Distribution Patterns of the Freshwater Oligochaete *Limnodrilus hoffmeisteri* Influenced by Environmental Factors in Streams on a Korean Nationwide Scale

Hyejin Kang 1, Mi-Jung Bae 2, Dae-Seong Lee 1, Soon-Jin Hwang 3, Jeong-Suk Moon 4 and Young-Seuk Park 1,5,*

1 Department of Biology, Kyung Hee University, Dongdaemun-gu, Seoul 02447, Korea; ghj2166@naver.com (H.K.); dleotjd520@naver.com (D.-S.L.)
2 Freshwater Biodiversity Research Division, Nakdonggang National Institute of Biological Resources, Sangju, Gyeongsangbuk-do 37242, Korea; mjbae@nnibr.re.kr
3 Department of Environmental Health Science, Konkuk University, Gwangjin-gu, Seoul 05029, Korea; sjhwang@konkuk.ac.kr
4 National Institute of Environmental Research, Seo-gu, Incheon 22689, Korea; waterfa@korea.kr
5 Nanopharmaceutical Sciences, Kyung Hee University, Dongdaemun-gu, Seoul 02447, Korea
* Correspondence: parkys@khu.ac.kr; Tel.: +82-2-961-0946

Received: 29 September 2017; Accepted: 14 November 2017; Published: 27 November 2017

**Abstract:** Aquatic oligochaetes are very common in streams, and are used as biological assessment indicators as well as in the biological management of organic-enriched systems. In this study, we analyzed the effects of environmental factors influencing the distribution of aquatic oligochaetes *Limnodrilus hoffmeisteri* in streams. We used 13 environmental factors in three categories (i.e., geography, hydrology, and physicochemistry). Data on the distribution of oligochaetes and environmental factors were obtained from 1159 sampling sites throughout Korea on a nationwide scale. Hierarchical cluster analysis (HCA) and nonmetric multidimensional scaling (NMDS) were performed to analyze the relationships between the occurrence of aquatic oligochaetes and environmental factors. A random forest model was used to evaluate the relative importance of the environmental factors affecting the distribution of oligochaetes. HCA classified sampling sites into four groups according to differences in environmental factors, and NMDS ordination reflected the differences of environmental factors, in particular, water depth, velocity, and altitude, among the four groups defined in the HCA. Furthermore, using a random forest model, turbidity and water velocity were evaluated as highly important factors influencing the distribution of *L. hoffmeisteri*.

**Keywords:** distribution patterns of species; environmental factor; multiple scale; multivariate analyses; machine learning model

1. Introduction

The distribution and abundance of organisms are governed by various environmental factors and, therefore, it is essential to understand the influence of environmental factors on ecological communities [1]. Freshwater oligochaetes are distributed in various freshwater habitats and have been widely applied as indicator species in environmental assessment [2–4]. The influences of hydrology and physicochemistry on stream oligochaetes have been reported in many studies [5–8]. For example, the distribution of oligochaetes has been shown to be related to salinity, temperature, dissolved oxygen (DO), current velocities, and habitat stability [9] and *Limnodrilus hoffmeisteri* has been observed to burrow deeper into sediment when exposed to hypoxic water [10]. Van Duinen et al. [11] assessed the effects of increased nutrient availability on aquatic oligochaetes, and Martinez-Ansemil and Collado [12] demonstrated that substrate type and current velocity are the principal variables...
explaining community structure. Other researchers reported that substrate type is important for oligochaete communities [13,14]. Sauter and Güde [14] presented that tubificids are more abundant in fine substrate. The proportions of functional feeding strategies did not follow the gradient of hydraulic conditions, and only a few oligochaete taxa were able to survive in hydraulically rough conditions, and most oligochaete taxa were found only in pools [15].

Their ecological functions and roles have been recognized in various studies, including bioprocessing ability for restoration of organic polluted freshwater habitats [16–19], sludge removal [20], improvement of water quality [21,22], behavioral changes in response to pollutants [23], sediment toxicity assessment [24–26], and response to several toxic substances [3,27,28]. In a recent literature review, Kang et al. [29] reported that three species (Lumbriculus variegatus, Tubifex tubifex, and Limnodrilus hoffmeisteri) among freshwater oligochaetes have been used most frequently in ecotoxicological studies, indicating the possibility of developing an early warning system using aquatic oligochaetes to monitor aquatic ecosystem disturbance.

