



**HAL**  
open science

## Microbiota associated with the skin, gills, and gut of the fish *Parachondrostoma toxostoma* from the Rhône basin

Emmanuel Guivier, Nicolas Pech, Rémi Chappaz, André Gilles

### ► To cite this version:

Emmanuel Guivier, Nicolas Pech, Rémi Chappaz, André Gilles. Microbiota associated with the skin, gills, and gut of the fish *Parachondrostoma toxostoma* from the Rhône basin. *Freshwater Biology*, 2020, 10.1111/fwb.13437 . hal-02505021

**HAL Id: hal-02505021**

**<https://hal.science/hal-02505021>**

Submitted on 11 Mar 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Microbiota associated with the skin, gills, and gut of the fish *Parachondrostoma toxostoma*  
from the Rhône basin**

Guivier E. <sup>1</sup>, Pech N. <sup>2</sup>, Chappaz R. <sup>1</sup>, Gilles A. <sup>1</sup>

1 Aix Marseille Univ, Irstea, RECOVER, Centre St Charles, 3 place Victor Hugo, 13331  
Marseille, France

2 Université Publique de France, Centre St Charles, 3 place Victor Hugo, 13331 Marseille,  
France

*Corresponding author:* Emmanuel Guivier

*e-mail address:* em.guivier@gmail.com

*Running title:* Microbiota associated with mucosal tissues in wild fish

keywords: mucosal tissues, bacterial communities, predicted functional pathways, holobiont,  
cyprinid species

## Summary

1. The close relationships between hosts and their microbiota underlie the current concept of holobionts. This biological entity enlarges our perception of the genomic diversity supported by individuals/species and the potential abilities of hosts to respond to the environment.

2. We investigated a total of 54 specimens of the sofie fish *Parachondrostoma toxostoma* (Pt) for the taxonomic and functional composition of microbiota associated with the skin, gills and gut tissues. Five wild populations from the Ardeche and Durance river basins in the Mediterranean part of the large watershed of the Rhône (Southern France) were compared.

3. We found that the skin microbiota was more diverse and variable in bacterial community composition than the gut microbiota. These differences suggest that different deterministic factors underlie the composition of these bacterial communities, resulting in tissue-specific variation of the microbiota composition between wild fish populations inhabiting two rivers. Microbiota diversity and composition differed between wild Pt populations at both regional and local spatial scales.

4. Using the Picrust tool, the taxonomic variation of the microbiota composition in the river fish populations could be associated with the enrichment of predicted metabolic pathways. In particular, carbohydrate metabolism was more represented and xenobiotic biodegradation pathways less represented in the gut microbiota of Pt populations from the Ardeche than of those from the Durance. This result suggests variability in the ability to acclimate to environmental features through the assimilation of food resources from plants and the detoxification of xenobiotic components.

5. This study opens perspectives in the detection of subtle shifts in environmental features through the description of the microbiota composition, and in the potential prediction of the functional capacities of species to acclimate.

## Introduction

The microbial communities associated with macro-organisms provide a large and diverse metagenome encoding proteins with essential physiological functions. These communities expand the genomic diversity supported by species and the ability of the host to respond rapidly to various environmental constraints (Alberdi *et al.*, 2016; Chiarello *et al.*, 2018). Indeed, many recent studies have shown that bacterial symbionts are critical biological components for host fitness traits, such as development and maturation of the immune system, or nutrient assimilation and energy metabolism (Fuhrman, 2009; Tremaroli & Bäckhed, 2012; Gould *et al.*, 2018). Microbial communities are, thus, thought to play key roles in the exploitation of environmental resources by the host, and in defining the ecological niche of the host and its overall adaptation and/or acclimation to local environmental conditions (Amato *et al.*, 2015, 2018; Alberdi *et al.*, 2016; Carrier & Reitzel, 2017). Conversely, the environment probably shapes or alters the composition and structure of the microbiota, resulting in differences between wild individuals/populations with consequences for host performance for various associated physiological functions (Uren Webster *et al.*, 2018). The close relationships between hosts and their microbiotas underlies the definition of a biological unit named holobiont (Bordenstein & Theis, 2015; Sevellec, Derome & Bernatchez, 2018; Simon *et al.*, 2019). In this new perspective, the definition of conservation unit that aims the preservation of the evolutionary processes and the viability of populations (Fraser & Bernatchez, 2001), is likely to benefit from the characterization of the microbiota composition differences and the associated predicted functions in wild threatened species.

Most investigations on bacterial symbionts have focused on the gut, but these microorganisms are present on all mucosal surfaces, including the skin, oral cavity and

respiratory systems (Ursell *et al.*, 2012). These body sites varies notably by the structure of epithelia, the type and activity of immune cells and the physico-chemical conditions such as temperature, pH and oxygenation (Wilson, 2008). They also differ by the bacterial colonization processes, which mainly depend from the extent of contact with surrounding environment. The aquatic organisms such as fish are in continuous contact with bacteria from the surrounding water, a rich environment for microbes and favorable to their dispersal. The freshwater ecosystems are particularly exposed to climate changes affecting temperature and the water flow regimes, as well as to physical/chemical alterations related to human activities such as stream channelization or pollutants releasing (Vörösmarty *et al.*, 2010). Microbial communities associated with mucosal tissues in fish species are also exposed to environmental changes and their sensitivity/resistance should play a critical role in the responses of organism to such disturbances.

The three main mucosal tissues in teleost fish are the skin, gills and gut tissues (Salinas, 2015). These tissues form different microhabitats within the body, and each harbors a specific bacterial assemblage (Lowrey *et al.*, 2015; Guivier *et al.*, 2018; Uren Webster *et al.*, 2018; Chiarello *et al.*, 2018) that may be associated with different host physiological functions. The skin and gills can be considered to be external mucosal tissues coming into direct contact with bacterioplankton and the surrounding freshwater, which has a number of specific physical and chemical properties, including pH and oxygen concentration (Sylvain *et al.*, 2016). These tissues also have essential immune functions, as they constitute the first line of defense against invading pathogens (Benhamed *et al.*, 2014; Boutin *et al.*, 2014). By contrast, the gut is considered an internal mucosal tissue, with both ingested freshwater and food serving as sources of bacteria (Liu *et al.*, 2016: meta-analysis on eight fish species belonging to different trophic levels, but see also review on fish gastrointestinal microbiota

in Wang *et al.* 2018). The gut microbiota has long been studied for its many metabolic functions, including vitamin biosynthesis and nutrient assimilation (Tremaroli & Bäckhed, 2012). Host physiology is also modulated by hormonal status that could explain the difference in microbiota composition between male and female, as observed in many studies (Bolnick *et al.*, 2014; Li *et al.*, 2016; Guivier *et al.*, 2018).

In this study, we investigated the variation of microbiota composition associated with skin, gills and gut of sofie fish, *Parachondrostoma toxostoma* (Pt), at five separate river sites. *Parachondrostoma toxostoma* is an endemic cyprinid species in southern France, which includes the Rhône, Garonne, and Adour watersheds, but also many other coastal Mediterranean streams (IUCN, 2018). This species has a generalist dietary regime consisting of variable proportions of invertebrates, algae and diatoms, depending on its habitat (Corse, 2010). The endemicity of this rheophilic cyprinid species, i.e. with preference for running water flows, suggests an adaptation to specific environmental conditions. It is particularly relevant in the Mediterranean region characterized by extreme and rapid changes in climatic conditions (Filipe, Lawrence & Bonada, 2013). We therefore focused on the Pt populations inhabiting the Ardeche and Durance river basins, which belong to the Mediterranean part of the Rhône watershed. These populations are characterized by weak neutral genetic differentiation over this large spatial scale (41 microsatellites,  $F_{st} < 0.01$ , in Šimková *et al.*, 2013) but are subject to very different environmental conditions in these two river systems. The Ardeche is a river basin with low levels of urbanization and variable and unregulated flow regimes influenced by Mediterranean weather conditions, with a long dry summer season and intense rainfall in the fall. Climatic conditions are similar for the Durance, but this river basin is heavily managed and drains large urbanized and agricultural areas, with a flow regime strongly regulated by successive dams and connected man-made canals (Guivier *et al.*, 2019). Fur-

thermore, the Pt populations of the Ardeche and Durance basins use different food resources, with the diet of individuals from the Ardeche containing a much higher proportion of invertebrates, and that of individuals from the Durance containing a higher proportion of diatoms (Corse *et al.*, 2015). We used high throughput sequencing of 16S rRNA gene to investigate the effects of these contrasting environmental conditions and specific dietary regimes on the taxonomic and functional variation of the bacterial communities composition associated with the skin, gills and gut in these wild Pt populations inhabiting the two river basins. We hypothesized that the taxonomic diversity and composition of microbiota differ according to the associated mucosal tissues. In particular, we expected stronger differences in microbiota composition associated with skin and gills due to higher influence of the external environment on these bacterial communities than gut associated microbiota. We also hypothesized that the Pt populations in the 2 river basins would harbor bacterial communities whose metabolic functions differ and reflect the specific environmental constraints faced in each basin.

## **Materials and Methods**

### Field sampling

We sampled Pt specimens from five sites in two rivers from the Rhône basin (Southern France): Buech (44.344N, 5.771E), Manosque (43.874N, 5.878E) and Pertuis (43.668N, 5.493E), (Durance river basin), Rosieres (44.475N, 4.264E) and Saint-Just (44.286N, 4.597E), (Ardeche river basin) (Figure 1). Field sampling was conducted in the summer of 2014 (Durance) and 2015 (Ardeche). Specimens were trapped by electrofishing method. Animals were euthanized by cervical dislocation and immediately dissected with sterile materials to sample the three main mucosal tissues: skin (caudal fin), gills and gut (removed

of contents and separated into midgut and hindgut). A piece of each tissue was stored in 95% ethanol at -80°C for further molecular analyses. The reproductive status of the fish was classified in three categories as follows: sexually immature, sexually mature male and sexually mature female. The study was conducted according to international guidelines for animal care national protocols, under authorizations 2014-156-0001 and 2015-1426DDT605.

### Molecular and bioinformatics pipelines

Bacterial DNA was extracted from each of the tissues sampled, with the Qiagen Food Mericon kit (Qiagen, Venlo, Netherlands), with a slightly modified version of the protocol described by Guivier *et al* (2018). Briefly, using sterile materials we isolated a piece of tissue (bout 5 mg) which was lysed by incubation in 700 µL of Food Lysis Buffer with 3 µl (20 mg/mL) proteinase K for 3 hours at 56°C. We added 500 µl chloroform, vortexed 15s by pulse, centrifuged tube sample at 14,000 x *g* for 15 min and we collected the supernatant. Then, we added 1.2 mL of PB buffer that was transferred onto silicate column. The sample was cleaned with 500 µL AW2 buffer and DNA was eluted with 100 µL EB buffer. We amplified a 251 bp fragment from the V4 region of the bacterial 16S rRNA gene by the dual-index method of Kozich *et al.* (Kozich *et al.*, 2013), with forward (16S-V4F [GTGCCAGCMGCCGCG-GTAA]) and reverse (16S-V4R [GGACTACHVGGGTWTCTAATCC]) primers including 8-bp index and Illumina adapters. The combinations of tagged-forward and tagged-reverse primers were specific to each DNA sample amplified, making it possible to reassign sequences at the end of the total procedure. We performed PCR amplification on 216 samples (four tissues from each of 54 individuals) from a larger sequencing project including 718 samples, four extraction-negative controls and eight amplification-negative controls. We also duplicated the amplification of a commercial mock community, the ZymoBiomics microbial community DNA

standard (Zymo Research, Irvine, CA, USA), including a mixture of eight known bacterial taxa. The conditions for DNA amplification, the purification and pooling of PCR products were as described by Galan *et al.* (Galan *et al.*, 2016). Briefly, the DNA amplification was performed in 5 µl of Multiplex PCR kit master mix (Qiagen, ref), 4 µl of combined forward and reverse primers (2.5 µM) and 2 µl genomic DNA. The initial denaturation at 95°C for 15min was followed by 40 cycles including denaturation at 95°C for 20 s, annealing at 55°C for 15 s, and extension at 72°C for 5 min, followed by a final extension at 72°C for 10 min. The final library was quantified with the Kapa quantification kit (Kapa Biosystems) and loaded into a MiSeq (Illumina) flow cell (expected cluster density: 120,000 to 140,000/mm<sup>2</sup>) with reagent kit v3 (Illumina, 2 × 300 cycles) including 5% Phix DNA. We performed 3 Illumina MiSeq sequencing runs, and every run contained all 216 samples that were analyzed in this study in order to increase sequencing depth. For two runs we performed the library preparation, and the third run was simply a sequencing rerun of one of these library preparations (see Supplementary File S1).

