Posted on Authorea 1 Jul 2022 | The copyright holder is the author/funder. All rights reserved. No reuse without permission. | https://doi.org/10.22541/au.165605345.51691863/v1 | This a preprint and has not been peer reviewed. Data may be preliminary.

Diet in phenotypically divergent sympatric species of African weakly electric fish (genus: Campylomormyrus) – a hybrid capture/NGS metabarcoding approach

Rahma Amen[1], Katja Havenstein[2], Frank Kirschbaum[2], and Ralph Tiedemann[3]

[1]Universität Potsdam
[2]Affiliation not available
[3]Universitat Potsdam

July 1, 2022

Abstract

Ecological speciation within the mormyrid genus Campylomormyrus resulted in sympatric species exhibiting divergence in their feeding apparatus and electric organ discharge (EOD). This study provides direct evidence to support the suggested ecological speciation scenario that Campylomormyrus radiation is caused by an adaptation to different food sources. We performed diet assessment of sympatric Campylomormyrus species with markedly different snout morphologies and EODs using hybrid capture/NGS DNA metabarcoding of their stomach contents. Our approach allowed for high taxonomic resolution of prey items, including benthic invertebrates, allochthonous invertebrates, and vegetation. Comparisons of the diet compositions using quantitative measures and diet overlap indices revealed that all species are able to exploit multiple food niches in their habitats, i.e. fauna at the bottom, the water surface, and the water column. The major part of the diet is larvae of aquatic insects, such as dipterans, coleopterans, and trichopterans, known to occur in holes and interstitial spaces of the substrate. The results showed that different snout morphologies and the associated divergence in the EOD translate into different prey spectra. This suggests that the diversification in EOD and the morphology of the feeding apparatus is under functional adaptation.

Introduction

African weakly electric fish (Mormyridae) comprise a species rich group of freshwater fish endemic to Africa with more than 200 described species in 20 genera (Lavoué et al. 2003). Within the genus *Campylomormyrus* , 15 species are described native to the Congo River and its tributaries (Feulner et al. 2007). Each species exhibits a species-specific electric organ discharge (EOD). Even closely related species markedly differ in their waveform shape and/or pulse duration (Tiedemann et al. 2010). Further, they exhibit species-specific morphological traits in their feeding apparatus, i.e. the snout, regarding the snout's length, thickness, and curvature (Feulner et al. 2008). This makes *Campylomormyrus* a prime model system to study the role of ecology in driving an adaptive radiation.

The adaptive radiation within the genus *Campylomormyrus* has been studied with regard to molecular genetics (Canitz et al. 2020; Lamanna et al. 2016), electrophysiology (Feulner et al. 2006), morphometry (Feulner et al. 2007; Lamanna et al. 2016), and behavior (Amen et al. 2020; Nagel et al.et al. 2018a,b). In combination, these studies suggest an ecological speciation scenario that *Campylomormyrus* radiation is caused by an adaptation to exploit different microhabitats and/or food sources, associated with diversification of the EOD. Indeed, behavioral experiments using sympatric *Campylomormyrus* species revealed an association between

Posted on Authorea 1 Jul 2022 | The copyright holder is the author/funder. All rights reserved. No reuse without permission. | https://doi.org/10.22541/au.165605345.51691863/v1 | This a preprint and has not been peer reviewed. Data may be preliminary.

differing snout morphologies and preferences for certain types of substrate structure (Amen et al. 2020). Specifically, in a choice experiment, the short snouted species (*C. tamandua*) favored a sandy substrate, while the long snouted species (*C. rhynchophorus*) preferred a stone substrate for feeding.

While these trait-specific substrate preferences appear plausible from a mechanical point of view (i.e., longer snouts allow for probing further into interstitial between stones, Amen et al. 2020), there is currently no information available as to whether different trunk shapes are associated with different diets. Furthermore, it is still not known whether the diversification of EOD serves as a prezygotic isolation factor only (Nagel et al. 2018 a,b) or also is related to foraging specialization (Feulner et al. 2009).

Information on diet composition of *Campylomormyrus* is so far limited to only two studies on single species: Roberts & Stewart (1976) reported briefly, based on field observations, some food items found in the stomach of a single specimen assigned to *C. rhynchophorus* . At that time, phylogeny and species delimitation of *Campylomormyrus*was not well established. Nwani et al. (2008) reported the diet composition of *C. tamandua* based on morphological determination. Hence, so far, to the best of our knowledge, there have been no controlled studies which compare the diet composition among*Campylomormyrus* species under natural conditions. Our study aims at contributing to fill that knowledge gap by performing a dietary study for some species of *Campylomormyrus* with markedly different EOD and snout morphologies. Our purpose is not only to document the dietary ranges and components for these species, but also to infer whether species with specific morphological traits prefer specific food items.

Direct observation of feeding in the natural habitat, i.e., the Congo river, seems unfeasible and microscopic examination of gut contents may yield incomplete results, as food items may be digested to various degree, compromising their microscopic identification (Pompanon et al. 2012). Therefore, we used DNA-based dietary analysis using hybrid capture and subsequent next generation sequencing (NGS) to analyze stomach contents of wild-caught fish. NGS-based DNA metabarcoding has been successfully used to investigate the DNA extracted from highly degraded diet samples (Deagle et al. 2006; Jarman et al. 2004), however, an initial PCR amplification necessitates DNA fragments of a certain length to allow both for primer annealing and a large-enough species-specific target sequence. As an alternative, DNA hybridization capture (target enrichment) has been successfully applied to enrich low-concentration and highly degraded DNA-fragments from environmental DNA (Shokralla et al. 2012) or ancient sedimentary DNA (Krueger et al. 2022). Here, we apply a combined hybrid capture/NGS shotgun approach to quantify diet composition in gut contents taken from wild-caught African electric fish to contribute to our understanding of their adaptive radiation.

Materials and Methods

Stomach samples collection

Stomach content samples ($n = 27$) were collected from five*Campylomormyrus* species with markedly different EOD and snout morphologies (*C. alces* , $n = 2$; *C.compressirostris* ,$n = 10$; *C.curvirostris* , $n = 2$;*C. numenius* , $n = 1$; *C. tshokwe* , $n = 9$; see Lamanna et al. 2016 for phylogenetic relationships, EODs, and snout characteristics). Additionally, stomach content samples of*Gnathonemus petersii* (*G. petersii,* $n = 3$), the sister genus of *Campylomormyrus* , were used for comparison. Unlike*Campylomormyrus* , *Gnathonemus petersii* has no snout, but a trunk-like protrusion on the head. The fish specimens from which these stomach content samples are extracted were collected during an expedition to the Republic of the Congo in fall 2012. All the analyzed stomach samples were extracted from sympatric fishes live in the Congo River rapids south of Brazzavilla [S4°18.788' E15deg13.790']. Before stomach extracting, all fish individuals were weighed and measured for total and standard length and were assigned unique and museum identifiers (see Table S1.1 in Supplementary File S1). After dissection, the stomach samples were stored in Queen's lysis buffer solution (Seutin et al.1991). Before DNA extraction, the stomach contents were visually checked and found to be heavily digested, with very few visible tissue parts remaining. Handling of animals was conducted at the

Posted on Authorea 1 Jul 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.165665345.51691863/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

University of Brazzaville in accordance with the relevant national guidelines and regulations guidelines for the care and use of animals for scientific purposes.

