bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Multilevel community assembly of the tadpole gut microbiome

Decio T. Correa[1*], David Rodriguez[2], Carine Emer[3], Daniel Saenz[4], Cory K. Adams[4], Luis Schiesari[5], Mikhail Matz[1], Mathew A. Leibold[6]

1. Department of Integrative Biology, The University of Texas at Austin, Austin, TX, USA
2. Department of Biology, Texas State University, San Marcos, TX, USA
3. Chico Mendes Institute for Biodiversity Conservation, National Centre for Wild Birds Research and Conservation, João Pessoa, PB, Brazil
4. Southern Research Station, US Forest Service, Nacogdoches, TX, USA
5. School of Arts, Sciences and Humanities, University of São Paulo, São Paulo, SP, Brazil
6. Department of Biology, University of Florida, Gainesville, FL, USA

* Corresponding author, decio.tadeu@gmail.com

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

The assembly of local communities is likely to reflect the effects of local environmental factors associated with filters that act at larger spatial scales. Dissecting these multiscale effects remains a timely challenge that is particularly important for host-associated microbiomes. We investigated the relative roles of local selection (due to host species identity) and regional effects (due to water body identity) on the community structure of bacteria in the gut of tadpoles from three biogeographic areas and used graph theory and metanetwork approaches to explore and illustrate the distribution of bacteria across different ponds. The pond of origin, which represents a regional species pool of bacteria, was in general more important in shaping the gut microbiome of tadpoles than host species identity. The resulting metanetworks are modular and indicate relatively few species of bacteria occurring in more than one pond. Thus, each pond represents a relatively distinct species pool of bacteria available for community assembly of the tadpole microbiomes. Our findings indicate that microbiome community assembly in amphibian larvae, as in many other communities, is a multiscale process with important regional effects that constrain how local (i.e. host-dependent) filters act to influence microbiome community composition.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

The origin and assembly of host-associated microbial communities (microbiomes) is a complex process of fundamental importance in helping us understand the symbiotic relationship between hosts and their microbes [1, 2]. Various factors can play a role in the community assembly of microbiomes at multiple levels. Perhaps the most obvious is the effect of host diet on the gut microbiome, which has been shown in many organisms from insects to humans [3–6]. Other local factors include host sex, health state, and genetics [7–11]. These effects can be thought of as local 'filters' to microbial community assembly. Here we evaluate these effects as they are determined by host species identity only, recognizing that there are likely to be other more subtle effects involving inter-individual factors. However, there can also be processes that constrain which microbes are available to colonize any given host. Here, we focus on factors that determine the microbial species pool available to colonize the host and evaluate these by the identity of the site of the host community.

We view community assembly of microbiomes as being driven by metacommunity processes [12]. This recognizes that the assembly of a local community depends strongly on the species available to colonize that community, known as the regional species pool [13, 14]. From that pool, if a species is able to disperse to a given community, its persistence in that community will depend on local demographic processes and on the ability of the species to tolerate the local biotic and abiotic conditions, which is known as the local selection component of community assembly

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[15]. Therefore, communities with identical environmental conditions could have a completely different species composition if colonized by different species pools [16]. While most work in metacommunity ecology has focused on macroorganisms, these ideas should apply equally well to microbes [1, 2, 12, 17, 18].

Host-associated microbial communities appear to be influenced by effects related to the individual host or individual species such as diet, physiology, and genotype [10, 19, 20], as well as regional and historical processes, as in the case of exposure to potentially different species pools of microorganisms [21–26], or vertical and horizontal transmission of the microbiome [27, 28]. Considering the importance that the microbiome can have to host health and well-being, understanding the relative importance of effects acting at different levels is fundamental, but studies explicitly tackling that question are still lacking [29].

We analyzed variation in tadpole microbiomes to evaluate the effects of local environmental filters by focusing on host species identity, and regional effects by focusing on the locality of the water body. We assume that different microbes would be differentially favored by selection associated with the physiology and diet of different tadpole species, and different microbes may be present at different sites due to either site-specific habitat variables or to isolation and other differences among water bodies. Our studies were conducted across three different geographic locations to test if they show consistent patterns. We also used graph theory and metanetwork approaches to

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

74 explore the degree to which each pond consists of a unique species pool, and how they
75 are connected to each other in terms of shared bacterial species.

76 The gut microbiome of tadpoles is critically important to their digestive function
77 [30], and some evidence shows gut microbiomes from the same host species are more
78 similar to each other than from different species [31]. Further, there are often multiple
79 species of tadpole per water body, which represent different hosts with different
80 characteristics that are exposed to the same pool of bacteria. Many of these tadpole
81 species can occupy a variety of different water bodies, each with potentially distinct
82 bacterial species pools, which makes them a suitable system to test questions regarding
83 the effects of environment and host identity on microbiome assembly [32]. We evaluate
84 two hypotheses: i) if the bacterial community of each pond is different across the
85 landscape, then the pond of origin would be more important in determining the gut
86 microbiome of tadpoles, i.e., the microbiomes of different species within each pond
87 would be more similar to each other than the microbiomes of the same species across
88 distinct ponds, and the metanetworks would be structured into modules according to
89 each pond; ii) if ponds have very similar bacterial communities across the landscape,
90 then the species of tadpole would be more important in determining their gut
91 microbiome, i.e., tadpoles of the same species across ponds would have more similar
92 microbiomes than tadpoles of different species from the same pond, and the
93 metanetworks would be structured into modules according to each species.
94

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Methods

Study area

We sampled tadpoles from 22 water bodies located in three different localities: i) six water bodies in the Boracéia Biological Station, located in the Atlantic Forest east of the state of São Paulo, Brazil, hereafter BR; ii) nine water bodies in the David Crockett and Angelina National Forests, located in the pine forest ecosystem in eastern Texas, United States, hereafter ET; and iii) seven water bodies in the Edwards Plateau, a mix of grassland and juniper/oak woodlands in the central part of Texas, United States, hereafter CT. We classified water bodies as either lentic or lotic. For ET ponds, we did a more thorough environmental characterization quantifying in each pond the dissolved oxygen (mg/L), pH, conductivity (uS/cm), chlorophyll a (ug/L), and water temperature (°C) using a Eureka multiparameter sonde, and the absence or presence of piscivorous (Bass) and/or insectivorous fish (Green Sunfish). Even though three water bodies in BR were lotic, we will refer to all of them as *ponds* for convenience.

Tadpole and microbiome sampling

In each pond we collected tadpoles with a dipnet that was swept through the pond. We then placed the larvae in a sterile plastic bag (Whirl Pak®) filled with water from their place of origin. Environmental microorganisms from water from CT and ET were sampled by filtering approximately 50 ml of water from three different places through a sterile syringe filter with a 0.45 μm cellulose acetate membrane (VWR). From

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

116 BR, approximately 2L of water was collected in a sterile plastic bag (Whirl Pak®), mixed,
117 and 100 mL was vacuum-filtered through a 0.45 μm cellulose acetate membrane. The
118 sediment in all ponds was sampled in three locations in each pond by removing the first
119 three centimeters with a sterile plastic straw of 1.3 cm of diameter and placed in a
120 sterile tube.

