Environmental heterogeneity ruling the action of ecological processes over the bacterial metacommunity assembly

Running title: High-level processes in bacterial metacommunity

Paula Huber 1, Sebastian Metz 2, Fernando Unrein 2, Gisela Mayora 1, Hugo Sarmento 3, Melina Devercelli 1

1 Instituto Nacional de Limnología (INALI), Consejo Nacional de Investigaciones Científicas y Técnicas, Universidad Nacional del Litoral, Ciudad Universitaria, Paraje El Pozo, C.P. 3000 Santa Fe, Argentina.

2 Instituto Tecnológico de Chascomús (INTECH), UNSAM-CONICET, Intendente Marino Km 8.2, CP 7130, Chascomús, Buenos Aires, Argentina.

3 Departamento de Hydrobiologia, Universidade Federal de São Carlos (UFSCar), Rodovia Washington Luiz, 13565-905 São Carlos, São Paulo, Brazil.

Keywords: Environmental heterogeneity; high-level processes; bacterial metacommunity; network associations; Paraná floodplain system.

Corresponding author:

Huber, Paula
Instituto Nacional de Limnología (CONICET-UNL)
Ciudad Universitaria - Paraje El Pozo SN.
Santa Fe, CP. 3000.
Argentina.
e-mail: mariapaulahuber@gmail.com
Tel: +54 342 4511645/48

29 ABSTRACT

30 Determining which processes take place in the spatial distribution of bacterioplankton
 31 metacommunities has been a central goal of microbial ecology. In freshwater ecosystems, selection
 32 has been spotted as the main driver shaping bacterial community. However, its relative importance
 33 compared with others processes (dispersal, drift, diversification) may depend on spatial
 34 heterogeneity and dispersal rates within a metacommunity. Still, a decrease in the role of selection
 35 is expected with increasing dispersal homogenization. Here, we investigate the main ecological
 36 processes modulating the bacterial assembly in contrasting scenarios of environmental
 37 heterogeneity. We carried out a spatiotemporal survey in the floodplain system of Paraná River. The
 38 bacterioplankton metacommunity was studied by a step-forward analysis combining different
 39 statistical inferences based on the phylogenetic and taxa turnover as well as co-occurrence
 40 networks. We found that selection was the main process even at both extremes of environmental
 41 heterogeneity and homogeneity, challenging the general view that the strength of selection is
 42 weakened due to dispersal homogenization. The ecological processes acting on the community also
 43 determined the complexity and randomness of bacterial networks. The heterogeneous selection
 44 promoted greater network complexity increasing the β -diversity, and bacterial associations were
 45 more random with the increasing influence of stochasticity. Finally, the spatiotemporal
 46 heterogeneity was an important factor determining the number and identity of keystone taxa in the
 47 system. Integrating all these empirical evidences we propose a new conceptual model that
 48 elucidates how the environmental heterogeneity guides the action of ecological processes shaping
 49 the bacterial metacommunity.

50

51 INTRODUCTION

52 Determining the processes behind community's assembly across Earth's ecosystems is a major
 53 research topic in ecology [1–4]. Two perspectives address this issue. On one hand, the traditional
 54 niche-based theory states that deterministic processes such as environmental filters and species
 55 interactions, govern community structure and determine species composition [5, 6]. On the other
 56 hand, the neutral theory considers that all species are ecologically equivalent and, therefore, the
 57 structure of communities is a consequence of stochastic processes of birth, death, colonization,
 58 immigration, speciation, and probabilistic dispersal [1, 7, 8]. Although these perspectives are often
 59 considered contradictory, they are not mutually exclusive and a central debate is on their relative
 60 importance in shaping natural communities.

61 The metacommunity concept provides an adequate framework to identify these processes
 62 considering multiple spatial scales [9, 10]. Under this approach, communities are spatially
 63 connected in a network by dispersal, and the processes involved in their assembly are the result of
 64 an interplay between local factors and regional dynamics.

65 More recently, Vellend [3, 11] proposed a conceptual framework in which the structure of
 66 communities could be explained by four high-level processes: selection, dispersal, ecological drift,
 67 and diversification. Briefly, selection involves deterministic factors that modify the community
 68 structure due to fitness differences among individuals. It can act under homogeneous or
 69 heterogeneous environmental conditions, leading to more similar or dissimilar structure among
 70 communities (homogeneous or heterogeneous selection) [12, 13]. Dispersal results from both
 71 deterministic and stochastic factors that either favor or limit the movement (active or passive) and
 72 establishment of organisms among communities. High dispersal rates have a homogenizing effect
 73 (homogeneous dispersal) leading to a low turnover within communities, whereas low dispersal rates
 74 (dispersal limitation), coupled with drift or weak selection, increase community turnover [12, 14].
 75 Drift influences changes in communities due to demographic events of birth, death and reproduction
 76 that occur at random, independently of species rates [15]. Finally, diversification refers to the

generation of new species by genetic variation mainly due to stochastic factors, and a larger spatial and temporal scale with respect to the other processes are needed to observe its effects [16].

Bacteria are the most conspicuous and diverse components of the ecosystems' microbiome, and the study of the processes shaping its structure contributes to gain proper comprehension of the functional organization of ecosystems. The ecological rules governing their metacommunity assembly change from land to water [17–21]. Differences among ecosystems are fundamentally due to environmental heterogeneity, connectivity, and spatial extent. Within aquatic ecosystems, the complexity of freshwater bodies is comparatively high comparing to oceans [22], consequently, striking idiosyncratic patterns in bacterial diversity were observed [23–27].

In freshwater ecosystems, deterministic processes (mainly selection) have been spotted as the main drivers shaping bacterial community structure [18, 26, 28–30]. However, it is not a general rule and the relative importance of selection compared with other processes may depend on the environmental heterogeneity and dispersion rates within a metacommunity [23, 27, 28, 31, 32]. For example, in systems with low environmental heterogeneity the stochastic processes seem to be more relevant, because environmental filters are not strong enough to act as selective forces allowing species sorting [31, 33]. In addition, bacterial communities in isolated systems should be mainly structured by dispersal limitation, whereas as environmental connectivity increases, homogeneous dispersal should become more important [29, 30, 33, 34].

