Fish skin and gut microbiomes show contrasting signatures of host species and habitat

François-Étienne Sylvain,[a]# Aleicia Holland,[b] Sidki Bouslama,[a] Émie Audet-Gilbert,[a] Camille Lavoie,[a] Adalberto Luis Val,[c] Nicolas Derome[a]

a. Institut de Biologie Intégrative et des Systèmes, Université Laval, 1030 avenue de la Médecine, Québec (QC), G1V 0A6, Canada

b. La Trobe University, School of Life Science, Department of Ecology, Environment and Evolution, Albury/Wodonga Campus, Vic, Australia

c. Instituto Nacional de Pesquisas da Amazônia (INPA), Laboratório de Ecofisiologia e Evolução Molecular, Manaus, AM, 69067-375, Brazil

Running head: Tissue-specific drivers of the fish microbiome.

* **Corresponding author**:
François-Étienne Sylvain
Email : francois-etienne.sylvain.1@ulaval.ca

Abstract

Teleost fishes represent an invaluable repertoire of host species to study the factors shaping animal-associated microbiomes. Several studies have shown that the phylogenetic structure of fish gut microbiome is driven by species-specific (e.g. host ancestry, genotype or diet) and habitat-specific (e.g. hydrochemical parameters and bacterioplankton composition) factors. However, our understanding of other host-associated microbial niches, such as the skin mucus microbiome, remains limited. The goal of our study was to explore simultaneously the phylogenetic structure of fish skin mucus and gut microbiome and compare the effect of species- and habitat-specific drivers on the structure of microbial communities in both tissues. We sampled 114 wild fish from 6 populations of 3 ecologically and phylogenetically contrasting Amazonian Teleost species. Water samples were collected at each site, and 10 physicochemical parameters were characterized. The skin mucus, gut, and water microbial communities were characterized using a metabarcoding approach targeting the V3-V4 regions of the 16S rRNA. Our results showed a significant distinction between the phylogenetic profile and diversity of the microbiome from each microbial niche. Skin mucus and bacterioplankton communities were significantly closer in composition than gut and free-living communities. Species-specific factors mostly modulated gut bacterial communities, while the skin mucus microbiome was predominantly associated to environmental physicochemistry and bacterioplankton community structure. These results suggest that the variable skin mucus community is a relevant target to develop microbial biomarkers of environmental status, while the more conserved gut microbiome

is better suited to study long-term host-microbe interactions over evolutionary time scales.

Importance

Whether host-associated microbiomes are mostly shaped by species-specific or environmental factors is still unresolved. Especially, it is unknown to what extent microbial communities from two different host tissues from the same host respond to these factors. Our study is one of the first to focus on the microbiome of teleost fishes to shed a light on this topic, as we investigate how the phylogenetic structure of microbial communities from two distinct fish tissues are shaped by species- and habitat-specific factors. Our study showed that in contrast to the teleost gut microbiome, skin mucus communities are highly environment dependent. This result has different implications: (1) the skin mucus microbiome should be used, rather than the gut, to investigate bacterial biomarkers of ecosystem perturbance in the wild; (2) the gut microbiome is better suited for studies on the drivers of phylosymbiosis, or the co-evolution of fish and their symbionts.

Introduction

Teleosts are represented by more than 28 000 species, which exhibit a variety of physiologies, natural histories and ecologies (1). Thus, they provide an invaluable repertoire of host species to study the nature of vertebrate microbial communities (2) and the factors shaping animal-associated microbiomes (3). Teleost's guts, skin mucus and gills support high concentrations of bacterial cells (4, 5) forming highly diverse and structured microbial communities (6). Like in humans, the gut microbial community of fishes has drawn a lot of attention from the scientific community in the last decade; dozens of studies have characterized the gut microbiota of various fish species (5, 7). These studies identified endogenous and exogenous factors that affect the establishment of bacteria in Teleosts guts (2, 7, 8). Some endogenous factors are species/population-specific (1)— they include: host ancestry (9), genotype (10), diet (11); others are individual-specific and include parasitic load (12), immunological state (13), and life history (14). Exogenous factors are habitat-specific and include: environmental water physicochemical parameters (15, 16) and bacterioplankton composition (17).

In contrast to the gut microbiota, our understanding of the mucosal microbiome of fish skin remains limited (18). Like the gut microbiota, the skin mucus microbial community plays crucial roles for fish holobionts (5). Skin mucus bacteria constitute the first barrier of defense against infections by environmental opportunistic pathogens (19). The

ecological niche provided by fish skin mucus is in constant contact with the external environment, rendering its associated microbiota susceptible to environmental physicochemical disturbances, compared to the stable conditions of the buffered gut environment (16, 17, 20). This sensitivity of fish skin mucus microbiota suggests that, in the future, skin mucus bacteria could be used as microbial biomarkers or proxies of fish health, which would help manage wild fish populations and improve captive fish performance and growth by enabling the rapid diagnosis of fish gut diseases through skin mucus microbiota profiling (11, 21). Characterizing fish skin mucus microbial communities in a range of different hosts and environments is critical for (i) identifying novel functions of animal-associated microbes, (ii) understanding the connectivity between environmental microbial communities and those of the host, and (iii) identifying the endogenous and exogenous factors shaping these microbial consortia. Furthermore, most studies which investigated the fish skin mucus microbiome did not concurrently report the bacterial communities from other host-associated tissues, such as the gut (4, 22). Therefore, we cannot determine to what extent skin mucus communities are driven by the same species-specific or habitat-specific factors shaping gut communities. The main goal of our study was to address this gap by (i) exploring the phylogenetic structure of fish skin mucus and gut microbial communities and (ii) comparing the effect of species-specific and habitat-specific drivers on the structure of microbial communities in both tissues.

Here we investigated the bacterial communities (gut and skin mucus) of three sympatric host species found in the Brazilian Amazon basin: the flag cichlid (*Mesonauta festivus*),

the pacu (*Mylossoma duriventre*), and the black piranha (*Serrasalmus rhombeus*). These three species are phylogenetically divergent: pacus and black piranhas belong to the *Serrasalmidae* family, which diverged from the flag cichlid's *Cichlidae* family 229.9 M years ago (23). These species also differ in their overall ecology (behavior, diet, and distribution). Flag cichlids are more sedentary and live on the surface of river margins underneath logs of wood or floating macrophytes (24). In contrast, black piranhas and pacus live in the middle of the water column and can tolerate strong river currents (25). Black piranhas are opportunistic carnivores, while flag cichlids are benthivores and pacus are planktivores (24-28). Flag cichlids and black piranhas are found in a great variety of habitats, and therefore have a larger overall distribution area than pacus (25). We used a metabarcoding approach targeting the 16S SSU V3-V4 region of the rRNA gene to analyze the gut and skin mucus bacterial communities of 33 wild flag cichlids, 47 pacus, 34 black piranhas, and six water samples, collected at six sampling sites. Our results show that the phylogenetic structures of fish gut, skin mucus and bacterioplankton communities significantly diverge. Most importantly, they also support the hypothesis that fish skin mucus and gut microbiomes do not show the same response to species- and habitat-specific factors.

Results

Host-associated versus free-living bacterial communities

Our results show a clear distinction between the taxonomical structure of the bacterial communities from skin mucus, gut and bacterioplankton (Fig. 1, Fig. 2, Suppl. Fig. 3, Suppl. Fig. 4). Permutational analyses of variances (PERMANOVAs) results based on

136 weighted Unifrac distances show smaller significant differences between
137 bacterioplankton and skin mucus communities ($p < 0.001$, $F = 3.58$, $df = 115$), than
138 between bacterioplankton and gut communities ($p < 0.001$, $F = 6.63$, $df = 113$). The high
139 similarity between skin mucus communities and bacterioplankton is also suggested in
140 the PCoA plots on Suppl. Fig. 3 —the sample type factor explains only 22.2 % of the
141 variance between skin mucus and free-living communities, while it explains 40.0 % of
142 the variance between gut and free-living communities. Finally, the Thetayc Dissimilarity
143 Index (TDI) (29), also showed that bacterioplankton structure was more similar to the
144 skin mucus microbiota (TDI = 0.690 ± 0.004 S.E.) than to the gut microbiota (TDI =
145 0.848 ± 0.003 S.E.). Faith's phylogenetic diversity measures (Fig. 1) also showed
146 significant differences in the α-diversity of the three sample types: overall α-diversity
147 was significantly higher in skin mucus than in gut (p-value < 0.001, T = 4.19, df = 232)
148 and bacterioplankton samples (p-value = 0.002, T = 5.77, df = 5). Average diversity of
149 gut was also significantly higher than bacterioplankton samples (p-value < 0.05, T =
150 2.43, df = 6). Finally, heatmaps in Suppl. Fig. 4 show the clustering of 10 ASVs from
151 each of the two most important genera in bacterioplankton (CL500-29 marine group and
152 hgcI clade), based on their mean relative abundance in fish skin mucus, gut and
153 bacterioplankton samples. They show that the ASVs from these abundant genera in
154 bacterioplankton share co-abundance with the same ASVs in the fish skin mucus
155 microbiome, but not with those in the gut microbiome.
156
157 **Skin mucus versus gut microbiome**

