., 2018. Aquatic Insects of Japan: Manual with Keys and
455 Illustrations [Japanese], The second edition. ed. Tokai University Press,
456 Hiratsuka, Kanagawa, Japan.
457 Macher, J.-N., Vivancos, A., Piggott, J.J., Centeno, F.C., Matthaei, C.D., Leese, F.,
458 2018. Comparison of environmental DNA and bulk-sample metabarcoding using
459 highly degenerate cytochrome c oxidase I primers. Mol. Ecol. Resour. 18, 1456–
460 1468. https://doi.org/10.1111/1755-0998.12940
461 Mächler, E., Deiner, K., Spahn, F., Altermatt, F., 2016. Fishing in the Water: Effect of
462 Sampled Water Volume on Environmental DNA-Based Detection of
463 Macroinvertebrates. Environ. Sci. Technol. 50, 305–312.
464 https://doi.org/10.1021/acs.est.5b04188
465 Mächler, E., Deiner, K., Steinmann, P., Altermatt, F., 2014. Utility of environmental DNA
466 for monitoring rare and indicator macroinvertebrate species. Freshw. Sci. 33,
467 1174–1183. https://doi.org/10.1086/678128
468 Mächler, E., Little, C.J., Wüthrich, R., Alther, R., Fronhofer, E.A., Gounand, I., Harvey,
469 E., Hürlemann, S., Walser, J.-C., Altermatt, F., 2019. Assessing different
470 components of diversity across a river network using eDNA. Environ. DNA 1,
471 290–301. https://doi.org/10.1002/edn3.33
472 Matsuoka, S., Sugiyama, Y., Sato, H., Katano, I., Harada, K., Doi, H., 2019. Spatial
473 structure of fungal DNA assemblages revealed with eDNA metabarcoding in a
474 forest river network in western Japan. Metabarcoding Metagenomics 3, e36335.
475 https://doi.org/10.3897/mbmg.3.36335
476 Nukazawa, K., Hamasuna, Y., Suzuki, Y., 2018. Simulating the Advection and
477 Degradation of the Environmental DNA of Common Carp along a River. Environ.
478 Sci. Technol. 52, 10562–10570. https://doi.org/10.1021/acs.est.8b02293
479 Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin,
480 P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E.,
481 Wagner, H., 2019. vegan: Community Ecology Package.
482 Pilliod, D.S., Goldberg, C.S., Arkle, R.S., Waits, L.P., 2014. Factors influencing
483 detection of eDNA from a stream-dwelling amphibian. Mol. Ecol. Resour. 14,
484 109–116. https://doi.org/10.1111/1755-0998.12159
485 R core team, 2018. R: A language and environment for statistical computing. R
486 Foundation for Statistical Computing, Vienna, Austria.
487 Rees, H.C., Maddison, B.C., Middleditch, D.J., Patmore, J.R.M., Gough, K.C., 2014.
488 REVIEW: The detection of aquatic animal species using environmental DNA – a
489 review of eDNA as a survey tool in ecology. J. Appl. Ecol. 51, 1450–1459.
490 https://doi.org/10.1111/1365-2664.12306
491 Reynoldson, T.B., Metcalfe-Smith, J.L., 1992. An overview of the assessment of aquatic
492 ecosystem health using benthic invertebrates. J. Aquat. Ecosyst. Health 1, 295–
493 308. https://doi.org/10.1007/BF00044171
494 Rodgers, T.W., Mock, K.E., 2015. Drinking water as a source of environmental DNA for
495 the detection of terrestrial wildlife species. Conserv. Genet. Resour. 7, 693–696.
496 https://doi.org/10.1007/s12686-015-0478-7

