Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

“Unraveling the Gut Microbiota Composition of the Genus Herichthys (Cichliformes:Pisces)”

Omar Mejía (hmejiag@ipn.mx)
Instituto Politecnico Nacional https://orcid.org/0000-0002-7021-8377
Andrés Sánchez Quinto
Universidad Nacional Autónoma de México Instituto de Ecología: Universidad Nacional Autonoma de Mexico Instituto de Ecologia
Elizabeth S Gómez Acata
Universidad Nacional Autónoma de México Instituto de Ecología: Universidad Nacional Autonoma de Mexico Instituto de Ecologia
Fabian Pérez Miranda
Instituto Politécnico Nacional Escuela Nacional de Ciencias Biológicas
Luisa I Falcón
Universidad Nacional Autonoma de Mexico Instituto de Ecologia

Research Article

Keywords: microbiota, genus, molluscivorous, Herichthys

Posted Date: March 30th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-348729/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

"Unraveling the gut microbiota composition of the genus *Herichthys* (Cichliformes:Pisces)"

Omar Mejía[1]*,Andrés Sánchez-Quinto[2,3], Elizabeth S. Gómez-Acata[2,3], Fabian Pérez-Miranda[1,4] and Luisa I. Falcón[2,3]

[1] Laboratorio de Variación Biológica y Evolución, Departamento de Zoología, Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, México

[2] Laboratorio de Ecología Bacteriana, Instituto de Ecología, UNAM, Mexico City 04510, Mexico

[3] Instituto de Ecología, Campus Yucatán, Parque Científico y Tecnológico de Yucatán, Mérida 97302, Mexico

[4] Instituto de Ciencias Biológicas, Universidad de Ciencias y Artes de Chiapas, Tuxtla Gutiérrez, Chiapas, Mexico

Emails addresses: hmejiag@ipn.mx, guitossanchezquinto@gmail.com, mistral_1984@hotmail.com, monodactilo@hotmail.com, falcon@ecologia.unam.mx

*Corresponding author: hmejiag@ipn.mx; falcon@ecologia.unam.mx

Key words

Abstract

Background

The coevolution between hosts and their gut microbiota arises as a promissory research program that could explain diversity patterns. The fishes of the cichlid family are an ideal model to evaluate coevolution due to their spectacular radiation. In particular, the neotropical genus *Herichthys* represents a great study case of study since it includes species with wide and narrow distributions, with several feeding habits as well as species that arose allopatrically and truly sympatrically. We used the hypervariable V4

region of the 16S rRNA gene from the 11 species of the genus *Herichthys* obtained from museum collections to evaluate the existence of phylosymbiosis between the fishes and their gut microbiota.

Results

The highest diversity values of gut microbiota diversity were found in the detritivorous species while the herbivorous, molluscivorous, and piscivorous showed the lowest diversity values. Differences in gut microbiota were found between species and trophic guilds, in particular for the sympatric species comparison. The phylosymbiosis test was significant showing that the evolution of the gut microbiota is different in species that arise in allopatric and sympatric conditions.

Conclusions

The most abundant phyla recovered from the gut microbiota were similar to those previously reported in other studies with cichlids supporting the idea that a gut microbial core is conserved in this group of fishes despite millions of years of evolution. Despite the caveats of working with museum specimens, our results provide evidence that gut microbiota divergence could occur even in sympatric conditions and reveals the potential use of museum collections in gut microbiota studies.

Background

The digestive tract of vertebrates is composed of a complex, diverse and dynamic microbiota with important nutritional, physiological and pathological roles (Nayak, 2010). In spite of being the most rich and diverse group of vertebrates, fishes have been left behind in the study of gut microbiota composition and significance in comparison with other groups of vertebrates (Clements et al. 2014). In recent years this gap has started

to close and the mechanisms of gastrointestinal colonization, microbial diversity changes during development, and specific microbial functions are starting to be comprehended (Wang et al. 2018). For example, it has been clearly stated that microorganisms in the surrounding aquatic environment are different from that recovered from fish digestive tracts (Härer et al. 2020). Hence, gastrointestinal microbiota is filtered by different factors such as diets and local selective pressures (Huber et al., 2004; Nayak, 2010; Sevellec et al., 2018). Nevertheless, there could be other ecological and evolutionary processes involved in gastrointestinal microbiota colonization such as dispersal, selection, ecological drift, diversification and interactions (Kohl, 2020). The aforementioned set of stochastic and deterministic forces are collectively known as phylosymbiosis, a concept that suggests that the microbial communities that structure the microbiota mirror the phylogeny of their hosts (Lim and Bordenstein 2020).

The Cichlidae family has been used as a study model since it harbors the greatest diversity within the bone fish, besides being a group with wide distribution in lentic and lotic environments (Klingenberg et al., 2003; Schott et al., 2014). Recent studies in cichlids have found that the bacterial microbiota between African and American genera are similar despite their presence in lakes with different physicochemical characteristics (Baldo et al., 2017). The slight differences between the proportions of the different bacterial groups and the presence of unique bacterial genera have led to suggest that different microbiota are associated with different trophic niches in cichlid species (Baldo 2019); however, this pattern could be absent in recently radiated cichlid species (Härer et al. 2020).

The genus *Herichthys*, comprises eleven recognized species distributed in rivers and lake environments in Northeastern Mexico, including sympatric and allopatric species with different trophic niches that turn it into an ideal model to evaluate different evolutionary and biogeographic scenarios (Pérez-Miranda et al. 2019). Recently, extensive research involving mitochondrial and nuclear molecular markers as well as a sampling genome approach using ddRAD has led to clearly support the phylogenetic relationships within the group (Pérez-Miranda et al. 2018).; Through the use of different calibration approaches it has been stated that the group originated between 10 to 17.1 Ma and comprise two assemblages of species, the clade *H. labridens* includes five species (*H. labridens*, *H.bartoni*, *H. pame*, *H.steindachneri* and *H. pantostictus*) where the speciation process is mainly sympatric; and the clade *H. cyanoguttatus* comprises six species (*H. cyanoguttatus*, *H. carpintis*, *H. tamasopoensis*, *H. deppii*, *H. minckleyi* and *H. tepehua*) whose speciation process was almost completely allopatric (Pérez-Miranda et al. 2020a, 2020b). Analysis of stomach content allow to state that the species included in the genus *Herichthys* comprise several trophic niches ranging from detritivorous (*H.carpintis*, *H.tamasopoensis* and *H. cyanoguttatus*), herbivorous (*H. minckleyi*, *H. pame* and *H. deppii*) and alguivorous (*H. labridens*) to invertivorous (*H. pantostictus* and *H. tepehua*), molluscivorous (*H. bartoni*) and piscivorous (*H. steindachneri*) species (Pérez-Miranda et al. 2019).

In the present work we focus on the evaluation of the gastrointestinal microbiome of the species included in the genus *Herichthys* in order to evaluate the microbiome composition within a single monophyletic group, compare the microbiome composition among different trophic groups, compare microbiome composition between sympatric

sibling species and the possible occurrence of phylosymbiosis between *Herichthys* species and their microbiome.

Results

Herichthys species genetic composition (V4 16S rRNA gene)

A total of 2,520,612 reads were obtained for the 16S rRNA gene V4 hypervariable region. After quality filtering, this number was reduced to 1,660,630 reads. The average number of sequences after QC per species were 55,944 for *H. deppii*, 40,570 for *H. tepehua*, 23,131 for *H. tamasopoensis*, 16,332 for *H. labridens*, 28,311 for *H. pame*, 34952 for *H. steindachneri*, 49,651 for *H. bartoni*, 53,673 for *H. cyanoguttatus*, 40,535 for *H. pantostictus*, 25,205 for *H. carpintis* and 4,598 for *H. minckleyi*. The genetic composition of the microbiota at the domain level indicates that 99.95% of all reads assigned were *Bacteria*, 0.025% were *Archaea* and 0.021% unassigned sequences.

