Diversity Analysis of Adult Chironomidae in the Lake Victoria Basin of Kenya

Walter Odhiambo OTIENO*, Reuben Oyoo MOSI, Peter BULLI
School of Agricultural and Food Sciences, Jaramogi Oginga Odinga University of Science and Technology, P. O. Box 210-40601, Bondo, Kenya

ABSTRACT

Chironomidae commonly inhabits most aquatic habitats and often dominate aquatic insect communities in abundance and species richness. Despite their ecological importance, the diversity and distribution of chironomids in the Lake Victoria ecosystem of Kenya have not been studied to date. Here we report on the diversity and distribution of adult Chironomidae in Usenge, Mbita and Ogal beaches of the Lake Victoria ecosystem in Kenya using morphological features and sequence data of Cytochrome c oxidase subunit I (COI) gene. Wing venation-based microscopic characterization identified four genera, Tanypus, Coelotanypus, Dicrotendipes and Chironomus. The COI gene barcoding further revealed several species, including Kiefferulus brevibucca, Chironomus flaviplumus, Polypedilum fuscovittatum, Polypedilum sp. and Dicrotendipes sp. The identified species were grouped into three clusters based on neighbor-joining phylogenetic approach. Differences in species richness were observed among the three study sites, with Mbita exhibiting the highest species richness. The evolutionary analysis revealed relatedness among all the identified species, suggesting a shared recent common ancestor. Unlike previous studies, this study represents the first report on detailed characterization of Chironomidae in the Lake Victoria ecosystem of Kenya. Moreover, this study serves as a first step towards a comprehensive understanding of the range of species of Chironomidae inhabiting this ecosystem.

Key words: Chironomidae, cytochrome c oxidase subunit I, DNA barcoding, wing venation, diversity

INTRODUCTION

Food demands of a rapidly growing world population (United Nations, 2017, Gu et al., 2021) are putting tremendous pressure on the available limited resources of water and land which also serve as important ecological habitats for many species. Climate change-driven variations in temperature and precipitation is further exacerbating the pressure on the limited resources through geographical shifts in habitats of many species. The impact of these challenges is pronounced in the Lake Victoria Basin of East Africa where human activities, coupled with climate change, have directly contributed to degradation of the lake and its surrounding terrestrial habitat, which is reflected in mideic responses to these impacts (Nyakeya et al., 2018). The habitat degradation, which is reflected in the observed low genetic variation in species such as Nile perch (Lates niloticus) (Basita et al., 2018) of the ecosystem, is worsening the already fragile balance of food webs in the lake ecosystem. For instance, changes in water quality and biological activities of the lake in the last five decades have caused shifts in populations of several species inhabiting the ecosystem, with the most notable change observed in the significant increase in biomass of chironomids which are commonly referred to as lake flies in the region (Gikuma-Njuru et al., 2005). Since ‘lake flies’ is a general term and can be misleading, we used the word ‘chironomids’ throughout in this publication.

Chironomids, which belong to the order Diptera and the family Chironomidae, are an important food resource for fish, amphibians, dragonfly larvae, birds, bats, and spiders (Niemi et al., 1999, Wardiatno and Krisanti, 2013, Sharifian Fard et al., 2014). The chironomids are also an important tool in ecological and paleo-ecological studies, as well as in environmental evaluation, agricultural entomology, and public-health research (Rosenberg, 1992, Ferrarese, 1992, Armitage et al., 1995). Although not widely reported, the chironomids are also one of the commonly consumed insects by the communities inhabiting the region in addition to their...
Diversity Analysis of Adult Chironomidae in the Lake Victoria Basin of Kenya

ecological importance (Ayieko and Oriaro, 2008). Recently, the chironomids are sighted less frequently (Ayieko et al., 2010b), raising fears among the local communities of decline in their population. In spite of the importance of chironomids as link between the aquatic and terrestrial food web (McCary et al., 2021), and as sources of protein to the local community (Ayieko and Oriaro, 2008), their diversity and distribution around the region has not been studied to date. Knowledge on the diversity of this important component of ecosystem will not only greatly enhance our understanding of the fauna and ecology of Lake Victoria, but will also enable formulation of evidence-based decisions on conservation of their biodiversity considering that the predicted exponential growth of human population in the region is expected to lead to increased consumption of the indigenous insects (Ayieko et al., 2010a). Therefore, to further enhance our understanding of the fauna and ecosystem of Lake Victoria, and contribute to the national and global database of Chironomidae in the region, this study is designed to characterize the diversity of Chironomidae in different sites of Lake Victoria.

MATERIAL AND METHODS

Study Area

The Lake Victoria Basin of Kenya (LVB) is located in the western part of the country (Fig. 1) with an approximate area of 4,113 km². The lake lies at an altitude of 1,134 m above sea level, and lies between latitudes 0° 20’ N, 3° 0’ S and longitudes 31° 39’ E, 34° 53’ E. Lake Victoria, the most notable landmark of the region, has an average depth of approx. 40 m, with a maximum depth of 84 m. LVB comprises seven districts, including parts of Siaya, Bondo, Kisumu, Nyando, Rachuonyo, Homa Bay, Migori, Suba and Busia, with a population of over 4 million people. Diverse ethnic groups of indigenous people sharing similar livelihoods within the region include the Luo, Luhya, Kisii, Kuria, Suba, Teso, Kalenjin and Maasai. The main economic activities in the lake region are fishing, farming, bee keeping, sand and gold mining and trading (UNEP, 2006).