Lysis and DNA extraction

Each stomach content sample was homogenized using a Tissue Tip Homogenizer before incubating with lysis reagents. The Qiagen DNeasy Mericon Food Kit for extraction of high-quality DNA from digested foods was used to carry out the DNA extraction from the homogenized samples. Since the stomach content samples were almost digested, the DNA was suspected to be highly fragmented. Therefore, the small fragment protocol of this kit (200 mg) was used according to the manufacturer's protocol. Initial attempts to directly amplify DNA with established DNA barcoding primers were unsuccessful (Lamanna & Tiedemann, unpubl. results).

Libraries preparation, indexing and amplification

After DNA extraction, all samples underwent single-stranded DNA library preparation using T4 DNA ligase following (Gansauge et al. 2017). The procedure is optimized to highly degraded DNA fragments, as expected in digested food from the stomach. An additional step of deoxyuracils removal using Uracil DNA glycolase was included. Deamination of cytosin to uracil is a typical damage pattern in degraded DNA and could cause erroneous mispairing to adenin, resulting in genotyping errors.

After library preparation, the optimal number of Polymerase Chain Reaction (PCR) cycles required for library indexing and amplification was estimated by Quantitative Polymerase Chain Reaction (qPCR). The rationale is to adjust the number of PCR cycles during the library amplification to different DNA concentrations (Basler et al. 2017). The reaction was performed by mixing 1ml of a 1:20 dilution of each library with 3µl of nuclease free water, 5µl of 2 x SYBRTM green qPCR master mix (ThermoFisher Scientic), 0.5µl of qPCR primer IS7, and 0.5µl of primer IS8 (both 10µM; from Gansauge & Meyer 2013). For each library three reaction replicates were analyzed. The PCR reactions were performed with 7min at 95°C, followed by 40 cycles of 10s at 95°C denaturation, 30s annealing at 60°C, and 1min extension at 72°C. Considering the differences in DNA concentration between the qPCR and the indexing PCR, the optimal number of cycles for the indexing PCR was calculated according to Basler et al. (2017).

The indexing primers P5 and P7 were chosen according to the library protocol of Gansauge et al. (2017). The reaction mix for each sample contained 44.8µl of nuclease free water, 8µl of 10 x AccuPrimeTM Pfx reaction mix (ThermoFisher Scientic), 0.8µl of 2.5U AccuPrimeTM Pfx polymerase, 3.2µl of P5 indexing primer, 3.2µl of P7 indexing primer (both 10 µM) and 20µl of each library. Thermocycling started with 2min at 95°C, followed by the number of cycles calculated according to Basler et al. (2017) consisting of 15s at 95°C denaturation, 30s annealing at 60°C and 1 min extension at 68°C. After indexing and amplification, all libraries were purified using the MinElute PCR Purification Kit (QIAGEN) following the manufacturer's instructions. The libraries were quantified with a Qubit 3.0 fluorometer after purification and the distribution of the DNA fragments lengths was measured using an Agilent 2100 bioanalyzer system.

Probe design and DNA hybridization capture

For probe design, we selected target species based on literature research on the potential taxa eaten by *Campylomormyrus* and the expected diversity of prey items in the Congo River region. We targeted 504 species across 6 animal taxa: Arthropoda, Annelida, Mollusca, Nemertea, Nematoda, and Rotatoria. Additionally, we targeted 29 species from 3 plant taxa: Streptophyta, Rhodophyta, and Chlorophyta. For these species, mitochondrial COI gene (animals) resp. chloroplast rbcL gene (plants) sequences were retrieved from the GenBank sequence database of the National Center for Biotechnology Information (NCBI). When a species of interest did not have sequence information available, sequence information from another species of the same genus was included. If this was not available, the taxon was excluded. The average sequence length of

Posted on Authorea 1 Jul 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.165605345.51691863/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

all collected sequences was 700 bp. A complete list of taxonomic names and NCBI accession numbers can be found in the Supplementary File S2. The sequence capture probes were custom-made at Arbor Biosciences (formerly MY croarray) and supplied in form of a MYBaits InSolution Custom Target Capture Kit (designs with 1-20k probes). The probes were designed to achieve a high overall probe count and an increased coverage in high GCcontent regions. The final probe set contained 15284 probes, each with a length of 80bp. The synthesized probes comprise short overlapping fragments representing the whole probe template sequences of COI resp. rbcL.

The hybridization capture enrichment reactions were performed according to the manufacturer's instructions (Biosciences 2016). Following these instructions, the recommended input of the library DNA was 100ng to 500ng. This was equivalent to 7µl of the library material with a DNA concentration of 14 to 72ng/µl. All materials from the library were used for two rounds of hybridization capture reactions, thereby increasing specificity (Krueger et al. 2021). The captured and amplified libraries of the first round of capture reaction were used as input libraries for the second round. For both rounds of capture, the libraries were not pooled since pooling might produce low quality results (Biosciences 2016). The second captured and amplified libraries were pooled and underwent sequencing on an Illumina MiSeq instrument using a protocol for 2x150bp paired end of double indexed libraries (Biosciences 2016).

Bioinformatics, data analysis, and statistics

Before further analyzing the high throughput sequencing reads, which were generated in FASTQ format, we performed a quality processing to remove the adapter sequences and the low-quality reads, using Cutadapt version 2.10 (Martin 2011). Quality cutoff, minimal sequence length, and minimal overlap between sequence and adapter were set to 20, 30, and 3, respectively. Then, we performed quality control checks to the filtered sequences using FastQC (Andrews 2010) to ensure their suitability for further analysis.

Taxonomic classification for the filtered sequences was done using Kraken version 2.0.9 (Kraken2; Wood et al. 2019). Kraken compares the reads of the metagenome to short sequences of length k, the so-called $k$mers, from a database that is associated with the sequence information underlying the tree of life phylogeny. The algorithm then places the read on the tree of life based on its similarities to these $k$mers. Before applying Kraken2, a custom marker database of available eukaryotic COI and RBCl sequences from NCBI was built to map the reads. The results of Kraken2 were visualized and inspected using the online tool Pavian metagenomics data explorer (Breitwieser & Salzberg, 2019).

The read counts were used to estimate the relative abundance of food taxa in the diet of each sample. For any sample, these count data were used to record both the occurrence (presence/absence) of a taxon and the percentage of reads assigned to that taxon. The frequency of occurrence (%FOO) is used to quantify occurrence across samples. The relative read abundance (RRA) is used as a proxy for relative biomass consumed. Details on the calculations of these terms are given in the Supplementary File S7.