121 In the lab, tadpoles were euthanized by overdose in a solution of Milli-Q® filtered
122 water and tricaine methanesulfonate (MS 222). After the death was confirmed, each
123 individual was dissected using instruments that were cleaned in 100% ethanol and
124 flame-sterilized. The entire tadpole gut was removed and placed directly in a MoBio
125 PowerSoil® bead tube. The sediment from each pond was homogenized, centrifuged to
126 remove the excess water, and an aliquot of approximately 0.2 g was used for DNA
127 extraction. DNA from the water sample was extracted from the entire filter membrane.
128 We performed DNA extraction using MoBio PowerSoil® DNA isolation kit following the
129 manufacturer's recommendation. The same procedures without a real sample were
130 repeated to count as negative control samples, which did not have DNA when quantified
131 using Qubit™ dsDNA high sensitivity assay.

132 We amplified the V4 hypervariable region of the 16s rRNA gene (515F, 806F). CT
133 samples were amplified using the version F: GTGYCAGCMGCCGCGGTA / R:
134 GGACTACHVGGGTWTCTAAT and ET and BR the version F: GTGYCAGCMGCCGCGGTAA /
135 R: GGACTACNVGGGTWTCTAAT. The library preparation and sequencing (Illumina MiSeq
136 2x250bp) of the CT samples was realized at the Genomic Sequencing and Analysis

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Facility (wikis.utexas.edu/display/GSAF) at The University of Texas at Austin following their protocols. The library preparation and sequencing (Illumina MiSeq 2x250bp) of the ET samples was realized at the Argonne National Laboratory (www.anl.gov) following their protocols. For BR samples, we prepared the libraries following the Earth Microbiome Project protocol (www.earthmicrobiome.org), with the exception of using 20uL PCR reactions and dual-index barcodes instead. Samples were sequenced in an Illumina MiSeq platform (2x300 bp paired-end sequences).

The sequences were processed using the dada2 pipeline [33] in the software Microsoft R Open v 3.5.1 [34]. Sequences from each site were processed separately. Briefly, primers, adapters, and barcodes were removed and the quality profile of the reads were visually inspected using the function plotQualityProfile aggregating over all the samples from a site. BR, CT, and ET reads were truncated at 200 and 120, 200 and 180, and 200 and 160 forward and reverse reads, respectively. Sequencing errors were calculated, sequences were clustered as Amplicon Sequence Variants based on the DADA2 algorithm [33] and paired ends were merged. Chimeras were removed using the consensus method in the function removeBimeraDenovo. Taxonomy was assigned based on the Silva database v128 [35]. ASV sequences were exported to Qiime2 2018.4 [36], aligned using the MAFFT algorithm [37], and a phylogenetic tree was constructed using FastTree [38]. We discarded all ASVs that were classified as Archaea, chloroplast, mitochondria, or that were not at least assigned to bacteria. Hereafter, we will refer to ASVs as species of bacteria.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

158

159 **Data analysis**

160 **Diversity**

161 For each sample we constructed a rarefaction curve and compared richness and
162 Faith's phylogenetic diversity [39]. To avoid sequencing depth effects, each sample was
163 rarefied 100 times to the number of reads of the sample with the lowest coverage using
164 the R package BAT [40]. For comparison, we considered the median richness from all
165 rarefactions from a sample as our richness measure.

166

167 **Species and pond effects**

168 We analyzed our data as compositional data (termed CODA) [41]. It requires a
169 centered log-ratio transformation of the read counts and therefore cannot have zeros,
170 so we added a pseudocount of one to all ASVs in all samples. We transformed the data
171 using the codaSeq.clr function from the R package CoDaSeq [42, 43]. We also analyzed
172 data that takes into account phylogenetic relatedness of microbes by using the PHILR
173 transformation from the R package philr [44]. This metric takes into account the
174 phylogenetic relationship between the ASVs, which is equivalent to the Unifrac metric
175 [45], but it considers the compositional nature of the data [44]. We then ran a Principal
176 Component Analysis (PCA) on the Euclidean distance matrix of the CODA- and PHILR-
177 transformed data to visualize the relationship between samples and tested if samples
178 from tadpoles, water, and sediment are different from each other using a Permutational

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Multivariate Analysis of Variance (perMANOVA). If a difference was detected we then used additional perMANOVAs for pairwise comparisons with Bonferroni correction.

To first test if either the identity of the species of tadpole or the pond of origin are more related to the gut microbiome composition we used a perMANOVA to fit a model with both pond and species identity to the CODA- and PHILR-transformed data using the function adonis2 from the R package vegan [46]. The significance of the marginal effects was tested based on 999 permutations. In addition, to test for the unique effects of the species of tadpole or the pond of origin as well their shared effects on the gut microbiome we performed variance partitioning and Redundancy Analysis (RDA) [47] using the function varpart in the R package vegan [46]. The significance of each unique component in the variation partitioning was tested using the vegan function anova.cca with 1000 permutations. We repeated perMANOVA and variation partitioning analyses to the data grouping the species of bacteria at Genus and Family level.

Inter-pond variation

We also investigated if there was evidence for dispersal limitation in the bacteria from water and sediment using multiple regression on distance matrices using the function MRM from the R package ecodist [48]. We used the Euclidean distance matrix based on the PHILR and CODA transformations as the dependent variable and a geographical distance matrix between sites as the independent variable. In addition, to

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

account for environmental effects, we also include in the model a distance matrix based on the type of the pond (lentic or lotic) for BR, and a distance matrix based on the environmental variables measured for ET. The latter was a Gower distance matrix [49] created by the function dist.ktab from the R package ade4 [50]. Quantitative variables were Gower-standardized (divided by maximum minus minimum). All CT ponds were lentic and there were no other variables besides spatial distance to include in the model. For models where distance was significant in the MRMs, we applied a Partial Mantel Correlogram approach using the function mpmcorrelogram from the R package with the same name [51]. That tests, at several distance classes, in our case estimated by Sturge's rule [52], the relationship between water and sediment microbiome with geographical distance while controlling for environmental effects.

Metanetworks

We further explored the structure of three different metanetworks within each locality (BR, CT, and ET) to understand the distribution of bacteria across ponds. Shared species (here bacteria) connect different ponds in a metacommunity perspective [53, 54]. We built three bacteria metanetworks for each locality: (i) the tadpole microbiome metanetwork, composed by the bacteria living in the gut of all tadpoles within a pond, independently of the tadpole species, (ii) the water metanetwork, with bacteria found in the water, and (iii) the sediment metanetwork, composed by bacteria found in the sediment of each pond. Then, we built an a_{mn} adjacency matrix for each metanetwork in

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

221 which *m* corresponds to a single pond (i.e., considering bacteria from water, sediment
222 or gut of tadpoles), and *n* to each bacterium. The *mn* element represents the
223 presence/absence of species of bacteria *n* in a pond *m,* and is represented by a link in
224 the graphical representation. Therefore, each node in our metanetwork is either a
225 bacterium or a pond; when the same bacterium is found in more than one pond, there
226 is a link connecting them, forming the edges of the metanetwork.