Morphological Characterization of Chironomidae in the Lake Victoria Basin of Kenya

Sampling of Chironomidae

Adult chironomids were collected from three sites: Usenge Beach (latitude 0° 24’ 54.7848” N; longitude 34° 12’ 14.3136” E), Mbita (latitude 0° 4’ 20.28” N; longitude 34° 3’ 39.4848” E) and Ogal Beach (latitude 0° 8’ 27.3264” N; longitude 34° 35’ 36.1428” E) in Siaya, Homabay and Kisumu Counties, respectively. The sites were selected based on frequent seasonal sightings of the chironomids. Samples of chironomids were collected twice, December 2020 to January 2021, and June 2021, coinciding with their seasonal emergence. Flying chironomids were trapped using a sweep net comprising a net tied to a round metallic ring and a 1.5 m-long wooden handle. To trap

Figure 1: Map of Lake Victoria Basin, Kenya
the flies, the net was whirled in the air. Shrubs and trees, along the shores of the lake, were occasionally shaken to disturb the flies and force them to fly during collection as described by Ayieko and Oriaro (2008). Collected samples were preserved in 70% ethanol in plastic containers, and transported to the laboratory for sorting and subsequent studies. For each collection site, there were nine samples (with each containing an average of approx. 151 flies), giving a total of 54 samples for the two sampling periods of December 2020 to January 2021, and June 2021.

**Morphological characterization**

Morphological characterization was performed at the National Museums of Kenya. Wings were dissected using fine needles under a dissecting Wild M3 microscope (Wild Heerbrugg Co., Switzerland), and then mounted on semi-permanent slides for characterization to group them into genera. Morphological characters that aided identification were first branch of anal vein (A1); basal-medial crossvein (bm-m); costal vein (C); calypter (calyp); cubitus (Cu); anterior branch of cubital vein (CuA); cubital vein fork (Cu-fork); posterior branch of cubital vein (CuP); media (M); fused first and second branch of media (M1+2); fused third and fourth branch of media (M3+4); medial-cubital crossvein (m-cu); anterior branch of radius (R1); upper branch of second branch of radius (R2); second branch of radius (R3); lower branch of second branch of radius (R4); third branch of radius (R5); and radial-medial crossvein (r-m) (Kirk-Springs and Sinclair, 2017).

**Diversity analysis**

The diversity indices of Shannon (H') and evenness index were used to assess insect diversity within and between sites. Shannon-Wiener diversity index $H' = \sum_{i=1}^{s} p_i \ln p_i$; where $s =$ number of genera; and $p_i =$ proportion of the i-th genus, $n_i/N$. Shannon evenness $= H'/\ln s$ where $s =$ total number of genera. The Dominance index ($D = P_{\text{max}}$) was used to express the proportion of individuals accounted for by the most abundant genera in each site; where $P_{\text{max}}$ is the maximum proportion of any one genus in a sample. All the above diversity indices were performed using PAST Software Version 4.10 (https://www.nhm.uio.no/english/research/infrastructure/past/).

**Genomic DNA extraction**

To help support taxonomic identifications, along with having a greater chance of determining novel taxa that were not identified by the microscopic-based morphological characterization, the specimens were separated into taxonomic groups based on morphology and then specimens were randomly selected from within those groups. Total genomic DNA was extracted from 4 of the random samples per site using AuPrep Genomic DNA Extraction Kit (Life Technologies, Deli, India) following the manufacturer's instructions. Each sample was ground and placed into 1.5 ml sterile tubes. 180 µl LYS buffer was added to each sample followed by 20 µl Proteinase K. The samples were mixed by vortexing for 20 s. To lyse the samples, they were incubated at 60°C for 1 hr 300 µl FX buffer was added to the samples and mixed by vortexing, followed by incubation at 70°C for 20 min. Two hundred microliter of 98% ethanol was added to the samples, followed by vortexing. The mix was pipetted into a B/T Genomic DNA Mini column placed onto a Collection Tube; and centrifuged at 8,000 rpm (6,000 xg) for 2 mins. The column was then placed onto a new Collection Tube. The column was washed twice with 0.5 ml WS Buffer by centrifuging at 8,000 rpm for 2 min. The column was centrifuged at full speed (10,000 xg) for another 2 mins to remove ethanol residue, and then placed onto a new 1.5 ml tube. DNA was eluted with 200 µl of the preheated elution solution. The column was left to stand for 5 mins, and then centrifuged for 2 mins to elute DNA. The extracted DNA was quantified, subjected to quality analysis, and stored at -20°C.