Independently for each of the three high-throughput runs, we performed denoising, and assembly independently with the R package Dada2 (Callahan *et al.*, 2016, see detailed R script in Supplementary File S2). The procedure began with a filtering and trimming of sequences on the basis of quality scores (specific filtering parameters: truncLen=c(240,160), maxN=0, maxEE=c(2,2), truncQ=2). We then considered independent error models defined in Dada2 on the basis of 10<sup>6</sup> forward and reverse sequences. Using these models we performed the correction procedure independently for forward and reverse sequences before merging sequences with a minimum 100 bp overlap and no mismatch accepted. We removed potential artefactual sequences due to poor quality of assembly and outside the expected length range of 241 to 261 bp (less and more 4% sequence length variation). In the

final step, we removed chimeric sequences with the *removeBimeraDenovo* function (see detailed of number of reads retained at each step in Supplementary File S3).

We checked the repeatability of the molecular and bioinformatics procedures, based on the correlation of alpha diversity estimates within and between sequencing runs. We therefore estimated Shannon and phylogenetic diversity indices for each run, based on a rarefaction procedure with a depth of 5,000 randomly sampled sequences. We detected strong correlations between runs for alpha diversity estimates (Shannon: Pearson's  $R > 0.98$ ,  $P < 0.01$ ; phylogenetic diversity: Pearson's  $R > 0.91$ ,  $P < 0.01$ ) and between PCR duplicates within runs (Shannon: Pearson's  $R > 0.99$ ,  $P < 0.01$ ; phylogenetic diversity: Pearson's  $R > 0.97$ ,  $P < 0.01$ ). we detected between 0 and 5053 sequences from extraction- and amplification-negative controls, and between 0 and 22 sequences were assigned to unused tagged-primers combinations but that have still been detecting at the demultiplexing step (Supplementary Files S4, S5, S6). We detected between nine and ten variants assigned to the eight bacterial taxa expected (100% of sequences matched the expected taxa, Supplementary Figure S1) in samples corresponding to the taxonomic composition of mock community DNA standard, with exception of one on eight mock samples showing ten sequences on 32,159 not assigned to the expected taxa. The relative abundance of taxa between observed and expected was correlated with  $0.81 < R^2 < 0.96$  (Supplementary Figure S1). Based on these validation procedures, we merged variants/sequences of replicated samples between the three sequencing runs. We used the silva v132 database to assign each variant (Quast *et al.*, 2013) using the naïve Bayes method (Wang *et al.*, 2007) implemented in Qiime2 (Bolyen *et al.*, 2018). We retained only variants assigned to the Eubacteria kingdom, and removed all variants classified as chloroplast or mitochondrial sequences. For the 216 samples, we obtained a total of 9,659,798 sequences corresponding to 19,995 amplicon sequence

variants (ASVs) (distribution of the number of reads per individual, Supplementary Figure S2). We generated a phylogenetic tree from the variants, using the FastTree algorithm implemented in Qiime2 (Price, Dehal & Arkin, 2010). We normalized the results by rarefaction of the variants to 10,000 sequences/sample, a value choice based on the sequencing depth distribution and the plateau observed for this value on alpha diversity rarefaction curves using the Shannon index and phylogenetic diversity (Supplementary Figure S2 and S3). This procedure retained 201 of the 216 samples (93%). Further taxonomic diversity estimates and analyses were based on this rarefied table.

We also investigated the functional pathways predicted to be associated with the taxonomically characterized microbial communities, with the Picrust bioinformatic pipeline (Langille *et al.*, 2013), deleting the 16S copy normalization step which remains a debated correction procedure inducing potential biases (Louca, Doebeli & Parfrey, 2018). We assigned ASV to the closest clustering OTU (97% similarity, as described by Langille *et al.* 2013) from the Greengenes v13.5 database. We filtered the OTU table obtained by retaining OTUs assigned to the Eubacteria kingdom but removing OTUs classified as chloroplast or mitochondrial sequences. We normalized the sequencing depth between samples by performing a rarefaction procedure as for the taxonomic analysis, resulting in a random 10,000 sequences/sample, which limited the dataset to 177 of the 216 samples (82%), as some ASVs have not been assigned to Greengenes OTUs. Using the open Galaxy platform, we performed Picrust analyses to assign OTUs to the closest known genome; we then multiplied OTU occurrences by the genome-associated KEGGs (Langille *et al.*, 2013). The pipeline estimated the quality of the prediction by calculating the nearest sequenced taxon index (NSTI) score based on the length of the branch between the OTU and the bacterial reference genome. The final procedure involved classifying KEGGs by level 2 function. We

focused on the 12 KEGG functional pathways classified as belonging to the “Metabolism” pathway (level 1) and accounting for a mean of 58% of KEGGs.

### Statistical methods

#### *Taxonomic variation of microbiota composition*

We estimated the alpha diversity of the microbial community by calculating the Shannon index and phylogenetic diversity (PD) for each sample. We hypothesized that environment would have different effects on microbial community diversity in different tissues. To examine the effect of environment on the community diversity, we built for each alpha index an independent linear mixed model (LMM) including river, tissue and the river x tissue interaction as fixed factors, and site and individual as random factors. We also included reproductive status in each model as an adjusted factor. We log-transformed PD data to meet the normality and heteroscedasticity assumptions of LMMs. The significance of the effect of explanatory variables was assessed with type III sums of squares and Satterthwaite approximation of the number of degrees of freedom.

We estimated beta diversity by calculating the dissimilarity distance between microbial communities in the form of the Bray-Curtis index based on the composition of ASVs, and unweighted UniFrac based on the fraction of the total phylogenetic branch length unique or common to the communities considered (Lozupone & Knight, 2005). For each index, we performed principal coordinate analyses (PCoA) in which microbiota were ordered along axes maximizing the variance between statistical units (Paliy & Shankar, 2016).

The shaping of microbial community composition by strong deterministic factors should result in a lower dispersion from the centroid of the factor group considered. We hypothesized that the strength of the deterministic factors shaping microbiota composition

would differ between mucosal tissues. We tested the homogeneity of dispersion from the centroid between tissues by performing analyses of multivariate homogeneity of group dispersions using the `betadisper` function of the `vegan` package (Oksanen *et al.*, 2016) of R (R Core Team, 2016) for Bray-Curtis and unweighted UniFrac indices, with significance based on 10,000 permutation tests. We performed pairwise post-hoc permutation tests for each comparison between tissues and applied the FDR correction procedure to adjust the  $p$ -value (Benjamini & Hochberg, 1995).

We investigated the environmental factors structuring these compositional dissimilarities by performing PerMANOVA (Adonis function in the R package `vegan`, Oksanen *et al.*, 2016) independently for each dissimilarity distance and including the factors reproductive status, tissue, river, site and the river x tissue interaction. The significance of pseudo-F estimates was determined with a 10,000-permutation procedure.

We investigated the differences in the taxonomic composition of the microbiota between the Pt populations from the two river basins. We explored taxonomic variation by assigning ASVs at phylum level, the highest level of taxonomic classification in the Eubacteria kingdom and, therefore, the highest level at which differences in the relative abundance of taxa could be detected. Independently for each tissue, we performed an analysis of composition (ANCOM) in Qiime2 to investigate differences in the microbiota compositions of Pt populations from the two (Bolyen *et al.*, 2018). ANCOM decreases the false discovery rate by calculating the abundance of each taxon relative (log ratios) to all the other taxa present within the microbial community, and can be used to assess differences in ratio between groups (Mandal *et al.*, 2015; Weiss *et al.*, 2017). The  $W$  value indicates, for each taxon, the number of log ratios significantly different between populations from the two rivers.

### *Predicted functional variation of microbiota composition*

We explored differences in the predicted functional metabolic pathways of the microbiota from different mucosal tissues between Pt populations in these environments. We hypothesized that the Pt populations of each river harbored host microbiotas with different metabolic profiles. For each tissue, we used STAMP software (Parks *et al.*, 2014) to perform Welch's *t* tests followed by correction by the FDR procedure for multiple testing (Benjamini & Hochberg, 1995), to detect metabolic functions differentially represented in the microbiotas of Pt populations between the two rivers.

## **Results**

### *Differences in the taxonomic diversity and composition of microbiotas*

We detected an effect of river, but this effect was dependent on the associated tissue considered (Shannon: LMM, river x tissue interaction,  $F=12.1$ ,  $p<0.01$ ; PD: LMM, river x tissue interaction,  $F=14.2$ ,  $p<0.01$ , Table 1). This interaction corresponds to higher levels of microbiota diversity associated with the skin, midgut and hindgut in Pt populations from the Durance than in those from the Ardeche, and the opposite pattern for the microbiota associated with gills (Supplementary Table S1). Interestingly, we observed a general trend towards an increase in microbiota diversity moving downstream along each river, with different degrees of variation in mucosal different tissues, except for the midgut, in which bacterial diversity was lower in Pt populations from Buech (upstream Durance) than in those from Pertuis (downstream Durance) (Figure 2). We assessed the significance of this effect by comparing the LMMs with and without site as a random factor by ANOVA. We detected a significant effect of site for the PD index (ANOVA,  $\chi^2=5.94$ ,  $p=0.01$ ) but not for Shannon index (ANOVA,  $\chi^2=0.86$ ,  $p=0.35$ ).

We explored the total range of environmental variation of microbiota structure for each of the associated mucosal tissues. Significant differences in distance to the centroid based on Bray-Curtis dissimilarity were found between tissues (Bray-Curtis: analyses of multivariate homogeneity of group dispersions,  $F=6.38$ ,  $p<0.01$ ). In particular, the microbiotas of external tissues (i.e. skin and gills) were characterized by a higher distance to their centroid than those of internal tissues (i.e. midgut and hindgut tissues; pairwise permutation test, skin/gills vs. midgut/hindgut  $p<0.05$ ). This pattern was also detected based on unweighted UniFrac distance to the centroid but only with the skin microbiota composition compared to gut microbiota composition (midgut and hindgut) (pairwise permutations test, skin vs. midgut/hindgut  $p<0.05$ ).

Microbiota composition dissimilarities differed significantly between rivers, according to the associated tissues (Bray-Curtis: Permanova, river x tissue interaction, pseudo- $F=6.60$ ,  $R^2=0.06$ ,  $p<0.01$ , Table 2, Figure 3). However, even if significant, we noted that the variance explained ( $R^2$  values) by the river x tissues interaction are low suggesting that others unknown factors affect the difference in microbiota composition. Observation of the PCoA graphs highlighted a strong effect of river on skin- and midgut-associated microbiotas, but a much weaker effect of river for the microbiotas of the gills and hindgut, whatever the dissimilarity index (Figure 3, S4). We also detected an important influence of site (Bray-Curtis: Permanova, site, pseudo- $F=5.48$ ,  $R^2=0.05$ ,  $p<0.01$ , Table 2), again highlighting the effect of local environment on microbiota composition and structure. Similar results were obtained based on unweighted Unifrac distance (Table S2). We therefore explored, for each tissue, the dissimilarity of microbiota structure between each pair of sites, in pairwise Permanova with FDR-based correction for multiple testing (Benjamini & Hochberg, 1995). For each tissue, we detected a consistent pattern of microbiota differentiation between pairs

of sites from different rivers (Bray-Curtis: Permanova,  $p$ -value  $<0.05$ , Table 3, Figure 4). We then distinguished between the two populations in the Ardeche for each tissue-associated microbiota (Bray-Curtis: Permanova, FDR-  $p$ -value  $<0.05$ ). Similar results were obtained based on unweighted unfrac distance (Supplementary Table S3 and Figure S5). In the Durance, each pair of sites differed significantly in terms of microbiota composition dissimilarity except for the gill-associated microbiotas of the neighboring populations from Buech and Manosque according to Bray-Curtis dissimilarity (pairwise Permanova, adjusted  $p=0.50$ , Table 3), and for the hindgut-associated microbiotas composition of the populations from Manosque and Pertuis, for which the two dissimilarity indices gave inconsistent results (hindgut, Bray-Curtis, pairwise Permanova, adjusted  $p=0.11$ , Table 3; midgut, unweighted UniFrac: pairwise Permanova, 10,000 permutations, adjusted  $p=0.06$ , Table S3).