The main bacterial phyla showed low heterogeneity among fish species: *Proteobacteria* (36.57%-73.12%) was the most abundant phylum in most fish species except for *H. labridens* and *H. deppii* (20.46%-19.11%) where *Firmicutes* (57.02%-53.92% respectively) was the most abundant phylum. In *H. bartoni* and *H. pantostictus*, *Fusobacteria* was the second most abundant phylum (28.20%-34.60% respectively). In addition, for the species *H. carpintis* and *H. tepehua*; *Planctomycetes* and *Actinobacteria* (6.00%-9.93%) were the second most abundant phyla. For the rest of fish species, the second most abundant phylum was *Firmicutes* (6.66%-28.95%). Only these three phyla comprised more than 70% of total gut microbial diversity in all fish species. Other abundant phyla excluding the three mentioned above were: *Planctomycetes* (0.30%-10.36%), *Actinobacteria* (0.23%-4.97%), *Cyanobacteria* (0.14–6.40%), *Deinococcus-*

Thermus (0.04%-9.16%), *Bacteroidetes* (0.41%–5.57%), *Chloroflexi* (0.02%–1.84%) and *Verrucomicrobia* (0.03%-1.28%). Phyla with low abundances in all sites included *Chlamydiae*, *Acidobacteria*, RsaHF231 and *Tenericutes*. Phyla with lower abundance of <1.0%, were included in a category called “others” (Figure 1).

The main bacterial class among *H.carpintis H. tepehua*, *H.steindachneri*, *H. pame* and *H. tamasopoensis* was *Gammaproteobacteria* (36.59%, 43.93%, 43.59%, 35.28% and 33.75% respectively). *Alphaproteobacteria* was the most abundant class in *H. cyanoguttatus* (26.78%); *Clostridia* in *H. deppii* (49.62%) and *H. labridens* (56.13%) and *Fusobacteria* in *H. pantostictus* (28.20%) and *H. bartoni* (34.60%). These four classes composed around 75% of relative abundances among all fish species. Therefore, around 15% were diversified in other classes. For *H. minckleyi,* only *Calditrichia* had a relative abundance of 51.71% and the remaining 46.54% distributed among low abundant classes (<1.0%). Classes with low abundance (<1.0%) in all fish species were clustered in a category named “others” (Figure 1, Additional file 4).

The analysis of alpha diversity revealed that the highest values of both Shannon and Simpson indices were found in the detritivorous species *H. cyanoguttatus* (H’=4.95, S=0.95, *H. tamasopoensis* (H’=4.87, S= 0.97), and *H. carpintis* (H’=4.81, S=0.94), meanwhile, the lowest values were found in the herbivorous species *H.deppii* (H’=2.84, S=0.88), the molluscivorous species *H. bartoni* (H’=2.83, S=0.87) and the piscivorous species *H. steindachneri* (H’=3.11 , S=0.87) (Figure 2).

A comparison of all the species and the trophic guilds using the Bray-Curtis index with a PERMANOVA analysis showed significant differences (F=1.91, *p*= 1e-04 for species and F=1.97 *p*=1e-04 for trophic guilds) (Figure 3A and 3B). Similar results were

found when for the Unweighted UniFrac distance (F=1.53, *p*= 1e-04; F=1.32 *p*=0.0023; Figure 3C and 3D) and the Weighted UniFrac distance but only at the species level (F=1.91, *p*= 0.001; F=1.23 *p*=0.017; Figure 3E and 3F) suggesting that at least two of the groups were different; a close inspection to the pairwise comparison test suggests significant differences among several species and trophic guild pairs for the three metrics used (Additional file 2 and Additional file 3). On the other hand, the analysis of beta diversity among sympatric species showed contrasting results. For the comparison between the pair *H. labridens* and *H. bartoni* no significant differences were found with a PERMANOVA analysis for any of the three distance metrics (p>0.05) (Figure 4 A-C). However, when the negative binomial was applied for this sympatric group, significant differences were found (*p*=0.00116) indicating differences in bacterial composition between *H. bartoni* and *H. labridens*. Further, sympatric species *H. pame*, *H. tamasopoensis* and *H. steindanchieri* showed significant differences for Bray-Curtis distance (F=2.12, p= 0.016) (Figure 4D) and UniFrac unweighted distance (F= 215, p= 0.007) (Figure 4E), in contrast, when the negative binomial was used for this species assemblage, only significant differences were found for the pair *H. steindachneri* and *H. tamasopoensis* (*p*= 0.0403), but not for the other tests; *H. pame vs H. tamasopoenesis* (p=0.316) and *H. pame* vs *H. steindachneri* (p= 0.295).

A phylosymbiosis comparison found significant results for the quasiswap analysis (P < 0.001), indicating that there is a relationship between the *Herichthys* species and its microbiota composition. A close inspection to the residuals indicates that the phylosymbiosis signal is stronger in the group of species included in the *Herichthys cyanoguttatus* group whose speciation process was completely allopatric; on the other

hand, the phylosymbiosis signal is weaker in the group of species included in the *Hericthys bartoni* group where four of the five species arose through a sympatric speciation process (Figure 4). In the same way, the analysis of residuals for the ten most abundant phyla in both group of species showed a stronger signal in the *Herichthys cyanoguttatus* group with similar median values and low dispersion in the interquartile range; in the *Herichthys labridens* species group, six of the phyla (*Actinobacteria*, *Chloroflexi*, *Cyanobacteria*, *Planctomycetes*, *Proteobacteria* and *Verrucomicrobia*) showed low median values compared with the other four phyla, but the interquartile range showed great dispersion around the median value (Additional file 5).

The analysis using the r0 model only found significant results in the Procrustes test with the Bray Curtis and Unifrac unweighted distances for the *Herichthys cyanoguttatus* species group (Table 1). Thus, based on the abovementioned results, we were able to suggest that the gut microbiota composition in the two different species assemblies evolved in two different ways: in sympatric conditions the complete microbiota core does not track the evolution of their hosts but signals of phylosymbiosis are provided by selected bacterial groups whereas in allopatric conditions the microbiota core completely tracks the host phylogeny.

Discussion

It has been documented that there is a relationship between feeding habits and gut microbiome composition in fishes. For example, Liu et al. (2016) compared eight species belonging to four different trophic guilds finding clear differences in community composition through the four trophic levels with special emphasis between carnivores and herbivores. In the same way, Li et al (2017) compared the gut microbiota of five

186 species of Cyprinids harvested from a farm and suggested a correlation between trophic
187 guilds and microbiota composition. Particularly, in the group of cichlids, Baldo et al.
188 (2017) analyzed fishes of lakes Tanganyka (radiated 9 to 12 mya) and Barombi Mbo
189 (radiated between 0.5.- 1 mya) from six different trophic guilds: carnivores, omnivores,
190 herbivores, planktivores, insectivores and scale eaters. Their results showed that
191 herbivores had the highest bacterial diversity, while carnivores and scale eaters had the
192 lowest values, contrary to the results found in this study where detritivorous showed the
193 highest values of Shannon diversity while the herbivorous, molluscivorous and
194 piscivorous showed the lowest values (Figure 2). It has been postulated that the pattern
195 in gut microbiota diversity in fishes occurs in other groups of vertebrates as mammals,
196 where herbivores show a higher diversity than omnivores, while carnivores show the
197 lowest diversity (He et al. 2013, Larsen et al. 2014, Ward et al. 2009). Feeding habits
198 are an important factor that determines the microbial gut diversity in fishes, but other
199 factors related with the host (genetics, gender, weight, age, immunity system and
200 intestinal motility), the environment (water temperature and salinity) and microbial
201 factors itself such as adhesion capacity and metabolic routes must be considered
202 (Wang et al. 2017).
203 According to Smith et al (2015), the metacommunity species composition of
204 vertebrate guts depends on. two processes: In a first instance it is mediated by the
205 colonization of external communities and in a second instance, the maintenance or loss
206 of the microbiota is determined by ecological filtering mediated by several factors such
207 as niche availability, the chance of lost via excretion, interaction with other microbes and
208 finally, the potential to survive the host immune response. The analysis of gut

209 microbiota among stickleback populations performed by Smith et al. (2015) attributed
210 the bacterial differences mainly to host genotype, although they were not able to
211 discriminate between long term and transient inhabitants. Similar results were found
212 recently in the gut microbiota of sympatric wood-eating catfishes (McCauley et al. 2020)
213 and between the skin and gut microbiota of three different species of fishes in
214 Amazonian rivers, where results suggest a strong correlation with host ancestry (Sylvain
215 et al. 2020). In opposition, Bledsoe et al. (2018) found that shared environment could
216 overcome the differences in the gut microbiota imposed by the host genotype, a
217 proposal that is also supported in other studies such as that performed by Riiser et al,
218 (2020) who found that the gut microbiota between different fish species in Norway is
219 driven by water temperature and diet and not with host selection. Thus, fish gut
220 microbiota composition could be driven by two main factors: host selection or external
221 selective pressures. The results found in this study showed differences in gut microbiota
222 among species (Additional file 2) and trophic guilds (Additional file 3), but these
223 apparent differences were masked in the PCoA comparison due to the inclusion of the
224 complete dataset of 11 species from several localities (Figure 3). Nevertheless, a close
225 inspection to the set of species that are sympatrically distributed clearly showed
226 differences in the gut microbiome, either in the Río Verde between the molluscivorous
227 *H. bartoni* and the alguivorous *H. labridens* (Figure 4A, 4B and 4C) and in the Río
228 Gallinas among the piscivorous *H. steindachneri*, the herbivorous *H. pame* and the
229 detritivorous *H. tamasopoensis* (Figure 4D, 4E and 4F). So, we can be able to conclude
230 that the gut microbiota composition in the species of the genus *Herichthys* is mediated

by both, host genotype and diet as has been proposed in other studies (Spor et al. 2011, Bolnick et al. 2014, Rennison et al. 2019).