**Amplification of cytochrome c oxidase subunit I gene**

The chironomid COI barcode approach (Kanget et al., 2022, Vasquez et al., 2022) is widely and effectively used to characterize chironomids at the species level. Cytochrome C Oxidase Subunit I (COI) gene was amplified following established protocols using the pairs of primers, LCO1490: 5’-GGTCAAAATCATAAAGATATTGG-3’ and HCO2198: 5’-TAAACTTCAGGGTGACCAAAAAATCA-3’. Amplification was performed using HotStarTaq Master Mix (QIAGEN, Valencia, CA) with standard buffer, and 7 µl of nuclease free water. To amplify the 5' region of mitochondrial COI gene, the following thermal cycling program was used: hot start at 94°C for 5 mins, 35 cycles of 94°C for 30 s, 50°C for 30 s and 68°C for 60 s, followed by a final extension at 68°C for 10 mins. The PCR products were held at 4°C. Integrity of the PCR amplicons was visualized on a 1% CSEL-AG500 agarose gel (Cleave Scientific Ltd, UK) stained with EZ-vision Bluelight DNA Dye.

**Purification of PCR product**

PCR products were purified using ExoSAP-IT™ PCR Product Cleanup Reagent (ThermoFisher Scientific, MA, USA) following the manufacturer’s protocol. ExoSAP master mix was prepared from 50 µl of Exonuclease I and 200 µl of Shrimp.
Alkaline Phosphatase in a 0.6 ml micro-centrifuge tube. Ten microliters (10 µl) of post-PCR reaction product was mixed with 2.5 µl ExoSAP Mix. The reaction mixture was well mixed and incubated at 37 °C for 15 mins to degrade remaining primers and nucleotides. To inactivate the ExoSAP-IT reagent, the mix was incubated at 80 °C for 15 mins. The purified PCR products were stored at -20 °C prior to sequencing.

**Sequencing of PCR products**

Amplified DNA fragments were sequenced using the BrilliantDyeTM Terminator (v3.1) Cycle sequencing Kit (NimaGen B.V., The Netherlands) according to manufacturer’s instructions. For labelling of the amplified PCR fragments, the cycle sequencing reaction contained 1 µl BrilliantDye v1.1 rr Premix, 3.5 µl 5x sequencing buffer, 1 µl template, 1 µl primer (3.2 - 5 pMol), and 13.5 µl water. The thermal cycling protocol was as follows: initial denaturation of 96 °C for 45 s, followed by 25 cycles of 96 °C for 10 s, 50 °C for 5 s and 60 °C for 4 mins. Before capillary electrophoresis the cycle sequencing products were purified to remove unincorporated fluorescent ddNTPs and salts using ZR-96 DNA Sequencing Clean-up Kit (Zymo Research Corporation, USA). The purified products were injected on an Applied Biosystems ABI 3500XL Genetic Analyzer (ThermoFisher Scientific Inc., USA) with a 50 cm array, using POP7. For control, 1 µl of pGEM plasmid DNA template and 1 µl of -21M13 primer provided in the BrilliantDyeTM Terminator (v3.1) Cycle sequencing Kit were included in the sequencing reaction to verify the performance of the total workflow and troubleshooting correlated to the amplified COI gene fragments and primers. Raw sequences were edited using BioEdit 7.2 software (Hall et al., 2011) to remove the ‘N’s and trim the beginning and end of the sequences to maintain high accuracy. Consensus sequences were derived from forward and reverse sequences.

**Analysis of sequences of COI fragments**

**Sequence similarity search**

Megablast program of the Nucleotide Basic Local Alignment Search Tool (BLASTn) was used to search for similarities between sequences of the amplified COI fragments and sequences in the Genome (NIH_TR_1.0 reference Annotation release 100) database of the National Center for Biotechnology Information (NCBI). Cut off for similarity between sequences in the database and the COI query sequences was 80%.

**Multiple sequence alignment**

Sequences of the COI fragments were aligned with their corresponding accessions retrieved from the NCBI database using MUSCLE (Multiple Sequence Comparison by Log-Expectation) tool in the software MEGA 11 (Edgar, 2004, Tamura et al., 2021) with the following default settings: gap open -400.00; gap extend 0.00; maximum memory 2,048 megabytes; maximum iterations 16; cluster method (iterations 1.2) UPGMA; cluster methods (other iterations) UPGMA; and minimum diagonal length (lambda) 24.

**Substitution model for genetic analysis**

Kimura-2 parameter (K2P) distance model (Kimura, 1980) was used for calculating intraspecies and interspecies sequence divergences, pairwise distances and overall distance. Assuming P is the fraction of nucleotide sites that show Type I difference or transition type (homologous sites occupied by different nucleotide bases but both are purines or both are pyrimidines), and Q the fraction of nucleotide sites showing Type II difference or transversion type (one of the two nucleotide bases is a purine and the other is a pyrimidine) between two sequences, the evolutionary distance per site is calculated using the formula $K = \frac{1}{2} \ln \left[ \frac{1 - 2P - Q}{4 - 2Q} \right]$. The evolutionary rate per year is given by $k = K/(2T)$, where T = time since the divergence of the two sequences.