By performing ANCOM between the two river basins independently for each tissue, we were able to detect the variation of the relative abundance of several phyla within the microbial community. This variation concerned different phyla in different mucosal tissues (Figure 5). Skin-associated microbiotas differed between rivers, with a higher relative abundance of *Deinococcus-Thermus* in the Ardeche ( $W=45$ ). Gill-associated microbiotas in the Ardeche were characterized by a higher relative abundance of Cyanobacteria than those of the populations in the Durance ( $W=32$ ). Interestingly, although there was no significant difference in abundance between river basins, the gill-associated microbiota displayed a high relative abundance of the pathogen *Candidatus piscichlamydia* from the phylum Chlamydiae (Figure S6). Both midgut and hindgut microbiotas had higher relative abundances of Cyanobacteria, Firmicutes and Fusobacteria ( $W>28$ ) in populations from the Ardeche than in populations from the Durance, whereas populations from the Durance displayed enrichment in Actinobacteria and Tenericutes ( $W>27$ ). The relative abundance of Planctomycetes

(W=25) was also higher in the hindgut of Pt populations from the Durance than in those from the Ardeche.

*Differences in the predicted metabolic function of microbiota between Pt populations from the two river basins*

We defined the predicted functional composition of the microbiota with the Picrust pipeline, resulting in a mean NSTI score of  $0.10 \pm 0.03$ , making it possible to provide accurate information about the predicted function of the various tissue-associated microbiotas (Langille *et al.*, 2013).

We identified several predicted metabolic pathways, categorized to level 2, as differentially represented in the microbial communities associated with the skin, midgut and hindgut (Figure 6) between the two rivers. We detected no significant difference in predicted metabolic pathways in gill-associated microbiotas between populations from the two rivers (adjusted  $p > 0.05$ ). For the skin microbiota, we detected enrichment in glycan biosynthesis and metabolism and energy metabolism in Pt populations from the Durance, and enrichment in amino-acid metabolism in Pt populations from the Ardeche. Observation of the PCoA plot obtained with STAMP (Figure S7), revealed the presence of an outlier corresponding to a midgut-associated microbiota from an individual from Pertuis. We removed this outlier from subsequent function-predicting analyses. We identified, consistently in midgut- and hindgut-associated microbiotas, nine functional pathways differing significantly in importance between Pt populations from the two rivers (adjusted  $p < 0.05$ ). Specifically, we detected an effect for carbohydrate metabolism enrichment in the Ardeche (midgut: Ardeche: mean=22.2 +/-std. dev=0.89, Durance: mean=20.3% +/- std. dev.=0.6, adjusted  $p < 0.01$ ) and amino-acid metabolism enrichment in the Durance (midgut:

Ardeche: mean=19.4 +/-0.4, Durance: 20.5 +/-0.4, adjusted  $p<0.01$ ). We also noted enrichment in predicted xenobiotic biodegradation and metabolism in Pt populations from the Durance relative to the Ardeche (midgut: Ardeche: mean=4.8 +/-0.1, Durance: mean=5.8 +/-0.7, adjusted  $p<0.01$ ). Finally, we detected an enrichment in the predicted metabolism of cofactors and vitamins in midgut microbiotas from populations in the Ardeche (Ardeche: mean=9.3 +/-0.1, Durance: mean=9.0 +/-0.3, adjusted  $p<0.05$ ), and glycan biosynthesis and metabolism in hindgut microbiotas from populations in the Durance (Ardeche: mean=4.5 +/-0.3, Durance: mean=4.8 +/-0.2, adjusted  $p<0.05$ ). Although significant, we note that the variation in the relative abundance of the predicted metabolic functions between populations from the two rivers were slight in magnitude (Figure 6).

## **Discussion**

Our study highlighted the variation of microbiota diversity, composition and structure among five wild populations of a cyprinid species, according to the nature of the associated mucosal tissues. Consistent with our hypotheses, we showed that the skin microbiota was more diverse and variable in composition than the gut microbiota, as previously reported by Guivier *et al.* (2018) for Pt and other fish species such as the rainbow trout *Oncorhynchus mykiss* (Lowrey *et al.*, 2015) or the tambaqui *Colossoma macropomum* (Sylvain *et al.*, 2016). This pattern suggests differences in the factors determining the composition of these bacterial communities. We found that the environmental conditions characterizing the spatial scales from the river basin to local sites structured the taxonomic diversity and composition of the microbiota, from ASV to phylum level. Moreover, as hypothesized, the differences in the microbiota composition of Pt populations from rivers were associated with specific predicted metabolic pathways, such as carbohydrate metabolism and xenobiotic

biodegradation in the midgut microbiota. We speculate that such predicted functional differentiation, although slight in magnitude, may reflect local adaptation to the assimilation of plant-diet resources and exposure to pollutants, respectively.

#### Tissue-specific microbiota compositional variation

Consistent with the first description of microbiota diversity in Pt (Guivier *et al.*, 2019), we were able to distinguish between the bacterial communities associated with the skin, gills and gut. In this study, over a larger spatial scale, the most important discriminating factor for microbiota composition remained the nature of the associated tissues, with strong differentiation between the external and internal tissues, which was particularly pronounced for the phylogenetic dissimilarity index based on the presence/absence of specific bacterial lineages. Such patterns could be explained mostly by the extent of contact with sources of bacteria in the surrounding environment. The skin and gills are in direct contact with the surrounding water flow, and the associated microbial communities probably reflect for one part the bacteriological composition of the water flowing over these surfaces (Pratte *et al.*, 2018; Uren Webster *et al.*, 2018; Tarnecki *et al.*, 2018). Indeed, Tamecki *et al.* (2018) showed that, in common snook, *Centropomus undecimalis*, the composition of the skin microbiota in released hatchery-reared juveniles rapidly converged on the bacterial community composition of the water at the release site. By contrast, the gut microbiota is known to be strongly influenced by dietary regime, directly by the source of bacteria colonizing the epithelium of the digestive tract, but also indirectly by nutrients favoring specific bacteria or by the modulation of host physiology (see review Wang *et al.*, 2018).

The bacterial communities associated with the skin were the most variable between the wild Pt populations studied. The skin microbiota was characterized by greater compositional variation than the mid- and hindgut microbiota, suggesting a high sensitivity

of cutaneous bacteria to environmental conditions. The plasticity of the skin microbiota has been investigated experimentally in the tambaqui fish, *Colossoma macropomum*, in response to a fall in pH (Sylvain *et al.*, 2016). Within a few days of the pH change, the authors detected a rapid and prolonged change in the taxonomic structure of the skin microbiota that was more pronounced than that in the gut microbiota, revealing both a higher sensitivity and a weaker resilience of these skin bacterial communities. By contrast, the composition of the gut microbiota results from selective processes involving host feeding behavior (i.e. the food resources ingested from those available in the environment) and stringent abiotic conditions, such as low pH and anaerobiosis, which constrain the range of bacteria able to colonize and grow in these tissues (Donaldson, Lee & Mazmanian, 2015; Yan *et al.*, 2016). However the relative roles of microbe-microbe, host-microbe and microbe-environment interactions, which shape the sensitivity/resistance of microbiota in response to the external environment, remain to be investigated.

#### *Microbiota compositional variation at different spatial scales*

The diversity and taxonomic composition of the microbiotas of wild Pt populations, regardless of the associated tissue considered, varied over a large spatial scale and was characteristic of river systems. We were able to distinguish between the microbiota composition of wild Pt populations from the Ardeche and Durance river basins. The taxonomic differentiation of the microbiota composition was particularly strong for skin and midgut tissues. The Ardeche and Durance differ in terms of many climatic, geological, hydrological and ecological features. In particular, the Ardeche river basin, in the Mediterranean region, is little managed by anthropogenic structures and is characterized by changes in flow regime, with sudden floods in the fall and long periods in which water volume remains low during the sum-

mer. By contrast, the Durance river basin is geographically situated in the Alpine and Mediterranean climatic regions, with a large positive temperature gradient from up- to downstream (Corse *et al.*, 2015). The Durance River is heavily managed, with its water flow regulated by successive dams and weirs that buffer climatic changes and maintain a minimum flow depth, particularly during long dry periods in summer. These environmental features probably influence the bacterioplankton, the renewed suspension of sediments and the local food resources affecting the sources of the bacteria colonizing tissues. We highlighted the greater abundance of *Deinococcus-Thermus* (genus *Deinococcus*, accounting for up to 90% of bacteria) in the skin microbiota of Pt from the Ardeche than in those from the Durance. *Deinococcus* is an aerobic bacterium known to be resistant to UV irradiation, with an optimal growth temperature of 25°C to 35°C (Brooks & Murray, 1981). These favorable conditions are achieved when the river is shallow during the hot summers of the Ardeche (until 29°C at Rosieres in 2015, personal observation), whereas the reserved flow regimes managed by dams in the Durance may limit the increase in water temperature during the summer. Moreover, the higher abundance of Cyanobacteria in the gill microbiota of fish from the Ardeche populations than in that of fish from the Durance, is also consistent with our hypothesis of thermal differences between the two rivers. Characterization of the gut microbiota provided clues to the dietary regimes and ecological features of wild Pt populations. Corse *et al.* (2015) showed that Pt populations from the Ardeche and Durance differed in their feeding habits, with a mixed regime in the Ardeche, including an important proportion of invertebrates (40% of the specimen are invertebrate eaters) that decrease to less than 5% in the diet of populations from the Durance. Consistent with these previous findings, we detected a clear differentiation of the midgut microbiota between populations from the two river basins, with a higher abundance of Fusobacteria (genus *Cetobacterium*) and

Firmicutes (family Erysipelotrichaceae) in the Ardeche than in the Durance. The presence of *Cetobacterium* has been observed dominant in carnivorous fish (Liu *et al.*, 2016; Hao *et al.*, 2017a; Kashinskaya *et al.*, 2018), and this bacterium is known to synthesize large amounts of vitamin B12 and to play an important role in host immune activation (Liu *et al.*, 2016; Hao *et al.*, 2017b a). By contrast, Erysipelotrichaceae is associated with a plant-based diet and the need to digest complex polysaccharides (Liu *et al.*, 2016; Hao *et al.*, 2017b a). The abundance of these two taxa in the gut microbiota is consistent with previous descriptions of a mixed dietary regime in Pt populations from the Ardeche. The reasons for the high abundance of Tenericutes (genus *Mycoplasma*) in the gut microbiota of populations from the Durance (Holben *et al.*, 2002; Bano *et al.*, 2007; Xing *et al.*, 2013), remains less clear as the presence of this taxon has not been associated with specific feeding habits (Holben *et al.*, 2002; Bano *et al.*, 2007; Xing *et al.*, 2013).

Microbiota diversity and composition in wild Pt populations varied within both rivers, and we observed a general trend towards an increase in microbiota diversity from up- to downstream. Read *et al.* (2015) reported a similar pattern across the basin of the River Thames (UK) with an increasing diversity of bacterioplankton with dendritic distance (i.e. network distance from upstream), and a gradual succession of the composition of the bacterial community in the water. The shift in the diversity of the bacterial community associated with the skin from Manosque to Pertuis may result at least partly from the confluence with the Verdon River, the main tributary of the Durance, between the two sampling sites. Further investigations are required to evaluate the relative importance of the selective processes acting between the bacterioplankton and the skin microbiota. The differences in microbiota composition between pairs of sites highlighted the effect of local environmental conditions on the shaping of bacterial communities. For instance, we

observed the pathogen *Candidatus piscichlamydia*, from the phylum Chlamydiae, in the gill microbiotas of individuals from different sites. This bacteria accounted for up to 84% of the bacteria associated with gills. *Candidatus piscichlamydia* causes epitheliocystis inclusions in the gills and has been already detected in salmonid and cyprinid species (Draghi *et al.*, 2004; Sellyei, Molnár & Székely, 2017). *Chlamydia*-like pathogens have recently been detected in wild *Salmo trutta* in the Upper Rhône (Guevara Soto *et al.*, 2016) and its detection in Pt opens the need for further epidemiological investigations.

#### Microbiota compositional variation with predicted metabolic function

We explored the relationship between the microbiota taxonomic dissimilarity observed between populations from the Ardeche and the Durance and enrichment in predicted functional metabolic pathways. Despite the taxonomic dissimilarity between the gill microbiotas of the populations from the two rivers, we detected no difference in metabolic profiles. By contrast, the skin microbiota was characterized by an enrichment in three metabolic pathways that differed between the river basins, and particularly with an enrichment in amino-acid metabolism in the Ardeche relative to the Durance. Interestingly, this metabolic pathway influences the activation of immune responses and would therefore be expected to be involved in key functions of host immune defense against pathogens colonizing cutaneous tissues (Li *et al.*, 2007). The midgut and hindgut microbiota displayed the highest degree of functional differentiation between the two rivers. The functional patterns of the two segments of gut were similar, with nine of 10 metabolic pathways displaying enrichment common to both tissues. In particular, we observed an enrichment in carbohydrate metabolism in the gut of populations from the Ardeche relative to the Durance. The variation of this metabolic pathway, which is critical for the degradation of complex polysaccharides present in plants and algae (Liu *et al.*, 2016), is consistent with the

proportion of plant material in the dietary regime of Pt populations from the Ardeche. It indicates that Pt from Ardeche could compensate the reduced proportion of plant in diet by assimilating and metabolizing more efficiently such foods. Conversely, the gut microbiota of populations in the Durance was characterized by an enrichment in amino-acid metabolism, which is involved in host immune functions (Li *et al.*, 2007), and higher levels of xenobiotic biodegradation. The response to xenobiotic components appears particularly relevant in the urbanized context of the Durance with many industrial and agricultural activities potentially exposing aquatic organisms to pollutants. However, the accuracy of the method used to predict the metabolic functions associated with microbiota remains limited by available reference genomes, particularly for poorly known microbiomes. Increasing characterization of the microbiomes, particularly in wildlife, should improve our ability to predict the functional characteristics of microbiotas and therefore of the holobiont to environmental constraints.