Most previous results have been based on a metacommunity snapshot. However, in complex network systems subject to hydrological influence, the relative importance of the structuring ecological processes could change over time according to the variability in connectivity [29, 35–40]. Moreover, very few studies actually applied Vellend's approach to its full extent [12, 30, 41–43]. Another gap lies in understanding how the ecological processes influence the interactions between taxa (e.g. co-occurrence networks). Changes in community composition should affect the microbial co-occurrence networks [21, 44–49], thus we might expect more random co-occurrence

and less complex networks when stochastic processes predominate. However, the empirical evidence of this dependency remains scarce.

The Paraná River floodplain constitutes a network of environments that resembles metacommunity organization [37], being an ideal system to address these issues. It is characterized by a wide range of temporal and spatial heterogeneity mediated by irregular hydrological fluctuations [50]. It comprises multiple shallow lakes, some of which are permanently connected to the main river or secondary channels, whereas others remain isolated still high water phases connects most the environments [51]. The floods have a homogenization effect on the environmental features [52, 53], as observed for other large rivers [54]. Here, we investigated the bacterial community structure through 16S amplicon sequencing in this complex floodplain system in order to (i) determine the relative importance of the High-level processes structuring the bacterial metacommunity, (ii) detect changes in bacterial co-occurrence networks, and (iii) identify keystone taxa in contrasting environmental heterogeneity scenarios mediated by landscape connectivity.

We hypothesized that the influence of different ecological processes shaping bacterioplankton metacommunities depend mainly on system's environmental heterogeneity determined by the degree of landscape connectivity. To test this hypothesis, we first analyzed the processes shaping the community structure of bacterioplankton using the approach based on phylogenetic and taxa turnover proposed by Stegen et al. [12]. Then, we constructed co-occurrence networks to explore how ecological processes affect putative interactions between taxa. We predicted that, as the degree of connectivity increases and the system becomes more homogeneous, non-selective processes would have major relative importance in structuring the bacterioplankton metacommunity leading to more random associations between bacteria taxa.

Material and Methods

Study site and sampling design

127 The Paraná River is the second largest river in South America and fifth in the world [51] (Fig. 1).
 128 The middle stretch is formed by a main channel and a large floodplain that encompasses a high
 129 number of temporary and permanent streams and lakes. The system is characterized by a complex
 130 spatiotemporal dynamic influenced by hydrosedimentological regime. It has a fluctuating
 131 hydrological dynamic with pulses of flood and drought that determine high and low water phases,
 132 and alternate in variable frequency with extreme hydrological phases [55]. During high water
 133 phases, water drains to the floodplain, connecting the environments in different degrees according
 134 to the magnitude of the flood. The sedimentary pulse that mainly depends on the suspended solids
 135 coming from the Bermejo Basin, is not coupled with the hydrological pulses, affecting abiotic
 136 factors of the environments connected with the main channel [56]. The dynamics of biological
 137 communities respond to a great extent to the hydrological fluctuations, since it has a direct influence
 138 on the environmental characteristics, dispersal processes, and habitats colonization [36–39, 57–59].

139 The temporal dynamic and the spatial heterogeneity of the system were captured sampling
 140 different hydrological phases and environment types characteristics of the floodplain River (Fig. 1).
 141 Four samplings that lasted 10 days each were performed during: a low-water phase (LW,
 142 November–December 2013), a low-water phase coupled with the sedimentological pulse (LWs,
 143 March–April 2014), a high water phase (HW, September 2105), and an extraordinary high water
 144 phase (eHW, March 2016). Four representative environment types with different hydrological
 145 connectivity and morphological features were selected (Fig. 1): the main channel and large
 146 secondary channels (MC), minor secondary channels (SC), connected lakes (CL), and isolated lakes
 147 and swamps (IL). Lentic environments were classified as IL or CL according to its hydrological
 148 connectivity degree during low water phase. Overall our dataset was constituted by 59 samples: 16
 149 from LW, 17 from LWs, 13 from HW and 13 from eHW.

150

151 *Environmental data collection*

152 Depth, water temperature, pH, conductivity, dissolved oxygen (DO) concentration (HANNA
153 checkers), and Secchi disk depth were measured in situ. Subsurface water samples were collected
154 and transported in polypropylene containers to the laboratory for turbidity, soluble reactive
155 phosphorus (SRP), nitrate (NO_3^-), ammonium (NH_4^+), chromophoric dissolved organic matter
156 (CDOM) concentration (A_{440}), CDOM molecular weight ($S_{275-295}$), and chlorophyll-*a* analyses.
157 Technical details of the abiotic variables' analyses are described in the Supplementary Methods.

158

159 *Bacterial samples collection and sequencing*

160 Subsurface water samples for DNA (100 to 140 ml) were prefiltered with a 50 μm pore mesh, and
161 filtered through a 0.22 μm pore-size polycarbonate filters (Millipore). The filters were frozen (-80
162 $^\circ\text{C}$) until DNA extraction. Genomic DNA was extracted using a CTAB protocol [60] as described in
163 the Supplementary Methods.

164 Tagged amplicons of the 16S rRNA gene (V3–V4 region) were obtained with the primers
165 341F and 805R [61] and sequenced using Illumina MiSeq 2x250 paired-end reads approach [62] by
166 MacroGen (Seoul, South Korea).

167

168 *Sequence data analysis*

169 Raw sequences were processed using a modified version of the pipeline proposed by Logares [63]
170 (<https://github.com/ramalok>), described in the Supplementary Methods. Operational taxonomic
171 units (OTUs) were defined with no clustering (zero-radius OTUs [zOTUs]) using UNOISE2
172 algorithm [64]. The zOTU table resulted in 10 253 zOTUs (2 560 053 reads). We constructed zOTU
173 rarefaction curves to evaluate richness saturation (Fig. S2).