The dietary niche width was assessed by calculating the Shannon diversity index (alpha diversity) for each sample using the *spaa*R package (Zhang 2016). Further, we used two complementary metrics to describe the diet overlap among species: Schöner index (Schoener 1970) and Pianka index (Pianka 1974). Pairwise calculations of these indices were performed using the *spaa* R package (Zhang 2016). The equations of the indices are given in the Supplementary File S7. We used the *EcoSimR* R package (Gotelli et al. 2015) to compare the results with reference to 1000 permutations of a null model that holds the same dietary niche width, while randomizing the values for the diet items. We further calculated Bray–Curtis dissimilarity indices (beta diversity) between each pair of samples using the*vegdist* -function in the *vegan* R package (Oksanen et al. 2019). Additionally, we performed nonmetric multidimensional scaling (NMDS) using the function *metaMDS* in the *vegan* R package to visualize the patterns of dietary dissimilarity among samples. We tested for differences in diet composition among species by performing a permutational multivariate analysis of variance (perMANOVA) with 999 permutations, using the *adonis* -function in the *vegan* R package. To further test for the influence of our two divergent phenotypic characters EOD and snout morphology on

Posted on Authorea 1 Jul 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.165605345.51691863/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

the food selection, we grouped the stomach samples as follows: Based on the EOD, we formed two groups: Long EOD (>2ms) and short EOD (<0.3ms; note the almost 10fold difference in EOD length among the two groups). The first group ($n = 12$) included the samples of *C. alces* ,*C. numenius* , and *C. tshokwe* . The second group ($n = 15$) included the samples of *C. compressirostris* , *C. curvirostris* , and *G. petersii* . EOD categorization was based on Lamanna et al. (2016). Further, the samples were grouped based on the snout length into three groups: Long snout (*C. curvirostris* ,*C. tshokwe* , and *C. numenius* , $n = 12$, medium snout (*C. compressirostris* , $n = 10$), and short snout (*C. alces* and *G. petersii* , $n = 5$). For this classification, we took advantage of a previous geometric morphometrics analysis, which stratified the respective species according to snout length along principal component 1 (Figure 3 in Lamanna et al. 2016). Differences in diet among groups were assessed with a perMANOVA as above.

Results

Sequencing results

The sequencing for all samples generated 6.9Gb raw sequencing data of FASTQ formatted reads. The raw data have been deposited to the open repository Zenodo (Amen et al. 2022). The total number of raw sequences from the 27 stomach samples is 19531074, ranging from 19226 to 867093 per individual sample. Out of them, a total of 18357160 sequences were obtained after quality filtering, ranging from 16703 to 828735 per individual.

Taxonomic classification

The taxonomic classification of the filtered sequences against the custom marker database using Kraken2 classified all the reads into seven phylogenetic levels (domain, kingdom, phylum, class, order, family, genus, and species) or unclassified reads. The percentage of the unclassified reads reached 33.22 % of the pooled sequences of all samples. On the individual level, the percentage of unclassified reads ranged from 8.38 % to 71.67 %. The high percentage of unclassified reads is attributed to sequences outside the range of our custom marker database (eukaryotic COI and rbcL only). When the entire nucleotide database (www.ncbi.nlm.nih.gov/nucleotide/) is used instead of our custom database, the percentage of unclassified reads dropped to 10.11 % of the pooled sequences of all samples. We limited the analysis to the taxonomic classifications using the eukaryote COI/rbcL custom database, however, to keep the focus on the diet analysis (instead of prokaryotes in the gut microbiomes).

The classified reads pooled for each species are visualized using the metagenomics data explorer Pavian and depicted in the supplement (Supplementary File S3). Comparing the distributions of the reads for the food taxa classified at each phylogenetic level reveals that most of the reads are assigned to few food taxa, while many of the classified taxa are represented by only a few reads. We excluded records from primates and birds, which were presumably contamination. Taxa represented by less than 0.01 % of reads were further excluded, as they may constitute contamination/background noise (Alberdi et al. 2018). The remaining identified prey items were assigned to 146 taxa of different taxonomic level: 20 classes, 34 orders, 97 families, 105 genera, and 90 species (Tables S4.1-S4.6 Supplementary File S4).

Relative abundance

Campylomormyrus had a broad spectrum of prey items. On the class level, Insecta dominated by far. They were found in all samples, representing more than 90% of the total reads (Figure 1). Other classes were also found in all samples but with less percentage of the total reads, such as Clitellata, Arachnida, Malacostraca, and Hexanauplia. At the order level, the most abundant prey items were Diptera, Coleoptera, and Hymenoptera (all Insecta) class, all of which were in all samples (Figure 1). Also, the Haplotaxida (Clitellata) and Araneae (Arachnida) were found in all samples. Further insect orders such as Lepidoptera,

Posted on Authorea 1 Jul 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.165665345.51691863/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Trichoptera, Ephemeroptera, and Hemiptera, as well as Decapoda (Crustacea) were found in more than 83 % of the samples.

The Relative Read Abundances (RRA) of the primary food taxa for the studied species are depicted in Figure 2. Most species share the same food taxa, albeit in different proportions. The most dominant prey taxon at class level is Insecta, followed by Clitellata. At the order level, the most abundant orders are all insects, i.e., Diptera, Coleoptera, Hymenoptera, and Lepidoptera.

The RRA data were used to assess the dietary niche width by calculating the Shannon diversity index (Figure 3 and Table S5.1 in the Supplementary File S5). The Shannon diversity index differed considerably amongthe studied species, pointing towards different dietary niche width. As an example, the lowest Shannon index values are found for *C. compressirostris* samples (1.62±0.02 SE), while the highest are found for *C. tshokwe* (2.21±0.03 SE).

Diet overlap

The diets of most of species significantly overlap at class and order phylogenetic levels (Table 1): At these levels, Pianka index values showed statistically significant niche overlap based on comparison with 1,000 null models (see the Supplementary File S8 for more details) and Schöner index scored more than 0.6 for all species comparisons except when compared to *C. numenius* . At lower taxonomic level (family, genus, species), there was much less overlap in the diet among species. The degree of diet overlap is further confirmed by the Bray–Curtis dissimilarity index (0: similar; 1: dissimilar), as shown in Table S5.2 in the Supplementary File S5.

As many reads could not be assigned to family, genus, or species level, we performed further statistical analyses on read assignments at the order level. A perMANOVA on the Bray–Curtis dissimilarity index data derived from the RRA values at order level indicates significant dietary differences among species (F =3.596,r^2 =0.275, p [?]0.001), excluding the species with small sample size (*C. alces* , *C. curvirostris* , and *C. numenius*). We performed a post hoc pairwise perMANOVA using a Bonferroni correction of the p -values to find out which species drive these results. *C. compressirostris* versus *C. tshokwe* and *G. petersii* versus *C. tshokwe* were statistically significantly different in their diet (p <0.05), while *C. compressirostris* and *G. petersii* did not show a significant difference.

A perMANOVA on the EOD groups (long *vs.* short EOD) indicated significant dietary differences between the two groups (F =8.5,r^2 =0.254, p [?]0.001, Bray–Curtis dissimilarity index data derived from the RRA values at order level). Similarly, for snout morphology (short *vs.* medium *vs.*long), a perMANOVA on the Bray–Curtis dissimilarity index data (derived from the RRA values at order level) showed significant dietary differences among the three groups (F =2.676,r^2 =0.182, p [?]0.001). The post hoc pairwise perMANOVA using a Bonferroni correction of the p -values indicated that only the long snout versus the medium snout is statistically significant (p<0.002).