Recently, it has been found that among the major clades of vertebrates, the group of fishes showed the lowest levels of phylosymbiosis signal, probably as a result of the occurrence of transitory microbes related with their feeding habits (Youngblut et al. 2019). In spite of the aforementioned, we found signals of phylosymbiosis association between the species of the genus *Herichthys* and the complete set of taxa identified in the gut microbiota using the quasiswap test, meanwhile we were only able to find signals of phylosymbiosis between the complete microbial core and their hosts in the *Herichthys cyanoguttatus* species group using the r0 model. The above mentioned models clearly differ in the assumptions between the association of host and associated phylogenies. The quasiswap is a more conservative model that does not assume that the associate tracks the phylogeny of the host, on the other hand, the r0 model strictly assumes that the microbiota tracks the phylogeny of the fishes. Parafit and Procrustes approach are both global-fit methods that allow to evaluate the degree of congruence between two tree topologies but differ in their null model. Parafit tests for random association between the trees, while PACo explicitly tests the dependence of the microbiota phylogeny onto the fish phylogeny, besides, the Procrustes approach performs better than Parafit in the type I error (reject the null hypothesis when is true) and allows a more direct inference of the potential coevolutionary relationship between hosts and their associated microbiota (Balbuena et al. 2013).

Similar results have been found in other studies with cichlids, for example, Franchini et al. (2014) compared the gut microbiota between benthic and limnetic species of the genus *Amphilopus* and found that in the older radiation that occurred in Lake Apoyo

(24000 ybp) bacterial communities are different, but in the younger radiation of Lake Xiloá (600 ybp) there were no significant differences in gut microbiota; meanwhile, Baldo et al. (2017) also found a phylogenetic signal for the complete microbiota in non-related species, and Baldo et al. (2019) found weak signals of gut microbiota and host story in cichlids of the genus *Amphilopus* from Nicaraguan crater lakes, but they were unable to detect differences among sympatric species within the same lake.

Nevertheless, unlike other study models such as African and Central American cichlids that are widely studied, the genus *Herichthys* comprises a unique study model to understand the relationship between cichlids and their gut microbiome. *Herichthys* comprises microendemic and widely distributed species that inhabit both lakes and rivers, as well as species that arose in allopatry and truly simpatry with a great variety of trophic guilds (Pérez-Miranda et al. 2019, 2020b). The six species included in the *Herichthys cyanoguttatus* species group arose almost strictly as a result of an allopatric mechanism between 7mya to 1 mya, where the arise of *Herichthys tamasopoensis* marks the separation of a parapatric population of *H. carpintis*. In contrast, the group of species included in the *Herichthys bartoni* species group comprise two pairs of truly sympatric species that arose near 1mya in the Río Verde and the Río Gallinas (Pérez-Miranda et al. 2020a, 2020b). The older radiation in the *H. cyanoguttatus* species group is reflected in a stronger phylosymbiosis signal compared with the *H. bartoni* species group (Figure 5). Within the *H. bartoni* species group, the phylosymbiosis signal is stronger in the Río Gallinas species pair in comparison with the Río Verde species pair, this later result could be related with the origin of the Río Verde species that probably arose in a paleolentic system and not in a lotic system as the species of the Río Gallinas (Pérez-Miranda 2020a).

According to Groussin et al. (2020) the coevolution between hosts and their associated microbiota requires three assumptions: 1. Host speciation must be allopatric, 2. After host isolation, the dispersal of symbionts must be limited promoting its divergence and speciation and finally, 3. The process of symbiont divergence is not reversible after the host species becomes sympatric. The occurrence of phylosymbiosis between sympatric species found in this study in the Rio Verde and the Río Gallinas contradicts the assumptions of coevolution mentioned earlier, however, similar results were found by Baldo et al. (2019) in cichlids of the Nicaraguan crater lakes that represent a process of incipient speciation where the microbiota of the gut are diverging with the host. In this way, different microbial pools are selected and could be the result of the immune system or physiological mechanisms related with gut epithelium (Mazel et al., 2018, Kohl 2020), and these differences could be magnified in older radiations where the species belong to different trophic guilds as in *Herichthys* compared with the young radiation and high levels of niche overlap observed in the cichlids of the genus *Amphilopus* in the Nicaraguan crater lakes (Baldo et al. 2019).

Proteobacteria, *Firmicutes* and *Fusobacteria* were, as in the current study, the most abundant phyla in the different *Amphilopus* cichlid species analyzed by Härer et al. (2020). One of the caveats of the present study is that most of the samples were taken in different years and seasons spanning a period of eight years (2009-2017, Additional file 1); and the results must be taken with caution. Nevertheless, the same restriction applies for the paper of Härer et al (2020) whom conclude that in spite of the limitations imposed by the sample strategy (their samples were taken between 2014 and 2015), the differences in microbial composition as well as the abundance of the three major

phyla *Proteobacteria*, *Firmicutes* and *Fusobacteria* are indicators that the host genotype drive the gut microbiota in this species, suggesting a long time persistence of main gut microbiota in cichlids. Additional support for this proceeds from the work of Riera and Baldo (2020) who compare the gut microbiome of cichlids from African and Central American lakes and found that microbial networks at both continents are similar in lineages that diverged between 60 to 75 mya (Matschiner 2019) concluding that the host gut environment constraints the evolution of gut microbiota in this group of fishes leading to predictable gut communities, as was reported in this study.

Finally, the other caveat of this study is the fact that all samples proceed from museum specimens. In this way, Heindler et al. (2018) analyzed the gut microbiome of museum fishes that were preserved either in ethanol or formaldehyde in a time period spanning from 20 to up to 100 years, reported high dropout rate and were able to successfully amplify only 15% of the samples with at least 1000 reads. In this study we were able to successfully amplify 100% of the samples with at least 4500 reads. It is necessary to point out that we were not able to include controls of the ethanol used to preserve the museum specimens.

Conclusions

The most abundant phyla recovered in this study were similar to that previously described in other cichlids suggesting a conserved gut microbial core in this group of fishes despite millions of years of evolution. Although the results found in this study must be validated in the future either with the inclusion of controls or with the comparison with other published data, they seem promising for the future inclusion of

museum specimens in gut microbiota analysis in fishes and in other biological groups (Bodawatta et al. 2018, Handler et al. 2018, Bodawatta et al. 2020).

Materials and Methods

Collection Data

A total of 48 fishes from 21 localities were included in this study representing 11 species recognized in the genus *Herichthys*. A range of individuals per species were included (ranging from three for the species with a restricted distribution up to twelve for the species with wide distribution). All samples proceed from museum specimens deposited in the "Colección Nacional de Peces Dulceacuícolas Mexicanos de la Escuela Nacional de Ciencias Biológicas del Instituto Politécnico Nacional" in Mexico. All samples were collected either by nets or electrofishing between 1966 to 2017 and were euthanized before they were fixed in 70% alcohol except for *H. minckleyi* samples that are older and were fixed in the field with 4% formalin and later transferred to 70% alcohol (Additional file 1). The entire digestive tract from the samples were removed and stored at -20°C until DNA extraction.