Our study highlighted the tissue-specific variation in the diversity and composition of the microbiota between five wild fish populations inhabiting two rivers. Different host and environmental determinants, such as dietary regimes or water temperature, shape differently the microbiota composition associated with tissues. In particular, as expected, we showed a higher taxonomic variation in the microbiota composition associated with skin compared to others tissues.

**Data accessibility:** Study, Samples and Sequence reads are available in the ENA public database (<https://www.ebi.ac.uk/ena/>). Accession number: Study (PRJEB33555), Samples (ERS3605204-ERS3605419), Sequences reads: Run1 (ERR3458554-ERR3458769), Run2 (ERR3460030-ERR3460245), Run3 (ERR3460247-ERR3460462).

**Acknowledgments:** We are grateful to Maxime Galan (UMR CBGP, Montferrier-sur-lez) for technical advices. We thank the platform Genotypage – Séquençage (Université Montpellier, IRD, CNRS, EPHE), and particularly Frédérique Cerqueira for her technical help. We are also grateful to Arnaud Ungaro (Aix Marseille University), Benjamin Hérodet (Fédération de l’Ain pour la pêche et la protection des milieux aquatiques) for help in the carrying out of the field sampling. We thank the associate editor and the two anonymous reviewers for their comments and suggestions that improved the quality of the manuscript. The study has been funded by Electricité de France within the project FACIES developing scientific advances in favor of Aquatic System Biodiversity. We declare no conflict of interest in this study.

**Authors’ contributions:** AG and NP conceived and designed the experiments. AG, NP and RC performed the fieldworks. EG carried out the molecular techniques and performed molecular and statistical data analyses. EG wrote the first draft of the manuscript. All authors read and approved the final manuscript.

## References

- Alberdi A., Aizpurua O., Bohmann K., Zepeda-Mendoza M.L. & Gilbert M.T.P. (2016). Do vertebrate gut metagenomes confer rapid ecological adaptation? *Trends in Ecology & Evolution* **31**, 689–699. <https://doi.org/10.1016/j.tree.2016.06.008>
- Amato K.R., G. Sanders J., Song S.J., Nute M., Metcalf J.L., Thompson L.R., *et al.* (2018). Evolutionary trends in host physiology outweigh dietary niche in structuring primate gut microbiomes. *The ISME Journal*. <https://doi.org/10.1038/s41396-018-0175-0>
- Amato K.R., Leigh S.R., Kent A., Mackie R.I., Yeoman C.J., Stumpf R.M., *et al.* (2015). The gut microbiota appears to compensate for seasonal diet variation in the wild black howler monkey (*alouatta pigra*). *Microbial Ecology* **69**, 434–443. <https://doi.org/10.1007/s00248-014-0554-7>
- Bano N., deRae Smith A., Bennett W., Vasquez L. & Hollibaugh J.T. (2007). Dominance of Mycoplasma in the guts of the long-jawed mudsucker, *Gillichthys mirabilis*, from five

- California salt marshes. *Environmental Microbiology* **9**, 2636–2641. <https://doi.org/10.1111/j.1462-2920.2007.01381.x>
- Benhamed S., Guardiola F.A., Mars M. & Esteban M.Á. (2014). Pathogen bacteria adhesion to skin mucus of fishes. *Veterinary Microbiology* **171**, 1–12. <https://doi.org/10.1016/j.vetmic.2014.03.008>
- Benjamini Y. & Hochberg Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289–300
- Bolnick D.I., Snowberg L.K., Hirsch P.E., Lauber C.L., Org E., Parks B., *et al.* (2014). Individual diet has sex-dependent effects on vertebrate gut microbiota. *Nature Communications* **5**. <https://doi.org/10.1038/ncomms5500>
- Bolyen E., Rideout J.R., Dillon M.R., Bokulich N.A., Abnet C., Al-Ghalith G.A., *et al.* (2018). QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. <https://doi.org/10.7287/peerj.preprints.27295v2>
- Bordenstein S.R. & Theis K.R. (2015). Host biology in light of the microbiome: ten principles of holobionts and hologenomes. *PLOS Biology* **13**, e1002226. <https://doi.org/10.1371/journal.pbio.1002226>
- Boutin S., Sauvage C., Bernatchez L., Audet C. & Derome N. (2014). Inter individual variations of the fish skin microbiota: host genetics basis of mutualism? *PLoS ONE* **9**, e102649. <https://doi.org/10.1371/journal.pone.0102649>
- Brooks B.W. & Murray R.G.E. (1981). Nomenclature for “*Micrococcus radiodurans*” and other radiation-resistant Cocci: Deinococcaceae fam. nov. and Deinococcus gen. nov., including five species. *International Journal of Systematic Bacteriology* **31**, 353–360. <https://doi.org/10.1099/00207713-31-3-353>
- 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. <https://doi.org/10.1038/nmeth.3869>
- Carrier T.J. & Reitzel A.M. (2017). The hologenome across environments and the implications of a host-associated microbial repertoire. *Frontiers in Microbiology* **8**. <https://doi.org/10.3389/fmicb.2017.00802>
- Chiarello M., Auguet J.-C., Bettarel Y., Bouvier C., Claverie T., Graham N.A.J., *et al.* (2018). Skin microbiome of coral reef fish is highly variable and driven by host phylogeny and diet. *Microbiome* **6**. <https://doi.org/10.1186/s40168-018-0530-4>
- Corse E. (2010). Le “barcoding alimentaire”: développement d’un nouvel outil d’écologie moléculaire pour les milieux d’eau douce. Exemple au travers de l’étude de la zone hybride Durance entre les deux Chondrostomes français.
- Corse E., Pech N., Sinama M., Costedoat C., Chappaz R. & Gilles A. (2015). When anthropogenic river disturbance decreases hybridization between non-native and endemic Cyprinids and drives an ecomorphological displacement towards juvenile state in both species. *Plos One* **10**, e0142592. <https://doi.org/10.1371/journal.pone.0142592>
- Donaldson G.P., Lee S.M. & Mazmanian S.K. (2015). Gut biogeography of the bacterial microbiota. *Nature Reviews Microbiology* **14**, 20–32. <https://doi.org/10.1038/nrmicro3552>
- Draghi A., Popov V.L., Kahl M.M., Stanton J.B., Brown C.C., Tsongalis G.J., *et al.* (2004). Characterization of “*Candidatus Piscichlamydia salmonis*” (Order Chlamydiales), a Chlamydia-Like bacterium associated with epitheliocystis in farmed atlantic salmon

- (*Salmo salar*). *Journal of Clinical Microbiology* **42**, 5286–5297. <https://doi.org/10.1128/JCM.42.11.5286-5297.2004>
- Filipe A.F., Lawrence J.E. & Bonada N. (2013). Vulnerability of stream biota to climate change in mediterranean climate regions: a synthesis of ecological responses and conservation challenges. *Hydrobiologia*, 331–351. <https://doi.org/10.1007/s10750-012-1244-4>
- Fraser D.J. & Bernatchez L. (2001). Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Molecular Ecology* **10**, 2741–2752
- Fuhrman J.A. (2009). Microbial community structure and its functional implications. *Nature* **459**, 193–199. <https://doi.org/10.1038/nature08058>
- Galan M., Razzauti M., Bard E., Bernard M., Brouat C., Charbonnel N., *et al.* (2016). 16S rRNA amplicon sequencing for epidemiological surveys of bacteria in wildlife. *mSystems* **1**, e00032-16. <https://doi.org/10.1128/mSystems.00032-16>
- Gould A.L., Zhang V., Lamberti L., Jones E.W., Obadia B., Korasidis N., *et al.* (2018). Microbiome interactions shape host fitness. *Proceedings of the National Academy of Sciences*, 201809349. <https://doi.org/10.1073/pnas.1809349115>
- Guevara Soto M., Vidondo B., Vaughan L., Seth-Smith H.M.B., Nufer L., Segner H., *et al.* (2016). The emergence of epitheliocystis in the upper Rhone region: evidence for Chlamydiae in wild and farmed salmonid populations. *Archives of Microbiology* **198**, 315–324. <https://doi.org/10.1007/s00203-016-1192-x>
- Guivier E., Gilles A., Pech N., Dufлот N., Tissot L. & Chappaz R. (2019). Canals as ecological corridors and hybridization zones for two cyprinid species. *Hydrobiologia* **830**, 1–16. <https://doi.org/10.1007/s10750-018-3843-1>
- Guivier E., Martin J.-F., Pech N., Ungaro A., Chappaz R. & Gilles A. (2018). Microbiota diversity within and between the tissues of two wild interbreeding species. *Microbial Ecology* **75**, 799–810. <https://doi.org/10.1007/s00248-017-1077-9>
- Hao Y.T., Wu S.G., Jakovlić I., Zou H., Li W.X. & Wang G.T. (2017a). Impacts of diet on hindgut microbiota and short-chain fatty acids in grass carp (*Ctenopharyngodon idellus*). *Aquaculture Research* **48**, 5595–5605. <https://doi.org/10.1111/are.13381>
- Hao Y.T., Wu S.G., Xiong F., Tran N.T., Jakovlić I., Zou H., *et al.* (2017b). Succession and fermentation products of grass carp (*Ctenopharyngodon idellus*) hindgut microbiota in response to an extreme dietary shift. *Frontiers in Microbiology* **8**. <https://doi.org/10.3389/fmicb.2017.01585>
- Holben W.E., Williams P., Saarinen M., Särkilahti L.K. & Apajalahti J.H.A. (2002). Phylogenetic analysis of intestinal microflora indicates a novel Mycoplasma phylotype in farmed and wild salmon. *Microbial Ecology* **44**, 175–185. <https://doi.org/10.1007/s00248-002-1011-6>
- IUCN (2018). The IUCN Red List of Threatened Species. Version 2018-1
- Kashinskaya E.N., Simonov E.P., Kabilov M.R., Izvekova G.I., Andree K.B. & Solovyev M.M. (2018). Diet and other environmental factors shape the bacterial communities of fish gut in an eutrophic lake. *Journal of Applied Microbiology* **125**, 1626–1641. <https://doi.org/10.1111/jam.14064>
- Kozich J.J., Westcott S.L., Baxter N.T., Highlander S.K. & Schloss P.D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology* **79**, 5112–5120. <https://doi.org/10.1128/AEM.01043-13>