The patterns of dietary difference among samples are visualized in Figure 4 by ordinating the Bray–Curtis dissimilarity index values in two dimensions using NMDS. The stress level for the NMDS was 0.133, which indicates a good representation (Clarke 1993). The NMDS plot shows the segregation of samples based on EOD and the degree of diet overlap/dissimilarity based on species and snout length.

Discussion

Diet profile

The DNA metabarcoding method provided a clear taxonomic resolution of the partially digested stomach content and potentially detected highly diverse food taxa. This high taxonomic resolution is in line with those of previous dietary studies based on DNA-based approaches (e.g., Harms-Tuohy et al. 2016;

Posted on Authorea 1 Jul 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.165665345.51691863/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Pan et al. 2021; Rees et al. 2020; Sakaguchi et al. 2017). The results of this study indicate that the diet of *Campylomormyrus* and *Gnathonemus* species is composed mainly of three types of prey items, i.e., benthic invertebrates, allochthonous invertebrates, and macrophyte material. The most dominant prey taxa found in the gut contents of these species belong to benthic invertebrates, especially aquatic insects. In particular, dipterans (Chironomidae, Simuliidae, Drosophilidae, and Tephritidae), coleopterans (Zopheridae, Carabidae, Histeridae, and Scarabaeidae), trichopterans (Hydropsychidae, Lycaenidae, and Elateridae), ephemeropterans (Leptophlebiidae, Baetidae, and Ephemerellidae), and odonatans (Coenagrionidae, Chlorocyphidae, and Chlorogomphidae) are important diet constituents of all the species. The larvae of these insects live usually in holes and interstitial spaces of the riverbed. Although the DNA approach used here cannot tell the stage of the aquatic insects found in the diet, it is reasonable to assume that the aquatic larvae, rather than the terrestrial imagines, were targeted by *Campylomormyrus* and *Gnathonemus* .

Beside aquatic insects, annelid worms (such as Glossoscolecidae, Naididae, and Megascolecidae) were also found in the diet of all *Campylomormyrus* and *Gnathonemus* species in this study. Similar to the insects' larvae, the annelid worms hide in mud and among aquatic vegetation in the substrate of the riverbed. Other benthic invertebrates found were freshwater snails (Gastropoda, orders Pachychilidae and Stylommatophora), and crustaceans (Malacostraca, orders Decapoda, Copepoda, Cladocera, and Amphipoda).

The second group of food items found in the diet of *Campylomormyrus* and *Gnathonemus* species is allochthonous invertebrates. The most abundant prey taxa from this group are Hymenoptera (including Formicidae, Mymaridae, and Braconidae) and Lepidoptera (including Nymphalidae, Lycaenidae, and Hepialidae). Additionally, Araneae (Arachnida) were frequently found in the diet.

The third group of food items is plants, including grasses, such as Poaceae of the Poales order, and flowering plants, such as Fabaceae and Asterales.

It must be noted that we cannot exclude some of these taxa having derived from the diet of the primary prey (secondary predation; Sheppard et al. 2005) or comprise small organisms and plant debris unintentionally ingested during grasp suction. However, the stomach contents of *Campylomormyrus* and *Gnathonemus* species found in this study using a DNA metabarcoding approach are compatible with *Campylomormyrus* mainly (about 90 %) feeding on aquatic insects (Nwani et al., 2008; Roberts & Stewart, 1976). A previous study, based on morphological observation, reported that stomach contents of some *Campylomormyrus* species contain larvae of chironomids, Povilla, trichopteran, ephemeropterans and odonates, dead plant debris, and decomposing animal debris (Roberts & Stewart, 1976). This study reported also that the stomach content of a specimen of *C. rhynchophorus* had Chironomidae, Simuliidae, and trichopterans, and a few small ephemeropterans. Another study on the stomach contents of *C. tamandua* using morphological observation reported similar food taxa (Nwani et al., 2008). The few available dietary studies on other fish species inhabiting the Congo River such as *Schilbe intermedius* (Dirat et al. 2019) and *Distichodus antonii* ,*D. affinis* and *D. lusosso* (Zebe et al. 2010) showed similar prey spectra.

Diet comparison among species

The dietary compositions of the *Campylomormyrus* and *Gnathonemus* species suggest that their feeding behavior is opportunistic, hence depending on food availability and accessibility. Based on our results, all species exploit diverse food niches in their habitats. For instance, the prevalence of benthic invertebrates, such as larvae of dipterans and coleopterans and annelid worms, in the diet of all species suggests that these species exploit the bottom of the riverbed, while the occurrence of allochthonous invertebrates, such as Formicidae, Nymphalidae and Arachnida spiders, may indicate a certain degree of surface feeding. Further, the diets also include food items from the water column, such as copepods. Accordingly, these species may exhibit high trophic flexibility and diverse feeding behaviors.

However, the RRA results revealed significant differences among the dietary compositions among the species, potentially associated with EOD and snout length. For example, the diet of some species, such as *C. compressirostris* , *C. curvirostris* , and G.*petersii* , contains more dipterans, while the diet of *C. tshokwe* contains

Posted on Authorea 1 Jul 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.165665345.51691863/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

more coleopterans. The diet of *C. numenius* was exceptional, as it contained large amount of grass. Unfortunately, this latter finding remains anecdotal, as we had only one sample of this species available.

Species with long EOD had preferentially fed on other taxonomic groups (i.e., coleopterans, ephemeropterans, spiders, annelids), compared to the species with short EOD, where dipterans dominated in the diet (Fig. S6.1 in the Supplementary File S6). Similar differences were found according to snout length (Fig. S6.2 in the Supplementary File S6). Note, however, that species with long EOD often also exhibit a long snout (Fig. 4; cf. Lamanna et al. 2016).

Dietary analysis and the radiation scenario

Our study was motivated by potentially providing further support for the hypothesis that radiation of *Campylomormyrus* is caused by an adaptation to different food sources, associated with diversification of EOD. The results proved that all the species tested in the current study are able to exploit diverse niches, especially the bottom fauna, regardless of their snout shape and EOD.

The current study still provides some evidence that different snout morphologies and the associated divergence in the EOD translate into different prey spectra. As a different morphology of the feeding apparatus among the species may constitute a functional adaptation to exploiting different substrates (Amen et al. 2020), the different prey spectra could simply reflect differential availability of different prey in the respective microhabitats. Unfortunately, no data on substrate-specific benthofauna are available for the Congo River. However, it has been argued that the duration of the EOD plays a crucial role during food detection by determining the prey items that can be detected best (Harlan Meyer 1982). In fact, prey detection via electrolocation is a frequency-based process. Our results cannot establish a causal link between EOD characteristics and prey spectra, yet they are compatible with the hypothesis that the divergence in EOD could be of adaptive value during feeding as well, beside its proven function as a prezygotic isolation mechanism. In this case, EOD would be a 'magic trait' triggering both adaptation and reproductive isolation (Feulner et al. 2009). One approach to further investigate a potential dual function of the EOD divergence (feeding specialization and reproductive isolation) would be to expand on the choice experiments performed by Amen et al. (2020) by presenting a variety of food items. This will though still be confounded by the association of long EOD with long snout. Ultimately, one would like to know prey-specific detection probabilities, relative to the physical properties of the different discharges. To achieve such knowledge remains a challenge.