158 The two host-associated microbial habitats show certain similarities. For instance, they
159 both contain a significantly higher abundance of *Oxyphotobacteria* (p-value < 0.001),
160 *Mollicutes* (p-value < 0.001), and of *Clostridia* (p-value < 0.001) than free-living
161 communities (Fig. 1). However, results of PCAs, PERMANOVAs (Fig. 2a,b,c), LEfSe
162 analyses (Fig. 2d,e,f), metagenomeSeq abundance heatmaps (Fig. 2g,h,i) and
163 differential abundance barplots based on linear discriminant analysis (LDA) log scores
164 (Fig. 2j,k,l) all show major differences in the phylogenetic structure of microbial
165 communities from both tissues. PERMANOVA results (Fig. 2a,b,c) first confirm that the
166 overall community structures were significantly different for both tissues for all three
167 species sampled (all p values < 0.001, see Fig. 2 for F and df values). Then, to
168 investigate further these differences, we performed LEfSe analyses (30) to identify the
169 bacterial taxa (at all taxonomic levels possible) significantly associated with each tissue
170 (Fig. 2d,e,f). The unpaired Wilcoxon rank-sum test embedded in LEfSe identified
171 several differentially abundant features in the skin mucus and gut of each species.
172 Common features in all three host species in this study include a significantly higher
173 abundance of *Clostridia* in gut samples, and of *Gammaproteobacteria* and
174 *Acidobacteria* in skin mucus samples. Three other taxonomic groups had a significantly
175 higher abundance in the skin mucus of at least two species: the class *Actinobacteria* for
176 flag cichlids and black piranhas, and the phylum *Proteobacteria* and the class
177 *Alphaproteobacteria* for flag cichlids and pacu. These results suggest the existence of
178 distinct and potentially consistent phylogenetic signatures of microbial communities
179 unique to both tissues investigated. At the ASV level, differences between communities
180 from both tissues were assessed with the metagenomeSeq LDA approach and

visualized with heatmaps and differential abundance barplots (Figure 2), which is more sensitive to significant differences, but less sensitive to potential confounding factors than the LEfSe approach (31). Heatmaps on Fig. 2g,h,I show that each host species microbial communities mostly cluster according to their tissue of origin. Clustering according to tissue of origin appeared more pronounced in flag cichlid and pacu samples than for black piranha samples. Differential abundance LDA scores barplots were made from the 20 ASVs with the highest differential abundance (Fig. 2j,k,l) between both tissues. *Enhydrobacter sp*. and *CL500-29 marine group sp*. were consistently found to be among the 10 most differentially abundant taxa for fish skin mucus for all three host species. Amongst these differentially abundant taxa, none were shared between all three host species for gut communities.

Species-specific factors

The site-specific PERMANOVAs of Fig. 3 always show a significant effect of the host phylogeny on the structure of microbial communities from both tissues (p-values < 0.02, see F and df values on Fig. 3) except for skin mucus communities of the site MPWS23. However, the factor "host species" (the axes of PCoAs of Fig. 3) consistently explained a higher percentage of the variance for gut samples (average of 60.9 %) than for skin mucus samples (average of 39.8 %) for fish assemblages of all six sampling sites. Also, F values of Fig. 3 PERMANOVAs, which tested the effect of the “host-species” variable, were invariably lower for skin mucus communities than for gut communities. These results suggest that species-specific factors particularly drive gut communities rather than skin mucus communities.

204

205 Interestingly, the Thetayc Dissimilarity Index values (TDI) (data not shown) suggest a
206 potential signal of host phylogenetic history for gut microbiotas, but not for skin mucus
207 microbiotas. Significant interspecific variations in the phylogenetic structures of gut
208 microbiotas (p-value < 0.001, F = 3.21.41, df = 108) were related to the host phylogeny:
209 the gut microbiota structures of the two *Serrasalmidae* species (pacus and black
210 piranhas) were more similar to each other (TDI = 0.661 ± 0.004 S.E.) than to the gut
211 microbiota structure of the *Cichlidae* species (flag cichlids): between flag cichlids and
212 black piranhas, TDI = 0.783 ± 0.004 S.E.; between flag cichlids and pacus, TDI = 0.786
213 ± 0.003 S.E.). This signal was not observed for skin mucus samples: between pacus
214 and black piranhas, TDI = 0.692 ± 0.009 S.E ; between flag cichlids and black piranhas,
215 TDI = 0.576 ± 0.003 S.E.; between flag cichlids and pacus, TDI = 0.66 ± 0.01 S.E.

216

217 **Habitat-specific factors**

218 The Spearman correlation network analysis on Fig. 4 highlights similarities between the
219 phylogenetic structure of samples, and enables visualization of global clustering
220 patterns among all samples. These networks show an increased sampling-site-
221 dependent clustering for skin mucus samples than for gut samples: there was a much
222 higher ratio of intra site edges (i.e. correlation between samples/nodes) among skin
223 mucus communities (81 % of edges) than among gut communities (36 % of edges). The
224 networks on Fig. 5 investigate deeper two known environmental factors driving
225 sampling-site-dependent clustering on teleosts: the structure of surrounding
226 bacterioplankton communities, and the environmental hydrochemical parameters (Table

1). The skin mucus and bacterioplankton communities are particularly interconnected: at all instances on Fig. 5, the percentage of taxa from the bacterioplankton and the host fish microbiome that are involved in significant Spearman correlations is always higher for skin mucus than for gut communities (for all three host species). Skin mucus communities also show an increased sensitivity to hydrochemical parameters: all three host species have interactions between several hydrochemical parameters and skin mucus taxa, but gut taxa only correlated with hydrochemical parameters for one host species (pacu). The most sensitive host species to hydrochemical parameters (all 10 parameters studied) was the flag cichlid. Environmental potassium concentration was the only parameter consistently associated with taxa from the skin mucus of all three host species.

These observations regarding the correlation between host habitat characteristics and skin mucus microbiome structure are also in accordance with the stacked barplots from Fig. 1, the PCoAs from Suppl. Fig. 3, and the Thetayc dissimilarity analyses, which show a higher resemblance between skin mucus and bacterioplankton communities than between gut and free-living communities. Thus, overall, the results consistently suggest that habitat-specific factors are differentially associated to the phylogenetic structure of skin mucus communities rather than to gut communities.

Discussion

The goal of our study was to explore the phylogenetic structure of fish skin mucus and gut microbial communities and measure the extent to which species- and habitat-

specific factors shape the structure of microbial communities from both tissues. Our results showed a significant distinction between the phylogenetic profile and diversity of the bacterial communities in each microbial niche. Skin mucus and bacterioplankton communities were significantly closer in resemblance than gut and free-living communities. Tissue-specific responses showed that species-specific factors mostly modulate gut bacterial communities, while the skin mucus microbiome is predominantly associated to environmental physicochemistry and bacterioplankton community structure.

Distinct phylogenetic structures in each microbial niche

Microbial assemblages differ along fine scale environmental gradients present in their host organism, with distinct communities associated with different parts of the host's body (6, 16, 17). Likewise, we observed that the three microbial habitats documented in this investigation (skin mucus, gut, and environmental water) showed significantly different microbial community taxonomic structures for all three species. We also showed a higher alpha diversity of skin mucus samples, than gut and bacterioplankton samples (Fig. 1) (except for the skin mucus of black piranhas at site MPWS2_3). Overall, the disparity between host-associated and free-living microbiotas, suggests that fish guts and skin mucus microbiota compositions are not simple reflections of the microbial assemblages in their habitat, but likely result from selective pressures which differ according to each tissue (1, 32-34). For instance, in fish guts, microbial symbionts with the best fitness are usually specialists, which thrive in a stable (buffered) environment and play key roles in the hosts nutrition by facilitating degradation and

273 assimilation of specific compounds found in the hosts' diet (reviewed in 35). While
274 community composition varies among fish species, the most abundant phyla typically
275 found in the gut of freshwater fishes are *Proteobacteria*, *Actinobacteria*, *Bacteroidetes*,
276 *Firmicutes*, and *Fusobacteria* (15, 20, 36-40). In our study, *Clostridia* from the
277 *Firmicutes* was an important discriminant feature of the gut microbiome for all three host
278 species. Most *Clostridia* are obligate anaerobes that ferment plant polysaccharides (41).
279 Interestingly, their presence in the guts of black piranha may suggest the ingestion of
280 plant material from this once taught strictly piscivorous species. Studies generally show
281 that gut community composition and diversity is trophic level dependent (42, 43): piscine
282 gut microbiome diversity tends to decrease from herbivores to omnivores, with the
283 lowest diversity in found in carnivores (44). However, our results did not show such
284 decrease in alpha diversity in relation to fish diet (Fig. 1). This result could be
285 associated to the fact that fishes were sampled in the dry season, a period of fasting for
286 several Amazonian fish species (45, 46). Thus, diet may not be a major driver of gut
287 microbiome structure during this period.
288
289 In contrast to gut symbionts, skin mucus bacteria are mostly generalists, which tolerate
290 a variable environment and play roles in the defense against opportunistic pathogens
291 (19) via colonization resistance mechanisms (47). Typically, skin mucus communities
292 are dominated by *Proteobacteria*, and other less abundant phyla such as *Bacteroidetes*,
293 *Actinobacteria*, *Firmicutes*, and *Verrucomicrobia* (15, 20, 48, 49). In our study,
294 *Proteobacteria* was the most abundant phyla in all skin mucus communities, and was
295 mostly composed of *Gammaproteobacteria* (Fig. 2d,e,f). Two taxa that were among the