Although there have been many studies on the distribution of oligochaetes, most studies have been conducted on the taxonomy [30–32]. However, limited studies have been conducted on the relationships between the distribution of oligochaetes and their habitats. Ecological data have non-linear and complex properties within variables as well as among variables. Therefore, nonlinear approaches are recommended to characterize the relationships among variables in ecological data [33,34]. Recently, machine learning algorithms, which are robust to nonlinear and complex data, are commonly used in environmental sciences and ecological studies [35,36]. Among machine learning techniques, the usage of random forest (RF), an ensemble machine learning technique, has been increasing in field of ecology, including in ecohydrological distribution modelling [37], species distribution modelling [38,39], and the relation between freshwater organisms and environmental factors [40,41], because in the computation of RF, no a priori assumptions about linear or nonlinear relationships between predictors and response variables are needed [42]. Furthermore, there are lots of reports that RF produces more accurate classification compared with other machine learning techniques (e.g., classification trees) because RF is computed based on the combinations of many decision trees [43]. In this study, we investigated the distribution patterns that improve, like that of the aquatic oligochaete L. hoffmeisteri in Korean streams, and the influence of environmental factors on their distribution patterns by using multivariate analyses and a machine learning algorithm.

2. Materials and Methods

2.1. Ecological Data

The data on aquatic oligochaetes used in this study was provided by the Nationwide Aquatic Ecological Monitoring Program (NAEMP), conducted by the Ministry of Environment and the National Institute of Environmental Research, Korea. We used a dataset compiled from 1159 sites sampled from 2009 to 2016 on a nationwide scale in Korea (Figure 1). The sampling was conducted from the headwater to downstream on a nationwide scale to reflect various habitat conditions, as well as to evaluate the status of freshwater habitats nationally. The altitude in sampling sites ranged from 0 to 1166 m a.s.l., the distance from source was from 0.2 to 498.9 km, and the stream order is from 1 to 7.

Benthic macroinvertebrates, including aquatic oligochaetes, were collected using a Surber net (30 × 30 cm², mesh size 1 mm) with three replicates at each site, twice a year in spring and summer, based on the guidelines of the National Survey for Stream Ecosystem Health in South Korea [44]. Oligochaeta samples were sorted and preserved in 70% ethanol in the laboratory. Then, for coarse identification, sorted individuals were separated into several possible genus/species groups based on morphological characters (e.g., body seta, body shape, etc.). The specimens were sorted into the lowest level (mostly species level) following the identification keys [45,46]. Among oligochaetes, we studied the distribution patterns of L. hoffmeisteri, which is the dominant species in Korean streams, and the
influence of environmental factors on their distribution patterns. According to the literature [31,47], we used the species name *L. hoffmeisteri* instead of *Limnodrilus gotoi*, which was used in the database of the NAEMP.

We used the following 13 environmental parameters grouped into three categories, namely, geography, hydrology, and physicochemistry: altitude and slope in geography; water depth, water width, and current velocity in hydrology; and dissolved oxygen (DO), biochemical oxygen demand (BOD), total nitrogen (TN), total phosphorus (TP), chlorophyll a (Chl-a), pH, conductivity, and turbidity in physicochemistry (Table 1). Geographical factors were extracted from a digital map using a geographic information system (ArcGIS®, ver.10.1, ESRI [48]). Hydrological and physicochemical factors were obtained from the NAEMP database [49], which were measured according to the NAEMP guidelines [46].

As the number of samples collected differed among sampling sites, the average value for each site was used in the study of environmental variables. Among the 1159 sampling sites, we used 1127 sites in the analyses owing to the unavailability of environmental data for the other 32 sites.

![Figure 1. A map of sampling sites in South Korea.](image-url)
Table 1. Differences (mean with standard error) in environmental factors among four different groups defined through a hierarchical cluster analysis. Different letters in the same row indicate significant differences between the groups based on Tukey’s multiple comparison tests ($p < 0.05$). Dissolved oxygen (DO); biochemical oxygen demand (BOD); total nitrogen (TN); total phosphorus (TP); chlorophyll a (Chl-a).