In general, the snout morphology of *Campylomormyrus* may enhance their grasp suction mode of feeding (Marrero & Winemiller 1993). This enhanced mode enables them to obtain the aquatic insects that burrow into the different substrate structures. This functional foraging specialization allow an efficient exploitation of the rich bottom fauna of benthic invertebrates, especially aquatic insect larvae, which other fishes may not reach. Such trophic specializations may have triggered the observed radiation of *Campylomormyrus* in the Congo river.

Acknowledgements

We thank Dr. Francesco Lamanna and Dr. Christiane Paul for participation in sample collection in Congo/Brazzaville and Dr. Victor Mamonekene, University of Brazzaville, for logistical support during field work. Support from Katrin Kiemel in data analysis and manuscript revision is also acknowledged. Financial support for this study was provided by the University of Potsdam. Open Access funding is provided by the DEAL project. R. Amen was funded by the Bridge Scholarship Program of the University of Potsdam and from the Ministry of Higher Education of the Arab Republic of Egypt.

Posted on Authorea 1 Jul 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.165665345.51691863/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

References

Alberdi, A., Aizpurua, O., Gilbert, M. T. P., & Bohmann, K. (2018). Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution* , *9* (1), 134–147. doi: 10.1111/2041-210X.12849

Amen, R., Havenstein, K., Kirschbaum, F., & Tiedemann, R. (2022). Raw sequencing data of NGS of stomach contents of sympatric species of weakly electric fish (genus: Campylomormyrus) [Data set].*Zenodo* . doi: 10.5281/zenodo.6569845

Amen, R., Nagel, R., Hedt, M., Kirschbaum, F., & Tiedemann, R. (2020). Morphological differentiation in African weakly electric fish (genus Campylomormyrus) relates to substrate preferences. *Evolutionary Ecology* , *34* (3), 427–437. doi: 10.1007/s10682-020-10043-3

Andrews, S. (2010). *FastQC* . Retrieved from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Basler, N., Xenikoudakis, G., Westbury, M. V., Song, L., Sheng, G., & Barlow, A. (2017). Reduction of the contaminant fraction of DNA obtained from an ancient giant panda bone. *BMC Research Notes* ,*10* (1), 754. doi: 10.1186/s13104-017-3061-3

Biosciences, A. (2016). *Mybaits: In-Solution Sequence Capture for Targeted High-Throughput Sequencing User Manual* .

Breitwieser, F. P., & Salzberg, S. L. (2019). Pavian: interactive analysis of metagenomics data for microbiome studies and pathogen identification. *Bioinformatics* , *36* (4), 1303–1304. doi: 10.1093/bioinformatics/btz715

Canitz, J., Kirschbaum, F., & Tiedemann, R. (2020). Transcriptome-wide single nucleotide polymorphisms related to electric organ discharge differentiation among African weakly electric fish species. *PLOS ONE* , *15* (10), e0240812. doi: 10.1371/journal.pone.0240812

Clarke, K. R. (1993). Non-parametric multivariate analyses of changes in community structure. *Austral Ecology* , *18* (1), 117–143. doi: 10.1111/j.1442-9993.1993.tb00438.x

Deagle, B. E., Eveson, J. P., & Jarman, S. N. (2006). Quantification of damage in DNA recovered from highly degraded samples - a case study on DNA in faeces. *Frontiers in Zoology* , *3* (1). doi: 10.1186/1742-9994-3-11

Dirat, I. M.-G., Obath, B. R. C. O., Tsoumou, A., Mikia, M., & Vouidibio, J. (2019). Food Habits of Schilbe Intermedius, Ruppel, 1832 (Siluriformes: Schilbeidae) of Right Bank of Pool Malebo (Congo River).*International Journal of Research Studies in Zoology* ,*5* (1). doi: 10.20431/2454-941X.0501001

Feulner, P. G. D., Kirschbaum, F., Mamonekene, V., Ketmaier, V., & Tiedemann, R. (2007). Adaptive radiation in African weakly electric fish (Teleostei: Mormyridae: Campylomormyrus): a combined molecular and morphological approach. *Journal of Evolutionary Biology* ,*20* (1), 403–414. doi: 10.1111/j.1420-9101.2006.01181.x

Feulner, P. G.D., Kirschbaum, F., Schugardt, C., Ketmaier, V., & Tiedemann, R. (2006). Electrophysiological and molecular genetic evidence for sympatrically occuring cryptic species in African weakly electric fishes (Teleostei: Mormyridae: Campylomormyrus).*Molecular Phylogenetics and Evolution* , *39* (1), 198–208. doi: 10.1016/j.ympev.2005.09.008

Feulner, P. G.D., Kirschbaum, F., & Tiedemann, R. (2008). Adaptive radiation in the Congo River: An ecological speciation scenario for African weakly electric fish (Teleostei; Mormyridae; Campylomormyrus).*Journal of Physiology-Paris* , *102* (4–6), 340–346. doi: 10.1016/j.jphysparis.2008.10.002

Feulner, P. G. D., Plath, M., Engelmann, J., Kirschbaum, F., Tiedemann, R. (2009). Magic trait electric organ discharge (EOD) - Dual function of electric signals promotes speciation in African weakly electric fish.*Communicative & Integrative Biology* **2** (4), 1-3.

Posted on Authorea 1 Jul 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.165665345.51691863/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Gansauge, M.-T., & Meyer, M. (2013). Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. Nature Protocols, 8(4), 737–748. doi: 10.1038/nprot.2013.038

Gansauge, M.-T., Gerber, T., Glocke, I., Korlević, P., Lippik, L., Nagel, S., ... Meyer, M. (2017). Single-stranded DNA library preparation from highly degraded DNA using T4 DNA ligase. *Nucleic Acids Research* , gkx033. doi: 10.1093/nar/gkx033

Gotelli, N., Hart, E., & Ellison, A. (2015). EcoSimR: Null Model Analysis for Ecological Data. doi: 10.1016/0304-3800(94)00117

Harlan Meyer, J. (1982). Behavioral responses of weakly electric fish to complex impedances. *Journal of Comparative Physiology* ,*145* (4), 459–470. doi: 10.1007/BF00612811

Harms-Tuohy, C., Schizas, N., & Appeldoorn, R. (2016). Use of DNA metabarcoding for stomach content analysis in the invasive lionfish Pterois volitans in Puerto Rico. *Marine Ecology Progress Series* ,*558* , 181–191. doi: 10.3354/meps11738

Jarman, S. N., Deagle, B. E., & Gales, N. J. (2004). Group-specific polymerase chain reaction for DNA-based analysis of species diversity and identity in dietary samples. *Molecular Ecology* , *13* (5), 1313–1322. doi: 10.1111/j.1365-294X.2004.02109.x