DNA Extraction and 16S rRNA gene amplification and sequencing

DNA was extracted from the digestive tract of 48 fishes using a salt extraction protocol (Aljanabi and Martínez 1997) with special care to avoid contamination of samples. We followed the protocol of Caporaso et al (2011) to obtain amplicons of the hypervariable V4 region of the 16S rRNA gene with universal bacteria/archaeal primers 515F/806R. Amplicon libraries consisted of triplicate PCR reactions of 25 µl that consisted of 2.5 µl Takara ExTaq PCR buffer 10X, 2 µl Takara dNTP mix (2.5 mM), 0.7 µl bovine serum albumin (BSA, 20 mg ml-1), 1 µl primers (10 µM), 0.125 µl Takara Ex Taq DNA

Polymerase (5 U µl-1) (TaKaRa, Shiga, Japan), 2 µl of DNA and 15.67 µl nuclease-free water. The PCR protocol included an initial denaturation step at 95°C (3 min), followed by 35 cycles of 95°C (30s), 52°C (40 s) and 72° C (90 s), with a final extension at 72°C for 12 min. Amplicons were pooled and purified using magnetic beads. The purified 16S rRNA fragments of the fish guts (~20 ng per sample) were sequenced on an IlluminaMiSeq platform (Yale Center for Genome Analysis, CT, USA), generating ~250 bp paired-end reads. All sequences obtained were uploaded to the NCBI database under the Bioproject number PRJNA715246.

Analysis of the sequence data

The 16S rRNA gene V4 sequences of 48 fish samples were processed in QIIME2 (v.2018.6) (Bolyen et al 2019). The FASTQ program (Andrews, 2010) was used to manually evaluate the quality of the sequences. Forward and reverse reads were both trimmed at position 10 from the 5'end and 200 from the 3'end. Sequences were then denoised using the DADA2 plugin *qiime dada2 denoised-paired* (Callahan et al 2016). All denoised, chimera and singleton sequences were removed and remaining sequences were assigned into ASVs (Callahan et al 2016). ASVs representing less than 0.01% of the sequences across the dataset were eliminated.

Representative ASV sequences were taxonomically assigned using the "classify consensus-vsearch pluggin", using the SILVA 132 database as a reference (Yilmaz et al. 2014). An alignment was performed with the MAFFT algorithm (Katoh et al. 2002). After masking positional conservations and gap filtering, a phylogeny was built in FastTree (Price et al. 2010) using the pipeline implemented in QIIME (Bolyen et al.

2019). Finally, the abundance table and phylogeny were exported to the R environment for further analysis.

Fish phylogeny

Fish species phylogenetic reconstruction was done using a random selected sequence dataset of the complete cytochrome b mitochondrial gene previously generated and built as described in Pérez-Miranda et al. (2020a).

Microbiota composition analysis

All sequences were analyzed using the R statistical environment (version 3.6.3) after QIIME2 analysis. The ASV table was used to construct the biological matrix of genetic diversity based on the 16S rRNA taxonomy. The abundance table and phylogeny were used to perform the statistical analysis and graphs with the phyloseq (McMurdie et al 2012), vegan (Oksanen et al 2007), and ggplot2 (Wickham, 2016) packages. Mitochondrial sequences were filtered out from the feature table before rarefaction. Samples were compared at a rarefaction of 4500 sequences and without rarefaction. Only rarefied sets were used for alpha diversity analysis, but for all further analysis we used the table without rarefaction.

Relative abundance barplots per fish species at phylum and family level were performed using the phyloseq package and all groups with less than 1% of relative abundance were clustered in the others group. The alpha diversity indices such as Shannon and Simpson were calculated for each fish species using vegan and phyloseq packages. Beta diversity was analyzed from PCoA ordinations calculated using Weighted UniFrac, Unweighted UniFrac and Bray-Curtis matrix distances in order to compare the complete set of *Herichthys* species by trophic guilds. Additionally, the same matrix

distances were analyzed to compare sympatric fish species *H. labridens* and *H. bartoni* from the Río Verde and *H. steindachneri, H. pame* and *H. tamasopoensis* from the Río Gallinas. Furthermore, the differences among species, trophic guilds and sympatric species datasets were compared through a PERMANOVA approach implemented in the ADONIS function using vegan with 9999 permutations and the differences between pairs were compared using the pairwise.adonis function as implemented in the pairwiseAdonis library over 999 permutations (Martinez Arbizu 2020) using a Bonferroni correction test for the species and trophic guilds dataset. Finally, for the sympatric species, the analysis was complemented with a negative binomial. The negative binomial allows to evaluate the differences of alpha diversity between fish populations. Since the measures of alpha diversity obtained for this project are species counts, a generalized linear model with negative binomial distance was necessary in addition to the negative binomial, that considers the over-dispersion of the residuals, which in this case is useful because of the reduced number of replicates.

Finally, in order to evaluate the existence of phylosymbiosis between *Herichthys* species and their microbiome, two different approaches were used. First, we pooled all the sequences recovered from a single fish species to obtain the most inclusive microbiome of each *Herichthys* species regardless of its geographic origin. Second, to compare the fish phylogeny against the microbial 16S rRNA phylogeny of the 4565 OTUs we used a presence/absence matrix among hosts and OTUs at the genus level using the quasiswap null model which does not assume that the symbionts are tracking the evolution of the host in the PACo library using a Cailliez correction for negative values with 1000 permutations with an asymmetric Procrustes statistic (Hutchinson et al. 2017).

Third, assuming that each fish species comprises a unique microbiome that has coevolved with fish phylogeny, we first calculated the beta diversity indices among fish species microbiome communities at genus level using Weighted UniFrac, Unweighted UniFrac and Bray-Curtis distances, after this we constructed a dendrogram using the Neighbor Joining method. Finally, we compared the fish phylogeny against beta distances dendrograms using the ro model that assumes that the associates track the host phylogeny using a Cailliez correction for negative values with 1000 permutations with a symmetric Procrustes statistic in Paco library (Hutchinson et al. 2017) and the Parafit function in ape library using a Cailliez correction for negative values with 1000 permutations with the test.links options as true (Paradis et al. 2004). The above mentioned phylosymbiosis tests were implemented at three different levels, for the eleven species included in the genus *Herichthys* and for each one of the two assemblages of species recognized within the genus, the species included in the clade *H. labridens* whose speciation process was almost entirely sympatric and the species included in the clade *H. cyanoguttatus* whose speciation process was completely allopatric.

Availability of data and materials

Raw sequencing data were deposited in the NCBI under the BioProject number PRJNA715246. Scripts used in the analysis of the data can be found at https://github.com/HOmarMejiaG/Herichthys-microbiome.

Abbreviations

ASV: Amplicon sequence variant

ddRAD: Double digest restriction-site associated DNA

mya: million years ago

OTU: Operational taxonomic unit

PcoA: Principal Coordiantes Analysis

PERMANOVA: Permutational analysis of variance

QC: Quality control

ybp: years before present

References

Aljanabi, S. M., & Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic acids research, 25, 4692-4693.

Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at:

http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Balbuena, J. A., Míguez-Lozano, R., & Blasco-Costa, I. (2013). PACo: a novel procrustes application to cophylogenetic analysis. PloS one, 8, e61048.

Baldo, L., Pretus, J. L., Riera, J. L., Musilova, Z., Nyom, A. R. B., & Salzburger, W. (2017). Convergence of gut microbiotas in the adaptive radiations of African cichlid fishes. The ISME journal, 11, 1975-1987.

Baldo, L., Riera, J. L., Salzburger, W., & Barluenga, M. (2019). Phylogeography and ecological niche shape the cichlid fish gut microbiota in Central American and African lakes. Frontiers in microbiology, 10, 2372.

Bledsoe, J. W., Waldbieser, G. C., Swanson, K. S., Peterson, B. C., & Small, B. C. (2018). Comparison of channel catfish and blue catfish gut microbiota assemblages shows minimal effects of host genetics on microbial structure and inferred function. Frontiers in microbiology, 9, 1073.

Bodawatta, K. H., Sam, K., Jønsson, K. A., & Poulsen, M. (2018). Comparative analyses of the digestive tract microbiota of New Guinean passerine birds. Frontiers in microbiology, 9, 1830.

Bodawatta, K. H., Puzejova, K., Sam, K., Poulsen, M., & Jønsson, K. A. (2020). Cloacal swabs and alcohol bird specimens are good proxies for compositional analyses of gut microbial communities of Great tits (*Parus major*). Animal Microbiome, 2, 1-13.