<table>
<thead>
<tr>
<th>Category</th>
<th>Variable</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geography</td>
<td>Altitude (m)</td>
<td>183.64 (9.80)</td>
<td>140.96 (9.83)</td>
<td>69.12 (4.10)</td>
<td>99.69 (5.37)</td>
</tr>
<tr>
<td></td>
<td>Slope (°)</td>
<td>2.76 (0.35)</td>
<td>1.99 (0.40)</td>
<td>0.59 (0.13)</td>
<td>0.75 (0.11)</td>
</tr>
<tr>
<td>Hydrology</td>
<td>Water width (m)</td>
<td>32.05 (1.89)</td>
<td>30.04 (2.93)</td>
<td>136.05 (11.12)</td>
<td>54.25 (5.81)</td>
</tr>
<tr>
<td></td>
<td>Water depth (cm)</td>
<td>26.15 (0.54)</td>
<td>22.97 (0.47)</td>
<td>49.53 (1.15)</td>
<td>25.51 (0.27)</td>
</tr>
<tr>
<td></td>
<td>Velocity (cm/s)</td>
<td>64.97 (0.75)</td>
<td>37.38 (0.46)</td>
<td>18.26 (0.83)</td>
<td>14.95 (0.42)</td>
</tr>
<tr>
<td>Physicochemistry</td>
<td>Biochemical oxygen demand (BOD) (mg/L)</td>
<td>1.62 (0.07)</td>
<td>1.70 (0.08)</td>
<td>2.10 (0.07)</td>
<td>1.89 (0.05)</td>
</tr>
<tr>
<td></td>
<td>Total nitrogen (TN) (mg/L)</td>
<td>2.83 (0.09)</td>
<td>2.52 (0.09)</td>
<td>2.67 (0.07)</td>
<td>2.40 (0.05)</td>
</tr>
<tr>
<td></td>
<td>Total phosphorus (TP) (mg/L)</td>
<td>0.10 (0.01)</td>
<td>0.10 (0.01)</td>
<td>0.10 (0.01)</td>
<td>0.08 (0.01)</td>
</tr>
<tr>
<td></td>
<td>Chlorophyll a (Chl-a) (mg/L)</td>
<td>2.65 (0.19)</td>
<td>3.17 (0.25)</td>
<td>5.05 (0.38)</td>
<td>3.26 (0.14)</td>
</tr>
<tr>
<td></td>
<td>Dissolved oxygen (DO) (mg/L)</td>
<td>9.29 (0.08)</td>
<td>8.86 (0.08)</td>
<td>8.74 (0.09)</td>
<td>8.41 (0.06)</td>
</tr>
<tr>
<td></td>
<td>pH</td>
<td>7.77 (0.02)</td>
<td>7.75 (0.03)</td>
<td>7.88 (0.03)</td>
<td>7.80 (0.02)</td>
</tr>
<tr>
<td></td>
<td>Electric conductivity (µS/cm)</td>
<td>195.05 (10.08)</td>
<td>239.53 (21.24)</td>
<td>764.06 (202.54)</td>
<td>285.68 (64.57)</td>
</tr>
<tr>
<td></td>
<td>Turbidity (Nephelometric Turbidity Units, NTU)</td>
<td>11.67 (0.77)</td>
<td>7.74 (0.80)</td>
<td>12.64 (0.60)</td>
<td>13.02 (0.85)</td>
</tr>
</tbody>
</table>

Note: Generally, alphabet "a" represent the highest value following "b", "c", "d", etc.
2.2. Data Analyses

Hierarchical cluster analysis (HCA) and nonmetric multidimensional scaling (NMDS) were performed to analyze the relationships between the occurrence (i.e., presence or absence) of aquatic oligochaetes and environmental factors at multiple spatial scales. Sampling sites were classified into several groups based on the similarities of their environmental conditions through a hierarchical HCA using the Ward linkage method with Euclidean distance measure. The same data used in HCA were used in NMDS, which is a method of scaling that places similarities between variables on a multidimensional scale [50]. In NMDS, a biplot was used to characterize the relationships among environmental factors, sampling sites, and occurrence of oligochaetes. Environmental factors such as water width, altitude, Chl-\(a\), conductivity, and turbidity were log-transformed prior to analyses to reduce the variation in each variable. One-way analysis of variance (ANOVA) was performed to evaluate differences in the occurrence frequencies of oligochaetes and environmental factors among the groups defined in the HCA. Tukey’s multiple comparison tests were performed if there were differences among groups in HCA.

A random forest (RF) model was used to evaluate the relative importance of environmental factors affecting the distribution of oligochaetes. After building the RF model with the 13 selected environmental factors as independent variables, and presence or absence of oligochaetes as a dependent variable, a measure of the total decrease in node impurity (IncNodePurity) was used to evaluate the relative importance of the 13 environmental factors. We evaluated the relationships between two of the most important factors and the occurrence of oligochaetes on a three-dimensional graph fitting a smooth surface to the points of oligochaete occurrence.

All analyses were performed in R [51]. HCA, ANOVA, and Tukey’s test were conducted using the stats package [51] in R. NMDS was carried out using the metaMDS function of the vegan package [52], RF was performed using the randomForest package [53], and generation of three-dimensional graphs with a fitting surface model was conducted with the car package [54] in R.