497 Roussel, J.-M., Paillisson, J.-M., Tréguier, A., Petit, E., 2015. The downside of eDNA as
498 a survey tool in water bodies. J. Appl. Ecol. 52, 823–826.
499 https://doi.org/10.1111/1365-2664.12428
500 Serrana, J.M., Yaegashi, S., Kondoh, S., Li, B., Robinson, C.T., Watanabe, K., 2018.
501 Ecological influence of sediment bypass tunnels on macroinvertebrates in dam-
502 fragmented rivers by DNA metabarcoding. Sci. Rep. 8, 10185.
503 https://doi.org/10.1038/s41598-018-28624-2
504 Smart, A.S., Tingley, R., Weeks, A.R., van Rooyen, A.R., McCarthy, M.A., 2015.
505 Environmental DNA sampling is more sensitive than a traditional survey
506 technique for detecting an aquatic invader. Ecol. Appl. 25, 1944–1952.
507 https://doi.org/10.1890/14-1751.1
508 Takenaka, M., Tojo, K., 2019. Ancient origin of a dipteromimid mayfly family endemic to
509 the Japanese Islands and its genetic differentiation across tectonic faults. Biol. J.
510 Linn. Soc. 126, 555–573. https://doi.org/10.1093/biolinnean/bly192
511 Takeuchi, A., Iijima, T., Kakuzen, W., Watanabe, S., Yamada, Y., Okamura, A., Horie,
512 N., Mikawa, N., Miller, M.J., Kojima, T., Tsukamoto, K., 2019. Release of eDNA
513 by different life history stages and during spawning activities of laboratory-reared
514 Japanese eels for interpretation of oceanic survey data. Sci. Rep. 9, 6074.
515 https://doi.org/10.1038/s41598-019-42641-9
516 Ushio, M., Fukuda, H., Inoue, T., Makoto, K., Kishida, O., Sato, K., Murata, K., Nikaido,
517 M., Sado, T., Sato, Y., Takeshita, M., Iwasaki, W., Yamanaka, H., Kondoh, M.,
518 Miya, M., 2017. Environmental DNA enables detection of terrestrial mammals
519 from forest pond water. Mol. Ecol. Resour. 17, e63–e75.
520 https://doi.org/10.1111/1755-0998.12690
521 Ushio, M., Hsieh, C., Masuda, R., Deyle, E.R., Ye, H., Chang, C.-W., Sugihara, G.,
522 Kondoh, M., 2018. Fluctuating interaction network and time-varying stability of a
523 natural fish community. Nature 554, 360–363.
524 https://doi.org/10.1038/nature25504
525 WWF, 2016. Living Planet Report 2016: Risk and Resilience in a New Era. WWW
526 International, Gland, Switzerland.
527

Figure 1

Study field

Sampling sites in Hirose River (from upmost site, H1, H2, H3, in orange circle) and in Natori River (N1, N2, N3, in blue circle) in the northeast part of Japan.

This map was modified by Uchida using a digital map provided by Geospatial Information Authority of Japan.