296 10 most significant discriminant features of the skin mucus tissue were generalists
297 found on all three host species: *Enhydrobacter* sp. and Cl500-29 marine group sp. (Fig.
298 2.j,k,l). *Enhydrobacter aerosaccus*, the only species of the *Enhydrobacter* genus, is
299 commonly found on fishes. This species is part of the butterflyfish gill core microbiome
300 (50), the rainbow trout swim-bladder (51) and gut (52) microbiome, and the gut
301 microbiome of the swamp eel (53). Our study showed that for three Amazonian fish
302 species, *Enhydrobacter* sp. was not associated with the gut, but rather with the skin
303 mucus tissue. Cl500-29 marine group sp. has been mostly reported in bacterioplankton
304 samples from a variety of freshwater and saltwater habitats (54-56), thus it can tolerate
305 a great range of hydrochemical parameters. This genus has also been shown to be a
306 generalist in terms of utilization of different carbon compounds (56). At this moment,
307 very little is known on the functional repertory of *Enhydrobacter aerosaccus* and Cl500-
308 29 marine group sp.. However, the ubiquity of these taxa in all skin mucus communities
309 of the three host species from our study suggests that their recruitment on skin mucus is
310 positively selected among the large pool of environmental taxa. Thus, they potentially
311 play essential roles in the skin mucus microbiome, which merit further investigation.
312
313 **Species-specific factors and the gut microbiome.**
314 Numerous studies have documented an effect of the host fish species on gut microbial
315 communities (7, 9). Our results show that phylogenetic structure of gut microbiotas were
316 especially correlated to host fish ancestry (skin mucus microbiota often showed a
317 significant, but weaker signal (Fig. 3)). The gut microbiota structure of the two
318 *Serrasalmidae* hosts (*S. rhombeus* and *M. duriventre*) were more similar to each other

than to the taxonomic structure of the *Cichlidae*'s (*M. festivus*) gut microbiota, even given strikingly different diets (i.e. planktivorous *versus* piscivorous). Similarly, some mammalians with atypical diets for their clade, such as the strictly herbivorous panda bear (57) and myrmecophagous mammalians (58), harbor gut microbial communities that are more similar to those of their close relatives than to those of other mammalians with analogous diets. This phylogenetic inertia highlights the role of the host genotype (phylogeny) in constraining the range of variation of the gut microbiota in response to different diets (59). This phenomenon, where closely related species tend to host more similar microbiomes than evolutionary distant species, is termed "phylosymbiosis" (60-62). Phylosymbiosis can be driven by various factors including phenotypic divergence between host species that are phylogenetically distinct (63), the host's diet (8, 9, 11, 64-66), the coevolution of specific bacteria with their host (cophylogeny) (67), or by vertical transmission of bacterial lineages across hosts generations (17).

Interestingly, a recent study by Riiser *et al*. (2020) (68) conducted on cod (*Gadus morhua*) and pout (*Trisopterus esmarkii*) populations in the North Atlantic show results that contrast with the species-specific and phylosymbiosis patterns observed for the gut microbiome in our study. First, they show that different host ecotypes have more influence than the host species factor in shaping gut microbiome. Then, they show high similarity in the microbiome of cod and pout lineages that diverged over 20M years ago. These results show that the fraction of microbiome variance that is explained by environmental versus species-specific factors can vary significantly according to the host species and the type of environment exploited by these species. For instance, the

342 availability of diverse nutrient sources, and the environmental heterogeneity
343 characteristic of tropical freshwater streams in the Amazon (25) are factors that may
344 contribute to enhance the specialization of ecological niches, and thus strengthen the
345 species- and habitat-specific responses observed in our study.
346
347
348
349 **Habitat-specific factors and the skin mucus microbiome.**
350 Ecologically mediated selection is a key driver of speciation (69-71), and it has been
351 shown that the host fish genotype is correlated to the phylogenetic structure of the host
352 microbiome (2, 7, 8, 10). Thus, the effect of habitat-specific factors on the structure of
353 the bacterial communities in this study could be related to the following factors: (**i**)
354 divergent selection pressures between sampling locations, leading to adaptive
355 divergence (72); (**ii**) possible low gene flow level between sampling locations (73), which
356 could lead to genetic drift of the host (74), which then modifies the microbiota of the fish
357 host under selection; or (**iii**) habitat-specific diet (especially for detritivorous fishes
358 feeding on sediments such as flag cichlids (11)).
359
360 However, we documented distinct responses to habitat-specific factors: skin mucus
361 microbial communities were more affected by the "Sampling site" factor than gut
362 communities (Fig. 4 and Fig. 5). These differences may be associated with the stability
363 of abiotic conditions on each tissue. The fish gut is a highly stable environment
364 compared to the skin mucus, which is constantly exposed to environmental water with

365 variable physicochemical parameters. The intestinal pH remains circumneutral although
366 being continuously exposed to acidic pH from gastric secretions (75), thus highlighting
367 the high efficiency of pH buffering from alkaline secretions in the lower gastro-intestinal
368 tract. Therefore, the stability of abiotic factors in the host gut promotes specialization of
369 the microbiota (76, 77). On the opposite, the high variability in terms of physicochemical
370 parameters in environmental water, to which the skin mucus is exposed, favors a more
371 dynamic skin mucus community (16). Sylvain *et al*. (2016) (16), exposed the tambaqui
372 fish (*Colossoma macropomum*) to extreme environmental conditions (pH 4.0) and found
373 that gut microbiotas showed a stronger resistance and resilience capacity than skin
374 mucus microbial communities. A review from Nemergut *et al.* (2013) (73) suggests that
375 the resilience of microbial communities to disturbances (biotic or abiotic) is related to the
376 duration and severity of the disturbance, community diversity, abiotic factors in the
377 environment and disturbance history (77-79). Neutral processes (i.e. drift and dispersal)
378 are expected to increase shortly following disturbance events (e.g. when a fish moves
379 from a river tributary to another with different physicochemical parameters), while there
380 is an increase in the effects of selection in stable environments (80) such as fish guts.
381 This association between disturbance, neutral and selective processes might explain
382 the distinct responses to habitat-specific factors that we observed for skin mucus and
383 gut samples.
384
385 Overall, our results have different implications. First, the skin mucus microbiome should
386 be used, rather than the gut, to investigate bacterial biomarkers of ecosystem
387 perturbance in the wild. Our study showed that in contrast to the teleost gut microbiome,

skin mucus communities are highly environment dependent. Thus, skin mucus samples will lead to a more sensitive and exact evaluation on the type of environment (e.g. perturbed or pristine, saltwater or freshwater) from which a fish sample originates. Second, due to its stability, its conserved structure, and its association to fish diets, the gut microbiome is better suited for studies on the drivers of phylosymbiosis, or on the co-evolution of fish and their symbionts over evolutionary scales.

Conclusion

Our study is one of the very first to explore the phylogenetic structure of microbiotas from wild fishes of the Amazon basin. Our analysis of microbial communities at multiple scales showed a clear distinction between the phylogenetic structure of the microbiome in different host tissues. They also highlighted a close resemblance and dense interaction network between skin mucus and bacterioplankton communities. We showed that skin mucus communities, in contrast to gut communities, were highly correlated to environmental hydrochemical parameters, and thus mostly relied on habitat-dependent factors, making them poor candidates for phylosymbiosis analyses but sensitive targets to develop database of environmental biomarkers. We suggest that future investigations use a shotgun metagenomic or metatranscriptomic approach to assess if species- and habitat-specific factors also affect the functional repertory of fish microbiomes.

Methods

Ethics approval

410 This project and protocol were approved by the Ethics Committee for the Use of
411 Animals of INPA (number 026/2015 as of Dec 18th, 2015). All methods were carried out
412 in accordance with the approved guidelines.
413
414 **Fish sampling**
415 The fish sampling was done between 05/11/2015 and 25/11/2015 at six different sites in
416 freshwater streams of the upper Brazilian Amazon, close to Manaus (AM, Brazil): (1)
417 Catalão Lake; (2) two sites in the Solimões River near Jacurutu Island; (3) the
418 Manacapuru River; and (4) two sites in Anavilhanas National Park. Map of sampling
419 sites is in Supplementary Material (Suppl. Fig. 1). 10 flag cichlids (FC) and 4 black
420 piranhas (BP) were collected at site BWS1; 3 FC and 7 BP at site BWS2; 10 FC, 2 BP,
421 and 18 pacus (PAC) at site MPWS2_3; 6 BP and 10 PAC at site WWS1; 5 BP and 10
422 PAC at site WWS2; 10 FC, 10 BP, and 10 PAC at site Catalão (N total = 115 fish). At
423 the moment, most of other studies on fish microbiota have only used 3-5 fish per site or
424 species (reviewed in 7). Fishing was done with fixed gillnets. If specimens of the
425 targeted species were in contact with other fishes in the net – e.g. if two fishes were in
426 adjacent net mesh – these specimens were discarded to avoid cross-contamination of
427 the skin mucus bacterial communities. Skin mucus bacterial composition can be
428 affected by net-fishing, due to a modification of mucus secretion and composition (81,
429 82). Thus, we carefully checked the nets every 15 minutes to minimize the time captive
430 fishes spent in the net and to reduce sources of contamination from the net itself. After
431 capture, the skin mucus of all fishes was immediately sampled by gently rubbing a
432 sterile cotton swab on ≈ 50% of the total surface on the right side of each fish. Whole

433 fishes and the skin mucus samples were then stored on ice with liquid Nitrogen until
434 arrival at the *Laboratório de Ecofisiologia e Evolução Molecular* of INPA, where all
435 samples were kept frozen at -80°C until dissection. There were about 4 hours between
436 fish collection/flash freezing and dissection at the laboratory, and all fishes were
437 processed the same way.
438
439 2 L of water was also sampled at each site in sterile Nalgene™ bottles to characterize
440 bacterioplankton community. Water samples were taken at 30 cm below the surface.
441 The species collected in this study are found at these depths (25). The water samples
442 were collected shortly before leaving the sampling sites to minimize time between
443 collection and filtration. They were immediately stored in a large cooler full of ice and
444 liquid nitrogen until arrival at the laboratory. There were about 4 hours between
445 collection and filtration. Water samples were filtered on 0.2 µm membranes
446 (Nucleopore©) using a Masterflex Easy-Load® II peristaltic pump from Cole-Parmer®.
447 Post-filtration, the membranes were stored dry at -80°C.
448
449 All fishes were measured, weighed, and dissected. Student's T tests showed no
450 significant difference (p-values > 0.05) between the average length and weight of all
451 three species between different sampling sites. All fish were dissected with sterile
452 instruments, under a flame, to isolate a section comprising midgut (right after pyloric
453 caeca) and hindgut. Complete midgut and hindgut were pooled together in one sample
454 for DNA extractions for each fish. Fish were sampled during dry season, which
455 corresponds to a period of fasting for Amazonian fishes (45, 46). Thus, the isolated gut