Bolnick, D. I., Snowberg, L. K., Hirsch, P. E., Lauber, C. L., Knight, R., Caporaso, J. G., & Svanbäck, R. (2014). Individuals' diet diversity influences gut microbial diversity in two freshwater fish (threespine stickleback and Eurasian perch). Ecology letters, 17, 979-987.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., ... & Bai, Y. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature biotechnology, 37, 852-857.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. Nature methods, 13, 581-583.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., ... & Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the national academy of sciences, 108 (Supplement 1), 4516-4522.

Clements, K. D., Angert, E. R., Montgomery, W. L., & Choat, J. H. (2014). Intestinal microbiota in fishes: what's known and what's not. Molecular ecology, 23, 1891-1898.

Franchini, P., Fruciano, C., Frickey, T., Jones, J. C., & Meyer, A. (2014). The gut microbial community of Midas cichlid fish in repeatedly evolved limnetic-benthic species pairs. PloS one, 9, e95027.

Härer, A., Torres-Dowdall, J., Rometsch, S. J., Yohannes, E., Machado-Schiaffino, G., & Meyer, A. (2020). Parallel and non-parallel changes of the gut microbiota during trophic diversification in repeated young adaptive radiations of sympatric cichlid fish. Microbiome, 8:149.

He, S., Wu, Z., Liu, Y., Wu, N., Tao, Y., Xu, L., ... & Ringø, E. (2013). Effects of dietary 60 g kg− 1 dried distiller's grains in least-cost practical diets on production and gut allochthonous bacterial composition of cage-cultured fish: comparison among fish species with different natural food habits. Aquaculture Nutrition, 19, 765-772.

Heindler, F. M., Christiansen, H., Frédérich, B., Dettaï, A., Lepoint, G., Maes, G. E., ... & Volckaert, F. A. (2018). Historical DNA metabarcoding of the prey and microbiome of trematomid fishes using museum samples. Frontiers in Ecology and Evolution, 6, 151.

Huber, I., Spanggaard, B., Appel, K. F., Rossen, L., Nielsen, T., & Gram, L. (2004). Phylogenetic analysis and in situ identification of the intestinal microbial community of rainbow trout (*Oncorhynchus mykiss*, Walbaum). Journal of Applied Microbiology, 96, 117-132.

Hutchinson, M. C., Cagua, E. F., Balbuena, J. A., Stouffer, D. B., & Poisot, T. (2017). paco: implementing Procrustean Approach to Cophylogeny in R. Methods in Ecology and Evolution, 8, 932-940.

Katoh, K., Misawa, K., Kuma, K. I., & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic acids research, 30, 3059-3066.

Klingenberg, C. P., Barluenga, M., & Meyer, A. (2003). Body shape variation in cichlid fishes of the *Amphilophus citrinellus* species complex. Biological Journal of the Linnean Society, 80, 397-408.

Kohl, K. D. (2020). Ecological and evolutionary mechanisms underlying patterns of phylosymbiosis in host-associated microbial communities. Philosophical Transactions of the Royal Society B, 375, 20190251.

Ley, R. E., Hamady, M., Lozupone, C., Turnbaugh, P. J., Ramey, R. R., Bircher, J. S., ... & Gordon, J. I. (2008). Evolution of mammals and their gut microbes. science, 320, 1647-1651.

Li, T., Long, M., Li, H., Gatesoupe, F. J., Zhang, X., Zhang, Q., ... & Li, A. (2017). Multi-omics analysis reveals a correlation between the host phylogeny, gut microbiota and metabolite profiles in cyprinid fishes. Frontiers in microbiology, 8, 454.

Liu, H., Guo, X., Gooneratne, R., Lai, R., Zeng, C., Zhan, F., & Wang, W. (2016). The gut microbiome and degradation enzyme activity of wild freshwater fishes influenced by their trophic levels. Scientific reports, 6, 1-12.

Lim, S. J., & Bordenstein, S. R. (2020). An introduction to phylosymbiosis. Proceedings of the Royal Society B, 287, 20192900.

Martinez Arbizu, P. (2020). pairwiseAdonis: Pairwise multilevel comparison using adonis. R package version 0.4.

526 Matschiner, M. (2019). Gondwanan vicariance or trans-Atlantic dispersal of cichlid
527 fishes: a review of the molecular evidence. Hydrobiologia, 832, 9-37.
528 McCauley, M., German, D. P., Lujan, N. K., & Jackson, C. R. (2020). Gut microbiomes
529 of sympatric Amazonian wood-eating catfishes (Loricariidae) reflect host identity and
530 little role in wood digestion. Ecology and Evolution, 10, 7117-7128.
531 Mazel, F., Davis, K. M., Loudon, A., Kwong, W. K., Groussin, M., & Parfrey, L. W.
532 (2018). Is host filtering the main driver of phylosymbiosis across the tree of life?.
533 Msystems, 3,5.
534 McMurdie PJ, Holmes S. (2013) phyloseq: An R Package for Reproducible Interactive
535 Analysis and Graphics of Microbiome Census Data. PloS one, 8, e61217.
536 Nayak, S. K. (2010). Probiotics and immunity: a fish perspective. Fish & shellfish
537 immunology, 29, 2-14.
538 Nishida, A. H., & Ochman, H. (2018). Rates of gut microbiome divergence in mammals.
539 Molecular ecology, 27, 1884-1897.
540 Nelson, T. M., Apprill, A., Mann, J., Rogers, T. L., & Brown, M. V. (2015). The marine
541 mammal microbiome: current knowledge and future directions. Microbiology Australia,
542 36, 8-13.
543 Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M. H. H., Oksanen, M. J., &
544 Suggests, M. A. S. S. (2007). The vegan package. Community ecology package,
545 10(631-637), 719.
546 Paradis, E., Claude, J., & Strimmer, K. (2004). APE: analyses of phylogenetics and
547 evolution in R language. Bioinformatics, 20, 289-290.
548

549 Pérez-Miranda, F Mejía, O, Soto-Galera, E., Espinosa-Pérez, H., Piálek, L., y O. Říčan.
550 2018.Phylogeny and species diversity of the genus *Herichthys* (Teleostei: Cichlidae).
551 Journal of Zoological Systematics and Evolutionary Research, 56, 233-247.
552
553 Pérez-Miranda, F., O. Mejia*, G. Zúñiga, E. Soto-Galera y O. Říčan.2019. Feeding
554 ecomorphologies in the genus *Herichthys* (Perciformes: Cichlidae): An approximation
555 using stomach content and lower pharyngeal jaw shapes. Revista de Biología Tropical
556 67, 643-653.
557 Pérez-Miranda, F., Mejia, O., López, B., & Říčan, O. (2020a). Molecular clocks,
558 biogeography and species diversity in Herichthys with evaluation of the role of Punta del
559 Morro as a vicariant brake along the Mexican Transition Zone in the context of local and
560 global time frame of cichlid diversification. PeerJ, 8, e8818.
561 Pérez-Miranda, F., O. Mejia, A. González-Díaz, N. Martínez-Méndez, E. Soto-Galera,
562 G. Zúñiga, y O. Říčan. (2020b). The role of head-shape ad trophic variation in the
563 diversification of the genus *Herichthys* in sympatry and allopatry. Journal of Fish Biology
564 96,1370-1378.
565 Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2–approximately maximum-
566 likelihood trees for large alignments. PloS one, 5, e9490.
567 Rennison, D. J., Rudman, S. M., & Schluter, D. (2019). Parallel changes in gut
568 microbiome composition and function during colonization, local adaptation and
569 ecological speciation. Proceedings of the Royal Society B, 286, 20191911.