**Evolutionary Analysis**

Disparity index test (Kumar & Gadagkar, 2001) was used to test the homogeneity of substitution patterns between the 11 sequences. Pairwise and within group distances were estimated using Kimura 2-parameter model. Monte Carlo test of 1,000 replicates to estimate the P-values and standard errors. Ambiguous positions were removed for each sequence pair (pairwise deletion option). P-values smaller than 0.05 are considered significant. The analysis was performed for codon positions 1st+2nd+3rd+Noncoding. A total of 730 positions were included in the final dataset. All the evolutionary analyses were conducted using MEGA 11.

The number of base substitutions per site from averaging over all sequence pairs within and between sites were shown. Standard error (SE) estimates were obtained by a bootstrap procedure of 1000 replicates. Analyses were conducted using the Kimura 2-parameter model. This analysis involved 11 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All ambiguous positions were removed for each sequence pair (pairwise deletion option). There was a total of 730 positions in the final dataset.

Neighbor-joining statistical method was used to construct phylogenetic tree based on the 11 nucleotide sequences of the COI gene fragments amplified from DNA of randomly selected chironomids from the three sites using MEGA version 11.0.13. Bootstrap was 1,000 replications with the following settings: substitution model Kimura 2-parameter, substitutions including transitions and transversions, uniform rates among sites, homogenous pattern among lineages, pairwise deletion for gaps removal, and 1st, 2nd and 3rd codon positions, and non-coding sites selected.
RESULTS

A total of 8,161 individual Chironomidae from 54 samples were characterized. The average number of flies per sample was approximately 151 (Table 1). In total, 4 genera, *Tanypus*, *Coelotanypus*, *Dicrotendipes* and *Chironomus* were identified based on wing venations using a dissecting microscope (Fig. 2a-d). The wing of *Coelotanypus* is characterized with vein R₂ (upper branch of second branch of radius) not connected to vein R₃ (lower branch of second branch of radius); while the wing of *Tanypus* was characterized with a distance between Cu-fork (cubital vein fork) and cross-vein of less than a third the length of vein CuA (anterior branch of cubital vein). Both *Dicrotendipes* and *Chironomus* have wings with cross-vein r-m (radial-medial crossvein) oblique to vein R₄+5 (third branch of radius). The number of Chironomidae per genus and their percent distribution across the three study sites are shown in Table 1 and Fig. 3a-c. At the taxonomic level, *Dicrotendipes* and *Chironomus* were the only genera identified in Usenge, with *Dicrotendipes* being the predominant genus. Three genera were identified at Ogal, represented with 54, 39 and 7% of *Tanypus*, *Coelotanypus* and *Chironomus*. At Mbita site, all four genera were identified, with *Tanypus*, *Coelotanypus*, *Dicrotendipes* and *Chironomus* accounting for 9, 8, 59 and 24% of the collected samples of *Chironomidae*. *Chironomus* was the only genus collected from all three sites.

Table 1: Chironomidae collected from sampling sites

<table>
<thead>
<tr>
<th>Genus/Site</th>
<th>Usenge beach</th>
<th>Mbita</th>
<th>Ogal beach</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Tanypus</em></td>
<td>0</td>
<td>124</td>
<td>1730</td>
<td>1854</td>
</tr>
<tr>
<td><em>Coelotanypus</em></td>
<td>0</td>
<td>112</td>
<td>1265</td>
<td>1377</td>
</tr>
<tr>
<td><em>Dicrotendipes</em></td>
<td>2312</td>
<td>804</td>
<td>0</td>
<td>3116</td>
</tr>
<tr>
<td><em>Chironomus</em></td>
<td>1268</td>
<td>331</td>
<td>215</td>
<td>1814</td>
</tr>
<tr>
<td>Total</td>
<td>3580</td>
<td>1371</td>
<td>3210</td>
<td>8161</td>
</tr>
</tbody>
</table>

Figure 2: Wing venation: (A) *Tanypus*; (B) *Coelotanypus*; (C) *Dicrotendipes*; and (D) *Chironomus*
Diversity analysis

The diversity indices are summarized in Table 2. All four genera were recorded in Mbita, with *Dicrotendipes* being the most abundant. An uneven distribution of the number of genera was observed in Mbita, with *Dicrotendipes* being the dominant genus (Figure 3). Similarly, the Shannon diversity index revealed a relatively higher level of diversity in Mbita, followed by Ogal and lastly Usenge. The individual Chironomidae for the four genera in Mbita are more evenly distributed compared to those of both Usenge and Ogal.

Table 2: Diversity indices for Chironomidae genera collected at Lake Victoria

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Indices</th>
<th>Usenge</th>
<th>Mbita</th>
<th>Ogal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taxa S (Genera)</td>
<td></td>
<td>2&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3&lt;sup&gt;ab&lt;/sup&gt;</td>
</tr>
<tr>
<td>Individuals</td>
<td></td>
<td>3580&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1371&lt;sup&gt;c&lt;/sup&gt;</td>
<td>3210&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Dominance D</td>
<td></td>
<td>0.5425&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.417&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.4502&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Shannon H</td>
<td></td>
<td>0.65&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.078&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.8812&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Evenness</td>
<td></td>
<td>0.0794&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.1492&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.1091&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Values with different superscripts in rows are significantly different at \( p < 0.05 \).