456 sections of all host species were small, empty of intestinal content, and therefore, gut
457 samples mostly represented mucosal flora rather than flora associated with digesta. Gut
458 samples were kept at -80°C along with skin mucus and water samples until DNA
459 extraction.
460
461 **Hydrochemical parameters**
462 Water parameters for each sampling site are found in Table 1. Temperature, pH, and
463 dissolved oxygen % were measured at each sampling site. Then, 2 L of water were
464 sampled at 30 cm below surface, and brought back to the laboratory on ice for further
465 analysis. These water samples were collected at the same time and depth than the
466 water samples used to characterize bacterioplankton. Carbonate hardness was
467 calculated from the Ca^{2+} and Mg^{2+} concentrations and alkalinity was assessed using the
468 method APHA Standard Methods for Examination of Water and Wastewater method
469 2320 (1992). The ionic composition (Na^{+}, Mg^{2+}, Ca^{2+}, K^{+}) of each sample was analyzed
470 using flame atomic absorption spectroscopy (Perkin-Elmer model 3100). The
471 concentration of Cl^{-} was measured using the colorimetric method as described by
472 Clarke (1950) (83). Samples for dissolved organic carbon analysis were first filtered on
473 0.45 um Sartorius filters and then were analyzed using a total carbon analyzer (Apollo
474 9000 combustion TOC analyzer: ©Teledyne Tekmar). The TOC machine was calibrated
475 according to the manufacturer's instructions, using primary standard grade potassium
476 hydrogen phthalate (KHP).
477
478 **Data availability**

The sequence files are available from the Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra), BioProjectID: PRJNA574456. The scripts used for the statistical analysis of sequence variants (dada2 pipeline), the output ASV table, taxonomy data, mapping (metadata) file and R markdowns used during this project are freely available from the Open Science Network platform (URL: https://osf.io/h7skx/).

Preparation of 16S amplicon libraries

DNA extraction of skin mucus samples, whole guts and 0.2 µm membranes from water samples was performed using DNeasy® Blood and Tissue Kit from QIAGEN according to the manufacturer's instructions. Extracted DNA from guts, skin mucus, and water was stored at -80°C until amplification. The fragment V3-V4 (≈ 500 base pairs) of the 16S rRNA was amplified by polymerase chain reaction (PCR) using the forward primer 347F of sequence 5'-GGAGGCAGCAGTRRGGAAT-3' and the reverse primer 803R of sequence 5'-CTACCRGGGTATCTAATCC-3' (84). All PCR reactions were performed according to the manufacturer's instructions of Q5® High-Fidelity DNA Polymerase from New England BioLabs® Inc. PCR program: (1) 30 sec 98°C; (2) 10 sec 98°C; (3) 30 sec 64°C; (4) 20 sec 72°C; (5) 2 min at 72°C; 35 amplification cycles total. To reduce PCR bias, and to increase precision in the assessment of microbial community composition and diversity, PCRs were done in triplicates. Amplified DNA was purified according to the manufacturer's instructions with AMPure beads (Beckman Coulter Genomics) to eliminate primers, dimers, proteins and phenols. All three PCR products for each sample were kept separate for post-PCR DNA purification but were pooled together before sequencing. Post-PCR DNA concentration and quality were assessed on

Nanodrop and by electrophoresis on [1.5%] agarose gels. After purification, Multiplex Sequencing was performed using the MiSeq platform from Illumina®, by the *Plateforme d'analyses génomiques* at the *Institut de Biologie Intégrative et des Systèmes* (IBIS) of *Université Laval*. A total of 232 samples were sequenced: 33 FC guts, 33 BP guts, 47 PAC guts, 33 FC skin mucus, 34 BP skin mucus, 46 PAC skin mucus, and 6 water samples.

Processing of 16S sequences

The analysis of amplicon sequences was done at the *Institut de Biologie Intégrative et des Systèmes* (IBIS) at *Université Laval*. After sequencing, 4 247 979 sequences were obtained (mean of 18 310 sequences per sample). The demultiplexed fastq sequence files were processed through QIIME2 (85), and the dada2 tool (86) was used for Amplicon Sequence Variant (ASV) picking. Quality control of reads was processed through the filterAndTrim function by using the following parameters : 270 for the read truncation length, 2 as the phred score threshold for total read removal, and a maximum expected error of 2 for forward reads and 4 for reverse reads. The filtered reads were then fed to the error rate learning, dereplication and ASV inference steps using the functions learnErrors, derepFastq and dada, which are all from the dada2 pipeline (86). Chimeric sequences were removed using the removeBimeraDenovo function (86) with the “consensus” method parameter. Taxonomic classification was done through the assignTaxonomy function (86) using the SILVA v. 132 reference database. A rarefaction analysis of observed "species" counts and Shannon diversity (according to sampling depth) for each sample type (Suppl. Fig. 2) showed that four samples had

conspicuously lower sampling depth than the others: WWS1 = 199 reads, BWS1 = 381 reads, 10.F = 474 reads, and 32.F = 9 reads. These four samples (< 5000 reds) were discarded, bringing total number of samples to 228, and lowest sampling depth to 5223 reads. The average Good's coverage index for all samples was 0.9917 ± 0.0005 (S.E.).

Statistical analysis of sequence variants

To compare the taxonomical structures of the bacterial communities from the three microbial niches considered (Fig. 1), we used stacked barplots built on phyloseq (87), from the relative abundance of the 12 most abundant classes (all ASVs were used for the analysis). We used the R package btools to calculate Faith's Phylogenetic Diversity metric (88) for each sample type (Fig. 1). Then, we used Principal Components Analyses (PCA) based on weighed Unifrac distances (77) using the R package phyloseq (87), to visualize sample clustering between the different tissues (Fig. 2a,b,c, and Suppl. Fig. 3) and the different host species (Fig. 3). Then, the significance of these clusterings were assessed with p-values from PERMutational ANalyses Of VAriances (PERMANOVAs), computed with 10 000 permutations using the R package *vegan* (89, 90) and a distance matrix of weighed Unifrac indexes between samples (Fig. 1a, Fig. 2a,b,c, and Fig. 3). Finally, to investigate further the dissimilarity between the taxonomic structures of bacterial communities from different tissues / host species, we also used Thetayc β-diversity measures, or Thetayc Dissimilarity Index (TDI) (29). We chose to use the Thetayc index as it is a function of species proportions from both the shared and non-shared species. In addition, in the Thetayc index, the shared species proportions in each community are compared one-to-one (instead of a sum of the

abundances of all shared species in the Bray-Curtis index, which gives no indication on which species are shared, and if the abundances of the shared species are similar or not). As a result, the Thetayc index places more weight on those shared species, which have similar species proportions in both communities.

We used LEfSe tests (Linear Discriminant Analysis Effect Size) (30) to identify the taxonomic groups which abundances varied significantly between different tissues (Fig. 2d,e,f). This test uses non-parametric factorial Kruskal-Wallis sum-rank test to detect discriminant features (taxonomic groups) with significant differential abundance between tissues. Biological significance of these features is subsequently investigated by completing pairwise tests between the abundance of the selected features, using an unpaired Wilcoxon rank-sum test. Finally, LEfSe uses Linear Discriminant Analysis (LDA score) to quantify the effect size of each differentially abundant feature. The threshold for the LDA parameter was 2.0 (the default value).

We also produced heatmaps based on the relative abundance of the 100 most abundant ASVs to investigate how samples from different tissues cluster on a phylogenetic tree (Fig. 2g,h,i). Then, we used the package metagenomeSeq (31) as another method (in addition to LEfSe) to identify discriminant features at the ASV level. metagenomeSeq's LDA approach is more sensitive to significant differences, but less sensitive to potential confounding factors than the LEfSe approach (31). The 10 discriminant ASVs with the highest LDA score for each tissue were highlighted in the barplots of Fig. 2j,k,l.

571

572 To assess the effect of habitat-specific factors on the clustering of samples between
573 both host-associated tissues, we computed Spearman correlations between all samples
574 for each tissue, based on co-abundance patterns of the ASVs in the samples.
575 Correlations kept for downstream analysis had a significant Spearman correlation value
576 < 0.05 after Bonferroni correction. Two correlations networks (skin mucus and gut
577 communities) were constructed using Cytoscape version 3.2.1 (91) to plot significant
578 interactions (Fig. 4). The nodes of the networks each represent one sample and the
579 edges (i.e. connections between nodes) are attributed to significant Spearman
580 correlations between nodes.

581

582 Then, to investigate further the correlations between bacterioplankton, hydrochemical
583 parameters, and host-associated communities, we computed Spearman correlations
584 based on co-abundance patterns of the ASVs in the samples. Correlations kept for
585 downstream analysis had a significant Spearman correlation value < 0.05 after
586 Bonferroni correction. Correlations networks were constructed using Cytoscape version
587 3.2.1 (91) to plot significant interactions (Fig. 5). The nodes of the networks represent
588 taxa or hydrochemical parameters and the edges (i.e. connections between nodes) are
589 attributed to significant Spearman correlations between nodes.