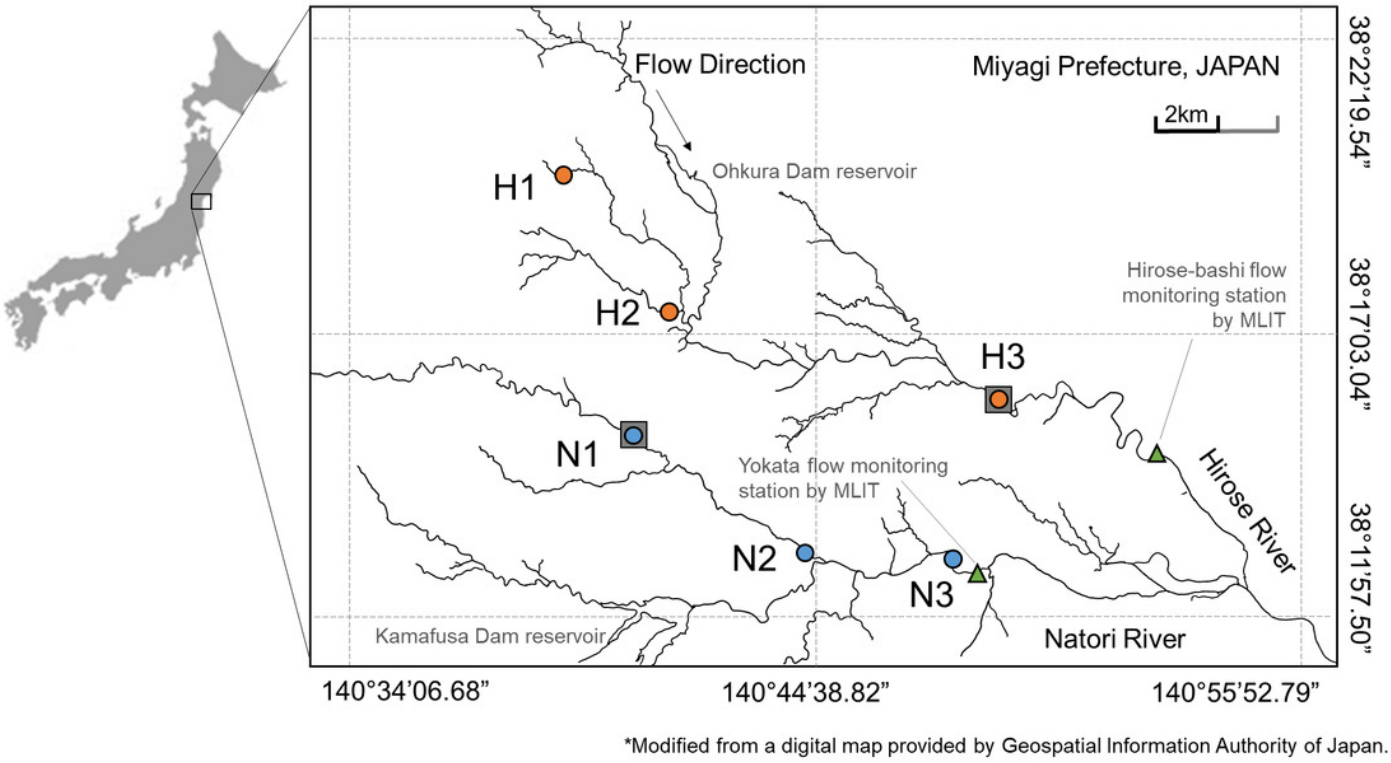


Figure 2

Non-metric multidimensional scaling (NMDS) using S ø rensen dissimilarity index (P/A data).

Each panel shows the communities derived from (A) eDNA at genus-level identification, (B) eDNA at family-level identification, (C) Surber net (at family-level identification). eDNA data were based on OTU richness with subsampled by 250 reads depth. The same month is enclosed by ellipses (orange: July, blue: November). The environmental parameters are shown by arrows, and the length of the arrow is proportional to the correlation between parameters and the community ordination.