174 A second zOTU table was generated discarding zOTUs with <10 reads as well those
175 assigned to Archaea domain or chloroplast. This table was normalized to an equal sampling depth to
176 generate the subsampled zOTU table (rrarefy function, Vegan v.2.0.9 package [65] in the R

177 environment [66]) which consisted of 8 927 OTUs and 18 747 reads per sample.

178 The sequence data obtained in this work were deposited at the European Nucleotide Archive
179 (ENA) public database with the references ERSXXXXXXX to ERSXXXXXXX.

180

181 *Environmental heterogeneity*

182 The environmental heterogeneity was evaluated by computing the average dissimilarity between
183 sites ($\overline{Ed}$) [67] based on nine abiotic variables: DO, turbidity, conductivity, pH, SRP, NO_3^- , NH_4^+ ,
184 CDOM concentration, and CDOM molecular weight. For each hydrological phase we computed a
185 Euclidean distance matrix (Vegan package, R) and calculated the dissimilarity between sites (Ed) as
186 follow:

$$Ed = \left(\frac{Euc}{Euc_{max}} \right) + 0.001$$

187 where Euc is the Euclidean distance between two sites and Euc_{max} corresponds to the maximum
188 Euclidean distance considering all the pairwise distances in the overall dataset. 0.001 was added to
189 account for zero similarity between sites [67]. Then, we calculated the mean Ed ($\overline{Ed}$) of each
190 similarity matrix computed and used it as an index of environmental heterogeneity in each
191 hydrological phase. Additionally, for each hydrological phase we calculated the coefficient of
192 variation (CV%), as the quotient of the standard deviation and mean of each variable.

193 To test if the environmental conditions in the four hydrological phases differed significantly,
194 we performed a permutational multivariate analysis of variance (PERMANOVA) [68], using
195 Euclidean similarity index, and Bonferroni pairwise post-hoc tests (Past software V3 [69]).

196

197 *Bacterial community structure and diversity*

198 To a general view of bacterial community, we analyzed the taxonomic structure in each
199 hydrological phase and constructed rank abundance curves (Fig. S3). We characterized the “rare

200 biosphere” and the “abundant fraction” (zOTUs relative abundances per sample <1% and >1%,
201 respectively) [70].

202 The effect of hydrological conditions on bacterial community structure was assessed by
203 calculating the community turnover in each hydrological phase. We computed a Bray-Curtis
204 distance matrix on the basis of zOTU normalized abundance for each hydrological phase and
205 calculated the dissimilarity (Xd) between the local communities as follow:

$$Xd = \left(\frac{Bray}{Bray_{max}} \right) + 0.001$$

206 where $Bray$ is the Bray-Curtis distance between two communities and $Bray_{max}$ corresponds to the
207 maximum Bray-Curtis distance considering the overall dataset. Then, we calculated the mean Xd
208 ($\overline{Xd}$) of each similarity matrix computed, which was used as a value of community turnover.

209 We also evaluated the significance of the effect of hydrological conditions on bacterial
210 turnover performing a PERMANOVA [68] with 9 999 permutations on Bray-Curtis dissimilarity
211 distance, and Bonferroni pair-wise comparisons were run to obtain P -values (Past software V3).

212 To visualize the taxonomic similarity across local communities, nonmetric multidimensional
213 scaling (NMDS) was performed using the Bray-Curtis distance metric (Vegan package, R).

214 zOTU richness, Shannon-Weaver diversity (H') and Whittaker β -diversity indices were
215 calculated from the normalized bacteria zOTU table (Past software V3). To test whether both
216 attributes differed significantly ($P < 0.01$) among hydrological phases we performed a Kruskal-
217 Wallis test with Mann-Whitney U test.

218

219 *Quantification of high-level process structuring the bacterial metacommunity*

220 To quantify the relative importance of ecological processes in structuring the bacterial
221 metacommunity we used the approach proposed by Stegen et al. [12]. This method analyzes the
222 influence of environmental filtering based on an intrinsic property of the community, independently

of the considered environmental variables [12, 71], sorting the problem of overestimating the effect of stochastic processes due to unmeasured environmental variables.

We first measured the influence of selection in each hydrological phase, comparing observed phylogenetic turnover to a random expectation using the β NTI Index (β -Nearest Taxon [72–74]). This metric is defined as the difference between the observed mean phylogenetic distance between each taxon and its closest relative in two communities (β MNTD metric) with the β MNTD obtained from the null distribution, divided by the standard deviation of phylogenetic distances in the null data. Absolute β NTI value greater than 2 ($|\beta$ NTI >2) indicates coexisting taxa are more closely related than expected by chance, so selection strongly influences community composition [12, 75]. Next, we estimated the percentage of homogeneous selection as the fraction of pairwise comparisons with a β NTI value of <-2 and heterogeneous selection as the fraction of pairwise comparisons with a β NTI value of $>+2$ [12].

As this approach considered that the habitat preferences of closely related taxa are more similar than the habitat preferences of distantly related taxa [12], we tested the phylogenetic signal performing a Mantel correlogram analysis between zOTU niche and zOTU phylogenetic distances [12, 23, 41, 76–78] (see Supplementary Methods). Phylogenetic signals were detected over short phylogenetic distances (Supplementary Fig. S1) consistent with previous work [12, 23, 41, 78, 79].

The communities that were not structured by selection (i.e. $|\beta$ NTI <2), were then analyzed in the second step, where the action of dispersal and drift was calculated based on the taxonomic (zOTU) turnover with the Raup-Crick metric [80] using Bray-Curtis dissimilarities (hereafter RC_{bray}) [12]. RC_{bray} compares the measured β -diversity from a metacommunity against the β -diversity obtained from a null model. RC_{bray} values between -0.95 and $+0.95$ indicate significant departures from the degree of turnover, occurring when drift is acting alone [12]. Additionally, RC_{bray} values $>+0.95$ indicate that communities are less similar than expected by chance as a result of dispersal limitation combined with drift, and RC_{bray} values <-0.95 indicate that communities are more similar than expected by chance as a result of homogeneous dispersion [12].