570 Riera, J. L., & Baldo, L. (2020). Microbial co-occurrence networks of gut microbiota
571 reveal community conservation and diet-associated shifts in cichlid fishes. Animal
572 Microbiome, 2, 1-13.
573 Riiser, E. S., Haverkamp, T. H., Varadharajan, S., Borgan, Ø., Jakobsen, K. S., Jentoft,
574 S., & Star, B. (2020). Metagenomic shotgun analyses reveal complex patterns of intra-
575 and interspecific variation in the intestinal microbiomes of codfishes. Applied and
576 environmental microbiology, 86, 6.
577
578 Sevellec, M., Derome, N., & Bernatchez, L. (2018). Holobionts and ecological
579 speciation: the intestinal microbiota of lake whitefish species pairs. Microbiome, 6, 47.
580 Schott, R. K., Refvik, S. P., Hauser, F. E., López-Fernández, H., & Chang, B. S. (2014).
581 Divergent positive selection in rhodopsin from lake and riverine cichlid fishes. Molecular
582 biology and evolution, 31, 1149-1165.
583 Smith, C. C., Snowberg, L. K., Caporaso, J. G., Knight, R., & Bolnick, D. I. (2015).
584 Dietary input of microbes and host genetic variation shape among-population
585 differences in stickleback gut microbiota. The ISME journal, 9, 2515-2526.
586 Spor, A., Koren, O., & Ley, R. (2011). Unravelling the effects of the environment and
587 host genotype on the gut microbiome. Nature Reviews Microbiology, 9, 279-290.
588 Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ & Rambaut A (2018)
589 Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10 Virus
590 Evolution 4, vey016.

Sylvain, F. É., Holland, A., Bouslama, S., Audet-Gilbert, É., Lavoie, C., Val, A. L., & Derome, N. (2020). Fish skin and gut microbiomes show contrasting signatures of host species and habitat. Applied and Environmental Microbiology, 86, 16.

Wang, A. R., Ran, C., Ringø, E., & Zhou, Z. G. (2018). Progress in fish gastrointestinal microbiota research. Reviews in Aquaculture, 10, 626-640.

Ward, N. L., Steven, B., Penn, K., Methé, B. A., & Detrich, W. H. (2009). Characterization of the intestinal microbiota of two Antarctic notothenioid fish species. Extremophiles, 13, 679-685.

Wickham H (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Youngblut, N. D., Reischer, G. H., Walters, W., Schuster, N., Walzer, C., Stalder, G., ... & Farnleitner, A. H. (2019). Host diet and evolutionary history explain different aspects of gut microbiome diversity among vertebrate clades. Nature communications, 10, 1-15.

Acknowledgements

Technical assistance is acknowledged to O Gaona (IE, UNAM) and J Ortiz (FC, UNAM). We also like to thank Eduardo Soto Galera for provide access to the museum collections.

Funding

This study was funded by Instituto de Ecología, UNAM (LIF). ASQ received a graduate studies scholarship from CONACyT. ESGA received a postdoctoral scholarship from DGAPA-UNAM.

Author information

Affiliations

1 Laboratorio de Variación Biológica y Evolución, Departamento de Zoología, Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, México
Omar Mejía & Fabian Pérez-Miranda
2 Laboratorio de Ecología Bacteriana, Instituto de Ecología, UNAM, Mexico City 04510, Mexico
3 Instituto de Ecología, Campus Yucatán, Parque Científico y Tecnológico de Yucatán, Mérida 97302, Mexico
Andrés Sánchez-Quinto, Elizabeth S. Gómez-Acata & Luisa I. Falcón
4 Instituto de Ciencias Biológicas, Universidad de Ciencias y Artes de Chiapas, Tuxtla Gutiérrez, Chiapas, Mexico
Fabian Pérez-Miranda

Contributions

OM, LIF: Developed the idea. FPM, ESGA: Carried out the molecular work. OM, ASQ, FPM: Carried out the data analyses. OM, LIF, ASQ, FPM, ESGA: Wrote the manuscript. All authors read and approved the final manuscript.

Corresponding author

Correspondence to Omar Mejía & Luisa I. Falcón

Ethics declarations

Ethics approval and consent to participate

No ethics approval was necessary due that all samples used in this study proceed from museum specimens.

Consent for publication

Not applicable.

Competing interests

There are no competing interests.

Supplementary information

Additional file 1. Table S1

Additional file 2. Table S2

Additional file 1. Table S3

Additional file 4 Figure S1

Additional file 5 Figure S2

Figures

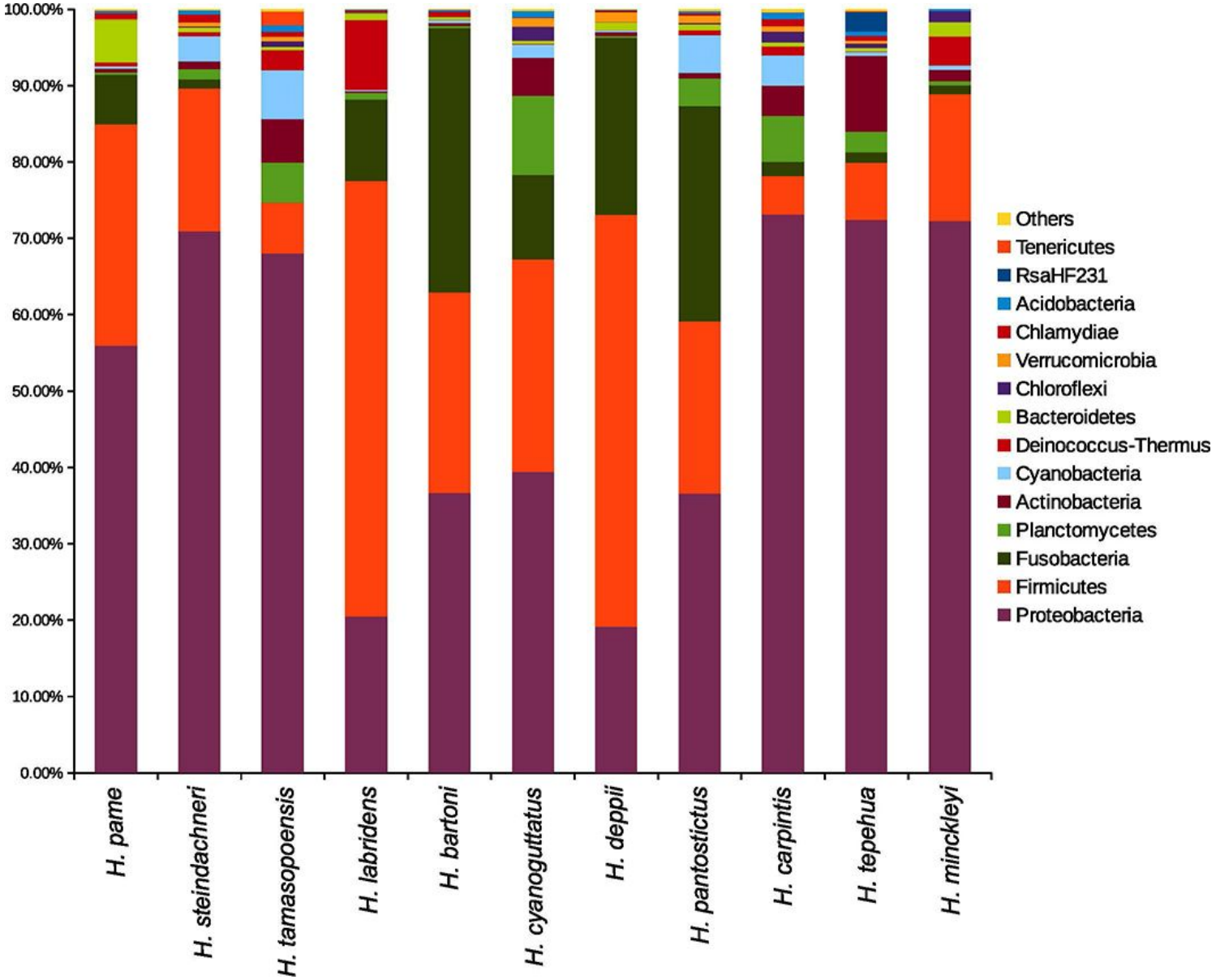

Figure 1

Classes with low abundance (<1.0%) in all fish species were clustered in a category named “others” (Figure 1, Additional file 4).