DNA Barcode analysis

Amplified fragments using primers targeting the COI gene region from 12 random samples of individual Chironomidae were sequenced (hereafter referred to as Seq #1 to Seq #12). Seq #8 was excluded from further analysis because the quality of the reverse sequence was poor. Sequence similarity search was performed using the NCBI Blastn database, and the results are shown in Table 3. The length of the 11 sequences ranged from 466 to 678 base pairs. The low Expected values (e-values) for all the searches indicated very high scores (S) for the matches between the queries and sequences in the NCBI database, with 96-100% of the query lengths included in the aligned sequences. The percent similarity ranged from 83.6 to 94.7. The sequence similarity search identified a total of 5 species, including *Kiefferulus brevibucca*, *Chironomus flaviplumus*, *Polypedilum fuscovittatum*, *Polypedilum* sp. and *Dicrotendipes* sp. *Kiefferulus brevibucca* was identified in Ogal and Mbita; *Chironomus flaviplumus* and *Polypedilum fuscovittatum* were identified in Usenge and Mbita; and *Polypedilum* sp., and *Dicrotendipes* sp. were identified in Ogal. The four sequences labeled Chironomidae could likely represent unknown or poorly sequenced genera since they are not identified at a finer resolution than family.

Evolutionary relationships of taxa

The evolutionary history was inferred using the Neighbor-Joining method (Saitou & Nei, 1987). The bootstrap consensus tree inferred from 1,000 replicates (Felsenstein, 1985) is taken to represent the evolutionary history of the taxa analyzed. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The evolutionary distances were computed using the Kimura 2-parameter method (Kimura, 1980) and are in the units of the number of base substitutions per site. Ambiguous positions were removed for each sequence pair (pairwise deletion option). The final dataset included a total of 730 positions. Three major clusters were identified (Fig. 4). Cluster 1 is the largest, and is comprised of *Kiefferulus brevibucca*, *Dicrotendipes* sp., *Polypedilum* sp., *Polypedilum fuscovittatum*, and Chironomidae sp. Cluster 2 consists of Chironomidae sp. (Sequence 10), Chironomidae sp. (Sequence 2) and Chironomidae sp. (Sequence 7); while *Chironomus flaviplumus* are grouped separately into Cluster 3. The
species in each of Clusters 1 and 2 are from the three study sites of Mbita, Ogal and Usenge. The sequences for *Kiefferulus brevibucca* from Ogal and Mbita are identical. Similarly, *Chironomus flaviplumus* clusters with *Chironomus flaviplumus*, both collected from Usenge. The Chironomidae sp. collected from Ogal and Chironomidae sp. collected from Mbita also cluster together suggesting that they share the most recent common ancestor at branch point with 56%. Three Chironomidae spp. share the most recent common ancestor at the node with 99%. *Polypedilum fuscovittatum* from Mbita and Chironomidae sp. from Usenge (Sequence 9) are distantly related as the *Kiefferulus brevibucca* from Ogal and Mbita based on the genetic distance between them.

**Evolutionary Relationships Between Species**

The Disparity Index tests the null hypothesis that sequences have evolved with the same pattern of substitution, as judged from the extent of differences in base composition biases between sequences (Kumar & Gadagkar, 2001). It is the probability of rejecting the null hypothesis that sequences have evolved with the same pattern of substitution. A Monte Carlo test is used to determine the statistical significance of the Disparity Index.
Diversity Analysis of Adult Chironomidae in the Lake Victoria Basin of Kenya

The evolutionary pattern of Polypedilum fuscovittatum from Mbita significantly differs from that of Chironomidae sp. from Mbita and Usenge. There are no significant differences in evolutionary patterns among the Chironomidae sp. from Mbita, Polypedilum sp. from Ogal, and Chironomidae sp. and Chironomus flaviplumus from Usenge. Kiefferulus brevibucca from Mbita is distantly related to the Chironomidae sp. and Chironomus flaviplumus. The evolutionary pattern of Polypedilum fuscovittatum from Mbita is similar to that of Chironomidae sp. from Usenge. The Chironomidae sp. from Usenge is distantly related to the Chironomidae sp. from Ogal and Usenge, respectively; while the evolutionary pattern of the Chironomidae sp. from Usenge is similar. The evolutionary patterns of the Chironomidae sp. and that of Chironomus flaviplumus from Usenge are similar. The Chironomus flaviplumus from Usenge share similar evolutionary patterns. Results from the rest of the homogeneity of substitution patterns between the identified species is further validated by the estimates of evolutionary distances between the species.

Estimates of Average Evolutionary Divergence

The average intraspecific divergence was 31.3% (Table 6). The highest rate of intraspecific diversity was observed in Usenge where three distinct species, including two species of Chironomidae and Chironomus flaviplumus were identified.