- Langille M.G.I., Zaneveld J., Caporaso J.G., McDonald D., Knights D., Reyes J.A., *et al.* (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology* **31**, 814–821. <https://doi.org/10.1038/nbt.2676>
- Li P., Yin Y.-L., Li D., Woo Kim S. & Wu G. (2007). Amino acids and immune function. *British Journal of Nutrition* **98**, 237. <https://doi.org/10.1017/S000711450769936X>
- Li X., Yan Q., Ringø E., Wu X., He Y. & Yang D. (2016). The influence of weight and gender on intestinal bacterial community of wild largemouth bronze gudgeon (*Coreius guichenoti*, 1874). *BMC Microbiology* **16**. <https://doi.org/10.1186/s12866-016-0809-1>
- Liu H., Guo X., Gooneratne R., Lai R., Zeng C., Zhan F., *et al.* (2016). The gut microbiome and degradation enzyme activity of wild freshwater fishes influenced by their trophic levels. *Scientific Reports* **6**, 24340. <https://doi.org/10.1038/srep24340>
- Louca S., Doebeli M. & Parfrey L.W. (2018). Correcting for 16S rRNA gene copy numbers in microbiome surveys remains an unsolved problem. *Microbiome* **6**. <https://doi.org/10.1186/s40168-018-0420-9>
- Lowrey L., Woodhams D.C., Tacchi L. & Salinas I. (2015). Topographical mapping of the rainbow trout (*oncorhynchus mykiss*) microbiome reveals a diverse bacterial community with antifungal properties in the skin. *Applied and Environmental Microbiology* **81**, 6915–6925. <https://doi.org/10.1128/AEM.01826-15>
- Lozupone C. & Knight R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology* **71**, 8228–8235. <https://doi.org/10.1128/AEM.71.12.8228-8235.2005>
- Mandal S., Van Treuren W., White R.A., Eggesbø M., Knight R. & Peddada S.D. (2015). Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microbial Ecology in Health & Disease* **26**. <https://doi.org/10.3402/mehd.v26.27663>
- Oksanen J., Blanchet G., Friendly R., Legendre P., McGlenn D., Minchin P.R., *et al.* (2016). vegan: community ecology package. R package version 2.4-1
- Paliy O. & Shankar V. (2016). Application of multivariate statistical techniques in microbial ecology. *Molecular Ecology* **25**, 1032–1057. <https://doi.org/10.1111/mec.13536>
- Parks D.H., Tyson G.W., Hugenholtz P. & Beiko R.G. (2014). STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* **30**, 3123–3124. <https://doi.org/10.1093/bioinformatics/btu494>
- Pratte Z.A., Besson M., Hollman R.D. & Stewart F.J. (2018). The gills of reef fish support a distinct microbiome influenced by host-specific factors. *Applied and Environmental Microbiology* **84**. <https://doi.org/10.1128/AEM.00063-18>
- Price M.N., Dehal P.S. & Arkin A.P. (2010). Fasttree 2 – approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490. <https://doi.org/10.1371/journal.pone.0009490>
- Quast C., Pruesse E., Yilmaz P., Gerken J., Schweer T., Yarza P., *et al.* (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* **41**, D590–D596. <https://doi.org/10.1093/nar/gks1219>
- R Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Salinas I. (2015). The mucosal immune system of teleost fish. *Biology* **4**, 525–539. <https://doi.org/10.3390/biology4030525>
- Sellyei B., Molnár K. & Székely C. (2017). Diverse *Chlamydia* -like agents associated with epitheliocystis infection in two cyprinid fish species, the common carp (*Cyprinus*

- carpio* L.) and the gibel carp (*Carassius auratus gibelio* L.). *Acta Veterinaria Hungarica* **65**, 29–40. <https://doi.org/10.1556/004.2017.003>
- Sevellec M., Derome N. & Bernatchez L. (2018). Holobionts and ecological speciation: the intestinal microbiota of lake whitefish species pairs. *Microbiome* **6**. <https://doi.org/10.1186/s40168-018-0427-2>
- Šimková A., Cíváňová K., Gettová L. & Gilles A. (2013). Genomic porosity between invasive *Chondrostoma nasus* and endangered endemic *Parachondrostoma toxostoma* (Cyprinidae): The evolution of MHC IIB genes. *PLoS ONE* **8**, e65883. <https://doi.org/10.1371/journal.pone.0065883>
- Simon J.-C., Marchesi J.R., Mougél C. & Selosse M.-A. (2019). Host-microbiota interactions: from holobiont theory to analysis. *Microbiome* **7**. <https://doi.org/10.1186/s40168-019-0619-4>
- Sylvain F.-É., Cheaib B., Llewellyn M., Gabriel Correia T., Barros Fagundes D., Luis Val A., *et al.* (2016). pH drop impacts differentially skin and gut microbiota of the Amazonian fish tambaqui (*Colossoma macropomum*). *Scientific Reports* **6**, 32032. <https://doi.org/10.1038/srep32032>
- Tarnecki A.M., Brennan N.P., Schloesser R.W. & Rhody N.R. (2018). Shifts in the skin-associated microbiota of hatchery-reared common snook *Centropomus undecimalis* during acclimation to the wild. *Microbial Ecology*. <https://doi.org/10.1007/s00248-018-1252-7>
- Tremaroli V. & Bäckhed F. (2012). Functional interactions between the gut microbiota and host metabolism. *Nature* **489**, 242–249. <https://doi.org/10.1038/nature11552>
- Uren Webster T.M., Consuegra S., Hitchings M. & Garcia de Leaniz C. (2018). Interpopulation variation in the Atlantic salmon microbiome reflects environmental and genetic diversity. *Applied and Environmental Microbiology* **84**. <https://doi.org/10.1128/AEM.00691-18>
- Ursell L.K., Clemente J.C., Rideout J.R., Gevers D., Caporaso J.G. & Knight R. (2012). The interpersonal and intrapersonal diversity of human-associated microbiota in key body sites. *Journal of Allergy and Clinical Immunology* **129**, 1204–1208. <https://doi.org/10.1016/j.jaci.2012.03.010>
- Vörösmarty C.J., McIntyre P.B., Gessner M.O., Dudgeon D., Prusevich A., Green P., *et al.* (2010). Global threats to human water security and river biodiversity. *Nature* **468**, 334–334. <https://doi.org/10.1038/nature09549>
- Wang A.R., Ran C., Ringø E. & Zhou Z.G. (2018). Progress in fish gastrointestinal microbiota research. *Reviews in Aquaculture* **10**, 626–640. <https://doi.org/10.1111/raq.12191>
- Wang Q., Garrity G.M., Tiedje J.M. & Cole J.R. (2007). Naive bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* **73**, 5261–5267. <https://doi.org/10.1128/AEM.00062-07>
- Weiss S., Xu Z.Z., Peddada S., Amir A., Bittinger K., Gonzalez A., *et al.* (2017). Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* **5**. <https://doi.org/10.1186/s40168-017-0237-y>
- Wilson M. (2008). *Bacteriology of humans: an ecological perspective*. Blackwell Pub, Malden, MA.
- Xing M., Hou Z., Yuan J., Liu Y., Qu Y. & Liu B. (2013). Taxonomic and functional metagenomic profiling of gastrointestinal tract microbiome of the farmed adult turbot (*Scophthalmus maximus*). *FEMS Microbiology Ecology* **86**, 432–443. <https://doi.org/10.1111/1574-6941.12174>

Yan Q., Li J., Yu Y., Wang J., He Z., Van Nostrand J.D., *et al.* (2016). Environmental filtering decreases with fish development for the assembly of gut microbiota: Gut microbiota assembly across fish development. *Environmental Microbiology* **18**, 4739–4754. <https://doi.org/10.1111/1462-2920.13365>

## Figures legend

**Figure 1.** Map of the sampling sites along the Durance and Ardeche River. Black bars represents major dams and blue lines the main canals connected to rivers. The number of individual sampled (N) is indicated for each site.

**Figure 2.** Alpha diversity A) Shannon B) Phylogenetic Diversity indices of microbiota associated with tissues. Each dot represents a sample (mean  $\pm$  SE). skin (red), gills (orange), midgut (darkblue) and hindgut (lightblue). Bar and error bars represents mean  $\pm$  SE for each tissues in each site.

**Figure 3.** Principal coordinates analysis exploring the effect of tissues and river on the Bray-Curtis distance of microbiota composition. We present the three dimension plot of PCoA analysis performed on Qiime2. Each point corresponds to one microbiota sample colored following the 4 mucosal tissues: skin (red), gills (orange), midgut (darkblue), and hindgut (lightblue). Filled dots correspond to samples from Ardeche River and rings to samples from Durance River.

**Figure 4.** Principal coordinates analysis exploring the effect of river and sites on the Bray-Curtis distance of microbiota composition. We present the two first axes of independent PCoA analyses performed A) skin B) gills C) midgut and D) hindgut. Each point corresponds to one microbiota sample colored following Buech (darkblue), Manosque (blue), Pertuis (lightblue), Rosieres (darkgreen) and St-Just (lightgreen). Filled dots correspond to samples from Ardeche River and rings to samples from Durance River.

**Figure 5.** Cumulative bar charts of main bacterial phyla present in mucosal tissues of *P. toxostoma* in each sites sampled. Percentages show the mean relative abundance of each phylum by tissues and sites.

**Figure 6.** Variable functional metabolic pathway of a) skin, b) midgut and c) hindgut microbiota between populations from Ardeche and Durance River. Only significant metabolic pathways are listed and sorted by their effect size. Standard deviation is represented on each bar subplot. Colors corresponds to samples from Pt populations of Ardeche (green) and Durance (blue).

### **Supplementary figures legend**

**Figure S1.** Mock microbial community Dna standard. Cumulative bar charts showing the mean relative abundance of expected taxon in Zymo Dna standard and observed in amplification 1 and 2 (Mock-1 and Mock-2) for each Run.

**Figure S2.** Sequencing depth distribution by sample. Bars represent the number of reads for each sample.

**Figure S3.** Rarefaction curves for A) Shannon and B) Phylogenetic Diversity indices from 1000 to 15000 reads by samples with 10 iteration by step.

**Figure S4.** Principal coordinates analysis exploring the effect of tissues and river on the unweighted Unifrac distance of microbiota composition. We present the three dimension plot of PCoA analysis performed on Qiime2. Each point corresponds to one microbiota sample colored following the 4 mucosal tissues: skin (red), gills (orange), midgut (darkblue), and hindgut (lightblue). Filled dots correspond to samples from Ardeche River and rings to samples from Durance River.

**Figure S5.** Principal coordinates analysis exploring the effect of river and sites on the unweighted Unifrac distance of microbiota composition. We present the two first axes of independent PCoA analyses performed A) skin B) gills C) midgut and D) hindgut. Each point corresponds to one microbiota sample colored following Buech (darkblue), Manosque

(blue), Pertuis (lightblue), Rosieres (darkgreen) and St-Just (lightgreen). Filled dots correspond to samples from Ardeche River and rings to samples from Durance River.

**Figure S6.** Individual relative abundance of the pathogen *Candidatus Piscichlamydia sp.* in microbiota associated with gills after a 10,000 rarefaction procedure. Each dot corresponds to one sample in the different sites.

**Figure S7.** Principal coordinates analysis on metabolic functional pathways of midgut microbiota. The figure includes only samples with more than 10000 reads assigned to Greengenes OTU. Three first axes are represented, each dot corresponds to a sample from Durance (blue) and Ardeche (green). We detect an outlier corresponding to a sample from Pertuis (Durance).

### Supplementary files

**File S1.** Metadata file. Tab-separated text file detailing information on each sample and controls.

**File S2.** R script of Dada2 pipeline.

**File S3.** Summary of reads retained for sample at each step of Dada2 pipeline for Run1, Run2 and Run3.

**File S4.** Qiime2 file corresponding to cumulative bar chart for controls of Run1. Interactive file .qzv is readable on <https://view.qiime2.org/>. Controls without reads are not represented. "Blanc-": extraction control, "T-": PCR control and "empty-": unused combinations of tagged-primers.

**File S5.** Qiime2 file corresponding cumulative bar chart for controls of Run2. Interactive file .qzv is readable on <https://view.qiime2.org/>. Controls without reads are not

represented. “Blanc-”: extraction control, “T-”: PCR control and “empty-”: unused combinations of tagged-primers.

**File S6.** Qiime2 file corresponding cumulative bar chart for controls of Run3. Interactive file .qzv is readable on <https://view.qiime2.org/>. Controls without reads are not represented. “Blanc-”: extraction control, “T-”: PCR control and “empty-”: unused combinations of tagged-primers.

**Table 1.** Summary of linear mixed models of alpha diversity variation of microbiota. The models including the fixed factor reproductive status, river, tissues and river : tissues and the random factors identity and sites were performed independently on Shannon and Phylogenetic diversity indices.

Fixed factors	Shannon			Phylogenetic Diversity		
	Df	F value	p value	Df	F value	p value
reproductive status	2	3.72	0.03	2	6.21	<0.01
river	1	0.62	0.48	1	0.47	0.54
tissues	3	53.44	<0.01	3	33.03	<0.01
river:tissues	3	12.12	<0.01	3	14.16	<0.01

**Table 2.** Permanova exploring the effect of reproductive status, tissues, river, site and tissues x river interaction on dissimilarity distances between microbiota compositions. The model was performed on Bray-Curtis distance matrices. The terms were added sequentially and tested on pseudo F value and random distribution estimated after 10,000 permutations.

<b>Bray-Curtis</b>				
Terms added sequentially	Df	pseudo-F	R <sup>2</sup>	p value
reproductive status	2	4.93	0.03	<0.01
tissues	3	12.57	0.13	<0.01
river	1	18.18	0.06	<0.01
site	3	5.48	0.05	<0.01
tissues : river	3	6.60	0.06	<0.01

**Table 3.** Post-hoc pairwise permanova of microbiota diisimilarity based on bray-curtis index between sites for each tissues. Matrices A) skin, B) Gills, C) Midgut, D) Hindgut, indicated p values associated to each pairwise permanova and corrected through Benjamini & Hochberg procedure.