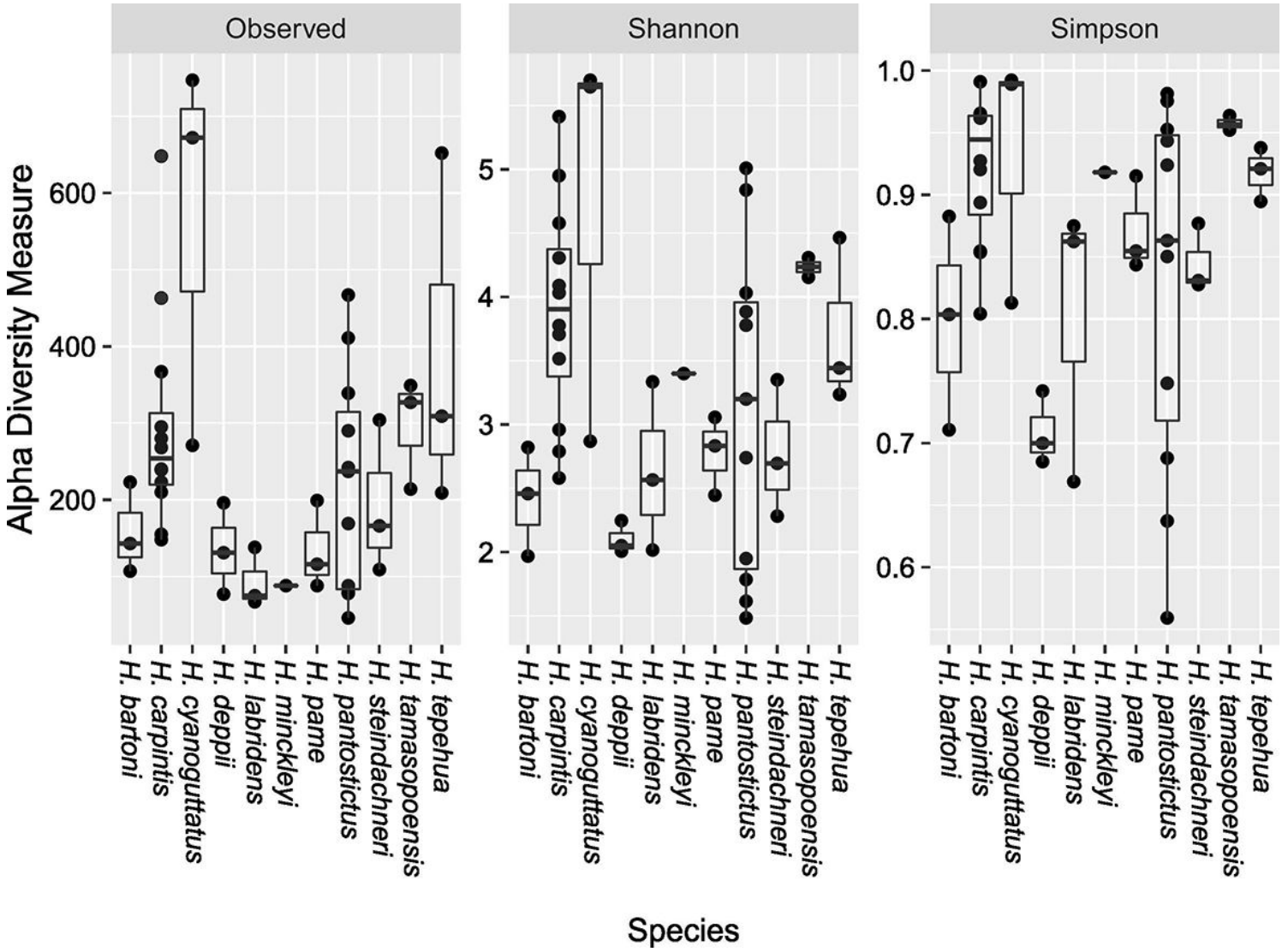

Figure 2

The analysis of alpha diversity revealed that the highest values of both Shannon and Simpson indices were found in the detritivorous species H. cyanoguttatus (H'=4.95, 132 S=0.95, H. tamasopoensis (H'=4.87, S= 0.97), and H. carpintis (H'=4.81, S=0.94), meanwhile, the lowest values were found in the herbivorous species H.deppii (H'=2.84, S=0.88), the molluscivorous species H. bartoni (H'=2.83, S=0.87) and the piscivorous species H. steindachneri (H'=3.11 , S=0.87) (Figure 2).

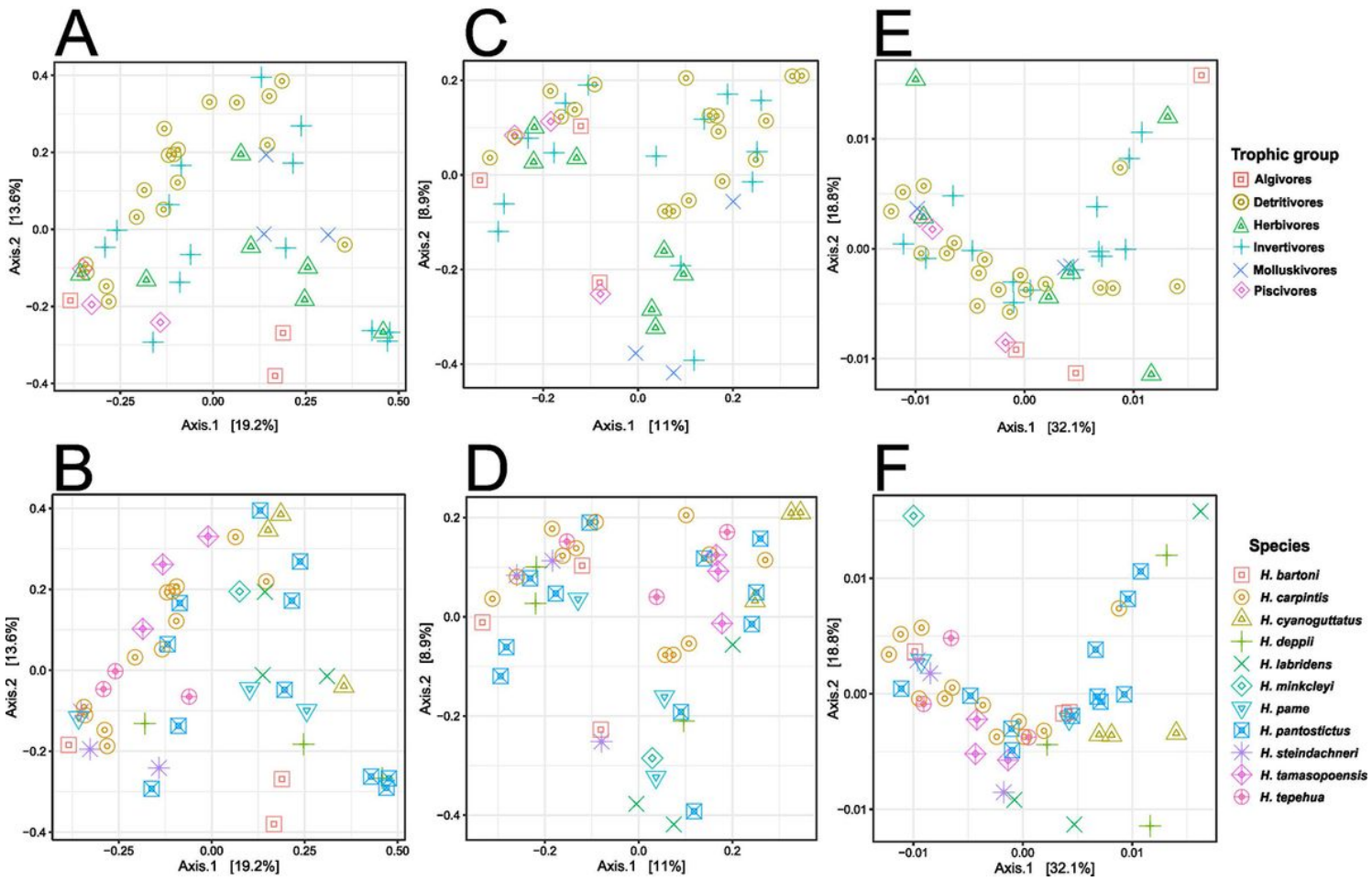

Figure 3

A comparison of all the species and the trophic guilds using the Bray-Curtis index with a PERMANOVA analysis showed significant differences (F=1.91, p= 1e-04 for species and F=1.97 p=1e-04 for trophic guilds) (Figure 3A and 3B). Similar results were found when for the Unweighted UniFrac distance (F=1.53, p= 1e-04; F=1.32 p=0.0023; Figure 3C and 3D) and the Weighted UniFrac distance but only at the species level (F=1.91, p= 0.001; F=1.23 p=0.017; Figure 3E and 3F) suggesting that at least two of the groups were different; a close inspection to the pairwise comparison test suggests significant differences among several species and trophic guild pairs for the three metrics used