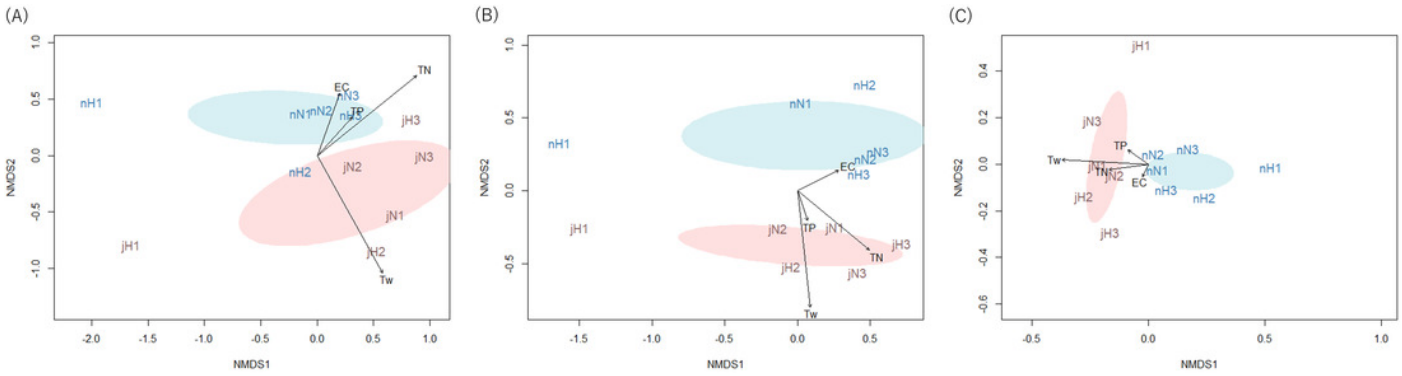
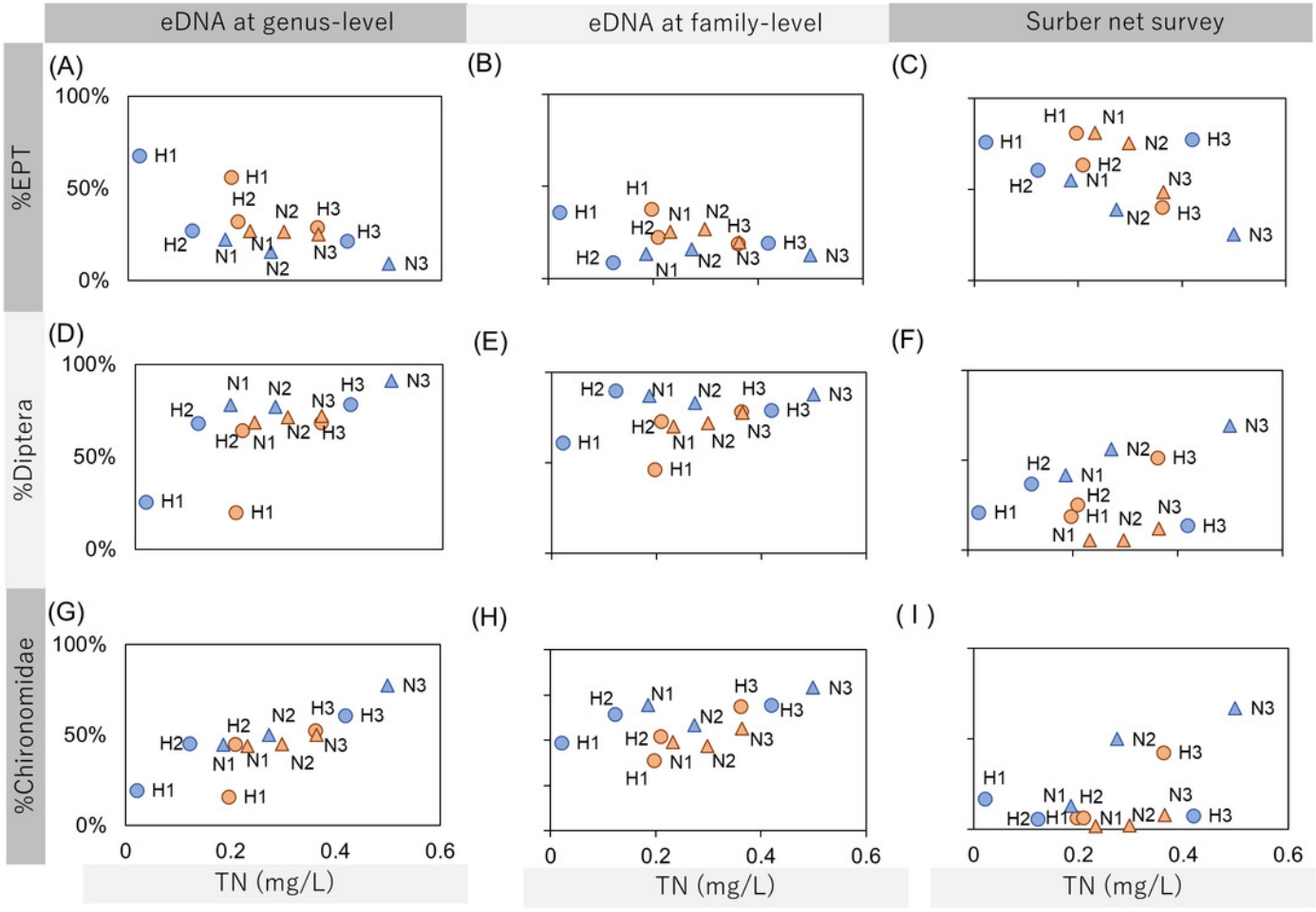


Figure 3

Relationships between biological assessment indices and TN concentration.

First raw show %EPT (A-C), second raw shows %Diptera (D-F), and third raw shows %Chironomidae (G-I). The first, second, third column shows results based on eDNA at genus-level identification, eDNA at family-level identification, and Surber net survey, respectively. The indices are calculated using OTU richness data for eDNA (subsampled) and abundance data for Surber net data. Seasonal differences are represented by colors (orange: July, blue: November) and river differences are represented by plot styles (circle: Hirose River, triangle: Natori River), respectively.