249

250 *Co-occurrence networks analysis*

251 To infer whether bacterial associations were influenced by the different ecological processes
 252 structuring the metacommunity, we performed a network analysis using the SparCC algorithm [81]
 253 implemented in FastSpar [82]. One specific network association for each hydrological phase was
 254 constructed considering zOTUs with >50 reads and present in at least 2 samples. FastSpar was run
 255 using 1 000 interactions and 1 000 bootstraps to calculate the *P*-values. Only the significant
 256 ($P < 0.01$) and strong positive and negative correlations ($|r| > 0.7$) were considered. The networks
 257 were visualized in Cytoscape V3.7.1 [83]. To validate the non-random co-occurrence patterns, we
 258 evaluated networks against their randomized versions using the Barabasi–Albert model available in
 259 Randomnet-works plugin in Cytoscape. NetworkAnalyzer tool [84] was used to calculate five
 260 network topology properties: number of nodes (bacterial zOTUs), number of edges (associations
 261 between bacteria), average number of neighbors, distribution of degree and clustering coefficient.
 262 These structural attributes were used to infer the complexity and chance in bacterial associations
 263 [46, 85–87]. The average number of neighbors is the average connectivity of a node in the network,
 264 thus higher values are expected according increase the complexity of the associations [88]. The
 265 distribution of degree indicates the number of nodes that have each degree, so the complexity of the
 266 associations can be inferred by analyzing the power-law distribution [46, 88]. Finally, the clustering
 267 coefficient indicates how nodes are embedded in their neighborhood and, thus, the degree to which
 268 they tend to cluster together [88, 89]. As consequence, low values of this metric are interpreted as
 269 more chance in the association patterns.

270

271 *Keystone bacterial taxa*

272 We studied those taxa that play central roles on community functioning, commonly referred as
 273 keystone species [90, 91]. For each network we identified the keystone taxa as those nodes with a

high degree (>30), closeness centrality (>0.36) and low betweenness centrality (<0.07) [46, 90, 92] (NetworkAnalyzer tool in Cytoscape). These metrics illustrate both the number of connections and how important those connections are to the overall network [90, 93, 94].

Redundancy Analysis (RDA) was performed to evaluate the effect of abiotic variables on the abundance of the keystone taxa (CANOCO software V5 [95]). Environmental variables tested as explanatory variables (DO, turbidity, conductivity, pH, SRP, NO_3^- , NH_4^+ , CDOM concentration CDOM molecular weight), were normalized as standard normal deviates. A forward selection procedure was run to find the subset of the significant explanatory variables using the Monte Carlo Permutation.

RESULTS

Environmental heterogeneity

The average environmental dissimilarity between sites ($\overline{\text{Ed}}$) as well coefficient of variation (CV) of the nine abiotic variables indicated a clear trend to spatial homogenization from low water (low hydrological connectivity) to high water (high hydrological connectivity) phases (Fig. 2). Additionally, the four hydrological phases were significantly different according to their environmental characteristics (PERMANOVA $P < 0.05$, Table S1 and S2).

Bacterial community structure and diversity

Overall in the 59 samples, a total of 8 927 zOTUs within the bacterial domain were defined. The metacommunity was mainly dominated by Proteobacteria (33%), Actinobacteria (23%), Bacteroidetes (18%), and Verrucomicrobia (11%). The community structure according to zOTUs abundance varied with hydrological conditions with significant differences among phases (PERMANOVA, $P < 0.001$, Fig. S4, Table S3) except between low water phases (LW and LWs). In the NMDS, samples were organized following a clear homogenization pattern from low water to

299 high water phases (Fig. 3). In the eHW phase, local communities were more similar displaying a
 300 distinct cluster. In contrast, those from low water phases were more spread, mainly due to
 301 differences in the community composition of isolated lakes (Fig. 3). The community turnover ($\overline{X_d}$)
 302 varied markedly among hydrological phases, being significantly lower (PERMANOVA, $P < 0.005$)
 303 in high water phases ($\overline{X_d}$: HW=0.58; eHW=0.29) than in low water phases ($\overline{X_d}$: LW=0.76;
 304 LWs=0.75). In line with this, the highest values of β -diversity were obtained in both low water
 305 phases (Whittaker: LW=3.50; LWs=3.27; HW=1.97; eHW=1.95).

306 Species richness and Shannon diversity (H') presented similar average values at the different
 307 hydrological phases (ANOVA, $P = 0.366$, Fig. 3). However, differences among local communities
 308 were pronounced at low water phases, when isolated lakes were more spread showing lower values
 309 than the mean (Fig. 3).

310

311 *High-level process structuring the bacterial metacommunity*

312 Having demonstrated that bacteria community was clearly linked to changes in the environmental
 313 heterogeneity, we quantified the relative importance of selection, dispersal, and drift processes in
 314 structuring the metacommunity.

315 The phylogenetic turnover analysis (βNTI) revealed that selection was the most important
 316 structuring process regardless hydrological periods. However, its relative importance as well as the
 317 type of selection, changed according to hydrological conditions and hence, to the environmental
 318 heterogeneity. During both low water phases, heterogeneous selection had the major role in
 319 structuring the metacommunity (76.66% LW and 84.55% LWs of the overall community turnover)
 320 being poorly significant the role of homogeneous selection and non-selection process. In HW phase,
 321 heterogeneous selection was also the most important process (39.74%), but the relative importance
 322 of homogenous selection and non-selection processes were twice as high compared to the low water
 323 phases (Fig. 4). In contrast, homogenous selection was the most important process in eHW

(88.46%) (Fig. 4). Regarding non-selection processes (RC_{bray}), dispersal limitation combined with drift had the prevalent role in all the hydrological phases, except in eHW period when homogeneous dispersion became more important (Fig. 4).