3. Results

The sampling sites were classified into four groups according to similarities in their environmental factors through HCA (Figures 2 and 3). The percentage occurrence of \(L.\) hoffmeisteri in each group was in the range of 82.7% (in group 1) to 96.0% (in group 4) (Table 1). The groups had significantly different environmental characteristics. The sampling sites in groups 1 and 2 were located in areas characterized by high altitude and slope, whereas those in groups 3 and 4 were in areas with low values of altitude and slope. Velocity was significantly different among the four groups (ANOVA, \(p < 0.05\)) with the highest value in group 1 (65.0 cm/s) and the lowest value in group 4 (14.9 cm/s). The water width of groups 3 and 4, which showed high occurrence frequencies of oligochaetes, were 136.1 and 54.2 m, respectively, whereas width was low in groups 1 and 2, at 32.0 and 30.0 m, respectively. Physicochemical factors reflected the differences in geological and hydrological factors among groups. Groups 1 and 2 displayed lower BOD, conductivity, and turbidity, whereas these factors were higher in groups 3 and 4, indicating the relatively polluted condition at sampling sites in groups 3 and 4. DO and TN showed the opposite trend. TP was relatively constant without any significant differences among the four groups.

Sampling sites in group 1 were mainly from the Han River (Figure 3). Sites in group 2 were mostly distributed near the mountain range, and groups 1 and 2 contained few samples from Nakdong River. Groups 3 and 4 showed high occurrence frequency of \(L.\) hoffmeisteri. Geographical distribution reveals that sampling sites in group 3 were polluted downstream of Nakdong River with high values of BOD, conductivity and turbidity. The sites in group 4 had the lowest flow rate and DO, high conductivity and turbidity.
was in the center, group 3 was on the upper right, and group 4 was on the lower right. Axis 1 was highly correlated with 1.0 of AUC (area under the curve), which is generally used to evaluate prediction models with presence or absence output. Therefore, we used this model to evaluate the importance of environmental variables affecting the distribution of oligochaetes. Turbidity, velocity, DO, and conductivity were accordingly identified as the most influential factors (Figure 5a). We evaluated the changes in oligochaete occurrence in response to changes in the two most important factors (the physicochemical factor turbidity and the hydrological factor velocity) based on a surface map using a smooth regression method (Figure 5b). As the velocity increased, the occurrence probability of oligochaetes decreased. In contrast, as turbidity increased, there was an increase in occurrence probability. Therefore, the highest occurrence percentage is the proportion of the number of sampling sites with presence of *Limnodrilus hoffmeisteri* among the number of sites assigned in each group, and the percentage is the proportion of the number of sampling sites with presence of *L. hoffmeisteri*.
probability of oligochaetes coincided with a low velocity and a high turbidity, whereas occurrence probability was low at very low turbidity with high velocity.

![Nonmetric multidimensional scaling (NMDS) ordination of sampling sites based on the similarities of 13 environmental factors. The stress in the first two axes was 0.034. Occ_freq: occurrence frequency of *Limnodrilus hoffmeisteri.*](image1)

**Figure 4.** Nonmetric multidimensional scaling (NMDS) ordination of sampling sites based on the similarities of 13 environmental factors. The stress in the first two axes was 0.034. Occ_freq: occurrence frequency of *Limnodrilus hoffmeisteri.*

![Nonmetric multidimensional scaling (NMDS) ordination of sampling sites based on the similarities of 13 environmental factors](image2)

**Figure 5.** (a) The importance of environmental factors affecting the occurrence of *Limnodrilus hoffmeisteri* based on a random forest model. (b) Relationship between the two most important factors (water velocity and turbidity) and the occurrence of *L. hoffmeisteri* based on the random forest model.

4. Discussion

We investigated the distribution patterns of aquatic oligochaetes *L. hoffmeisteri* in relation to environmental factors in three categories. Sampling sites were classified according to altitude, and this geographical factor influenced hydrological and water quality factors. Our results showed that *L. hoffmeisteri* was mostly present at sites with low altitude and slope, wide water width, and a low velocity. In addition, on the basis of an RF model, current velocity and turbidity were identified as the most important factors influencing the distribution of *L. hoffmeisteri*. Therefore, local environmental conditions are important in determining the distribution of *L. hoffmeisteri* in our study even though large-scale environmental conditions (e.g., altitude) are also influential to its occurrence, supporting the environmental filtering theory for the distribution of species [55,56].
The habitat conditions for oligochaetes can be explained in terms of their ecological traits. Aquatic oligochaetes are negatively phototactic [57]. The rear end of the body burrows in the water, whereas the front of the body remains buried in the sediment [57,58]. Moreover, they are deposit-feeders, so they eat soft bottoms which are mainly present in the most lowland section of rivers, where water velocity is low with high turbidity and Chl-α. In addition, it is well known that oligochaetes can resist to oxygen depletion [59] and sediment contamination. Therefore, they can survive at sites with high anthropogenic contamination (e.g., high turbidity, high organic matter and presence of toxic substances).