Table 4: Test of the homogeneity of substitution patterns between sequences

<table>
<thead>
<tr>
<th>Site</th>
<th>Seq #</th>
<th>Blast hit</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ogal</td>
<td>1</td>
<td>Polypedilum sp.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ogal</td>
<td>2</td>
<td>Chironomidae sp.</td>
<td>0.030*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ogal</td>
<td>3</td>
<td>Kiefferulus brevibucca</td>
<td>0.118***</td>
<td>0.001**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ogal</td>
<td>4</td>
<td>Dicrotendipes sp.</td>
<td>1.000***</td>
<td>0.006**</td>
<td>1.000***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mbita</td>
<td>5</td>
<td>Kiefferulus brevibucca</td>
<td>0.118***</td>
<td>0.001**</td>
<td>1.000***</td>
<td>1.000***</td>
<td>1.000***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mbita</td>
<td>6</td>
<td>Polypedilum fuscovittatum</td>
<td>1.000***</td>
<td>0.003**</td>
<td>0.035*</td>
<td>0.140***</td>
<td>0.038*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usenge</td>
<td>7</td>
<td>Chironomidae sp.</td>
<td>0.066**</td>
<td>0.010*</td>
<td>0.005*</td>
<td>0.037*</td>
<td>0.002**</td>
<td>0.000***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usenge</td>
<td>8</td>
<td>Chironomidae sp.</td>
<td>0.122***</td>
<td>0.000***</td>
<td>1.000***</td>
<td>0.056**</td>
<td>1.000***</td>
<td>0.001**</td>
<td>0.005*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usenge</td>
<td>9</td>
<td>Chironomidae sp.</td>
<td>0.012***</td>
<td>1.000***</td>
<td>0.024*</td>
<td>0.122***</td>
<td>0.046*</td>
<td>0.009*</td>
<td>0.111***</td>
<td>0.001**</td>
<td>0.000***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usenge</td>
<td>10</td>
<td>Chironomidae sp.</td>
<td>0.163***</td>
<td>0.038*</td>
<td>0.042*</td>
<td>0.122***</td>
<td>0.038*</td>
<td>1.000***</td>
<td>0.041*</td>
<td>0.038*</td>
<td>0.145***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usenge</td>
<td>11</td>
<td>Chironomus flaviplumus</td>
<td>0.163***</td>
<td>0.038*</td>
<td>0.042*</td>
<td>0.122***</td>
<td>0.038*</td>
<td>1.000***</td>
<td>0.041*</td>
<td>0.038*</td>
<td>0.145***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usenge</td>
<td>12</td>
<td>Chironomus flaviplumus</td>
<td>0.240***</td>
<td>0.068**</td>
<td>0.063**</td>
<td>0.180**</td>
<td>0.054**</td>
<td>1.000***</td>
<td>0.063**</td>
<td>0.013*</td>
<td>0.184**</td>
<td>1.000**</td>
<td></td>
</tr>
</tbody>
</table>

*, ** and *** = significant values at $p < 0.05$, $p < 0.01$, and $p < 0.001$; NS = non-significant value

Table 5: Estimates of evolutionary distances between species

<table>
<thead>
<tr>
<th>Site</th>
<th>Seq #</th>
<th>Blast hit</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ogal</td>
<td>1</td>
<td>Polypedilum sp.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ogal</td>
<td>2</td>
<td>Chironomidae sp.</td>
<td>0.160</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ogal</td>
<td>3</td>
<td>Kiefferulus brevibucca</td>
<td>0.138</td>
<td>0.239</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ogal</td>
<td>4</td>
<td>Dicrotendipes sp.</td>
<td>0.139</td>
<td>0.206</td>
<td>0.151</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mbita</td>
<td>5</td>
<td>Kiefferulus brevibucca</td>
<td>0.129</td>
<td>0.232</td>
<td>0.002</td>
<td>0.140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mbita</td>
<td>6</td>
<td>Polypedilum fuscovittatum</td>
<td>0.163</td>
<td>0.208</td>
<td>0.163</td>
<td>0.076</td>
<td>0.160</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usenge</td>
<td>7</td>
<td>Chironomidae sp.</td>
<td>0.165</td>
<td>0.005</td>
<td>0.237</td>
<td>0.214</td>
<td>0.228</td>
<td>0.202</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usenge</td>
<td>8</td>
<td>Chironomidae sp.</td>
<td>0.250</td>
<td>0.241</td>
<td>0.322</td>
<td>0.288</td>
<td>0.310</td>
<td>0.253</td>
<td>0.239</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usenge</td>
<td>9</td>
<td>Chironomidae sp.</td>
<td>0.173</td>
<td>0.002</td>
<td>0.248</td>
<td>0.214</td>
<td>0.237</td>
<td>0.205</td>
<td>0.006</td>
<td>0.232</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usenge</td>
<td>10</td>
<td>Chironomidae sp.</td>
<td>0.776</td>
<td>0.777</td>
<td>0.833</td>
<td>0.791</td>
<td>0.830</td>
<td>0.821</td>
<td>0.797</td>
<td>0.862</td>
<td>0.737</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usenge</td>
<td>11</td>
<td>Chironomus flaviplumus</td>
<td>0.772</td>
<td>0.765</td>
<td>0.822</td>
<td>0.780</td>
<td>0.815</td>
<td>0.814</td>
<td>0.772</td>
<td>0.866</td>
<td>0.733</td>
<td>0.002</td>
<td></td>
</tr>
</tbody>
</table>
The species in Mbita and Ogal, respectively, exhibited relatively lower levels of intraspecific diversity. We observed relatively higher levels of genetic dissimilarities between species in Ogal and Usenge, and between species in Mbita and Usenge, respectively. Overall, a moderately low evolutionary divergence ($d = 0.38$ with a standard error of $0.02$) was observed over all the species identified in the three sites.