**A) Skin**

Sites	Rosieres	Buech	Manosque	Pertuis
Buech	<0.010	-	-	-
Manosque	<0.010	<0.010	-	-
Pertuis	<0.010	<0.010	<0.010	-
St-Just	<0.010	<0.010	<0.010	<0.010

**B) Gills**

Sites	Rosieres	Buech	Manosque	Pertuis
Buech	<0.010	-	-	-
Manosque	<0.010	0.501	-	-
Pertuis	<0.010	0.013	<0.010	-
St-Just	<0.010	<0.010	<0.010	<0.010

**C) Midgut**

Sites	Rosieres	Buech	Manosque	Pertuis
Buech	<0.010	-	-	-
Manosque	<0.010	<0.010	-	-
Pertuis	<0.010	<0.010	0.015	-
St-Just	<0.010	<0.010	<0.010	<0.010

**D) Hindgut**

Sites	Rosieres	Buech	Manosque	Pertuis
Buech	<0.010	-	-	-
Manosque	<0.010	<0.010	-	-
Pertuis	<0.010	<0.010	0.107	-
St-Just	0.043	<0.010	<0.010	<0.010

Figure 1

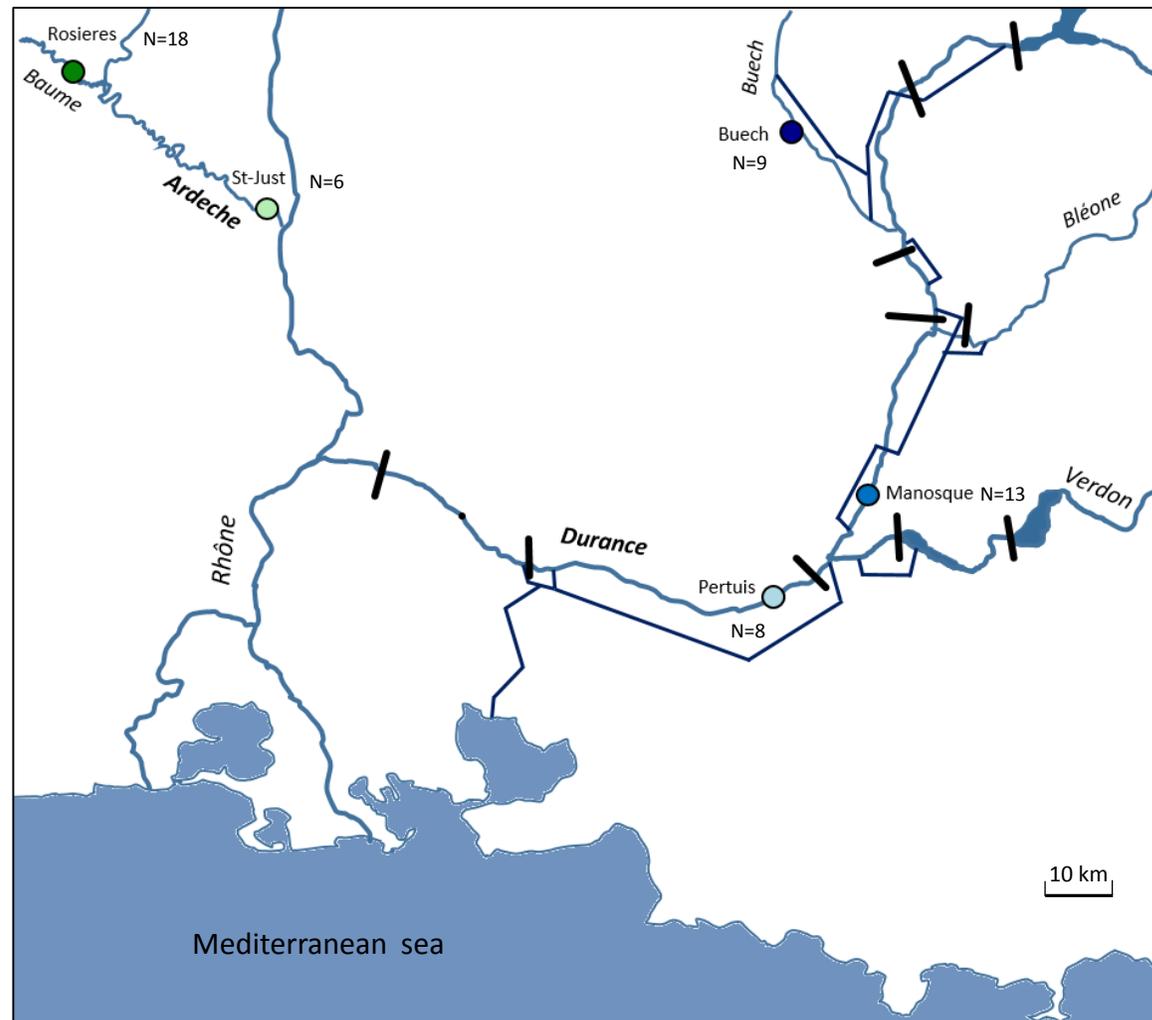


Figure 2

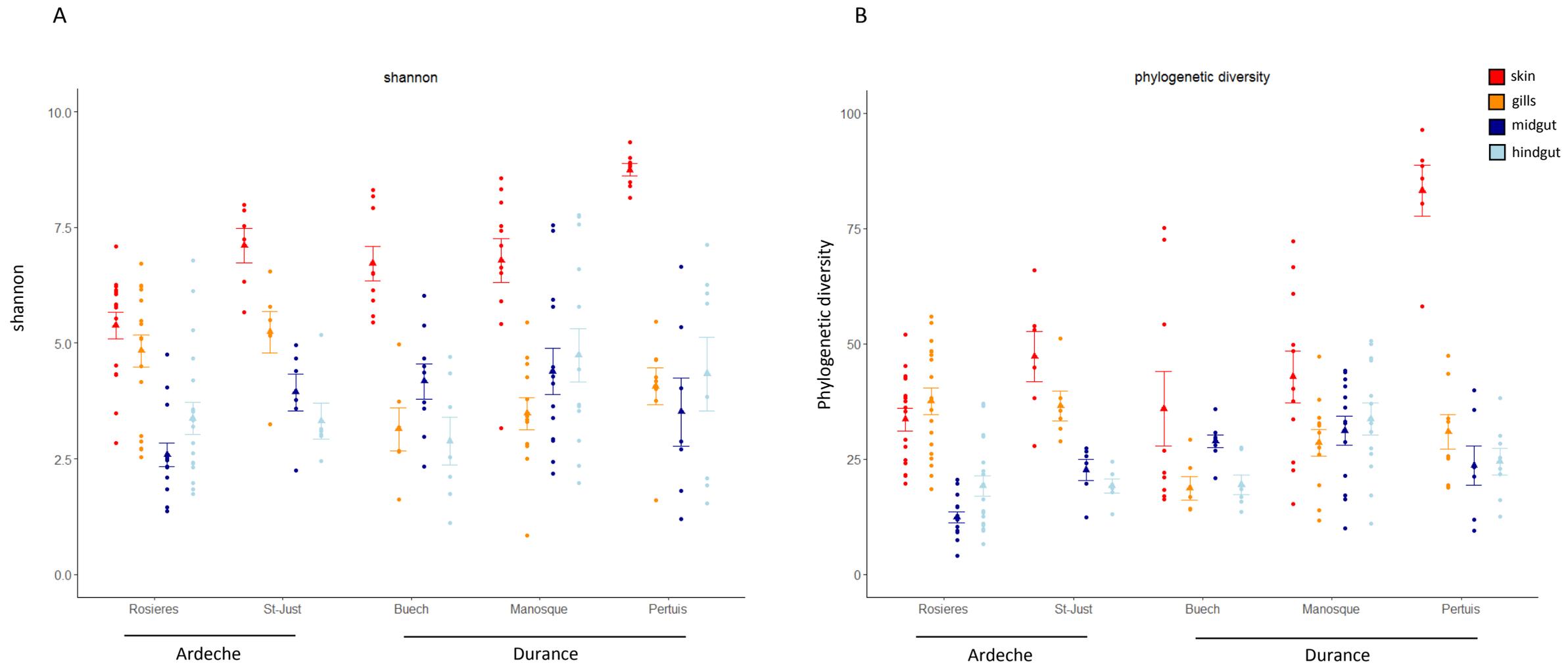


Figure 3

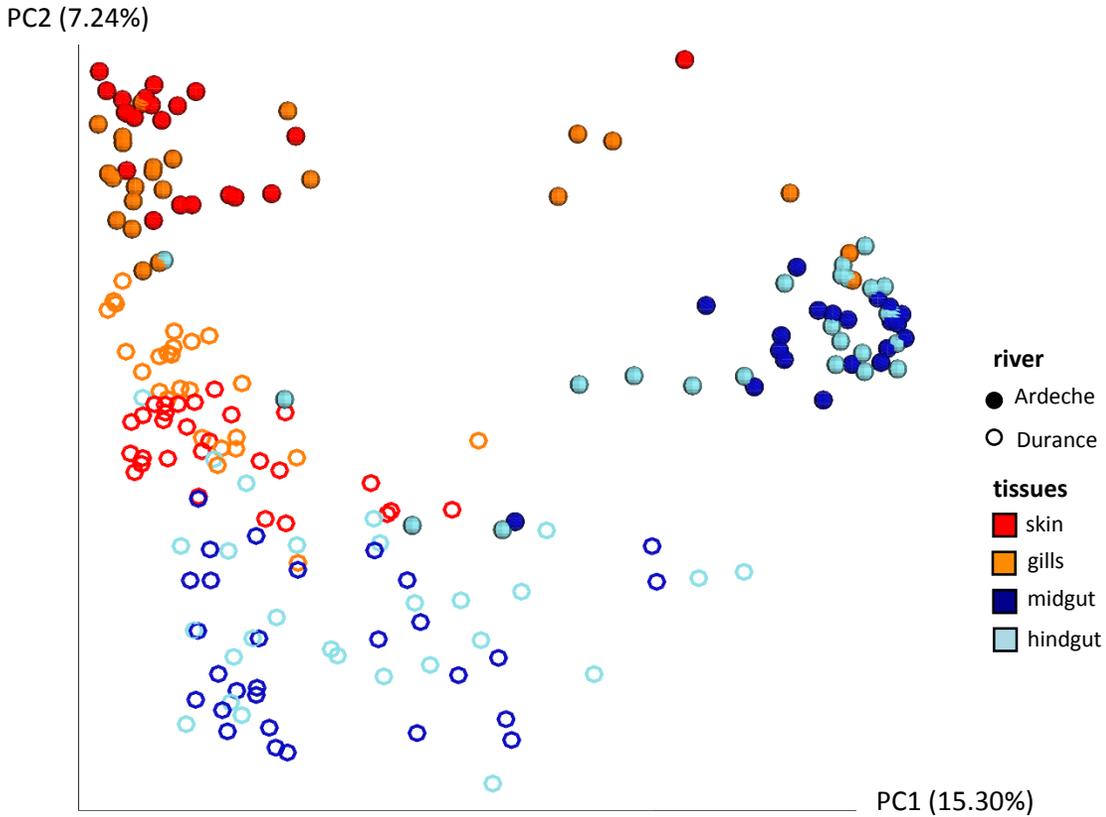


Figure 4

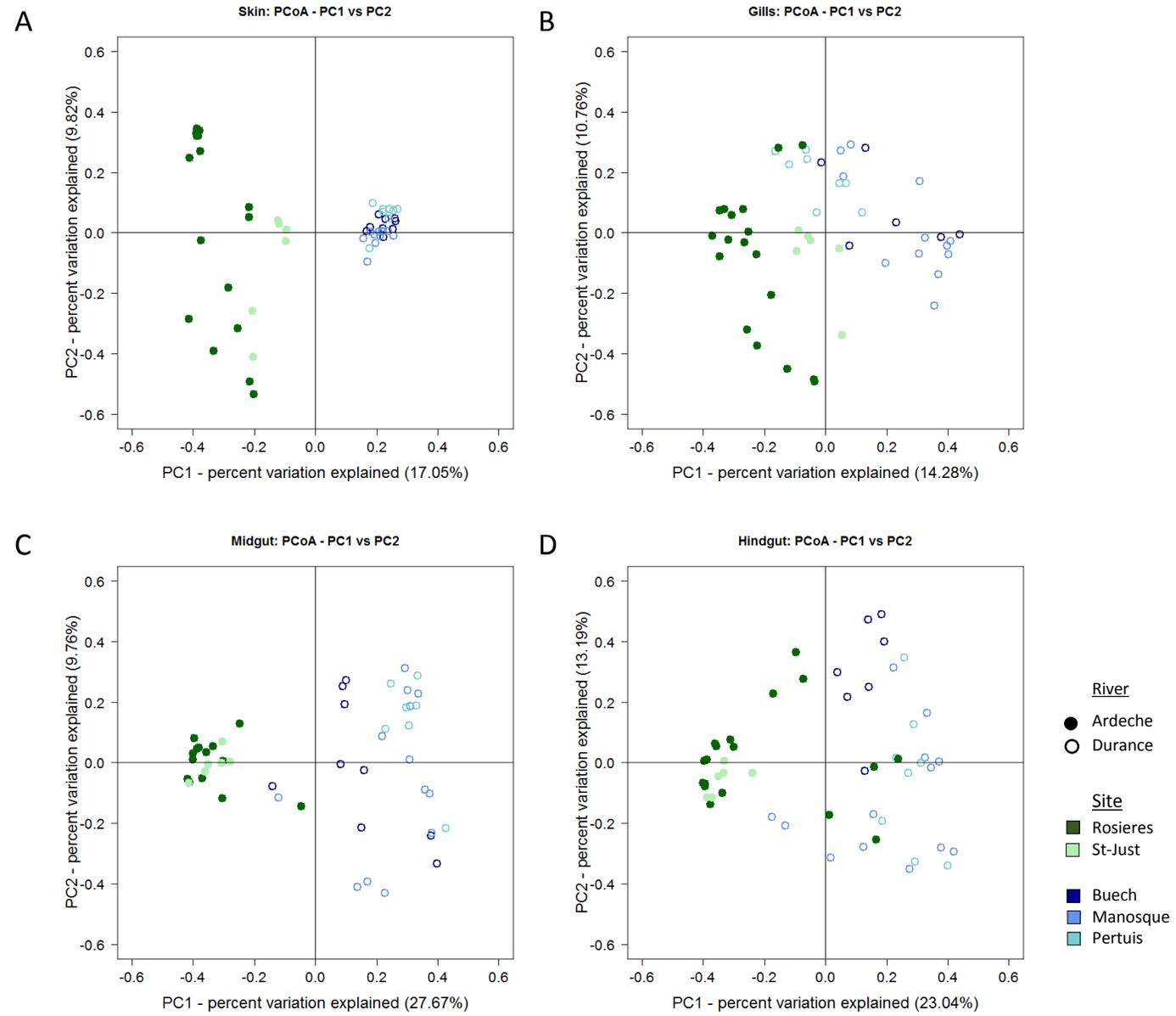
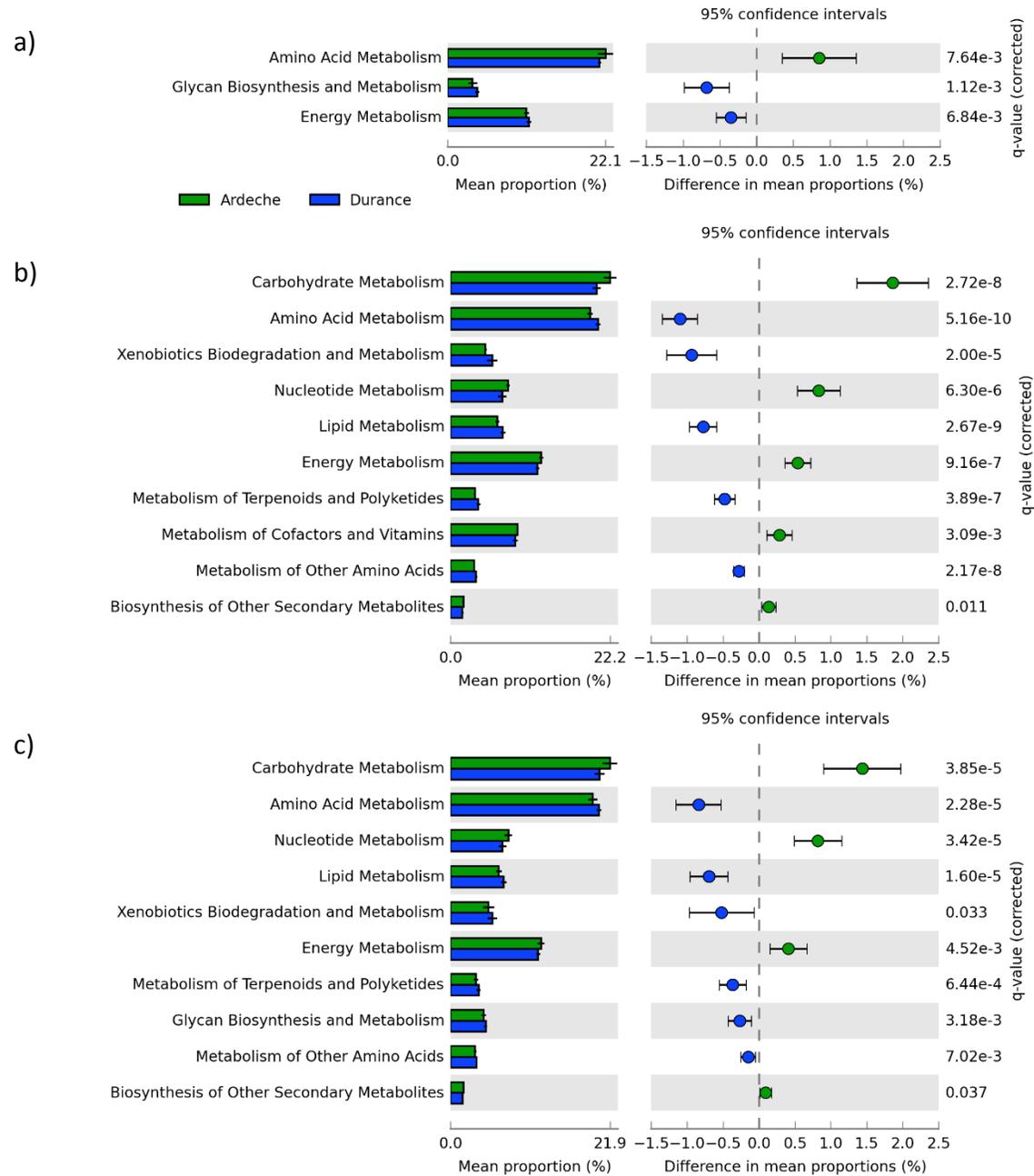




Figure 6



**Table S1.** Summary of coefficient effect of fixed factors of linear mixed models on A) Shannon and B) Phylogenetic diversity variation of microbiota. The models including the fixed factor reproductive status, river, tissues and river : tissues and the random factors identity and sites.

**A) Shannon**

	<b>Estimate</b>	<b>Std. Error</b>	<b>df</b>	<b>t value</b>	<b>Pr(&gt; t )</b>
(Intercept)	6.702	0.527	10.820	12.713	<0.001
rep.status sexually_immature	-0.706	0.456	50.990	-1.547	0.128
rep.status male	-1.043	0.393	48.660	-2.652	0.011
river Durance	1.403	0.596	6.860	2.355	0.051
tissues gills	-0.883	0.388	142.400	-2.275	0.024
tissues hindgut	-2.458	0.388	142.400	-6.333	<0.001
tissues midgut	-2.761	0.409	143.680	-6.758	<0.001
river Durance : tissues gills	-2.917	0.532	148.010	-5.489	<0.001
river Durance : tissues hindgut	-0.714	0.526	146.920	-1.357	0.177
river Durance : tissues midgut	-0.405	0.539	146.020	-0.751	0.454

**B) Phylogenetic Diversity**

	<b>Estimate</b>	<b>Std. Error</b>	<b>df</b>	<b>t value</b>	<b>Pr(&gt; t )</b>
(Intercept)	3.822	0.179	6.950	21.388	<0.001