Bacterial community associations

The associations between bacterial zOTUs remarkably differed regarding the environmental heterogeneity (Fig. 5). The networks varied in both, number of nodes (bacterial zOTUs) and edges (associations between bacteria), being the highest under the dominance of heterogeneous selection (low water phases) (Fig. 5). The higher network complexity was observed during the dominance of heterogeneous selection, as the average number of neighbors was considerably higher in both low water phases with respect to high water phases, especially HW (Fig. 5). In agreement, the histograms of degree distribution showed a stronger power-law in low than in high water phases (Fig. 6), indicating the presence of nodes highly connected when the heterogeneous selection had the major role in structuring the metacommunity.

Finally, the values of clustering coefficient revealed more random associations during the HW phase (Fig. 5) when heterogeneous and homogeneous selection had similar importance in structuring the metacommunity, and the non-selection processes had higher contribution compared to the other phases.

Metacommunity keystone taxa

The number as well as the taxonomic composition of keystone taxa were notably different in the four hydrological phases. In LW, 50 zOTUs were defined as keystones that were reduced to 11 in LWs and HW, and 13 in eHW. Most of these keystone taxa were exclusive of each hydrological phase and no one was present at the four phases. Keystone taxa accounted differential contributions

to each phylum in each hydrological phase, being Actinobacteria and Bacteroidetes the better represented in all phases (Fig. S5).

In the RDA, keystone taxa were arranged according to the hydrological phase in which they were defined (Fig. 7). The first two axes accounted for 43.04% of the variance (axis 1: 33.09%; axis 2: 9.95%) being CDOM concentration (A_{440}), CDOM molecular weight ($S_{275-295}$), DO, and nutrients (SRP, NH_4^+ and NO_3^-) the main explanatory variables ($P < 0.01$). The first axis was mainly defined by the NO_3^- concentration and CDOM molecular weight (intra-set correlation coefficients: -0.75 and 0.65, respectively). The second axis was principally associated with CDOM concentration and molecular weight (intra-set correlation coefficients: 0.70 and -0.70, respectively), separating low from high water samples. The keystone taxa from both low water phases were positioned towards the upper part of the graph (Fig. 7). Taxa from the LWs grouped together by the left of the graph, while those from LW showed a clear segregation in 2 groups: taxa with higher abundances in lotic systems were positively related to NO_3^- and NH_4^+ by the left panel; whereas taxa with higher abundance in lentic environments were inversely related with CDOM molecular weight by the right panel. The keystone taxa from high water phases were mainly positioned towards the lower part of the graph (Fig. 7): taxa from eHW located in the left panel were mainly associated with high CDOM molecular weight, meanwhile those from the HW positioned throughout the right panel were more related with high CDOM concentration (A_{440}).

Discussion

This study provides the first empirical evidence that selection is the most important process structuring the bacterial metacommunity at both extremes of environmental heterogeneity and homogeneity. This finding challenges the general view that the strength of the selection is weakened due to dispersal homogenization. Additionally, at intermediary environmental heterogeneity, stochastic processes became more important leading to more random metacommunity network.

374 Based on the underpinning of Vellend's approach [3], we predicted that the influence of
375 non-selective processes would increase as the system becomes more homogeneous. Contrarily, we
376 found that selection was the main process shaping the bacterial metacommunity not only at
377 environmental heterogeneity but also at environmental homogeneity.

378 Several studies found evidence of the pivotal role of selection determining bacterial
379 communities' structure in different environments, and a decrease in its relative importance as
380 dispersal rates increase and the systems become more homogeneous in their environmental
381 conditions [29, 30, 78, 96]. Particularly, Wang et al. [23] in a comparative survey of a broad range
382 of ecosystems (i.e. soil, stream biofilm, and lake), observed a clearly dominant role of deterministic
383 processes on bacterial assemblages, whereas stochastic processes became more relevant as
384 environmental heterogeneity decreased and dispersal rates increased. The authors proposed a
385 theoretical threshold of selective strength above which communities should be mainly structured by
386 deterministic processes as a consequence of high environmental heterogeneity and low dispersal
387 rates.

388 However, these previous attempts only considering spatial scales. To address the full scope
389 of the question, we extended these findings considering the spatial and temporal variability in a
390 complex ecosystem. We summarized our findings on three possible scenarios: 1) With high
391 environmental heterogeneity characterized by low connectivity, the structure of bacterial
392 assemblage was mainly driven by heterogeneous selection (Fig. 8a). In this scenario, species with
393 different fitness may be strongly filtered by different selective forces in each local community
394 leading to a high β -diversity, whereas the influence of drift and dispersion should be irrelevant; 2)
395 At intermediate environmental heterogeneity, both heterogeneous and homogeneous selections had
396 similar weight and the relative influence of stochastic processes increased (Fig. 8a). Here, the
397 magnitude of ordinary floods is insufficient to thoroughly mix water among all the floodplain, and
398 some environments begin to homogenize in their environmental filters while others retain their
399 identity [97]. The selection may act with similar strength both homogenizing and segregating local

communities, leading to an overall reduction of community turnover and β -diversity. Simultaneously, the metacommunity should be relatively random due to intermediate dispersal rates; 3) At extremely environmental homogeneity, the homogeneous selection was the main ecological rule structuring the metacommunity (Fig. 8a). Despite stochasticity is expected to increase due to the high dispersal rates, the bacterial species should be strongly filtered by common environmental factors and the strength of homogeneous selection may dampen the influence of drift and dispersion. For instance, low community turnover and β -diversity are expected. These evidences give support to the early assumptions made by Stegen et al. [41] and Dini-Andreote et al. [78] and to a new view of bacterial community assembly in which the action of selective and non-selective processes are considered.

The understanding of ecological processes acting on metacommunities was extended studying the bacterial co-occurrence patterns. Network analyses-based approaches have the potential to infer inter-taxa correlations and can be applied to investigate structure complexity [47, 98], randomness [99], and the ecological rules guiding community assembly [44]. Particularly, the influence of ecological processes on bacterial interactions is a poorly explored field. Emerging studies in soil communities have revealed that systems under organic farming harbor more complex networks than conventional farming, mainly associated to higher habitat heterogeneity [46]. Additionally, it has been recently demonstrated that greater microbial diversity ensures greater association complexity [47].