Krueger, J., Foerster, V., Trauth, M. H., Hofreiter, M., & Tiedemann, R. (2021). Exploring the Past Biosphere of Chew Bahir/Southern Ethiopia: Cross-Species Hybridization Capture of Ancient Sedimentary DNA from a Deep Drill Core. *Frontiers in Earth Science* , *9* , 775. doi: 10.3389/feart.2021.683010

Lamanna, F., Kirschbaum, F., Ernst, A. R. R., Feulner, P. G. D., Mamonekene, V., Paul, C., & Tiedemann, R. (2016). Species delimitation and phylogenetic relationships in a genus of African weakly-electric fishes (Osteoglossiformes, Mormyridae, Campylomormyrus). *Molecular Phylogenetics and Evolution* , *101* , 8–18. doi: 10.1016/j.ympev.2016.04.035

Lavoué, S., Sullivan, J. P., & Hopkins, C. D. (2003). Phylogenetic utility of the first two introns of the S7 ribosomal protein gene in African electric fishes (Mormyroidea: Teleostei) and congruence with other molecular markers. *Biological Journal of the Linnean Society* , *78* (2), 273–292. doi: 10.1046/j.1095-8312.2003.00170.x

Marrero, C., & Winemiller, K. O. (1993). Tube-snouted gymnotiform and mormyriform fishes: convergence of a specialized foraging mode in teleosts. *Environmental Biology of Fishes* , *38* (4), 299–309. doi: 10.1007/BF00007523

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal* , *17* (1), 10. doi: 10.14806/ej.17.1.200

Nagel, R., Kirschbaum, F., Engelmann, J., Hofmann, V., Pawelzik, F., & Tiedemann, R. (2018a). Male-mediated species recognition among African weakly electric fishes. *Royal Society Open Science* , *5* (2), 170443. doi: 10.1098/rsos.170443

Nagel, R., Kirschbaum, F., Hofmann, V., Engelmann, J., & Tiedemann, R. (2018b). Electric pulse characteristics can enable species recognition in African weakly electric fish species. *Scientific Reports* ,*8* (1), 10799. doi: 10.1038/s41598-018-29132-z

Nwani, C. D., Eyo, J. E., & Udeh, E. F. (2008). Food and Feeding Habits of Campylomormyrus tamandua in Anambra River, Nigeria. *Animal Research International* , *3* (1). doi: 10.4314/ari.v3i1.40760

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., ... Others. (2019). vegan: Community Ecology Package.*R Package Version 2.5-6* .

Pan, W., Qin, C., Zuo, T., Yu, G., Zhu, W., Ma, H., & Xi, S. (2021). Is Metagenomic Analysis an Effective Way to Analyze Fish Feeding Habits? A Case of the Yellowfin Sea Bream Acanthopagrus latus (Houttuyn) in Daya Bay. *Frontiers in Marine Science* , *8* . doi: 10.3389/fmars.2021.634651

Posted on Authorea 1 Jul 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.165665345.51691863/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Pianka, E. R. (1974). Niche Overlap and Diffuse Competition.*Proceedings of the National Academy of Sciences* , *71* (5), 2141–2145. doi: 10.1073/pnas.71.5.2141

Pompanon, F., Deagle, B. E., Symondson, W. O. C., Brown, D. S., Jarman, S. N., & Taberlet, P. (2012). Who is eating what: diet assessment using next generation sequencing. *Molecular Ecology* , *21* (8), 1931–1950. doi: 10.1111/j.1365-294X.2011.05403.x

Rees, G. N., Shackleton, M. E., Watson, G. O., Dwyer, G. K., & Stoffels, R. J. (2020). Metabarcoding demonstrates dietary niche partitioning in two coexisting blackfish species. *Marine and Freshwater Research* , *71* (4), 512. doi: 10.1071/MF18491

Roberts, T. R., & Stewart, D. J. (1976). An ecological and systematic survey of fishes in the rapids of the Zaire or Congo River.*Bulletin of the Museum of Comparative Zoology* , *147* (6), 239–317.

Sakaguchi, S. O., Shimamura, S., Shimizu, Y., Ogawa, G., Yamada, Y., Shimizu, K., ... Takishita, K. (2017). Comparison of morphological and DNA-based techniques for stomach content analyses in juvenile chum salmon Oncorhynchus keta: a case study on diet richness of juvenile fishes. *Fisheries Science* , *83* (1), 47–56. doi: 10.1007/s12562-016-1040-6

Schoener, T. W. (1970). Nonsynchronous Spatial Overlap of Lizards in Patchy Habitats. *Ecology* , *51* (3), 408–418. doi: 10.2307/1935376

Seutin, G., White, B. N., & Boag, P. T. (1991). Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology* , *69* (1), 82–90. doi: 10.1139/z91-013

Sheppard, S. K., Bell, J., Sunderland, K. D., Enlon, J., Skervin, D., & Symondson, W. O. C. (2005). Detection of secondary predation by PCR analyses of the gut contents of invertebrate generalist predators.*Molecular Ecology* , *14* (14), 4461–4468. doi: 10.1111/j.1365-294X.2005.02742.x

Shokralla, S., Spall, J. L., Gibson, J. F., & Hajibabaei, M. (2012). Next-generation sequencing technologies for environmental DNA research.*Molecular Ecology* , *21* (8), 1794–1805. doi: 10.1111/j.1365-294X.2012.05538.x

Tiedemann, R., Feulner, P. G. D., & Kirschbaum, F. (2010). Electric Organ Discharge Divergence Promotes Ecological Speciation in Sympatrically Occurring African Weakly Electric Fish (Campylomormyrus). In *Evolution in Action* (pp. 307–321). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-12425-9_15

Wood, D. E., Lu, J., & Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biology* , *20* (1), 257. doi: 10.1186/s13059-019-1891-0

Zebe, V., Nseu, B., & Micha, J.-C. (2010). Diet comparison in three species of *Distichodus* Müller & Troschel, 1845 (Distichodontidae, Pisces) of Pool Malebo, Congo River (Democratic Republic of Congo). *International Journal of Biological and Chemical Sciences* , *4* (1). doi: 10.4314/ijbcs.v4i1.54240

Zhang, J. (2016). *Spaa: SPecies association analysis* . Retrieved from https://cran.r-project.org/web/packages/spaa/spaa.pdf

Data Accessibility and Benefit-Sharing Section

Data Accessibility Statement

Raw sequence reads are deposited in the open repository Zenodo (Amen et al., 2022). The codes used to analyze the data and generate the results can be found on GitHub repository DATMetaB:

https://github.com/RahmaAmen/DATMetaB

Benefit-Sharing Statement

Posted on Authorea 1 Jul 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.165665345.51691863/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

Benefits from this research arise from the sharing of our data and results on public databases as described above as well as the codes developed to analyze the data.

Author Contributions

RA participated in the design of the study, carried out the experiments, analyzed the data and drafted the manuscript; KH participated in carrying out the experiments and revised the manuscript; RT and FK conceived the study, organized and participated in sample collection, participated in the design of the study, revised the manuscript, and coordinated the study. All authors have read and approved the manuscript for publication.