rep.status sexually_immature	-0.132	0.124	50.970	-1.061	0.294
rep.status male	-0.322	0.106	49.280	-3.039	0.004
river Durance	0.167	0.213	5.160	0.784	0.468
tissues gills	0.022	0.118	143.560	0.191	0.849
tissues hindgut	-0.689	0.118	143.560	-5.848	<0.001
tissues midgut	-0.901	0.124	145.300	-7.266	<0.001
river Durance : tissues gills	-0.633	0.161	149.270	-3.934	<0.001
river Durance : tissues hindgut	0.096	0.159	148.040	0.604	0.547
river Durance : tissues midgut	0.386	0.163	147.410	2.367	0.019

**Table S2.** Permanova exploring the effect of reproductive status, tissues, river, site and tissues x river interaction on dissimilarity distances between microbiota compositions. The model was performed independently on unweighted Unifrac distance matrice. The terms were added sequentially and tested on pseudo F value and random distribution estimated after 10,000 permutations.

Terms added sequentially	Df	unweighted Unifrac		
		pseudo-F	R <sup>2</sup>	p value
reproductive status	2	4.60	0.03	<0.01
tissues	3	9.19	0.10	<0.01
river	1	8.26	0.03	<0.01
site	3	4.37	0.05	<0.01
tissues : river	3	3.55	0.04	<0.01

**Table S3.** Post-hoc pairwise permanova of microbiota diisimilarity based on unweighted unfrac index between sites for each tissues. Matrices A) skin, B) Gills, C) Midgut, D) Hindgut, indicated p values associated to each pairwise permanova and corrected through Benjamini & Hochberg procedure.

**A) Skin**

Sites	Rosieres	Buech	Manosque	Pertuis
Buech	<0.010	-	-	-
Manosque	<0.010	0.017	-	-
Pertuis	<0.010	<0.010	<0.010	-
St-Just	<0.010	<0.010	<0.010	<0.010

**B) Gills**

Sites	Rosieres	Buech	Manosque	Pertuis
Buech	<0.010	-	-	-
Manosque	<0.010	<0.010	-	-
Pertuis	<0.010	<0.010	<0.010	-
St-Just	<0.010	<0.010	<0.010	<0.010

**C) Midgut**

Sites	Rosieres	Buech	Manosque	Pertuis
Buech	<0.010	-	-	-
Manosque	<0.010	<0.010	-	-
Pertuis	<0.010	<0.010	0.06019	-
St-Just	<0.010	<0.010	<0.010	<0.010

**D) Hindgut**

Sites	Rosieres	Buech	Manosque	Pertuis
Buech	<0.010	-	-	-
Manosque	<0.010	<0.010	-	-
Pertuis	<0.010	<0.010	<0.010	-
St-Just	<0.010	<0.010	<0.010	<0.010