227 To test whether ponds consist of different bacterial communities, we calculated
228 the modularity of the metanetworks using the FastGreedy algorithm [55, 56] in the
229 software Modular v 0.1 [57]. We tested the significance of the observed modularity
230 against two different null models with 100 permutations each: the Erdős-Rényi [58], in
231 which pond-bacteria interactions are connected randomly, and the 'null model 2' [59],
232 in which the probability of a pond-bacteria interaction is proportional to their number of
233 links in the observed matrix. To identify the bacteria with higher potential to connect
234 the metanetwork through shared occurrence in a higher number of ponds, we
235 calculated the degree (k) for each metanetwork using the software Pajek v. 4.10 [60].
236 Degree measures how many links each node establishes in its correspondent network.
237 For the bacteria, it represents the number of ponds in which it occurs; for the pond, it
238 represents the number of bacteria it has (from tadpoles, water or sediment). For each
239 metanetwork we also calculated the proportion of bacteria found uniquely in p_1, p_2, …,
240 p_n, where n is the total number of ponds in one locality.
241

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Results

Diversity

We registered 19 species of tadpole in all ponds: three in CT, six in ET, and eleven in BR (Table S1). After the filtering steps, there were a total of 533 777, 1 289 662, and 2 064 298 16s rRNA gene fragment reads for CT, ET, and BR, respectively. The rarefaction curve shows that all samples reached an asymptote, even the ones with the lowest number of reads (Figure S1). For richness and phylogenetic diversity measures of bacteria, CT, ET, and BR samples were rarefied to 3 950, 1 791, and 3 916 reads, respectively. There was no significant difference among CT samples in terms of richness, but there was in terms of phylogenetic diversity, with sediment samples having the highest diversity and *Rana berlandieri* tadpoles the lowest (Figure S2). Sediment samples had the highest richness and phylogenetic diversity in ET and *Hylodes phyllodes* the highest in BR (Figures S3 and S4). A great part of the tadpole microbiome was composed by Fusobacteria, Firmicutes, Proteobacteria, and Bacteroidetes, while sediment and water samples were mostly composed by Proteobacteria (Figures S5-S7).

Species and pond effects

The composition of the tadpole gut microbiome is different from both water and sediment samples in all localities (Figure 1, Figure S8, Table S2). Also, in most cases, tadpole gut microbiome samples cluster more with other samples from the same pond than with samples from the same species from a different pond (Figure 1, Figures S8-

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

263 S10). However, within a pond, there can be separation by species and, in fact, the
264 perMANOVA shows that both species identity of tadpole and pond of origin are
265 significant predictors of the tadpole gut microbiome in all localities when using either
266 PHILR- or CODA-transformed data, except for CT CODA, in which species identity was
267 only marginally significant (Table 1). In CT and ET, pond of origin was a stronger
268 predictor of the tadpole gut microbiome than tadpole species identity. The overall
269 results are consistent over all the taxonomic levels considered (Tables S3-S6).
270 For CT and ET, the pond of origin explained considerably more of the variance in
271 tadpole gut microbiome than the species of tadpole (Figure 2). For BR, the tadpole
272 species explained a slightly higher proportion of the variance in their gut microbiome.
273 However, most of the variation in that environment is explained by the joint effects of
274 tadpole species and pond of origin, i.e., the variation that could not be separated in
275 unique components (Figure 2). Again, the overall results are comparable across all the
276 taxonomic levels considered (Tables S7-S8).
277
278 **Inter-pond variation**
279 Only the bacterial communities of water from CT were significantly related to the
280 spatial distance between ponds (Table S9). However, that was the only dataset where
281 environmental data is not available and could not be accounted for, allowing for the
282 possibility that such distance effects are due to spatially structured environmental
283 factors. Partial Mantel Correlograms showed that, for both CODA- (p_{adj} = 0.044) and

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

PHILR-transformed data (p_{adj} = 0.022), there is a spatial signal only for the first distance range, from 174 to 6 223 m. ET water and sediment communities were significantly related to environmental variables for PHILR-transformed data and BR water communities were significantly related to both PHILR- and CODA-transformed data (Table S9).

Metanetworks

All metanetworks are significantly modular according to the two null models used ($p < 0.001$ for all null models; Figure 3, Figure S11). In most cases, each module corresponds to a single pond (Figure 3, Figure S11 and Table S10). Metanetworks of water and sediment are more modular than those of tadpoles for all localities (Figure 3, Figure S11). In all cases, the vast majority of bacteria occur in a single pond (k = 1), varying from 63.2% in East Texas to 91.3% in Brazil (Figure 4, Figures S12-S16). Only 0.1 to 1.1% of bacteria were present in all ponds (Figure 4, Figure S12-S16). From those, for BR and ET they were mostly from the phyla Bacteroidetes and Proteobacteria and Orders Bacteroidales and Desulfovibrionales. CT had only three bacteria present across all ponds. From the ponds' perspective, their degree, which corresponds to the number of unique links they have in the metanetwork, varied from 118 to 615 for water, 90 to 478 for sediment, and 241 to 2 751 for bacteria from tadpoles (Figures S17-S19).

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Discussion

We found that microbial community assembly of tadpole microbiomes was structured by processes happening at different levels and broadly resemble findings from macroorganisms [1, 15, 61–63]. At higher levels, there are inter-pond differences in microbial communities due to site-specific variables, isolation, and/or stochasticity. It was not our goal to understand the drivers of inter-pond differences, which is ultimately responsible for the different bacterial species pool tadpoles are exposed to. However, in some cases, for the ponds where we had environmental data there is a relationship between environmental variables and the bacterial community of water and/or substrate. In these cases, each pond could act as a regional scale environmental filter for bacteria. For CT, we found some evidence for a spatial correlation in bacterial communities among nearby ponds, which could indicate dispersal limitation structuring the pond communities, but that was the only dataset for which we did not have environmental variables. Nevertheless, independently of the causes of the differences among ponds, each one of them constitutes a unique species pool of colonizers for the microbiomes of tadpoles. As tadpoles develop, the community assembly of their microbiome is restricted by that species pool of microbes. Each tadpole comprises a smaller-scale filter to a subset of the microbes within the species pool. Therefore, even if similar habitats (same species of tadpole in our case) are exposed to different species pools, they will not end up with the same microbial community.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The composition and diversity of the gut microbiome of tadpoles can be variable across species. The microbiome of tadpoles from this study is slightly different than previous studies. Proteobacteria and Firmicutes dominated in other studies [31, 64–66] but here we found that Fusobacteria were also important. Alpha and phylogenetic diversity were also mostly higher in our study than in previous ones [31, 64–66]. The knowledge about the gut microbiome of tadpoles is still limited to a few species and quantitative comparisons of diversity across studies is problematic because of differences in sampling procedures, sample processing, data cleaning, and the algorithms and databases. However, overall differences in microbiome composition could be due to the exposure of tadpoles to different pools of colonist microbes.