590

591 ## References

1. Wong SD, Rawls JF. 2012. Intestinal microbiota composition in fishes is influenced by host ecology and environment. *Mol Ecol* **21**: 3100-3102. doi:10.1111/j.1365-294X.2012.05646.
2. Nayak SK. 2010. Role of gastrointestinal microbiota in fish. *Aquaculture Research* **41**: 1553-1573. doi:10.1111/j.1365-2109.2010.02546.
3. Lescak EA, Milligan-Myhre KC. 2017. Teleosts as model organisms to understand host-microbe interactions. *J Bacteriol* **199**:e00868-16. doi:10.1128/JB.00868-16.
4. Smriga S, Sandin SA, Azam F. 2010. Abundance, diversity, and activity of microbial assemblages associated with coral reef fish guts and feces. *Fems Microbiology Ecology* **73**: 31-42. doi:10.1111/j.1574-6941.2010.00879.
5. Llewellyn MS, Boutin S, Hoseinifar SH, Derome N. 2014. Teleost microbiotas: the state of the art in their characterization, manipulation and importance in aquaculture and fisheries. *Frontiers in Microbiology* **5**: 17. doi:10.3389/fmicb.2014.00207.
6. van Kessel MA, Dutilh BE, Neveling K, Kwint MP, Veltman JA, Flik G, Jetten MS, Klaren PH, Op den Camp HJ. 2011. Pyrosequencing of 16S rRNA gene amplicons to study the microbiota in the gastrointestinal tract of carp (*Cyprinus carpio* L.). *Amb Express* **1**. doi:10.1186/2191-0855-1-41.
7. Ghanbari M, Kneifel W, Domig KJ. 2015. A new view of the fish gut microbiota: Advances from next-generation sequencing. *Aquaculture* **448**: 464-475. doi:10.1016/j.aquaculture.2015.06.033.

8. Sullam KE, Essinger SD, Lozupone CA, O'Connor MP, Rosen GL, Knight R, Kilham SS, Russell JA. 2012. Environmental and ecological factors that shape the gut bacterial communities of fish: a meta-analysis. *Molecular Ecology* **21**: 3363-3378. doi:10.1111/j.1365-294X.2012.05552.
9. Miyake S, Ngugi DK, Stingl U. 2015. Diet strongly influences the gut microbiota of surgeonfishes. *Molecular Ecology* **24**: 656-672. doi:10.1111/mec.13050.
10. Boutin S, Sauvage C, Bernatchez L, Audet C, Derome N. Inter Individual Variations of the Fish Skin Microbiota: Host Genetics Basis of Mutualism? *Plos One* **9**. doi:10.1371/journal.pone.0102649.
11. Wu S, Wang G, Angert ER, Wang W, Li W, Zou H. 2012. Composition, Diversity, and Origin of the Bacterial Community in Grass Carp Intestine. *Plos One* **7**. doi:10.1371/journal.pone.0030440.
12. Llewellyn MS, Leadbeater S, Garcia C, Sylvain FE, Custodio M, Ang KP, Powell F, Carvalho GR, Creer S, Elliot J, Derome N. 2017. Parasitism perturbs the mucosal microbiome of Atlantic Salmon. *Sci Rep* **7**:43465. doi: 10.1038/srep43465.
13. Pérez T, Balcázar JL, Ruiz-Zarzuela I, Halaihel N, Vendrell D, de Blas I, Múzquiz JL. 2010. Host-microbiota interactions within the fish intestinal ecosystem. Mucosal Immunol **3**(4):355-60. doi: 10.1038/mi.2010.12
14. Stephens WZ, Burns AR, Stagaman K, Wong S, Rawls JF, Guillemin K, Bohannan BJ. 2016. The composition of the zebrafish intestinal microbial community varies across development. *ISME J* **10:** 644–654. doi: 10.1038/ismej.2015.140.

15. Llewellyn MS, McGinnity P, Dionne M, Letourneau J, Thonier F, Carvalho GR, Creer S, Derome N. 2016. The biogeography of the atlantic salmon (*Salmo salar*) gut microbiota. *Isme Journal* **10**: 1280-1284. doi:10.1038/ismej.2015.189.

16. Sylvain FÉ, Cheaib B, Llewellyn M, Gabriel Correia T, Barros Fagundes D, Luis Val A, Derome N. 2016. pH drop impacts differentially skin and gut microbiota of the Amazonian fish tambaqui (*Colossoma macropomum*). *Scientific Reports* **6**: 10. doi:10.1038/srep32032.

17. Sylvain FE, Derome N. 2017. Vertically and horizontally transmitted microbial symbionts shape the gut microbiota ontogenesis of a skin-mucus feeding discus fish progeny. *Scientific Reports* **7**: 14. doi:10.1038/s41598-017-05662-w.

18. Chiarello M, Auget JC, Bettarel Y, Bouvier C, Claverie T, Graham AJ, Rieuvilleneuve F, Sucré E, Bouvier T, Villéger S. 2018. Skin microbiome of coral reef fish is highly variable and driven by host phylogeny and diet. *Microbiome* **6:** 147. doi: 10.1186/s40168-018-0530-4.

19. Boutin S, Bernatchez L, Audet C, Derome N. 2013. Network Analysis Highlights Complex Interactions between Pathogen, Host and Commensal Microbiota. *Plos One* **8**: 16. doi:10.1371/journal.pone.0084772.

20. Sylvain FE, Holland A, Audet-Gilbert É, Luis Val A, Derome N. 2019. Amazon fish bacterial communities show structural convergence along widespread hydrochemical gradients. *Molecular Ecology* **28**(15):3612-3626. doi: 10.1111/mec.15184.

21. Legrand TPRA, Catalano SR, Wos-Oxley ML, Stephens F, Landos M, Bansemer MS, Stone DAJ, Qin JG, Oxley APA. 2018. The Inner Workings of the Outer

Surface: Skin and Gill Microbiota as Indicators of Changing Gut Health in Yellowtail Kingfish. *Frontiers in Microbiology* **8**. doi:10.3389/fmicb.2017.02664.

22. Pratte ZA, Besson M, Hollman RD, Stewart FJ. 2018. The gills of reef fish support a distinct microbiome influenced by species-specific factors. *Appl Environ Microbiol* 84:e00063-18.
23. Hedges SB, Marin J, Suleski M, Paymer M, Kumar S. 2015. Tree of life reveals clock-like speciation and diversification. *Mol Biol and Evol* **32**(4): 835–845.
24. Pires THS, Campos DF, Ropke CP, Sodre J, Amadio S, Zuanon J. 2015. Ecology and life-history of *Mesonauta festivus*: Biological traits of a broad ranged and abundant Neotropical cichlid. *Environmental Biology of Fishes* **98**: 789–799.
25. Van der Sleen A, Albert JS. 2018. Field guide to the fishes of the Amazon, Orinoco and Guianas. Oxford University Press.
26. Ropke CP, Ferreira E, Zuanon J. 2014. Seasonal changes in the use of feeding resources by fish in stands of aquatic macrophytes in an Amazonian floodplain, Brazil. *Environmental Biology of Fishes* **97**: 401-414. doi:10.1007/s10641-013-0160-4.
27. Goulding M, Carvalho ML, Ferreira EG. 1988. Rio Negro, Rich Life in Poor Water. The Hague: SPB Academic Publishing.
28. Oliveira C, Avelino GS, Abe KT, Mariguela TC, Benine RC, Ortí G, Vari RP, Corrêa e Castro RM. 2011. Phylogenetic relationships within the speciose family *Characidae* (*Teleostei: Ostariophysi: Characiformes*) based on multilocus analysis and extensive ingroup sampling. *Bmc Evolutionary Biology* **11**. doi:10.1186/1471-2148-11-275.

29. Yue JC, Clayton MK. 2005 A similarity measure based on species proportions. *Communications in Statistics-Theory and Methods* **34**: 2123-2131. doi:10.1080/sta-200066418.

30. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. 2011. Metagenomic biomarker discovery and explanation. *Genome Biology* **12**: R60. doi: 10.1186/gb-2011-12-6-r60.

31. Paulson JN, Stine OC, Bravo HC, Pop M. 2013. Differential abundance analysis for microbial marker-gene surveys. *Nat Methods* **10**(12): 1200–1202. doi:10.1038/nmeth.2658.

32. Leser TD, Molbak L. 2009. Better living through microbial action: the benefits of the mammalian gastrointestinal microbiota on the host. *Environmental Microbiology* **11**: 2194-2206. doi:10.1111/j.1462-2920.2009.01941.

33. Bevins CL, Salzman NH. 2011. The potter's wheel: the host's role in sculpting its microbiota. *Cellular and Molecular Life Sciences* **68**: 3675-3685. doi:10.1007/s00018-011-0830-3.

34. Bolnick DI, Snowberg LK, Hirsch PE, Lauber CL, Org E, Parks B, Lusis AJ, Knight R, Caporaso JG, Svanbäck R. 2014. Individual diet has sex-dependent effects on vertebrate gut microbiota. *Nature Communications* **5**. doi:10.1038/ncomms5500.

35. Ray AK, Ghosh K, Ringo E. 2012. Enzyme-producing bacteria isolated from fish gut: a review. *Aquaculture Nutrition* **18**: 465-492. doi:10.1111/j.1365-2095.2012.00943.

36. Desai AR, Links MG, Collins SA, Mansfield GS, Drew MD, Van Kessel AG, Hill JE. 2012. Effects of plant-based diets on the distal gut microbiome of rainbow trout (*Oncorhynchus mykiss*). *Aquaculture* **350**: 134–142. doi: 10.1016/j.aquaculture.2012.04.005.

37. Nielsen S, Wilkes Walburn J, Verges A, Thomas T, Egan S. 2017. Microbiome patterns across the gastrointestinal tract of the rabbitfish *Siganus fuscescens*. *PeerJ* **5**. doi: e3317.10.7717/peerj.3317.