Figure S1

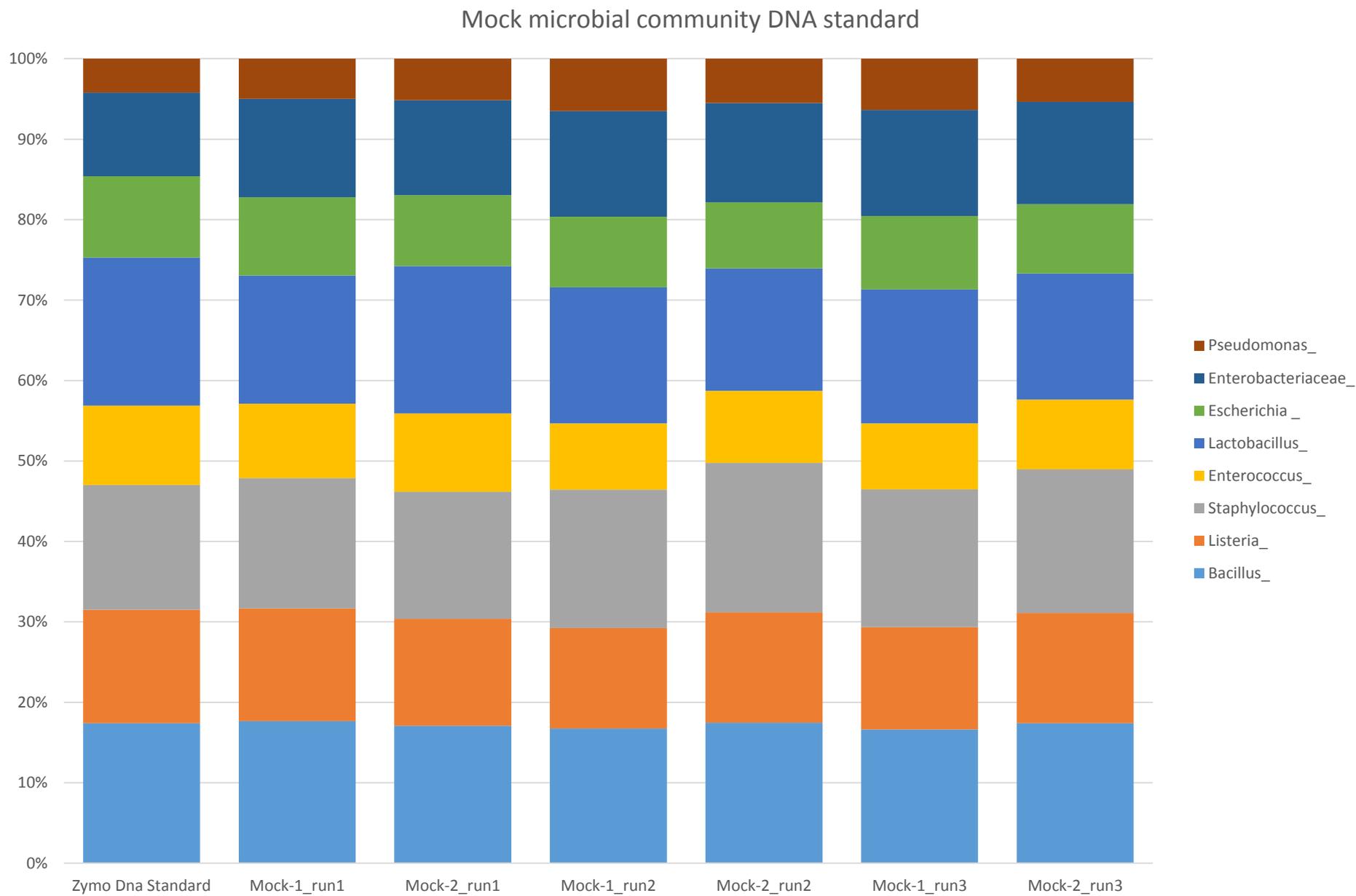
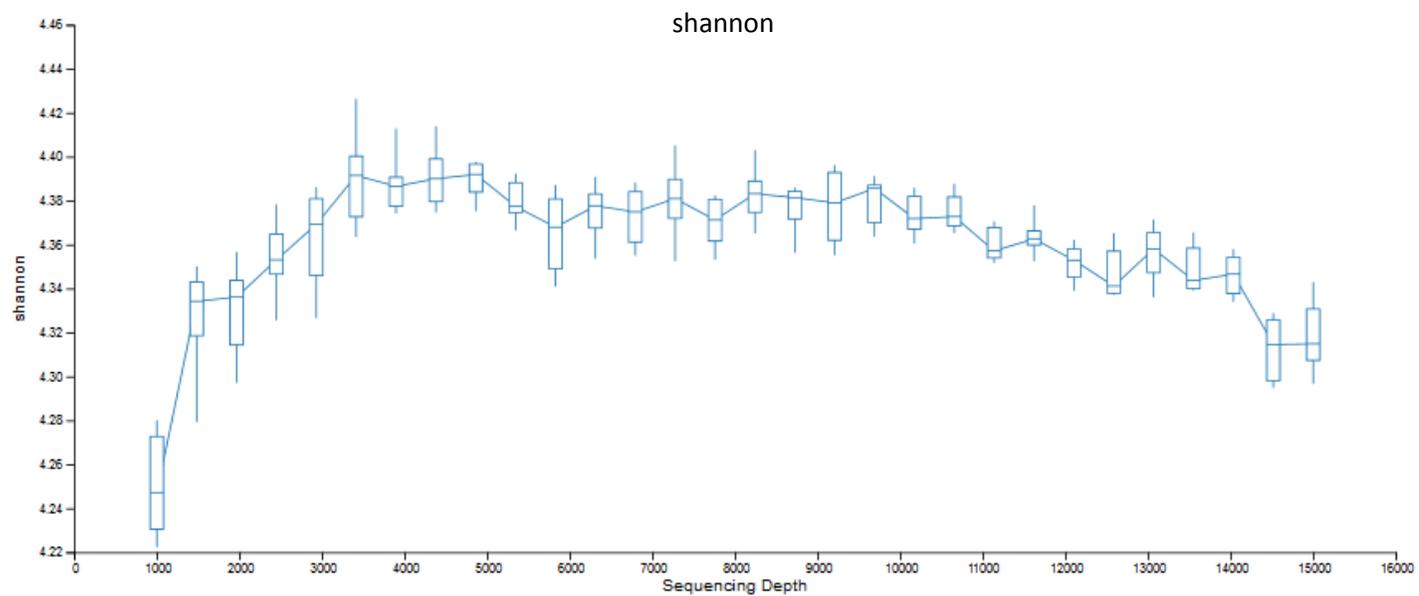




Figure S3

A



B

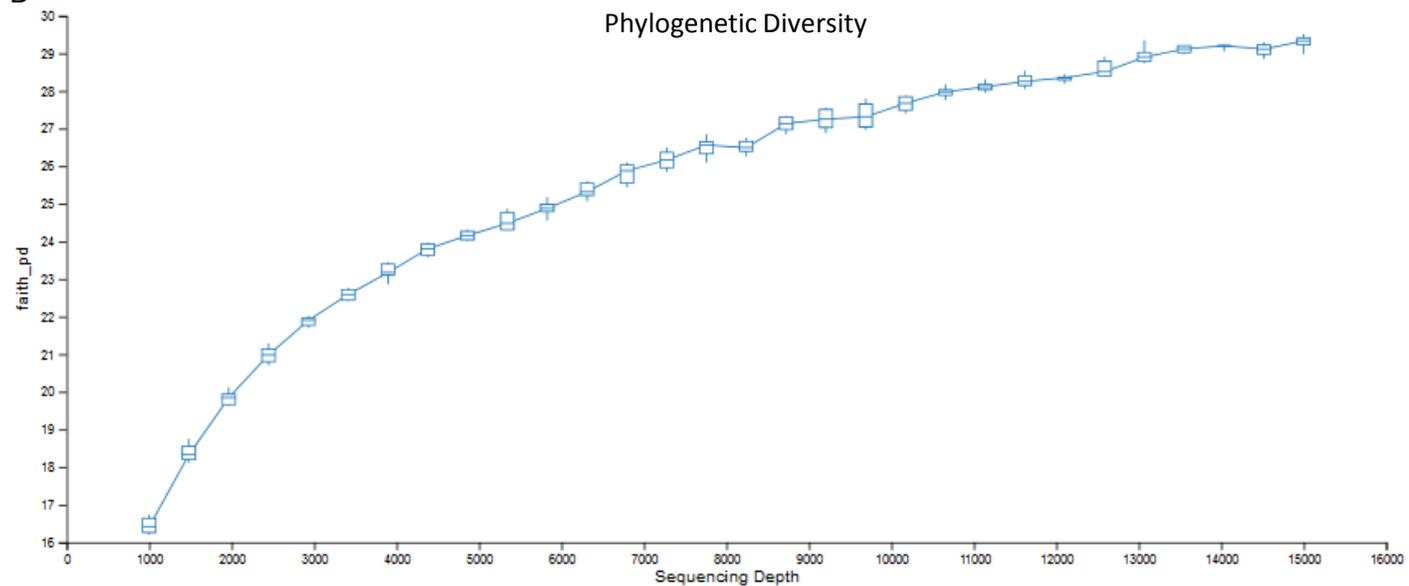


Figure S4

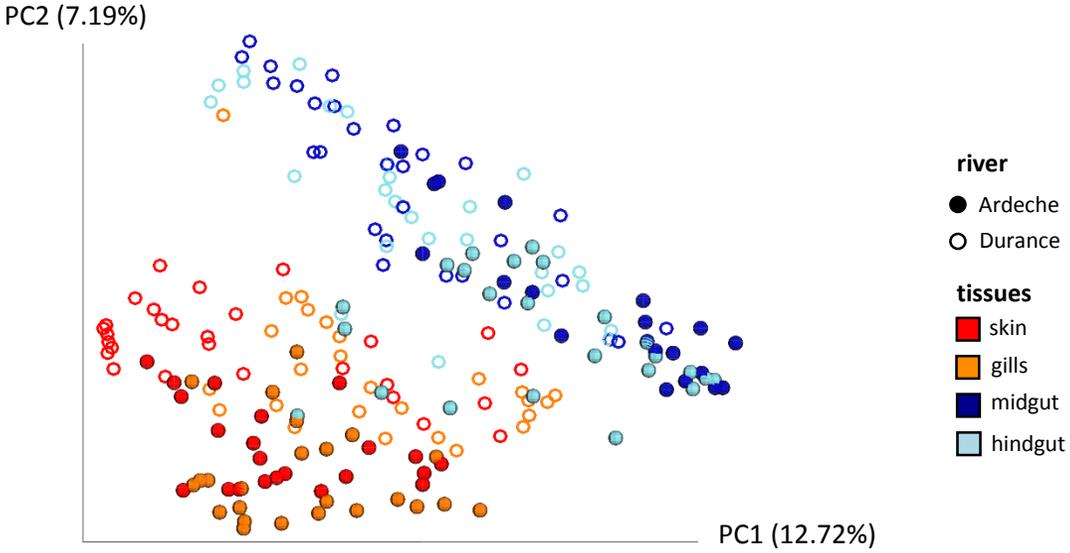


Figure S5

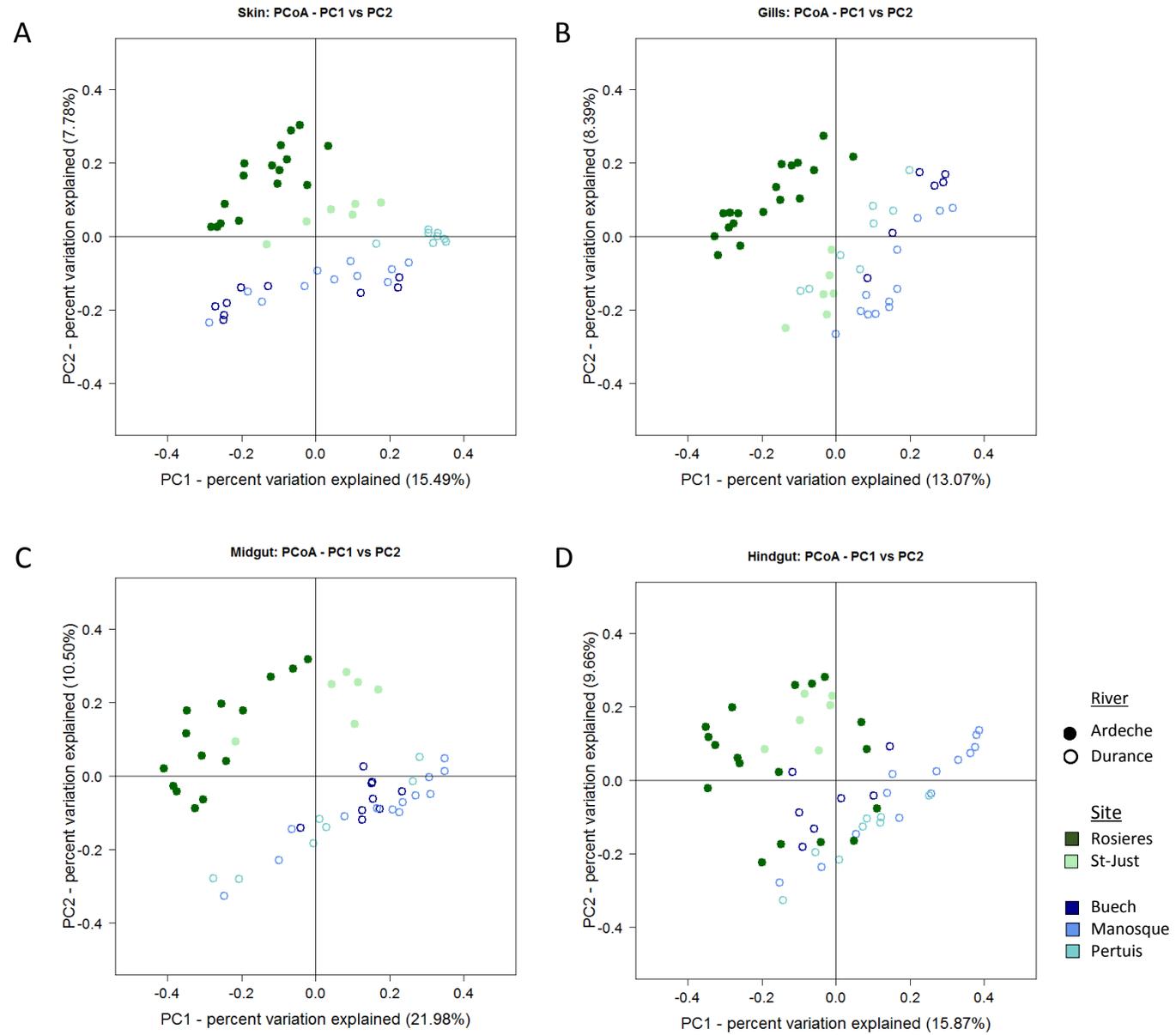




Figure S7