The pond of origin of the tadpoles was more important in determining their gut microbiome than the species of the tadpole in both of the Texas locations. We assume the same species of tadpole has similar requirements in terms of microbes and/or constitutes a similar patch for colonization by microbes. There were only a few bacteria in common across all ponds in all three localities, therefore each pond constitutes a unique bacterial species pool from which the tadpoles can be colonized. In Brazil, the pond of origin and species of tadpole showed similar relationships with the gut microbiome of tadpoles. However, most of variation in Brazil was due to variation that could not be separately attributed to either because of the uniquely distinct tadpole faunas of ponds.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The limited and unique pool of bacteria tadpoles are exposed to and the higher similarity within a pond can be visualized on the metanetworks of bacteria from water, sediment, and tadpole gut. The metanetworks are modular and each pond constitutes a unique community of bacteria. The vast majority of bacteria occur only in a single pond, and just a few are present in all of them. In fact, for water and sediment, the maximum number of bacteria found in all ponds was only eight. That number would probably increase if more sediment or water was sampled. On the other hand, we likely sampled all species of tadpoles from each water body and the proportion of bacteria from their gut found in all water bodies is still very low (0.1-1.1%) despite similar phyla found in all of them. The major bacteria in tadpole guts belong to orders Bacteroidales and Desulfovibrionales, which do not overlap with the most widespread taxa found in water or sediment. It is common to not find much overlap between bacteria from environment and gut of aquatic organisms [67, 68]. Likewise, we show that the gut microbiome of tadpoles is distinct from both water and sediment from the ponds suggesting that the gut constitutes a strong environmental filter that favors certain taxa. Desulfovibrionales, for example, is a group of sulfur-reducing bacteria that is mostly anaerobic [69]. Therefore, at least certain parts of the digestive tract of tadpoles might be anaerobic [31, 70]. Pryor and Bjorndal [30] detected significant levels of fermentation in the hindgut of bullfrog tadpoles. It is unknown to which extent other species of tadpoles are hindgut fermenters. For those that are, even though they might not have a core

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

364 microbiome, at least not at fine-scale taxonomy levels, there may be a functional core
365 gut microbiome.

366 In terrestrial and aquatic environments, the importance of the regional species
367 pool in shaping the microbial communities can be noticed even when not tested
368 directly. In many species, from insects to humans, individuals likely exposed to the same
369 regional species pool, such as being from the same environment, have more similar
370 microbiomes [21–26]. There are many difficulties in quantifying and testing how
371 regional effects affect community assembly [29, 62], but comparing results with
372 predictions from theory can lead to important insights. If hosts from different
373 environments have similar microbiomes it could indicate that (i) there is not much
374 difference in the regional species pools between environments, (ii) the hosts move
375 between species pools, or (iii) there are other methods for microbiome acquisition such
376 as transmission from conspecifics. For example, Kueneman *et al*. [71] and McKenzie *et*
377 *al*. [32] studying the skin microbiome of tadpoles found that the identity of the species
378 was the strongest predictor of the skin microbiome, with the water body of origin
379 explaining additional [71] or no variance at all [32]. Vertical transmission of the
380 microbiome is unlikely in amphibians [72], and they probably acquire their microbiome
381 from the environment at every generation [73]. Therefore, the results from Kueneman
382 *et al*. [71] and McKenzie *et al*. [32] indicate there is at least a partially shared species
383 pool of microbes across their sampling sites.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

384 The enormous intra and inter-specific variation in the microbiomes of different
385 species of animals and plants can be at least partially explained by differences in
386 regional species pool during the community assembly of the microbiome. The gut
387 microbiome of animals can be stable over time [63, 74] and can also be resistant to
388 invasions and colonization by other bacteria [75–77]. Therefore, if the community
389 assembly of individuals of the same species occurred while they were exposed to
390 different species pools, historical effects can cause lasting effects in the development of
391 the microbiome [78]. Moreover, recent efforts to manipulate the microbiome of plants
392 and animals to improve health and/or increase productivity might not be as effective if
393 the manipulation is done only at the level of the individual host, i.e., at the local scale.
394 As stated by Chase [16] "if local and regional processes determine the community
395 composition, then both processes need to be restored to achieve the desired
396 community". Our results indicate that the same logic is valid for microbes.
397 In summary, we showed that local assembly of host-associated bacterial
398 communities are affected by regional scale processes, more specifically changes in the
399 regional species pool of colonizers. The inter-pond variation in bacterial community can
400 be due to stochastic, historical, spatial, and/or environmental processes or even the
401 result of local evolution [79]. Nevertheless, microbiome assembly, as any other
402 community, is a multiscale process. Microbial and microbiome ecology could be better
403 linked to ecological theory by considering the multiscale dynamics of community
404 assembly [12, 17, 80, 81]. Because it is much easier to quantify effects at the scale of the

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

405 individual host or host species, i.e., local effects [29, 62], the role of regional effects is
406 probably underestimated. Regional effects, however, appear to be a fundamental piece
407 of the community assembly puzzle that can be used to understand variation in
408 microbiomes.
409

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Acknowledgements

We thank Ananda Brito de Assis and Carlos Arturo Navas for laboratory access during field work in Brazil and Christopher Schalk and James Childress for help in the field in East Texas. We thank the Zoology Museum of University of São Paulo (MZUSP), The National Forests and Grasslands of Texas, and several private landowners for granting permission to access sampling sites and ICMBio (47263-1), Texas Parks and Wildlife (SPR-0414-067), and IACUC (AUP-2014-00286) for collecting and animal care permits. Funding was provided by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brasil, grant number 458796/2014-0) and Texas Ecolab. DTC received a Science Without Borders scholarship during the research (CAPES 1176/13-7). Finally, we thank Daniel Bolnick, Michel Ryan, Ulrich Mueller, and members of Leibold and Matz labs for comments throughout the project.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1. Costello EK, Stagaman K, Dethlefsen L, Bohannan BJM, Relman DA. The application of ecological theory toward an understanding of the human microbiome. *Science* 2012; **336**: 1255–62.
2. Christian N, Whitaker BK, Clay K. Microbiomes: unifying animal and plant systems through the lens of community ecology theory. *Front Microbiol* 2015; **6**: 1–15.
3. Broderick NA, Raffa KF, Goodman RM, Handelsman J. Census of the Bacterial Community of the Gypsy Moth Larval Midgut by Using Culturing and Culture-Independent Methods. *Appl Environ Microbiol* 2004; **70**: 293–300.
4. Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R, Gordon JI. The Effect of Diet on the Human Gut Microbiome: A Metagenomic Analysis in Humanized Gnotobiotic Mice. *Sci Transl Med* 2009; **1**: 6ra14-6ra14.
5. Bolnick DI, Snowberg LK, Hirsch PE, Lauber CL, Knight R, Caporaso JG, et al. Individuals' diet diversity influences gut microbial diversity in two freshwater fish (threespine stickleback and Eurasian perch). *Ecol Lett* 2014; **17**: 979–87.
6. Gomez A, Rothman JM, Petrzelkova K, Yeoman CJ, Vlckova K, Umaña JD, et al. Temporal variation selects for diet–microbe co-metabolic traits in the gut of Gorilla spp. *ISME J* 2016; **10**: 514–526.
7. Fierer N, Hamady M, Lauber CL, Knight R. The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proc Natl Acad Sci U S A* 2008; **105**: 17994–9.
8. Pflughoeft KJ, Versalovic J. Human Microbiome in Health and Disease. *Annu Rev Pathol Mech Dis* 2012; **7**: 99–122.
9. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. Diversity, stability and resilience of the human gut microbiota. *Nature* 2012; **489**: 220–30.
10. Goodrich JK, Waters JL, Poole AC, Sutter JL, Koren O, Blekhman R, et al. Human genetics shape the gut microbiome. *Cell* 2014; **159**: 789–799.
11. Bolnick DI, Snowberg LK, Hirsch PE, Lauber CL, Org E, Parks B, et al. Individual diet has sex-dependent effects on vertebrate gut microbiota. *Nat Commun* 2014; **5**: 4500.
12. Miller ET, Svanbäck R, Bohannan BJM. Microbiomes as Metacommunities: Understanding Host-Associated Microbes through Metacommunity Ecology. *Trends Ecol Evol* 2018; **33**: 926–935.
13. Ricklefs RE. Community Diversity: Relative Roles of Local and Regional Processes. *Science (80-)* 1987; **235**: 167–171.
14. Leibold MA, Holyoak M, Mouquet N, Amarasekare P, Chase JM, Hoopes MF, et al. The metacommunity concept: a framework for multi-scale community ecology. *Ecol Lett* 2004; **7**: 601–613.
15. Vellend M. Conceptual synthesis in community ecology. *Q Rev Biol* 2010; **85**: 183–206.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