Our results demonstrate that networks' topology was markedly different according to the weight of the ecological processes acting on the bacterial community assemblage. At heterogeneous selection dominance (LW and LWs phases), the network associations had higher complexity (Fig. 8b). We assign this finding to the fact that the heterogeneous selection promotes a greater taxonomic β -diversity [47]. Contrarily, the homogeneous selection drove less complex networks decreasing the β -diversity. Additionally, under multiple ecological processes acting with similar strength, the network also tended to be less complex (Fig. 8b).

Another important finding of our study was the influence of the ecological processes on the randomness of networks. We observed less chance of bacterial associations when selection was stronger (Fig 8b). This is a general pattern mainly associated to the action of selection as taxa with similar ecological requirements tend to co-occur more than by chance (e.g. [44, 99, 100]). However, this is one of the first studies that have empirically linked bacterioplankton networks to structuring processes. Thus, our results add a new dimension revealing that both heterogeneous and homogeneous selection promotes less random associations. Additionally, as stochastic processes become more relevant, more random networks would be expected.

A useful feature of network analysis is that it allows to identify the strongly interconnected keystone taxa which have a core effect on the community assemblage [87, 90] and are optimal predictors of overall community changes [91]. Each hydrological phase presented a particular set of keystone taxa that were determined by different selective factors. It was not surprisingly considering that environmental conditions significantly varied among the hydrological phases (Fig. 1). Perhaps the most intriguing fact of our findings was the huge number of keystone taxa found under the highest environmental heterogeneity (50 vs 11 to 15 taxa). Even under similar networks' complexity and structuring processes at both low water phases, the number of keystone taxa was clearly higher. We attribute this finding to a higher niche segregation and number of significant environmental factors related with keystone taxa from LW. While the inorganic dissolved nitrogen was the main driver at both low water periods, in LW the molecular weight of CDOM had also a significant role in determining keystone taxa, particularly those abundant in isolated lakes. This result brings new evidence to previously assumption about the role of environmental heterogeneity in determining the number and identity of keystone taxa [46, 90].

Concluding remarks

Integrating all the empirical evidence obtained here we propose a conceptual model that synthetizes how environmental heterogeneity guides the action of ecological processes assembling the bacterial

metacommunity (Fig 8c). In systems with high environmental heterogeneity, the heterogeneous selection plays the major role in structuring the community, promoting a greater β -diversity, and more complex association networks. At intermediate values of environmental heterogeneity, the action of stochastic processes reaches more importance, and both heterogeneous and homogeneous selections have similar contribution; this leads to a decrease of β -diversity and to more randomness in networks associations. In systems with high environmental homogeneity, the strength of homogeneous selection dampens the influence of drift and dispersion, and local communities become more similar and the association networks less complex.

The strength of this model is that it was conceived relying on different metacommunity features (e.g. taxonomic and phylogenetic turnover, network associations) taking into account the spatial and temporal scales. Furthermore, we used a step-forward analysis combining different statistical inferences, which allowed us to take advantage of the strengths of each method alone and in synergy with the others. Therefore, the proposed model represents a significant improvement on the fundamental knowledge of bacterial community's assembly across freshwater ecosystems and provides a new framework for future studies in other communities.

Acknowledgements

This study was supported by Agencia Nacional de Promoción Científica y Tecnológica (PICT 2012-2095 PI M. Marchese, and PICT 2016-0465 PI M. Devercelli), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). H. Sarmiento work was supported by Conselho Nacional de Desenvolvimento e Pesquisa Tecnológica (CNPq process 309514/2017-7) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP process 2014/14139-3). We thank C. Debonis, E. Creus, and M. Piacenza for their field assistance, and A. Dabin for bioinformatic support. We thank the members of Metacommunity Project particularly M. Marchese, P. Scarabotti, D. Borzone, M. Licursi, for the long discussions that largely improved the manuscript.

478 **Author contributions**

479 PH and MD conceived and designed the study; MD conducted the samplings with the help of GM;
480 PH and SM performed molecular analyzes; GM performed chemical analyses; PH ran statistical
481 analyses; PH and MD wrote the manuscript with substantial contributions from FU and HS; all
482 authors edited the manuscript.

483
484 The authors declare that they have no conflict of interest.
485

486 Supplementary information is available.

487

488 **References**

- 489 1. Hubbell SP. The Unified Neutral Theory of Biodiversity and Biogeography, 1st ed. Princeton
490 University Press: Woodstock, United Kingdom, 2001.
- 491 2. Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, et al.
492 Microbial biogeography: Putting microorganisms on the map. *Nat Rev Microbiol* 2006; **4**:
493 102–112.
- 494 3. Vellend M. The Theory of Ecological Communities, 1st ed. Princeton University Press:
495 Woodstock, United Kingdom, 2016.
- 496 4. McGill BJ. Towards a unification of unified theories of biodiversity. *Ecol Lett* 2010; **13**:
497 627–642.
- 498 5. Gilbert B. Joint consequences of dispersal and niche overlap on local diversity and resource
499 use. *J Ecol* 2012; **100**: 287–296.
- 500 6. Chase JM, Leibold MA. Ecological Niches: Linking Classical and Contemporary
501 Approaches. *Biodivers Conserv* 2004; **13**: 1791–1793.
- 502 7. Rosindell J, Hubbell SP, Etienne RS. The Unified Neutral Theory of Biodiversity and