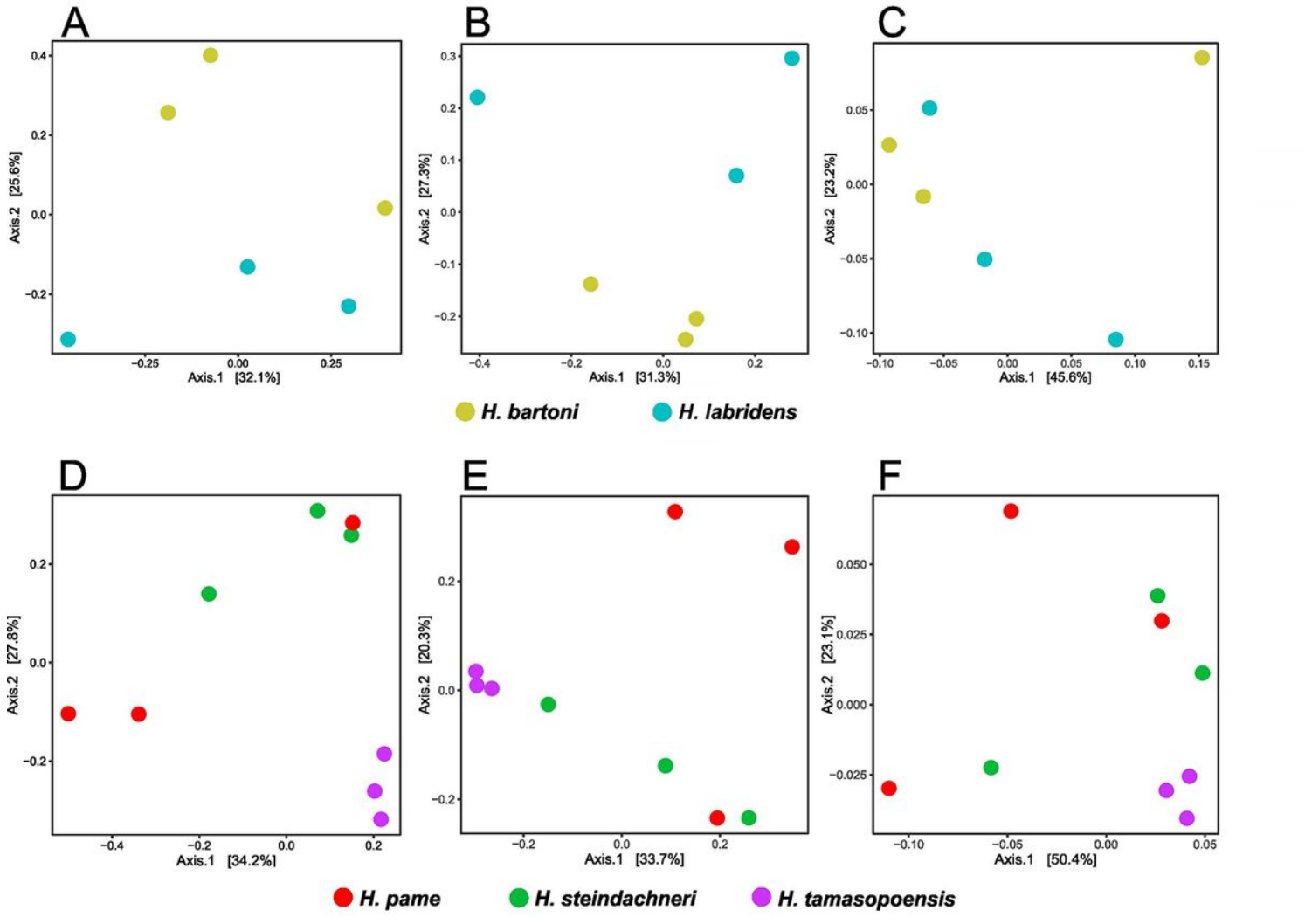

Figure 4

A close inspection to the set of species that are sympatrically distributed clearly showed differences in the gut microbiome, either in the Río Verde between the molluscivorous H. bartoni and the alguivorous H. labridens (Figure 4A, 4B and 4C) and in the Río Gallinas among the piscivorous H. steindachneri, the herbivorous H. pame and the detritivorous H. tamasopoensis (Figure 4D, 4E and 4F).

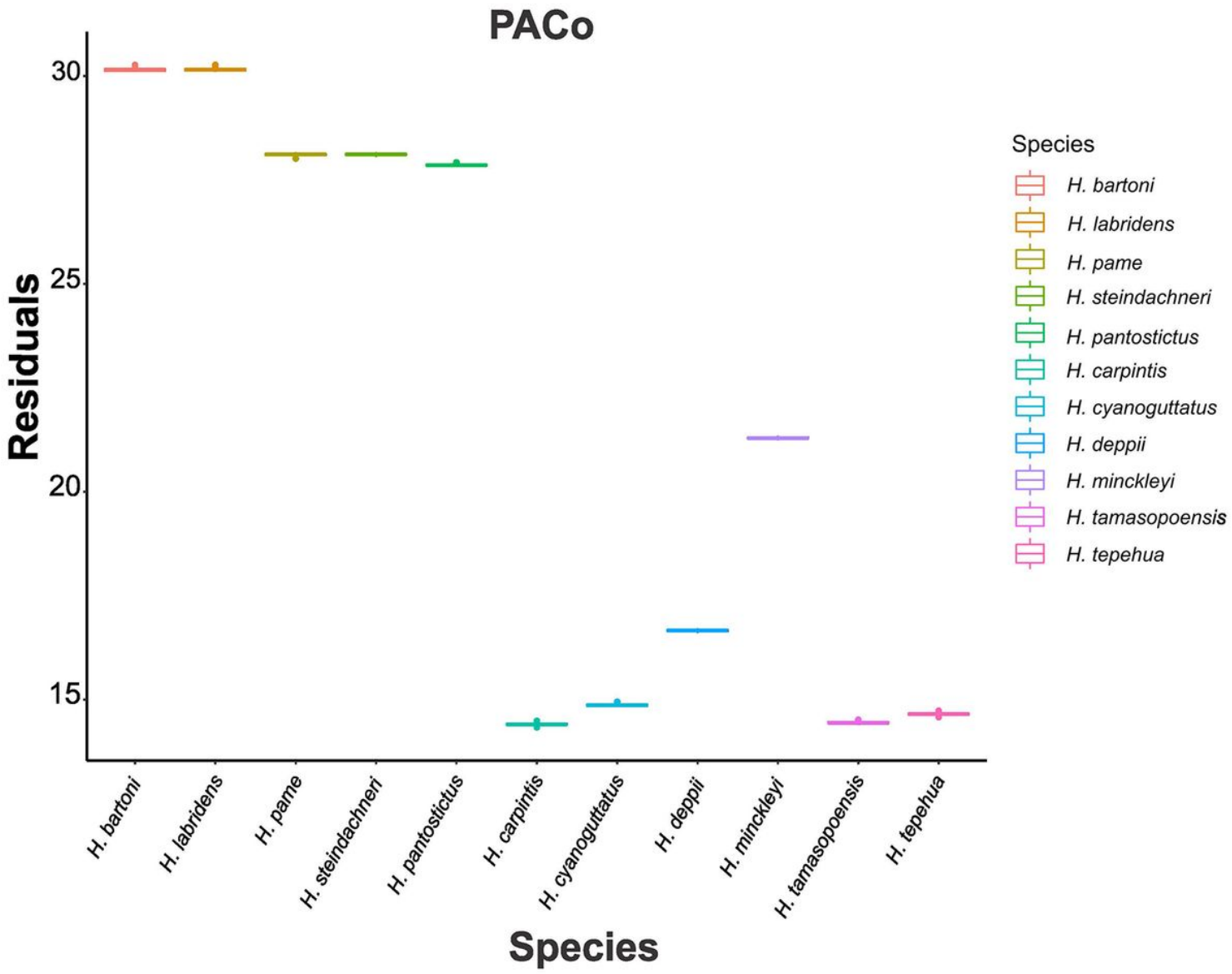

Figure 5

The older radiation in the H. cyanoguttatus species group is reflected in a stronger phylosymbiosis signal compared with the H. bartoni species group (Figure 5).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1.xlsx
- Additionalfile1.xlsx
- Additionalfile2.xlsx
- Additionalfile3.xlsx
- Additionalfile4.pdf

- Additionalfile5.pdf