38. Burgos FA, Ray CL, Arias CR. 2018. Bacterial diversity and community structure of the intestinal microbiome of channel catfish (*Ictalurus punctatus*) during ontogenesis. *Syst. Appl. Microbiol* **41**: 494–505. doi: 10.1016/j.syapm.2018.04.006.

39. de Bruijn I, Liu Y, Wiegertjes GF, Raaijmakers JM. 2018. Exploring fish microbial communities to mitigate emerging diseases in aquaculture. *FEMS Microbiol Ecol* **94**:fix161.

40. Wang AR, Ran C, Ringo E, Zhou ZG. 2018. Progress in fish gastrointestinal microbiota research. *Rev Aquacu* **10**: 626–640. doi: 10.1111/raq.12191.

41. Boutard M, Cerisy T, Nogue PY, Alberti A, Weissenbach J, Salanoubat M, Tolonen AC. 2014. Functional Diversity of Carbohydrate-Active Enzymes Enabling a Bacterium to Ferment Plant Biomass. *PLoS Genet* **10**(11): e1004773. doi: 10.1371/journal.pgen.1004773

42. Givens CE, Ransom B, Bano N, Hollibaugh JT. 2015. Comparison of the gut microbiomes of 12 bony fish and 3 shark species. *Mar Ecol Prog Ser* **518**: 209–223. doi: 10.3354/meps11034.

43. Miyake S, Ngugi DK, Stingl U. 2015. Diet strongly influences the gut microbiota of surgeonfishes. *Mol Ecol* **24**: 656–672. doi: 10.1111/mec.13050.

44. Wang AR, Ran C, Ringo E, Zhou ZG. 2018. Progress in fish gastrointestinal microbiota research. *Rev Aquacu* **10**: 626–640. doi: 10.1111/raq.12191

45. Lowe-McConnell RH. 1975. Fish Communities in Tropical Freshwaters. Their Distribution, Ecology and Evolution. Longman, London, 337 p.

46. Lowe-McConnell RH. 1987. Ecological Studies in Tropical Fish Communities. Cambridge University Press, Cambridge.

47. Vanderwaaij D, Vanderwaaij BD. 1990. The colonization resistance of the digestive-tract in different animal species and in man - a comparative-study. *Epidemiology and Infection* **105**: 237-243. doi: 10.1017/s0950268800047841.

48. Merrifield DL, Rodiles A. 2015. "The fish microbiome and its interactions with mucosal tissues" in *Mucosal Health in Aquaculture*, eds Beck BH, Peatman E, editors. (Oxford, UK: Elsevier).

49. Mohammed HH, Arias CR. 2015. Potassium permanganate elicits a shift of the external fish microbiome and increases host susceptibility to columnaris disease. *Vet Res* **46**:82. doi: 10.1186/s13567-015-0215-y.

50. Reverter M, Sasal P, Tapissier-Bontemps N, Lecchini D, Suzuki M. 2017. Characterisation of the gill mucosal bacterial communities of four butterflyfish species: a reservoir of bacterial diversity in coral reef ecosystems. *FEMS Microbiol Ecol*. 93(6). doi: 10.1093/femsec/fix051.

51. Villasante A, Ramírez C, Rodríguez H, Catalán N, Díaz O, Rojas R, Opazo R, Romero J. 2019. In-depth analysis of swim bladder-associated microbiota in rainbow trout (*Oncorhynchus mykiss*). *Sci Rep* **9:** 8974. doi: 10.1038/s41598-019-45451-1.

52. Betiku OC, Yeoman CJ, Gaylord TG, Americus B, Olivo S, Duff GC, Sealey WM. 2018. Water system is a controlling variable modulating bacterial diversity of gastrointestinal tract and performance in rainbow trout. *PLoS One* **13**(4): e0195967. doi: 10.1371/journal.pone.0195967.

53. Chen X, Fang S, Wei L, Zhong Q. 2019. Systematic evaluation of the gut microbiome of swamp eel (*Monopterus albus*) by 16S rRNA gene sequencing. *PeerJ* 7: e8176. doi: 10.7717/peerj.8176.

54. Yu S, He R, Song A, Huang Y, Jin Z, Liang Y, Li Q, Wang X, Müller WEG, Cao J. 2019. Spatial and temporal dynamics of bacterioplankton community composition in a subtropical dammed karst river of southwestern China. *MicrobiologyOpen* **8**: e849. doi: 10.1002/mbo3.849.

55. Yang C, Wang Q, Simon PN, Liu J, Liu L, Dai X, Zhang X, Kuang J, Igarashi Y, Pan X, Luo F. 2017. Distinct Network Interactions in Particle-Associated and Free-Living Bacterial Communities during a *Microcystis aeruginosa* Bloom in a Plateau Lake. *Front Microbiol* **8**:1202. doi: 10.3389/fmicb.2017.01202.

56. Gomez-Consarnau L, Lindh MV, Gasol JM, Pinhassi J. Structuring of bacterioplankton communities by specific dissolved organic carbon compounds. *Env Microbiol* **14**:2361–2378. doi: 10.1111/j.1462-2920.2012.02804.x.

773 57. Zhu LF, Wu Q, Dai JY, Zhang SN, Wei FW. 2011. Evidence of cellulose
774 metabolism by the giant panda gut microbiota. *Proceedings of the National*
775 *Academy of Sciences of the United States of America* **108**: 17714-17719.
776 doi:10.1073/pnas.1017956108.
777 58. Delsuc F, Metcalf JL, Parfrey LW, Song SJ, González A, Knight R.
778 2014. Convergence of gut microbiotas in myrmecophagous mammals. *Molecular*
779 *Ecology* **23**: 1301-1317. doi:10.1111/mec.12501.
780 59. Soverini M, Quercia S, Biancani B, Furlati S, Turroni S, Biagi E, Consolandi C,
781 Peano C, Severgnini M, Rampelli S, Brigidi P, Candela M. 2016. The bottlenose
782 dolphin (*Tursiops truncatus*) faecal microbiota. *Fems Microbiology Ecology* **92**.
783 doi: 10.1093/femsec/fiw055.
784 60. Ochman H, Worobey M, Kuo CH, Ndjango JB, Peeters M, Hahn BH, Hugenholtz
785 P. 2010. Evolutionary relationships of wild hominids recapitulated by gut
786 microbial communities. *PLoS Biol* **8**. doi:10.1111/e1000546.
787 61. Phillips CD, Phelan G, Dowd SE, McDonough MM, Ferguson AW, Delton
788 Hanson J, Siles L, Ordóñez-Garza N, San Francisco M, Baker RJ. 2012.
789 Microbiome analysis among bats describes influences of host phylogeny, life
790 history, physiology and geography. *Mol Ecol* **21**:2617–27. doi: 10.1111/j.1365-
791 294X.2012.05568.
792 62. Groussin M, Mazel F, Sanders JG, Smillie CS, Lavergne S, Thuiller W, Alm EJ.
793 2017. Unraveling the processes shaping mammalian gut microbiomes over
794 evolutionary time. *Nat Commun* **8**: 14319. doi: 10.1038/ncomms14319.

63. Brooks AW, Kohl KD, Brucker RM, van Opstal EJ, Bordenstein SR. 2016. Phylosymbiosis: Relationships and Functional Effects of Microbial Communities across Host Evolutionary History. PLOS Biol **14.** doi: 10.1111/e2000225.

64. Givens CE, Ransom B, Bano N, Hollibaugh JT. 2015. Comparison of the gut microbiotas of 12 bony fish and 3 shark species. *Marine Ecology Progress Series* **518**: 209-223. doi:10.3354/meps11034.

65. Franchini P, Fruciano C, Frickey T, Jones JC, Meyer A. 2014. The Gut Microbial Community of Midas Cichlid Fish in Repeatedly Evolved Limnetic-Benthic Species Pairs. *Plos One* **9**. doi:10.1371/journal.pone.0103923.

66. Liu H, Guo X, Gooneratne R, Lai R, Zeng C, Zhan F, Wang W. 2016. The gut microbiota and degradation enzyme activity of wild freshwater fishes influenced by their trophic levels. *Scientific Reports* **6**. doi:10.1038/srep24340.

67. Miyake S, Ngugi DK, Stingl U. 2016. Phylogenetic diversity, distribution, and cophylogeny of giant bacteria (*Epulopiscium*) with their surgeonfish hosts in the Red Sea. *Front Microbiol* **7.** doi: 10.3389/fmicb.2016.00285.

68. Riiser ES, Haverkamp THA, Varadharajan S, Borgan Ø, Jakobsen KS, Jentoft S, Star B. 2020. Metagenomic shotgun analyses reveal complex patterns of intra- and interspecific variation in the intestinal microbiomes of codfishes. *Appl Environ Microbiol* **86**:e02788-19.

69. Schluter D. 2001. Ecology and the origin of species. *Trends in Ecology & Evolution* **16**: 372-380. doi:10.1016/s0169-5347(01)02198.

70. Schluter D, Nagel LM. 1995. Parallel speciation by natural-selection. *American Naturalist* **146**: 292-301. doi:10.1086/285799.

71. Shafer ABA, Wolf JBW. 2013. Widespread evidence for incipient ecological speciation: a meta-analysis of isolation-by-ecology. *Ecology Letters* **16**: 940-950. doi:10.1111/ele.12120.

72. Saint-Laurent R, Legault M, Bernatchez L. 2003. Divergent selection maintains adaptive differentiation despite high gene flow between sympatric rainbow smelt ecotypes (*Osmerus mordax* Mitchill). *Molecular Ecology* **12**: 315-330. doi:10.1046/j.1365-294X.2003.01735.