16. Chase JM. Community assembly: when should history matter? *Oecologia* 2003; **136**: 489–98.
17. Prosser JI, Bohannan BJM, Curtis TP, Ellis RJ, Firestone MK, Freckleton RP, et al. The role of ecological theory in microbial ecology. *Nat Rev Microbiol* 2007; **5**: 384–392.
18. Fierer N, Ferrenberg S, Flores GE, González A, Kueneman J, Legg T, et al. From Animalcules to an Ecosystem: Application of Ecological Concepts to the Human Microbiome. *Annu Rev Ecol Evol Syst* 2012; **43**: 137–155.
19. Phillips CD, Phelan G, Dowd SE, McDonough MM, Ferguson AW, Delton Hanson J, et al. Microbiome analysis among bats describes influences of host phylogeny, life history, physiology and geography. *Mol Ecol* 2012; **21**: 2617–2627.
20. Fitzpatrick BM, Allison AL. Similarity and differentiation between bacteria associated with skin of salamanders (*Plethodon jordani*) and free-living assemblages. *FEMS Microbiol Ecol* 2014; **88**: 482–94.
21. Staubach F, Baines JF, Künzel S, Bik EM, Petrov DA. Host Species and Environmental Effects on Bacterial Communities Associated with Drosophila in the Laboratory and in the Natural Environment. *PLoS One* 2013; **8**: e70749.
22. Schloss PD, Iverson KD, Petrosino JF, Schloss SJ. The dynamics of a family's gut microbiota reveal variations on a theme. *Microbiome* 2014; **2**: 25.
23. McCord AI, Chapman CA, Weny G, Tumukunde A, Hyeroba D, Klotz K, et al. Fecal microbiomes of non-human primates in Western Uganda reveal species-specific communities largely resistant to habitat perturbation. *Am J Primatol* 2014; **76**: 347–354.
24. Eichmiller JJ, Hamilton MJ, Staley C, Sadowsky MJ, Sorensen PW. Environment shapes the fecal microbiome of invasive carp species. *Microbiome* 2016; **4**: 44.
25. Hernández-Gómez O, Hoverman JT, Williams RN. Cutaneous Microbial Community Variation across Populations of Eastern Hellbenders (*Cryptobranchus alleganiensis alleganiensis*). *Front Microbiol* 2017; **8**: 1–16.
26. Muletz Wolz CR, Yarwood SA, Campbell Grant EH, Fleischer RC, Lips KR. Effects of host species and environment on the skin microbiome of Plethodontid salamanders. *J Anim Ecol* 2018; **87**: 341–353.
27. Sanders JG, Powell S, Kronauer DJC, Vasconcelos HL, Frederickson ME, Pierce NE. Stability and phylogenetic correlation in gut microbiota: Lessons from ants and apes. *Mol Ecol* 2014; **23**: 1268–1283.
28. Moeller AH, Suzuki TA, Phifer-Rixey M, Nachman MW. Transmission modes of the mammalian gut microbiota. *Science (80-)* 2018; **362**: 453–457.
29. Lindström ES, Langenheder S. Local and regional factors influencing bacterial community assembly. *Environ Microbiol Rep* 2012; **4**: 1–9.
30. Pryor GS, Bjorndal KA. Symbiotic fermentation, digesta passage, and gastrointestinal morphology in bullfrog tadpoles (*Rana catesbeiana*). *Physiol Biochem Zool* 2005; **78**: 201–215.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

31. Vences M, Lyra ML, Kueneman JG, Bletz MC, Archer HM, Canitz J, et al. Gut bacterial communities across tadpole ecomorphs in two diverse tropical anuran faunas. *Sci Nat* 2016; **103**: 25.
32. McKenzie VJ, Bowers RM, Fierer N, Knight R, Lauber CL. Co-habiting amphibian species harbor unique skin bacterial communities in wild populations. *ISME J* 2012; **6**: 588–96.
33. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 2016; **13**: 581–583.
34. Microsoft and R Core Team. Microsoft R Open. 2018. Redmond, Washington.
35. Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, et al. The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res* 2014; **42**: D643–D648.
36. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet C, Al-Ghalith GA, et al. QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. *PeerJ Prepr* 2018; **6**: e27295v2.
37. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol Biol Evol* 2013; **30**: 772–780.
38. Price MN, Dehal PS, Arkin AP. FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS One* 2010; **5**.
39. Faith DP. Conservation evaluation and phylogenetic diversity. *Biol Conserv* 1992; **61**: 1–10.
40. Cardoso P, Rigal F, Carvalho JC. BAT - Biodiversity Assessment Tools, an R package for the measurement and estimation of alpha and beta taxon, phylogenetic and functional diversity. *Methods Ecol Evol* 2015; **6**: 232–236.
41. Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. Microbiome Datasets Are Compositional: And This Is Not Optional. *Front Microbiol* 2017; **8**: 2224.
42. Gloor GB, Reid G. Compositional analysis: a valid approach to analyze microbiome high-throughput sequencing data. *Can J Microbiol* 2016; **62**: 692–703.
43. Gloor GB, Wu JR, Pawlowsky-Glahn V, Egozcue JJ. It's all relative: analyzing microbiome data as compositions. *Ann Epidemiol* 2016; **26**: 322–329.
44. Silverman JD, Washburne AD, Mukherjee S, David LA. A phylogenetic transform enhances analysis of compositional microbiota data. *Elife* 2017; **6**: 1–20.
45. Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* 2005; **71**: 8228–35.
46. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, et al. vegan: Community Ecology Package. 2018.
47. Legendre P, Legendre L. Numerical ecology, 3rd ed. 2012. Elsevier.
48. Goslee SC, Urban DL. The ecodist Package for Dissimilarity-based Analysis of Ecological Data. *J Stat Softw* 2007; **22**.
49. Pavoine S, Vallet J, Dufour A-B, Gachet S, Daniel H. On the challenge of treating

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

various types of variables: application for improving the measurement of functional diversity. *Oikos* 2009; **118**: 391–402.