Table 6: Estimates of average evolutionary divergence

<table>
<thead>
<tr>
<th>Site</th>
<th>Within sites ($d$)</th>
<th>Between sites</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ogal</td>
<td>Mbita</td>
<td>Usenge</td>
</tr>
<tr>
<td>Ogal</td>
<td>0.17±0.01</td>
<td>0.012</td>
<td>0.031</td>
</tr>
<tr>
<td>Mbita</td>
<td>0.20±0.02</td>
<td>0.153</td>
<td>0.031</td>
</tr>
<tr>
<td>Usenge</td>
<td>0.57±0.05</td>
<td>0.503</td>
<td>0.507</td>
</tr>
</tbody>
</table>

Values above the diagonal (between sites) are standard error estimates

**DISCUSSION**

**Characterization of Chironomidae**

Chironomids (Diptera: Chironomidae) are the most globally diverse and widely distributed aquatic insects. Despite their presence in the Lake Victoria ecosystem of Kenya, little is known about the diversity of chironomids relative to other aquatic insect taxa. The morphological-based characterization using wing venations identified only four genera. Tanybus, Coelotanypus, Dicrotendipes and Chironomus from the many specimens. Complementing the morphological approach with analysis of the Cytochrome c oxidase subunit I sequence fragment yielded more detailed results, identifying eight groups at the generic level, and 5 at the species level. The molecular results clearly suggest the presence of many more taxa occurring in the samples which were likely masked by Dicrotendipes and Chironomus. Although the DNA barcode data has been extremely important to this study, our work likely scratched the surface of diversity of the chironomids as we used a relatively smaller sample size for sequencing. DNA barcode data from larger samples of chironomids collected from the shores of the lake as well as adjacent habitats using sweep nets could have yielded a more detailed and comprehensive information on the diversity. The usefulness of integrating DNA barcoding with morphological-based characterization of chironomids was demonstrated by Stur and Ekrem (2020) who could only detect morphological differences of Chironomidae of Svalbard and Jan Mayen to closely related species in the Afrotropical Region (i.e. the biogeographic region that includes Europe, Asia north of the Himalayas, and Africa north of the Sahara), and thus may represent similar African species that

**Diversity of Chironomidae in the Lake Victoria**

The most widespread genus, occurring in all three study sites, was Chironomus, a genus widely distributed worldwide with 19 species reported in the Afrotropical Region (Kirk-Spriggs & Sinclair, 2017). Eggermont et al. (2005) also reported five morphotypes of Chironomus larvae in the East African lakes. Ayieko & Oriaro (2008) previously reported Chironomus as the only genus in the Lake Victoria region of Kenya using personal observation methods and non-structured interviews with relevant local authorities. The findings reported by Ayieko and Oriaro (2008) clearly demonstrated the limitation of morphological characterization based on visual observation in identification of Chironomidae at the species level. They used color as a morphological marker, with black colored insects characterized as Chironomus. Additionally, the information collected from the local community leaders was likely biased as the local community residing in the region commonly refer to all insects resembling the chironomids as ‘kitambo’ in the local language. Moreover, the use of wing vein patterns-based characterization in this study may have also created a situation where specimens were lumped into a limited number of categories that are not always accurate. Although Ayieko and Oriaro (2008) published a very interesting paper, they approaching their study of chironomids of the Lake Victoria region as agriculturalist and/ or cultural anthropologists rather not biologist or taxonomists. Thus, as interesting as their findings are to the growing literature on chironomids, they have a very limited influence on our taxonomic work. As suggested by Stur and Ekrem (2020), suitable identification keys for specific regions are essential for correct identification of chironomids to species using morphology since reliance on identification keys designed for other geographical regions is prone to errors. Although the microscopic-based morphological approach identified the genera Coelotanypus and Tanybus, analysis of the sequence data did not reveal the two genera. However, this does not rule out their presence as the samples collected for sequencing of the COI region are likely not representative of all species of the flies inhabiting the shores of the lake. Twenty-one species of Coelotanypus have been reported, with some found in the Afrotropical Region. The genus Tanybus is characterized with relatively fewer species (only five have been reported so far) (Kirk-Spriggs & Sinclair, 2017). In addition to Chironomus, Coelotanypus and Tanybus, species of the genus Dicrotendipes identified in this study through both the morphological and molecular-based approaches, have also been previously reported in the Afrotropical Region. Other species identified in the Lake Victoria region based on sequence information of the COI region include *Polypedilum* sp., *Kieferulus brevibucca*, *Polypedilum fuscovittatum* and *Chironomus flavipilum*. The presence of *K. brevibucca* is highly likely given that it has been reported previously in some parts of sub-Saharan Africa (Cranston and Martin, 1989). Moreover, *K. brevibucca*, *P. fuscovittatum* and *C. flavipilum* are known from the Palearctic region (i.e. the biogeographic region that includes Europe, Asia north of the Himalayas, and Africa north of the Sahara), and thus may represent similar African species that...
are not yet described or are not yet separable from similar species. Based on the test of the homogeneity of substitution patterns between sequences, there are likely two different species of Chironomidae, with the two occurring in Usenge beach, and one each in the other two sites.