The particle size of substrates is assumed to play a major role in the distribution and abundance of tubificids [16]. Generally, aquatic oligochaetes prefer silt clay, which has the smallest sediment particle size [60], and tubificid species in particular have preference for a particle size of less than 63 µm [61,62]. In a study on the hydrological and substrate determinants of oligochaete distribution, Verdonschot [7] showed that hydro-morphological parameters explained differences in the major distribution patterns.

On the basis of the aforementioned properties, aquatic oligochaetes, including *L. hoffmeisteri*, are used for bioprocessing ability in the restoration of organic-polluted freshwater habitats [16–19], and for the improvement of water quality [21, 22]. Tubificid species are useful for the denitrification of NO₃⁻ [18] and the reduction of sludge [16,19] in the overlying water. In terms of water chemistry, major changes in the structure of oligochaete assemblages have been shown to be related to water mineralization. Martínez-Ansemil and Collado [12] reported that substrate and water velocity are the most important factors influencing the distribution of oligochaetes, but that water mineralization is also positively related to oligochaete distribution. Marchand [63] has also reported that DO and organic matter affect the distribution of oligochaetes. Organic matter and dissolved minerals are components of turbidity [64], and whereas oligochaetes are affected by water mineralization, these worms in turn also have an influence on water mineralization [18].

*L. hoffmeisteri* is one of the cosmopolitan oligochaetes [47] because it has tolerance to both organic and inorganic pollution, and exhibits variation in life history (duration of breeding period) characteristics according to the gradients of variously scaled environmental conditions [65,66]. Furthermore, unlike other tubificids, *L. hoffmeisteri* does not have a preference for silt and clay substrates [16], which accordingly contributes to this species’ widespread distribution and abundance [65].

5. Conclusions

*L. hoffmeisteri* is distributed throughout South Korea, widely applied as biological indicators and a highly recommended candidate to mitigate organic-enriched freshwater ecosystem. However, up to now there has been limited studies on their ecological traits and favorable environmental conditions. In our study, we found out that its distribution is mainly determined by local conditions such as turbidity and water velocity based on the results of HCA, NMDS and RF. Our study, based on a nationwide scale database, showed that other local environmental factors such as water quality are also important. However, further studies including local to large scaled surveyed dataset are necessary to figure out its ecology and ecological traits deeply.

Acknowledgments: This work was supported by the National Research Foundation of Korea (NRF) funded by the Korean government (MSIP) (grant number NRF-2016R1A2B4011801). The data were collected under the project of “National Aquatic Ecosystem Health Survey and Assessment” supported by the Ministry of Environment and the National Institute of Environmental Research, Korea. The authors are grateful to survey members involved in the project, and also thank the anonymous reviewers for their helpful comments in improving the scientific content of the manuscript.

Author Contributions: H.K. prepared the dataset, performed data analyses, and wrote the manuscript. M.-J.B. performed data analyses, and wrote the manuscript. S.-J.H., and J.-S.M. provided the dataset, and D.-S.L. assisted data processing. Y.-S.P. developed the concept of the study, performed data analysis and wrote the manuscript. All the authors contributed to the review of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.
References


9. Pascar-Gluzman, C.; Dimentman, C. Distribution and habitat characteristics of Naididae and Tubificidae in the inland waters of Israel and the Sinai Peninsula. *Hydrobiologia* 1984, 115, 197–205. [CrossRef]


27. Lucan-Bouché, M.L.; Biagianti-Risbourg, S.; Arsac, F.; Vernet, G. An original decontamination process developed by the aquatic oligochaete Tubifex tubifex exposed to copper and lead. Aquat. Toxicol. 1999, 45, 9–17. [CrossRef]


42. Prasad, A.M.; Iverson, L.R.; Liaw, A. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. Ecosystems 2006, 9, 181–199. [CrossRef]

44. Ministry of Environment (MOE); National Institute of Environmental Research (NIER). The Survey and Evaluation of Aquatic Ecosystem Health in Korea; NIER: Incheon, Korea, 2008; pp. 43–72.
46. Brinkhurst, R.O. Guide to the Freshwater Aquatic Microdrile Oligochaetes of North America; Canadian Special Publication of Fisheries and Aquatic Sciences: Ottawa, ON, Canada, 1986; Volume 84, p. 259.