50. Dray S, Dufour A. The ade4 package: implementing the duality diagram for ecologists. *J Stat Softw* 2007; **22**: 1–20.

51. Matesanz S, Gimeno TE, de la Cruz M, Escudero A, Valladares F. Competition may explain the fine-scale spatial patterns and genetic structure of two co-occurring plant congeners. *J Ecol* 2011; **99**: 838–848.

52. Borcard D, Gillet F, Legendre P. Numerical Ecology with R. 2011. Springer New York, New York, NY.

53. Hagen M, Kissling WD, Rasmussen C, De Aguiar MAM, Brown LE, Carstensen DW, et al. Biodiversity, Species Interactions and Ecological Networks in a Fragmented World. *Advances in Ecological Research*, 1st ed. 2012. Elsevier Ltd., pp 89–210.

54. Emer C, Galetti M, Pizo MA, Guimarães PR, Moraes S, Piratelli A, et al. Seed-dispersal interactions in fragmented landscapes – a metanetwork approach. *Ecol Lett* 2018; **21**: 484–493.

55. Clauset A, Newman MEJ, Moore C. Finding community structure in very large networks. *Phys Rev E* 2004; **70**: 066111.

56. Wakita K, Tsurumi T. Finding Community Structure in Mega-scale Social Networks. *WWW*. 2007. pp 1275–1276.

57. Marquitti FMD, Guimarães PR, Pires MM, Bittencourt LF. MODULAR: Software for the autonomous computation of modularity in large network sets. *Ecography (Cop)* 2014; **37**: 221–224.

58. Erdős P, Rényi A. On random graphs I. *Publ Math* 1959; **6**: 290–297.

59. Bascompte J, Jordano P, Melian CJ, Olesen JM. The nested assembly of plant-animal mutualistic networks. *Proc Natl Acad Sci* 2003; **100**: 9383–9387.

60. Batagelj V, Mrvar A. Pajek - Analysis and Visualization of Large Network. In: Juenger M, Mutzel P (eds). *Graph Drawing Software*. 2003. Springer, Berlin, pp 77–103.

61. Emerson BC, Gillespie RG. Phylogenetic analysis of community assembly and structure over space and time. *Trends Ecol Evol* 2008; **23**: 619–630.

62. Cornell H V., Harrison SP. What Are Species Pools and When Are They Important? *Annu Rev Ecol Evol Syst* 2014; **45**: 45–67.

63. Coyte KZ, Schluter J, Foster KR. The ecology of the microbiome: Networks, competition, and stability. *Science (80-)* 2015; **350**: 663–666.

64. Kohl KD, Cary TL, Karasov WH, Dearing MD. Restructuring of the amphibian gut microbiota through metamorphosis. *Environ Microbiol Rep* 2013; **5**: 899–903.

65. Chai L, Dong Z, Chen A, Wang H. Changes in intestinal microbiota of *Bufo gargarizans* and its association with body weight during metamorphosis. *Arch Microbiol* 2018; **0**: 1–13.

66. Lyra ML, Bletz MC, Haddad CFB, Vences M. The Intestinal Microbiota of Tadpoles Differs from Those of Syntopic Aquatic Invertebrates. *Microb Ecol* 2018; **76**: 121–

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

124.

67. Sullam KE, Essinger SD, Lozupone CA, O'Connor MP, Rose GL, Knight R, et al. Environmental and ecological factors that shape the gut bacterial communities of fish: a meta-analysis. *Mol Ecol* 2012; **21**: 3363–3378.

68. Parris DJ, Brooker RM, Morgan MA, Dixson DL, Stewart FJ. Whole gut microbiome composition of damselfish and cardinalfish before and after reef settlement. *PeerJ* 2016; **4**: e2412.

69. Madigan MT, Martinko JM, Bender KS, Buckley DH, Stahl DA. Brock Biology of Microorganisms. *International Microbiology* . 2015. Pearson.

70. Pryor G. Anaerobic bacteria isolated from the gastrointestinal tracts of bullfrog tadpoles (*Rana catesbeiana*). *Herpetol Conserv Biol* 2008; **3**: 176–181.

71. Kueneman JG, Parfrey LW, Woodhams DC, Archer HM, Knight R, McKenzie VJ. The amphibian skin-associated microbiome across species, space and life history stages. *Mol Ecol* 2014; **23**: 1238–1250.

72. Hughey MC, Delia J, Belden LK. Diversity and stability of egg-bacterial assemblages: The role of paternal care in the glassfrog *Hyalinobatrachium colymbiphyllum*. *Biotropica* 2017; **49**: 792–802.

73. Walke JB, Becker MH, Loftus SC, House LL, Cormier G, Jensen R V, et al. Amphibian skin may select for rare environmental microbes. *ISME J* 2014; **8**: 2207–2217.

74. Faith JJ, Guruge JL, Charbonneau M, Subramanian S, Seedorf H, Goodman AL, et al. The Long-Term Stability of the Human Gut Microbiota. *Science (80-)* 2013; **341**: 1237439–1237439.

75. Britton RA, Young VB. Interaction between the intestinal microbiota and host in *Clostridium difficile* colonization resistance. *Trends Microbiol* 2012; **20**: 313–319.

76. Risely A, Waite D, Ujvari B, Klaassen M, Hoye B. Gut microbiota of a long-distance migrant demonstrates resistance against environmental microbe incursions. *Mol Ecol* 2017; **26**: 5842–5854.

77. Zmora N, Zilberman-Schapira G, Suez J, Mor U, Dori-Bachash M, Bashiardes S, et al. Personalized Gut Mucosal Colonization Resistance to Empiric Probiotics Is Associated with Unique Host and Microbiome Features. *Cell* 2018; **174**: 1388-1405.e21.

78. Fukami T. Historical Contingency in Community Assembly: Integrating Niches, Species Pools, and Priority Effects. *Annu Rev Ecol Evol Syst* 2015; **46**: 1–23.

79. Urban MC, De Meester L. Community monopolization: local adaptation enhances priority effects in an evolving metacommunity. *Proc Biol Sci* 2009; **276**: 4129–38.

80. Mihaljevic JR. Linking metacommunity theory and symbiont evolutionary ecology. *Trends Ecol Evol* 2012; **27**: 323–9.