- 503 Biogeography at Age Ten. *Trends Ecol Evol* 2011; **26**: 340–348.
- 504 8. Bell G. Neutral macroecology. *Science* 2001; **293**: 2413–2418.
- 505 9. Leibold MA, Holyoak M, Mouquet N, Amarasekare P, Chase JM, Hoopes MF, et al. The
506 metacommunity concept: A framework for multi-scale community ecology. *Ecol Lett* 2004;
507 **7**: 601–613.
- 508 10. Wilson DS. Complex interactions in metacommunities, with implications for biodiversity
509 and higher levels of selection. *Ecology* 1992; **73**: 1984–2000.
- 510 11. Vellend M. Conceptual synthesis in community ecology. *Q Rev Biol* 2010; **85**: 183–206.
- 511 12. Stegen JC, Lin X, Fredrickson JK, Chen X, Kennedy DW, Murray CJ, et al. Quantifying
512 community assembly processes and identifying features that impose them. *ISME J* 2013; **7**:
513 2069–2079.
- 514 13. MacArthur RH, Wilson EO. The Theory of Island biogeography. 1st ed. Princeton University
515 Press: Woodstock, United Kingdom, 2001.
- 516 14. Lowe WH, McPeck MA. Is dispersal neutral? *Trends Ecol Evol* 2014; **29**: 444–450.
- 517 15. Vellend M, Srivastava DS, Anderson KM, Brown CD, Jankowski JE, Kleynhans EJ, et al.
518 Assessing the relative importance of neutral stochasticity in ecological communities. *Oikos*
519 2014; **123**: 1420–1430.
- 520 16. Rundle HD, Nosil P. Ecological speciation. *Ecol Lett* 2005; **8**: 336–352.
- 521 17. Graham EB, Knelman JE, Schindlbacher A, Siciliano S, Breulmann M, Yannarell A, et al.
522 Microbes as engines of ecosystem function: When does community structure enhance
523 predictions of ecosystem processes? *Front Microbiol* 2016; **7**.
- 524 18. Lindström ES, Langenheder S. Local and regional factors influencing bacterial community
525 assembly. *Environ Microbiol Rep* 2012; **4**: 1–9.
- 526 19. Isabwe A, Ren K, Wang Y, Peng F, Chen H. Community Assembly Mechanisms Underlying

- 527 the Core and Random Bacterioplankton and Microeukaryotes in a River–Reservoir System.
528 *Water* 2019; **11**: 1–17.
- 529 20. Fodelianakis S, Lorz A, Valenzuela-Cuevas A, Barozzi A, Booth JM, Daffonchio D.
530 Dispersal homogenizes communities via immigration even at low rates in a simplified
531 synthetic bacterial metacommunity. *Nat Commun* 2019; **10**: 1314.
- 532 21. Jiao S, Yang Y, Xu Y, Zhang J, Lu Y. Balance between community assembly processes
533 mediates species coexistence in agricultural soil microbiomes across eastern China. *ISME J*
534 2020; **14**: 202–216.
- 535 22. Grosberg RK, Vermeij GJ, Wainwright PC. Biodiversity in water and on land. *Curr Biol*
536 2012; **22**: R900–R903.
- 537 23. Wang J, Shen J, Wu Y, Tu C, Soininen J, Stegen JC, et al. Phylogenetic beta diversity in
538 bacterial assemblages across ecosystems: Deterministic versus stochastic processes. *ISME J*
539 2013; **7**: 1310–1321.
- 540 24. Zinger L, Amaral-Zettler LA, Fuhrman JA, Horner-Devine MC, Huse SM, Welch DBM, et
541 al. Global patterns of bacterial beta-diversity in seafloor and seawater ecosystems. *PLoS One*
542 2011; **6**: e24570 .
- 543 25. Barberán A, Casamayor EO. Global phylogenetic community structure and β -diversity
544 patterns in surface bacterioplankton metacommunities. *Aquat Microb Ecol* 2010; **59**: 1–10.
- 545 26. Nemergut DR, Schmidt SK, Fukami T, O'Neill SP, Bilinski TM, Stanish LF, et al. Patterns
546 and Processes of Microbial Community Assembly. *Microbiol Mol Biol Rev* 2013; **77**: 342–
547 356.
- 548 27. Ruiz-González C, Niño-García JP, del Giorgio PA. Terrestrial origin of bacterial
549 communities in complex boreal freshwater networks. *Ecol Lett* 2015; **18**: 1198–1206.
- 550 28. Comte J, Lovejoy C, Crevecoeur S, Vincent WF. Co-occurrence patterns in aquatic bacterial

- 551 communities across changing permafrost landscapes. *Biogeosciences* 2016; **13**: 175–190.
- 552 29. Heino J, Melo AS, Siqueira T, Soininen J, Valanko S, Bini LM. Metacommunity
553 organisation, spatial extent and dispersal in aquatic systems: Patterns, processes and
554 prospects. *Freshw Biol* . 2015. , **60**: 845–869
- 555 30. Logares R, Deutschmann IM, Giner CR, Krabberød AK. Different processes shape
556 prokaryotic and picoeukaryotic assemblages in the sunlit ocean microbiome. *Environ*
557 *Microbiol* 2018; **20**: 37–49.
- 558 31. Östman Ö, Drakare S, Kritzberg ES, Langenheder S, Logue JB, Lindström ES. Importance of
559 space and the local environment for linking local and regional abundances of microbes.
560 *Aquat Microb Ecol* 2012; **67**: 35–45.
- 561 32. Niño-García JP, Ruiz-González C, del Giorgio PA. Interactions between hydrology and
562 water chemistry shape bacterioplankton biogeography across boreal freshwater networks.
563 *ISME J* 2016; **1**: 1–12.
- 564 33. Wang J, Shen J, Wu Y, Tu C, Soininen J, Stegen JC, et al. Phylogenetic beta diversity in
565 bacterial assemblages across ecosystems: deterministic versus stochastic processes. *ISME J*
566 2013; **7**: 1310–1321.
- 567 34. Ofîteru ID, Lunn M, Curtis TP, Wells GF, Criddle CS, Francis CA, et al. Combined niche
568 and neutral effects in a microbial wastewater treatment community. *Proc Natl Acad Sci U S*
569 *A* 2010; **107**: 15345–15350.
- 570 35. Fernandes IM, Henriques-Silva R, Penha J, Zuanon J, Peres-Neto PR. Spatiotemporal
571 dynamics in a seasonal metacommunity structure is predictable: The case of floodplain-fish
572 communities. *Ecography (Cop)* 2014; **37**: 464–475.
- 573 36. José de Paggi SB, Paggi JC. Hydrological connectivity as a shaping force in the zooplankton
574 community of two lakes in the Paraná River floodplain. *Int Rev Hydrobiol* 2008; **93**: 659–
575 678.