Tables and Figures

Tables

Table 1: Schöner index (bottom left) and Pianke index (top right) for diet overlap (0: no diet overlap, 1: complete diet overlap) of 5 species of Campylomormyrus and one species of its sister genus Gnathonemus based on RRA data considering food taxa at different phylogenetic levels based on pooled samples

Species	*C. alces*	*C. compres.*	*C. curvirost.*
Index* at class/order level	Index* at class/order level	Index* at class/order level	Index* at class/order level
C. alces		**0.99**/**0.69**	**0.99**/**0.74**
C. compres.	**0.93**/**0.60**		**0.99**/**0.98**
C. curvirost.	**0.97**/**0.64**	**0.92**/**0.80**	
C. numenius	0.37/0.37	0.36/0.36	0.37/0.35
C. tshokwe	**0.97**/**0.69**	**0.92**/**0.61**	**0.98**/**0.75**
G. petersii	**0.95**/0.56	**0.93**/**0.88**	**0.96**/**0.81**
Index* at family/genus level	Index* at family/genus level	Index* at family/genus level	Index* at family/genus level
C. alces		**0.52**/**0.47**	0.49/**0.42**
C. compres.	0.46/0.36		**0.93**/**0.40**
C. curvirost.	0.43/0.31	**0.66**/0.32	
C. numenius	0.14/0.11	0.18/0.10	0.19/0.09
C. tshokwe	0.52/0.48	0.44/0.34	0.59/0.48
G. petersii	0.24/0.03	0.55/0.09	0.50/0.04

*Bold numbers indicate statistically significant niche overlap (i.e., Pianke index is greater than that expected by chance based on comparison with 1,000 null models, $\alpha = 0.05$, and Schöner index is greater than 0.6).

Figures

Posted on Authorea 1 Jul 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.165665345.51691863/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

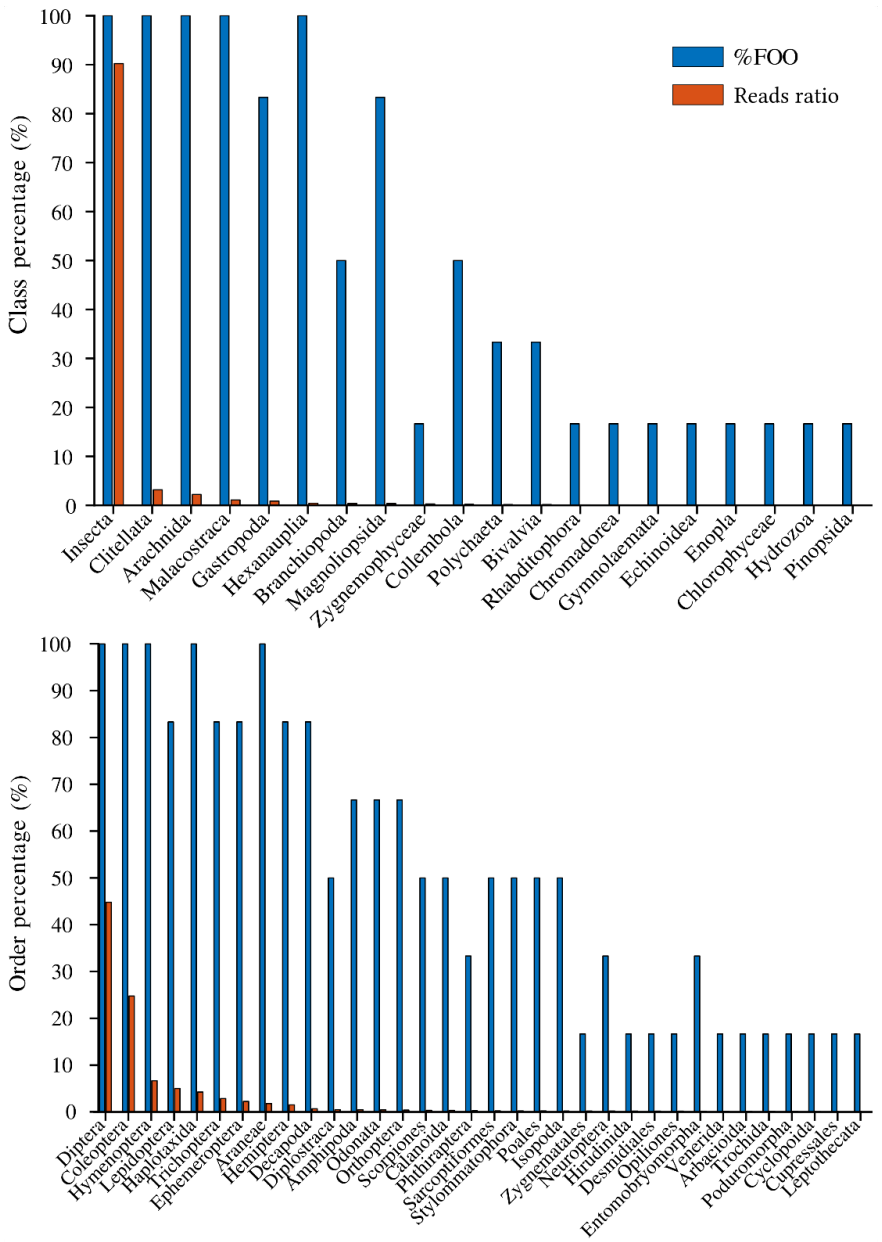

Figure 1: Overall diet composition among all samples (5 species of Campylomormyrus and one species of Gnathonemus) based on frequency of occurrence (%FOO, shown in blue bars) at class (upper panel) and order (lower panel) levels. Taxa are ordered based on the reads percentage abundance (shown in orange bars). Only primary prey items are considered.

Posted on Authorea 1 Jul 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.165665345.51691863/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