81. Shade A, Dunn RR, Blowes SA, Keil P, Bohannan BJM, Herrmann M, et al. Macroecology to Unite All Life, Large and Small. *Trends Ecol Evol* 2018; **33**: 731–744.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

627

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

628 **Tables**

629 Table 1. Results of a Permutational Analysis of Variance (perMANOVA) testing if the
630 species of tadpole or the pond of origin are related to the gut microbiome composition
631 of tadpoles using the bacteria as Amplicon Sequence Variants (ASVs). Models were
632 fitted to CODA and PHILR transformed data. The significance of the marginal effects was
633 tested based on 999 permutations. BR = Brazil, CT = Central Texas, ET = East Texas.
634

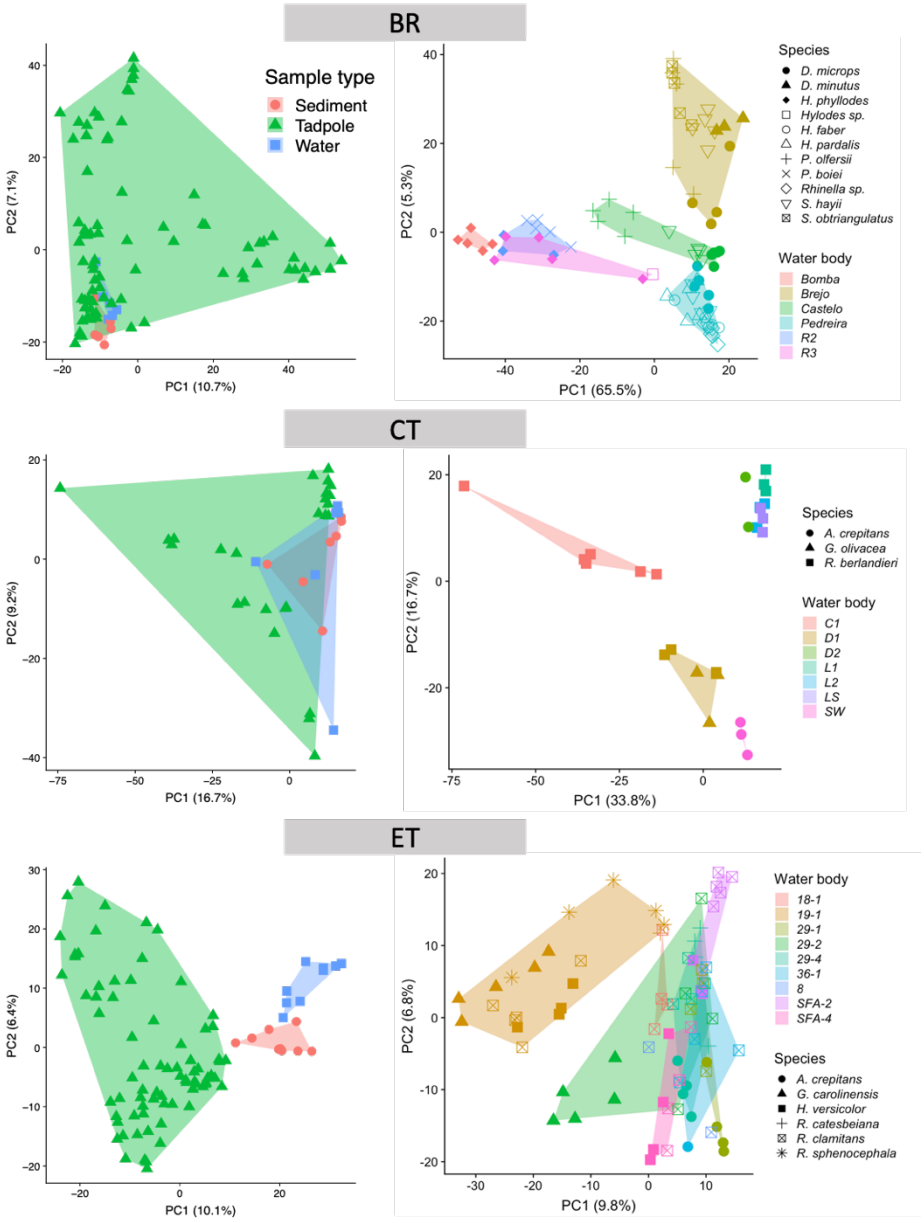
		PHILR			CODA		
		R^2	F	p	R^2	F	p
BR	Water body$_{(4,80)}$	0.08	3	0.005	0.1	2.7	0.001
	Species$_{(9,80)}$	0.15	2.5	0.002	0.16	2.1	0.001
CT	Water body$_{(5,25)}$	0.48	6	0.001	0.41	3.6	0.001
	Species$_{(1,25)}$	0.05	2.8	0.010	0.04	1.8	0.056
ET	Water body$_{(8,76)}$	0.27	4.7	0.001	0.24	3.2	0.001
	Species$_{(5,76)}$	0.18	5.1	0.001	0.12	2.6	0.001

635

636

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

637 **Figures**



638

639 Figure 1. Plot of the first two axes of a Principal Component Analysis on CODA
640 transformed microbiome data (amplicon sequence variants) from tadpole samples from
641 Brazil (BR), Central Texas (CT), and Eastern Texas (ET). Panels on the left show sediment,
642 water, and tadpole samples. Panels on the right show samples from tadpoles only. Left
643 and right panels are results from separate analyses. Shapes represent different species
644 of tadpoles and colors represent water body of origin. See Table S1 for abbreviations.
645

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

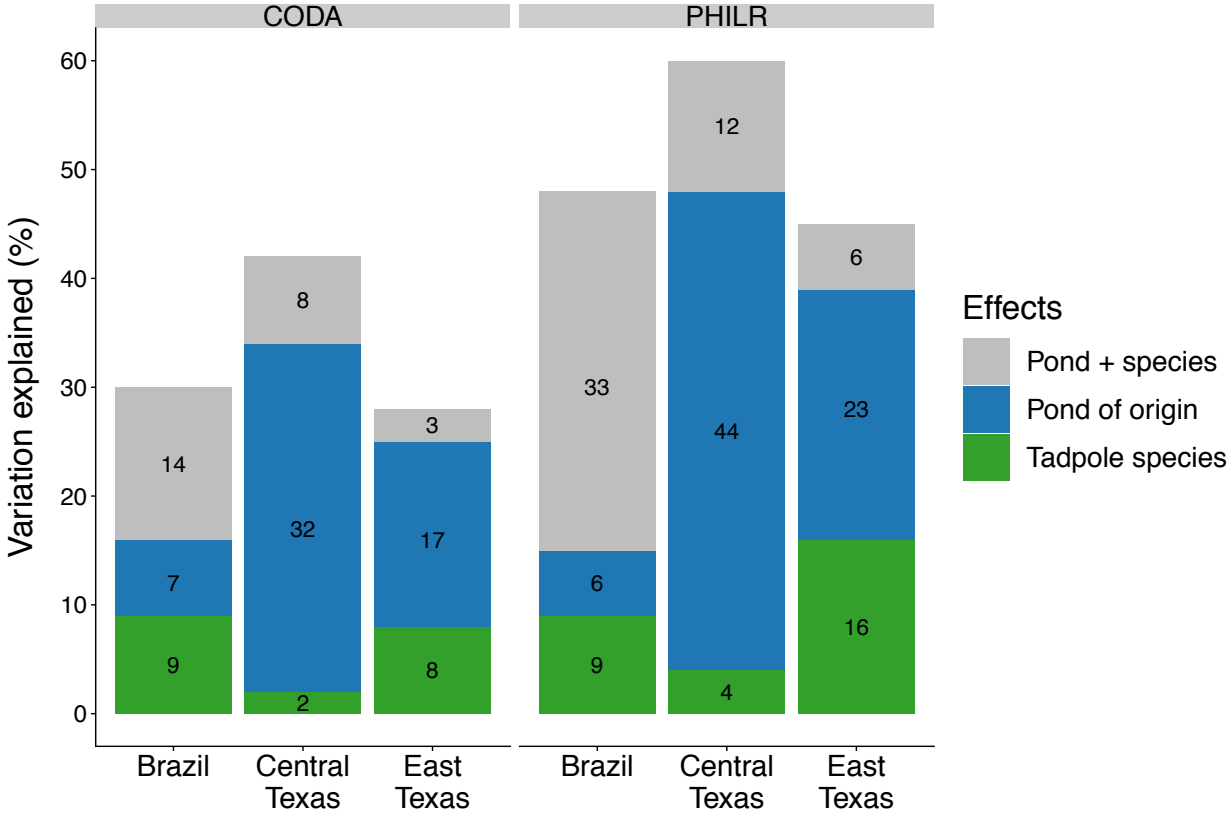


Figure 2. Results of a variation partitioning analysis testing the unique effects of the species of tadpole or pond of origin as well their shared effects on the tadpole gut microbiome considering the bacteria as Amplicon Sequence Variants (ASVs) for CODA- and PHILR-transformed data. The effects of pond of origin and species of tadpole were all significant ($p < 0.05$) based on 1000 permutations, except for the shared component, which cannot be tested.