**Evolutionary relationship among species of Chironomidae**

Mitochondrial genomes (mitogenomes) have been widely used for studying the taxonomy and phylogeny of insects (Li et al., 2022). One mitochondrial gene, Cytochrome c oxidase subunit I (COI), was as phylogenetic marker in this study to provide an insight into their evolutionary relationships. The nucleotide sequence of 596 bp from COI was used to determine 11 DNA samples of 11 chironomids collected from Usenge, Mbita and Ogal beaches of Lake Victoria; and the nucleotide sequence alignments were used for construction of phylogenetic trees based on neighbor-joining methods. One key finding is that sampling site is not important in grouping the species into clusters. Rather, the grouping of species from the different sites appeared random. A likely explanation is that there is limited environmental variations among the three sites. In addition, grouping of different species into the same cluster groups is an indication of a shared common ancestor among the species, a fact supported by similarities in nucleotide substitution rate between species, and low intraspecific divergence. An important remaining question is the determination of the position of *Chironomus flaviplumus* in relation to the other species of the identified Chironomidae. *C. flaviplumus* (probably Palearctic) is a global species complex, and the species in Lake Victoria is likely a similar but inseparable species.

**CONCLUSION**

In this study, we found more diversity among chironomids in Lake Victoria Region of Kenya than previously reported. The use of microscopic-based morphological and molecular-based approaches allowed us to point out the limitation of the former which is only efficient at the genus level of identification. However, in spite of the limitations of the morphological-based characterization, there are still excellent keys to the species level that encompass a majority of the diversity present in many bio-regions. The DNA barcoding using Cytochrome c oxidase subunit I (COI) gene as a marker revealed 5 species of the chironomids along the shores of the Lake Victoria. Although the number of the species identified is relatively smaller, it is an important start to what is likely present in the lake. To the best of our knowledge, this study represents the first detailed characterization of Chironomidae in the Lake Victoria region of Kenya. The currently available data and reports tend to group all the available species into lake flies. Sequencing of the partial genes ascertained the high degree of genetic divergence of Chironomidae in the lake region, which is a starting point for further research. Although we recorded 7 species of Chironomidae in the Lake Victoria ecosystem of Kenya for the first time, the species richness of Chironomidae is likely much higher due to the small sample sizes collected in our study. Although for some genera such as Chironomus and Polypedilum, molecular approaches are very important, they do not always solve taxonomic problems.

Chironomidae are not considered a nuisance by the rural communities along the shores of Lake Victoria in Kenya. Rather, these insects are considered important sources of protein for the rural local communities residing in the Lake Victoria region of Kenya. Therefore, it necessary to study the nutritional profile of all species prevalent in the region. Future studies should also establish the relationship between diversity of the Chironomidae with respect to the environmental variables of Lake Victoria and its surroundings so as to design effective conservation strategies that target the insects. Several studies have linked correlation of environmental variables with the composition of chironomid communities (Belle and Goedkoop, 2021, Leszczyńska et al., 2021, Čerba et al., 2022). Climate change induced changes in precipitation patterns and temperatures are affecting the ecosystem of the Lake Victoria region (Akurut et al., 2014, Luhunga and Songoro, 2020, Ototo et al., 2022) which in turn can have potentially profound effects on the diversity and distribution of the chironomids. In addition, landscape inputs from urban and/or agricultural areas are probably also important in influencing the diversity and distribution of the chironomids. For example, the locals harvesting these insects for food have reported less sightings of the chironomids, and this is likely attributable to the changing climate and other factors such as human activities. Thus, it is necessary to monitor the population dynamics of the chironomids over a longer period of time to further enhance our understanding of the long-term impacts of climate change on the diversity and distribution of these edible insects. As a key component to future monitoring activities, a peer reviewed reference collection of the specimens collected in this study is established, and will be used to confirm populations over time and across different localities. Moreover, cleared specimens of adults (not just wings) will need to be permanently slide mount to create a reference collection that will be housed at an insect museum to help establish a good collection with long term curation for future research and will benefit researchers in the long run. Also, future studies should focus on the ecological impacts of chironomids in the Lake Victoria region by investigating how subsidies of the chironomids affect litter processing and microbial communities, and evaluating how those belowground effects related to changes in inorganic nitrogen, plant composition and net primary productivity.
Acknowledgements

This study was funded by the Africa Center of Excellence in Sustainable Use of Insects as Food and Feeds (INSEFOODS) of Jaramogi Oginga Odinga University of Science and Technology, and the World Bank.

REFERENCES


IZVLEČEK


Ključne besede: Chironomidae, citokrom c-oksidaza podenota 1, črno kodiranje DNA, venacija kril, raznolikost