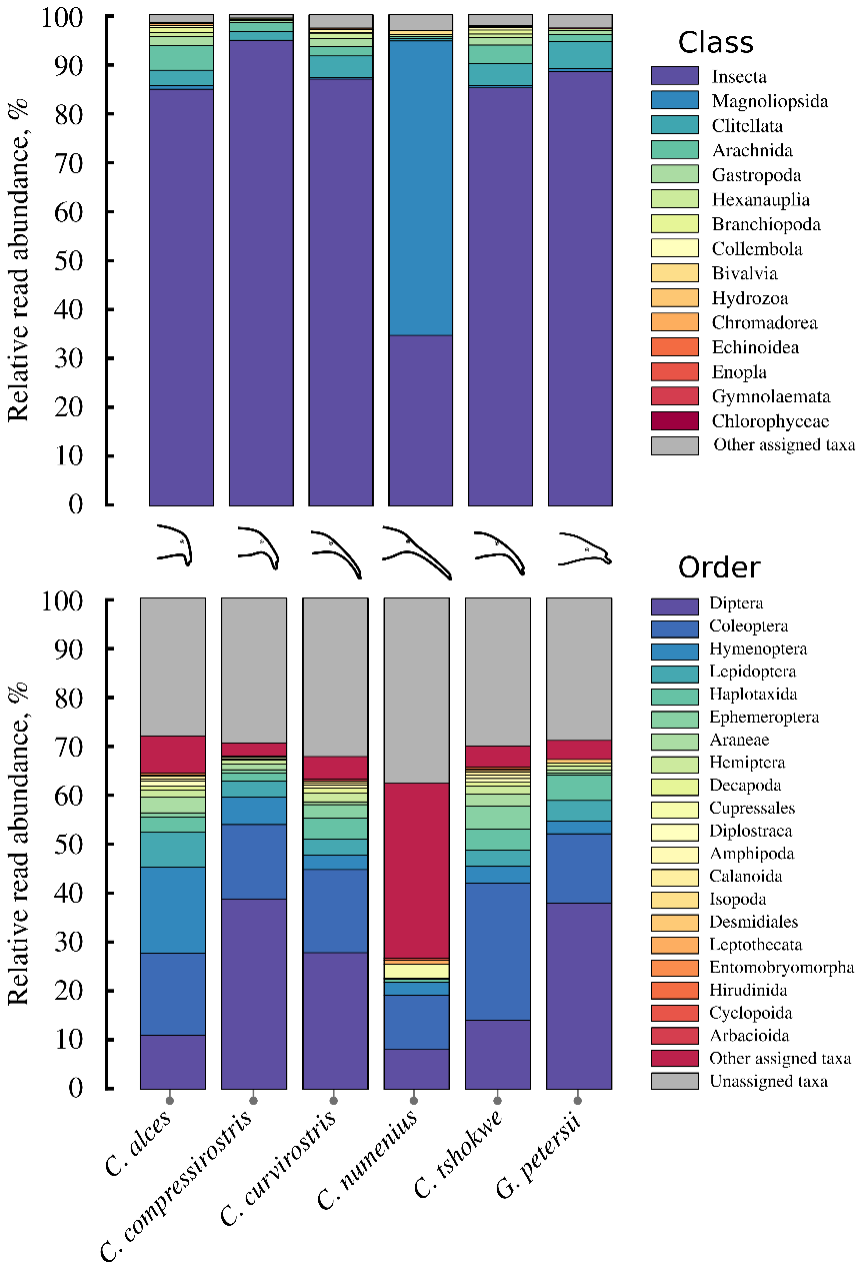

Figure 2: Relative read abundance (RRA) of food taxa from dietary metabarcoding data sets of stomach contents of 5 species of Campylomormyrus and one species of its sister genus Gnathomemus at class level (upper panel) and order level (lower panel). Fish pictograms provide a sketch on species-specific snout morphology.

Posted on Authorea 1 Jul 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.165665345.51691863/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.

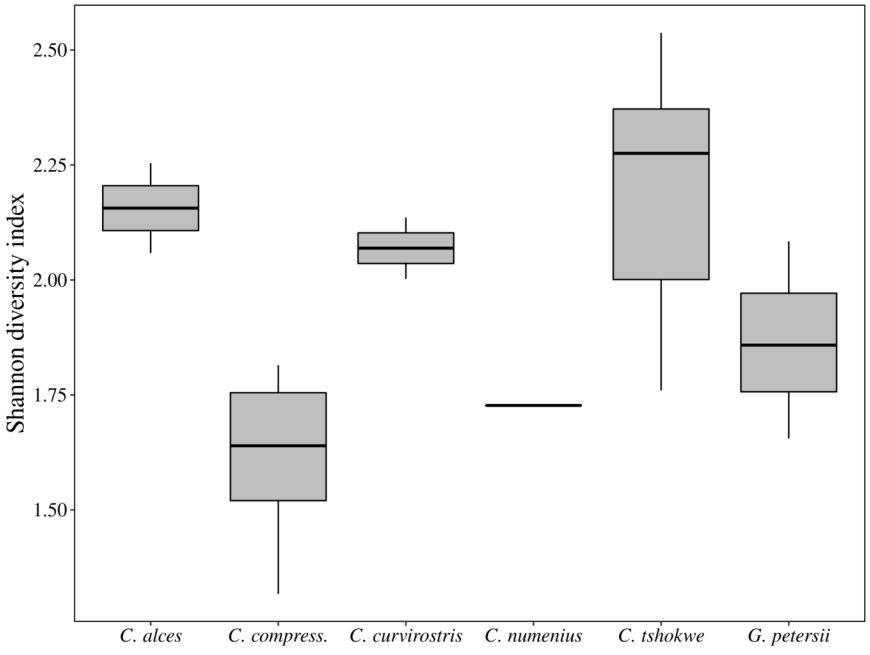

Figure 3: Boxplot of Shannon diversity indices of 5 species of Campylomormyrus and one species of its sister genus Gnathonemus (based on RRA at order level).

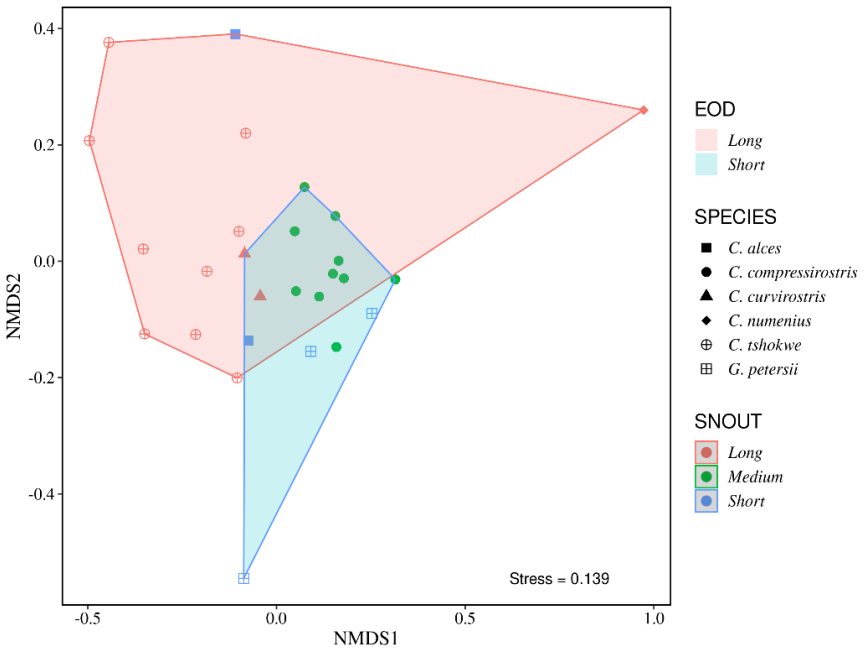

Figure 4: Dietary niche partitioning within and among 5 species of *Campylomormyrus* and one species of its sister genus *Gnathomemus* by nMDS of RRA-based Bray–Curtis dissimilarity indices of samples on the order level (adonis F = 3.04, R^2 = 0.243, p [?] .001). The symbols represent the species and the colors

represent snout length. The shaded convex hulls indicate short vs. long EOD. The stress level of 0.133 is under the cut-off value of 0.2 as recommended by Clarke (1993) to indicate an interpretable ordination.

Posted on Authorea 1 Jul 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.165665345.51691863/v1 — This a preprint and has not been peer reviewed. Data may be preliminary.