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

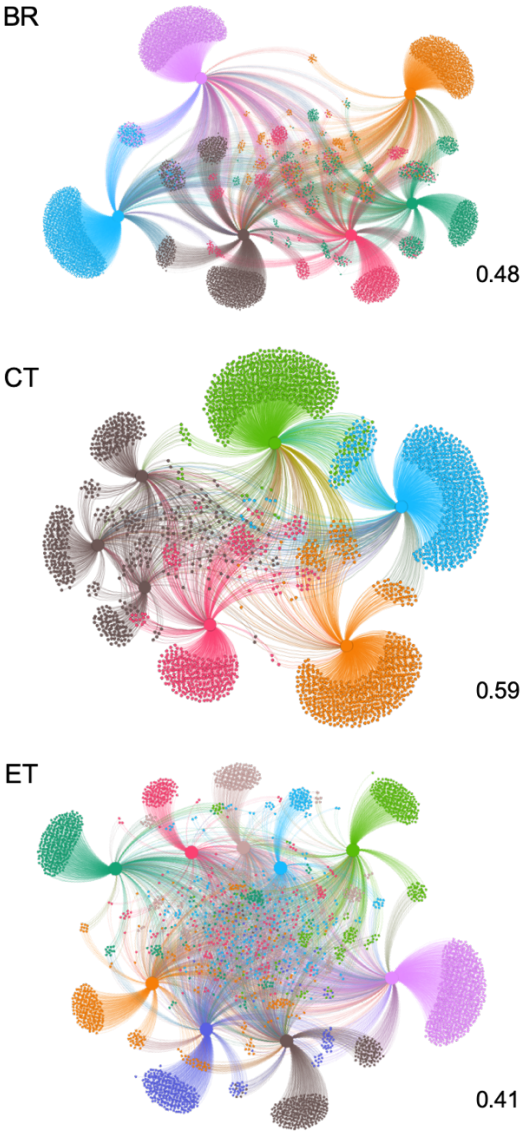


Figure 3. Microbiome metanetworks of the shared bacteria (amplicon sequence variant - ASV) found in the gut of tadpoles. Larger nodes represent each water body for each locality, *i.e.*, Brazil (BR), Central Texas (CT), and East Texas (ET). Smaller nodes represent a unique bacterium that can be found in a single water body (link between smaller and larger nodes) or shared among ponds thus connecting them. Numbers on the bottom-right are the modularity values for each metanetwork. Distinct colors represent different modules which, with a few exceptions, correspond to a single water body (Table S10).

bioRxiv preprint doi: https://doi.org/10.1101/2020.07.05.188698; this version posted July 6, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

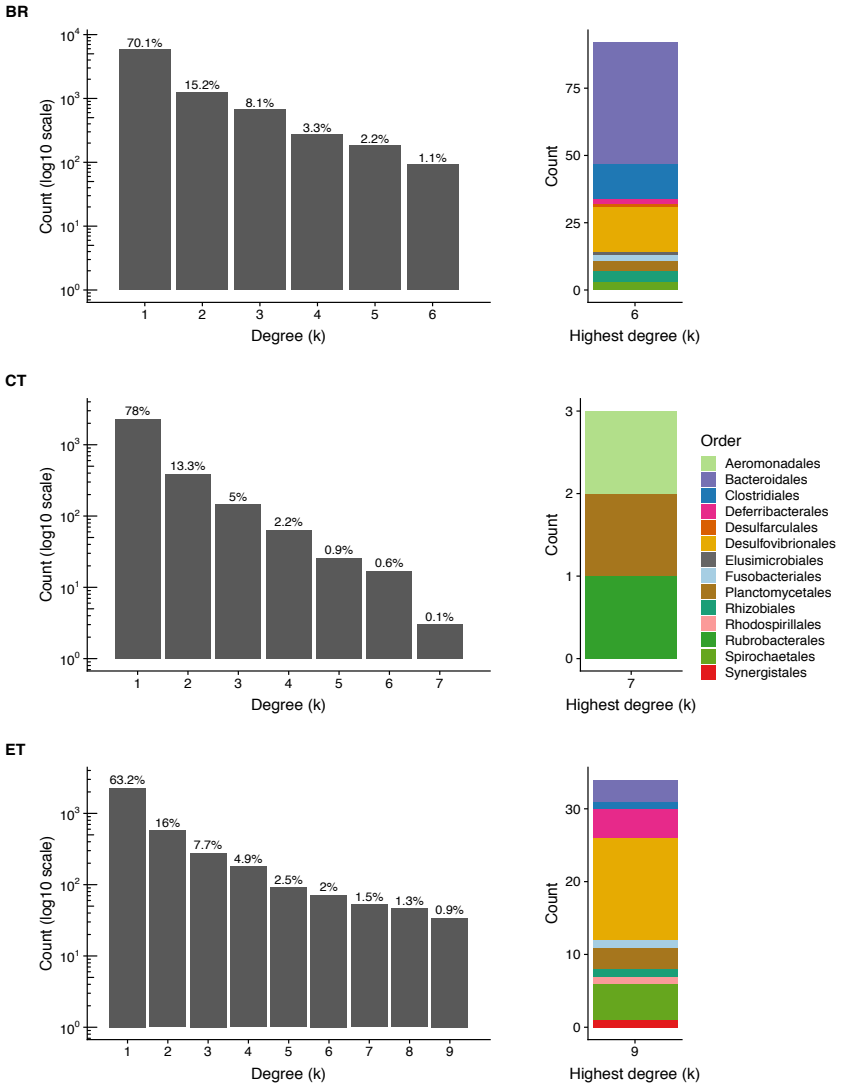


Figure 4. Number of bacteria from the gut of tadpoles for each degree (k) on the left and Order of bacteria present in the highest degree on the right. Each degree represents one pond. Bacteria found in only one pond are represented in the degree category 1. Likewise, bacteria present in tadpoles across all ponds are represented in the highest category of each histogram. The Order of bacteria present in tadpoles across all ponds from each locality is shown on the barplot on the right. Same color across barplots represent the same Order. See Figures S11-S16 for water and sediment metanetworks. BR = Brazil, CT = Central Texas, ET = East Texas.