73. Nemergut DR, Schmidt SK, Fukami T, O'Neill SP, Bilinski TM, Stanish LF, Knelman JE, Darcy JL, Lynch RC, Wickey P, Ferrenberg S. 2013. Patterns and Processes of Microbial Community Assembly. *Microbiology and Molecular Biology Reviews* **77**: 342-356. doi:10.1128/mmbr.00051-12.

74. Nosil P, Funk DJ, Ortiz-Barrientos D. 2009. Divergent selection and heterogeneous genomic divergence. *Molecular Ecology* **18**: 375-402. doi:10.1111/j.1365-294X.2008.03946.

75. Payne AI. 1978. Gut ph and digestive strategies in estuarine grey mullet (*Mugilidae*) and tilapia (*Cichlidae*). *Journal of Fish Biology* **13**: 627-629. doi:10.1111/j.1095-8649.1978.tb03476.

76. Levine JM, D'Antonio CM. 1999. Elton revisited: a review of evidence linking diversity and invasibility. *Oikos* **87**: 15-26. doi:10.2307/3546992.

77. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. 2012. Diversity, stability and resilience of the human gut microbiota. *Nature* **489**: 220-230. doi:10.1038/nature11550.

78. Allison SD, Martiny JBH. 2008. Resistance, resilience, and redundancy in microbial communities. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 11512-11519. doi:10.1073/pnas.0801925105.

79. Griffiths BS, Philippot L. 2013. Insights into the resistance and resilience of the soil microbial community. *Fems Microbiology Reviews* **37**: 112-129. doi:10.1111/j.1574-6976.2012.00343.

80. Ferrenberg S, O'Neill SP, Knelman JE, Todd B, Duggan S, Bradley D, Robinson T, Schmidt SK, Townsend AR, Williams MW, Cleveland CC, Melbourne BA, Jiang L, Nemergut DR. 2013. Changes in assembly processes in soil bacterial communities following a wildfire disturbance. *Isme Journal* **7**, 1102-1111. doi:10.1038/ismej.2013.11.

81. Smith AC, Ramos F. 1976. Occult haemoglobin in fish skin mucus as an indicator of early stress. *J of Fish Biol* **9**(6): 537–541.

82. Shephard KL. 1994. Functions for fish mucus. *Rev in Fish Biol and Fish*, **4**(4): 401–429.

83. Clarke FE. 1950. Determination of chloride in water - improved colorimetric and titrimetric methods. *Analyt Chem* **22**: 553–555.

84. Nossa CW, Oberdorf WE, Yang L, Aas JA, Paster BJ, Desantis TZ, Brodie EL, Malamud D, Poles MA, Pei Z. 2010. Design of 16S rRNA gene primers for 454 pyrosequencing of the human foregut microbiota. *World Journal of Gastroenterology* **16**: 4135-4144. doi:10.3748/wjg.v16.i33.4135.

85. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K,

Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu YX, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS 2nd, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG. 2018. QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. *PeerJ Preprints* **6**: 852-857. doi: 10.1038/s41587-019-0209-9.

86. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nature methods. **13**(7): 581-3. doi: 10.1038/nmeth.3869.

87. McMurdie PJ, Holmes S. 2014. Shiny-phyloseq: Web Application for Interactive Microbiome Analysis with Provenance Tracking. *Bioinformatics* **31**(2): 282–283. doi: 10.1093/bioinformatics/btu616.

88. Faith DP. 1992. Conservation evaluation and phylogenetic diversity. *Biol Conserv* **61**:1–10. doi: 10.1016/0006-3207(92)91201-3.

89. Dixon P. 2003. VEGAN, a package of R functions for community ecology. Journal of Vegetation Science. **14**: 927-930. doi: https://doi.org/10.1111/j.1654-1103.2003.tb02228.

90. Oksanen JF, Blanchet G, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Henry M, Stevens H, Szoecs E, Wagner H. 2019. vegan: Community Ecology Package. R package version 2.5-4. https://CRAN.R-project.org/package=vegan

91. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research* **13**: 2498-2504. doi:10.1101/gr.1239303.

Acknowledgements

We gratefully acknowledge help from all members of the Derome and Val laboratory. Thank you to Maria de Nazare Paula da Silva, Fernanda Dragan, Rogerio Pereira and Raquel Abecassis (INPA) for organizing sampling expeditions. Thank you to Thiago L. A. Nascimento and Derek Campos (INPA) for continuous help during sampling

907 expeditions. Thank you to the three fishermen that helped us on the field: Fabiano Mota,
908 Francisco Fonseca, and Frank Queirzo Coelho Thank you to Jansen Zuanon, Derek
909 Campos, Dinho Heinrichs and Natalia Wagner (INPA) for help with fish identification.
910 Thank you to ICMBIO/Instituto Chico Mendes de Conservação da Biodiversidade for in
911 situ support for fish collection and for issuing the permit to transport biological samples
912 (permit number 29837/11). Thank you to Pierre-Luc Mercier and Camille Desrosiers
913 from U. Laval for support with samples processing in the laboratory. Thank you to Brian
914 Boyle, Luciana Fé, Érica Souza, and Daniel Fagundes (INPA) for molecular analysis
915 and bioinformatics support. This research was part of the ADAPTA project at INPA and
916 was supported by *Ressources Aquatiques Québec*, the *Fonds de recherche du Québec*
917 *– Nature et technologies* (FRQNT) grant to F-ÉS, INCT ADAPTA (CNPq/FAPEAM) and
918 INPA/MCTI grants to ALV, and the Natural Sciences and Engineering Research Council
919 of Canada (NSERC) Discovery grant (grant #6333) to ND. Finally, we would like to
920 thank two anonymous reviewers for their constructive comments on a previous version
921 of this manuscript.

922

923

924 ## Author contributions

925 F-ÉS, ND and ALV designed the experiment. F-ÉS, AH, ND and ALV organized
926 sampling expeditions. F-ÉS, AH and ND sampled fish during field expeditions. F-ÉS,
927 AH, EAG, CL and ND processed samples in the laboratory (fish dissections, DNA

928 extractions and PCRs). F-ÉS, SB, EAG and CL performed 16S sequence analysis. F-
929 ÉS, and ND wrote the manuscript. All authors reviewed the manuscript.

930

931 Additional information

932 Competing Interests Statement

933 The authors declare that they have no competing interests.

934

935 Table

Sites	GPS coordinates		Water parameters									
	S	W	pH	O_2 %	Temp. °C	Na	Cl	Ca	Mg mg/L	K	DOC	Hard.
Catalao	03° 10' 09"	059° 54' 43"	6.71	42.00	31.5	12.56	3.19	2.22	1.96	2.14	7.89	14
BW1	02° 42' 55.1"	060° 44' 52.2"	4.7	89.32	31.2	0.64	0.03	0.01	0.072	1.19	7.26	0
BW2	02° 41' 29.5"	060° 45' 34.1"	4.65	88.50	30.3	0.68	0.02	0.00	0.073	1.19	8.02	0
WW1	03° 16.322'	060° 15.078'	6.9	81.00	32.7	6.02	7.49	0.75	0.665	2.06	3.50	5
WW2	03° 15.490'	060° 13.833'	6.86	76.30	30.7	5.96	7.69	0.74	0.635	2.02	3.44	4
MPWS2_3	03° 18' 54.7"	060° 35' 37.4"	6.15	79.10	32.5	9.66	24.10	1.04	1.035	1.96	3.46	7

936

937 **Table 1**: GPS coordinates and hydrochemical parameters of the six sampling sites of
938 this study. "Temp." stands for temperature, "DOC" for dissolved organic carbon, and
939 "Hard." for carbonate harness.

940

941 Figure legends

942 **Figure 1**: The phylogenetic structure of the skin mucus community is closer to
943 bacterioplankton than to gut community. Stacked barplots display the relative
944 abundance of the 12 most abundant classes in all sample types collected. Boxplot of
945 Faith's phylogenetic alpha diverity index were also added below barplots for each
946 sample type. These analyses were performed on the 4 sampling sites from which we
947 have bacterioplankton samples with adequate sequencing depth (> 5000 sequences):
948 Catalão (a), MPWS2_3 (b), BWS2 (c), and WWS2 (d). All specimens from each species
949 were pooled together for these analyses.
950
951 **Figure 2**: The gut and skin mucus have unique community structures, as shown on the
952 three species sampled. Principal components analyses show the clustering of gut
953 versus skin communities for flag cichlids (a), black piranhas (b), and pacus (c). Results
954 of PERMANOVA tests between groups are indicated on the top left corner of each PCA
955 plot. LEfSe analyses highlight the phyla and classes significantly more abundant in one
956 of the two tissues considered for flag cichlids (d), black piranhas (e), and pacus (f).
957 Heatmaps produced with metagenomeSeq show that samples cluster according to the
958 relative abundance of the 100 most abundant ASVs for flag cichlids (g), black piranhas
959 (h), and pacus (i). Barplots produced with metagenomeSeq highlight the 10 discriminant
960 ASVs with the highest LDA score for each tissue for flag cichlids (j), black piranhas (k),
961 and pacus (l).
962
963 **Figure 3**: Host species has a more important impact on the structure of gut
964 communities than of skin mucus communities. The assessment of host species impact

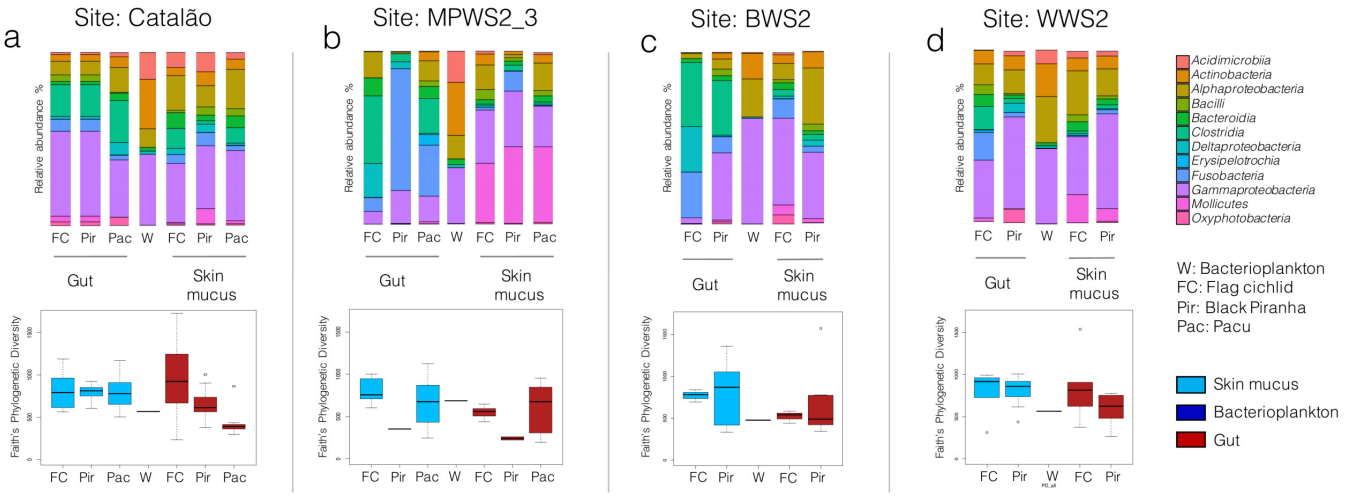
was based on PERMANOVA F and p values, and the total % of variance explained by the 2 first axes of PCA plots. The PCAs show the clustering of bacterial communities according to the host fish species at each sampling site. PERMANOVAs with 10 000 permutations tested the significance of the "Host Species" factor at each site. At least two species were found at every site, and all three species were found at the sites Catalão and MPWS2_3.

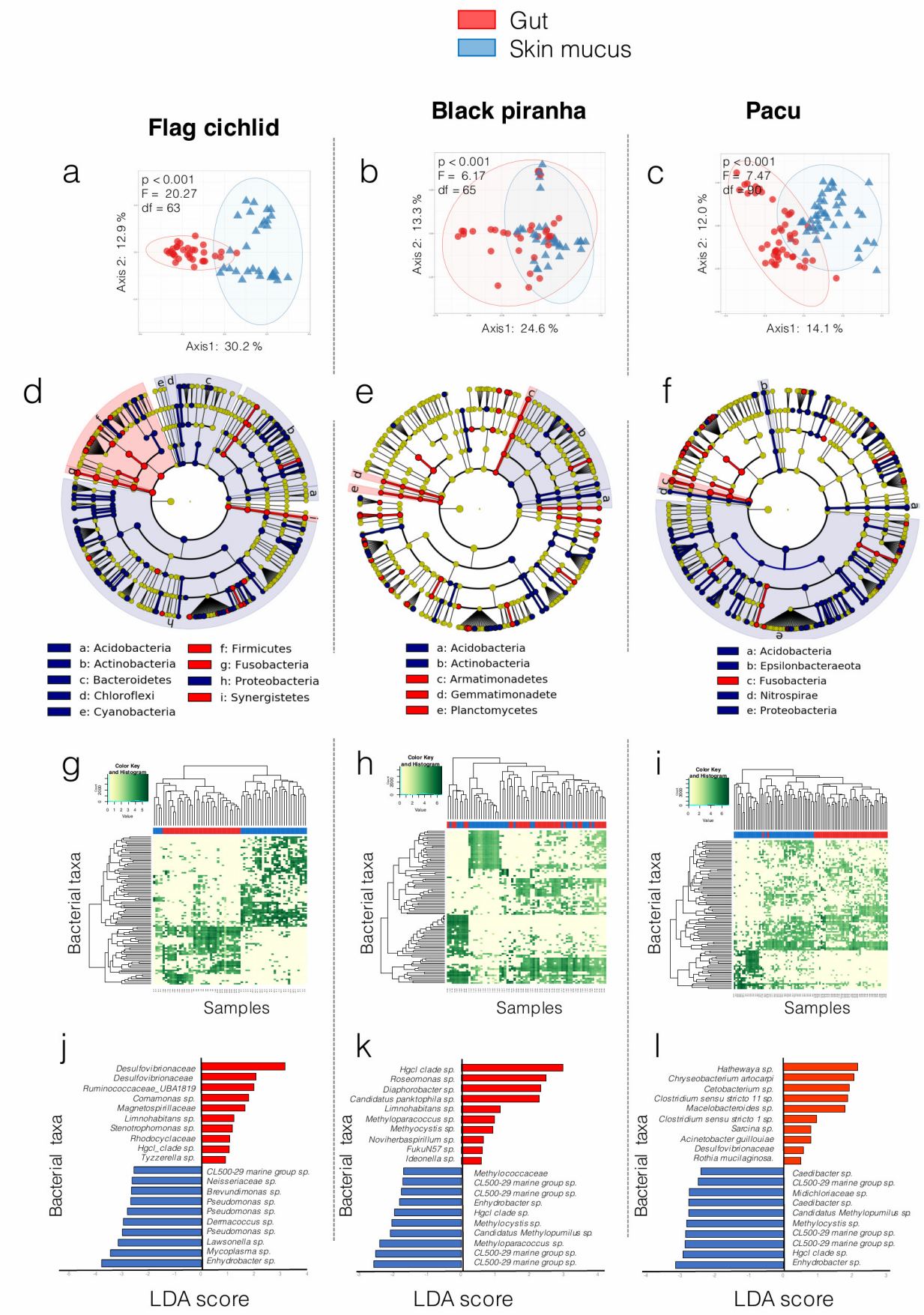
Figure 4: Overall, skin mucus communities are more associated to habitat-specific factors than gut communities. The figure shows the clustering of nodes (i.e. samples) according to the pattern of Spearman correlations when all samples (for communities of the same tissue) are compared to one another. Nodes sharing and edge (link) have a Spearman correlation > 0.5 with a p-value < 0.05 (corrected with Bonferroni). The color of the node indicates the host species, while the shape indicates the sampling site of the sample.

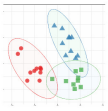
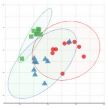
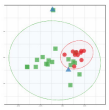
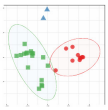
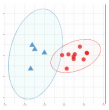
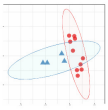
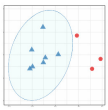
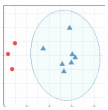
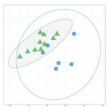
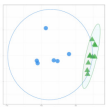
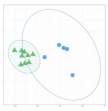
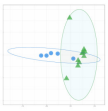
Figure 5: Interaction networks show high connectivity between skin communities, environmental bacterioplankton and hydrochemical parameters. Nodes sharing an edge (link) have a Spearman correlation > 0.5 with a p-value < 0.05 (corrected with Bonferroni). The color of the node indicates the type of sample: blue nodes represent bacterioplankton taxa, orange nodes are hydrochermical parameters, and grey nodes are host fish taxa. No significant correlations were detected between gut taxa and environmental hydrochemical parameters for black piranha and flag cichlid samples.

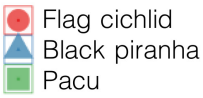
987 Numbers under each network indicate the number and % of ASVs implicated in
988 significant interactions in each community.
989
990

Applied and Environmental Microbiology





	Skin			Gut		
	PCA	Main axes	PERMANOVA	PCA	Main axes	PERMANOVA
* Site 1: Catalão		Axis 1: 18.9% Axis 2: 12.3% **Total** = 31.2%	df: 26 F-value: 4.81 **p-value** < 0.001		Axis 1: 32.0% Axis 2: 24.3% **Total** = 56.3%	df: 24 F-value: 6.19 **p-value** < 0.001
* Site 2: MPWS23		Axis 1: 22.1% Axis 2: 12.4% **Total** = 34.5%	df: 27 F-value: 1.30 **p-value** = 0.22		Axis 1: 37.6% Axis 2: 15.2% **Total** = 52.8%	df: 26 F-value: 6.47 **p-value** < 0.001
Site 3: BWS1		Axis 1: 33.1% Axis 2: 13.6% **Total** = 46.7%	df: 12 F-value: **3.56** **p-value** < 0.002		Axis 1: 37.8% Axis 2: 20.7% **Total** = 58.5%	df: 12 F-value: 4.37 **p-value** < 0.001
Site 4: BWS2		Axis 1: 29.5% Axis 2: 20.5% **Total** = 50.0%	df: 8 F-value: 2.87 **p-value** = 0.012		Axis 1: 44.5% Axis 2: 16.6% **Total** = 61.1%	df: 8 F-value: 5.26 **p-value** = 0.007
Site 5: WWS1		Axis 1: 18.6% Axis 2: 17.2% **Total** = 35.8%	df:14 F-value: 1.87 **p-value** = 0.011		Axis 1: 51.8% Axis 2: 16.7% **Total** = 68.5%	df: 14 F-value: 11.33 **p-value** < 0.001
Site 6: WWS2		Axis 1: 25.5% Axis 2: 14.9% **Total** = 40.4%	df: 12 F-value: 3.51 **p-value** < 0.001		Axis 1: 52.9% Axis 2: 15.4% **Total** = 68.3%	df: 13 F-value: 7.39 **p-value** = 0.002
Average total		**39.8 %**			**60.9 %**	



* All three host species were found at these sites.