- 576 37. Devercelli M, Scarabotti P, Mayora G, Schneider B, Giri F. Unravelling the role of
577 determinism and stochasticity in structuring the phytoplanktonic metacommunity of the
578 Paraná River floodplain. *Hydrobiologia* 2016; **764**: 139–156.
- 579 38. Devercelli M. Changes in phytoplankton morpho-functional groups induced by extreme
580 hydroclimatic events in the Middle Paraná river (Argentina). *Hydrobiologia* 2010; **639**: 5–19.
- 581 39. Scarabotti PA, López JA, Pouilly M. Flood pulse and the dynamics of fish assemblage
582 structure from neotropical floodplain lakes. *Ecol Freshw Fish* 2011; **20**: 605–618.
- 583 40. Declerck SAJ, Winter C, Shurin JB, Suttle CA, Matthews B. Effects of patch connectivity
584 and heterogeneity on metacommunity structure of planktonic bacteria and viruses. *ISME J*
585 2013; **7**: 533–542.
- 586 41. Stegen JC, Lin X, Konopka AE, Fredrickson JK. Stochastic and deterministic assembly
587 processes in subsurface microbial communities. *ISME J* 2012; **6**: 1653–1664.
- 588 42. Vass M, Székely AJ, Lindstrom ES, Langenheder S. Using null models to compare bacterial
589 and microeukaryotic metacommunity assembly under shifting environmental conditions
590 under shifting environmental conditions. *Ev. bioRxiv Prepr first* 2019;
591 **doi**: <https://doi.org/10.1101/752303>.
- 592 43. Li Y, Gao Y, Zhang W, Wang C, Wang P, Niu L, et al. Homogeneous selection dominates
593 the microbial community assembly in the sediment of the Three Gorges Reservoir. *Sci Total*
594 *Environ* 2019; **690**: 50–60.
- 595 44. Barberán A, Bates ST, Casamayor EO, Fierer N. Using network analysis to explore co-
596 occurrence patterns in soil microbial communities. *ISME J* 2012; **6**: 343–351.
- 597 45. Banerjee S, Thrall PH, Bissett A, Van Der Heijden MGA, Richardson AE. Linking microbial
598 co-occurrences to soil ecological processes across a woodland-grassland ecotone. *Int J Bus*
599 *Innov Res* 2018; **17**: 8217–8230.

- 600 46. Banerjee S, Walder F, Büchi L, Meyer M, Held AY, Gatteringer A, et al. Agricultural
601 intensification reduces microbial network complexity and the abundance of keystone taxa in
602 roots. *ISME J* 2019; **13**: 1722–1736.
- 603 47. Wagg C, Schlaeppli K, Banerjee S, Kuramae EE, van der Heijden MGA. Fungal-bacterial
604 diversity and microbiome complexity predict ecosystem functioning. *Nat Commun* 2019; **10**:
605 4841.
- 606 48. Graham EB, Crump AR, Resch CT, Fansler S, Arntzen E, Kennedy DW, et al. Deterministic
607 influences exceed dispersal effects on hydrologically-connected microbiomes. *Environ*
608 *Microbiol* 2017; **19**: 1552–1567.
- 609 49. Liu J, Meng Z, Liu X, Zhang X-H. Microbial assembly, interaction, functioning, activity and
610 diversification: a review derived from community compositional data. *Mar Life Sci Technol*
611 2019; **1**: 1–17.
- 612 50. Neiff JJ. Large rivers of South America: toward the new approach. *Verh Internat Verein*
613 *Limnol* 1996; **26**: 167–180.
- 614 51. Drago EC. The physical dynamics of the river-lake floodplain system. *The Middle Paraná*
615 *River: Limnology of a Subtropical Wetland*. 2007. pp 83–122.
- 616 52. Mayora G, Devercelli M, Giri F. Spatial variability of chlorophyll-a and abiotic variables in a
617 river-floodplain system during different hydrological phases. *Hydrobiologia* 2013; **717**: 51–
618 63.
- 619 53. Mayora G, Scarabotti P, Schneider B, Alvarenga P, Marchese M. Multiscale environmental
620 heterogeneity in a large river-floodplain system. *J South Am Earth Sci* 2020; (in press).
- 621 54. Thomaz SM, Bini LM, Bozelli RL. Floods increase similarity among aquatic habitats in
622 river-floodplain systems. *Hydrobiologia* 2007; **579**: 1–13.
- 623 55. Camilloni IA, Barros VR. Extreme discharge events in the Paraná River and their climate

- 624 forcing. *J Hydrol* 2003; **278**: 94–106.
- 625 56. Depetris PJ, Probst JL, Pasquini AI, Gaiero DM. The geochemical characteristics of the
626 Paraná River suspended sediment load: An initial assessment. *Hydrol Process* 2003; **17**:
627 1267–1277.
- 628 57. Marchese M, de Drago IE. Benthos of the lotic environments in the middle Paraná River
629 system: transverse zonation. *Hydrobiologia* 1992; **237**: 1–13.
- 630 58. Zilli FL, Paggi AC. Ecological responses to different degrees of hydrologic connectivity:
631 Assessing patterns in the bionomy of benthic chironomids in a large river-floodplain system.
632 *Wetlands* 2013; **33**: 837–845.
- 633 59. José de Paggi SB, Devercelli M, Molina FR. Zooplankton and their driving factors in a large
634 subtropical river during low water periods. *Fundam Appl Limnol / Arch für Hydrobiol* 2014;
635 **184**: 125–139.
- 636 60. Fernandez Zenoff V, Sineriz F, Farias ME. Diverse Responses to UV-B Radiation and
637 Repair Mechanisms of Bacteria Isolated from High-Altitude Aquatic Environments. *Appl*
638 *Environ Microbiol* 2006; **72**: 7857–7863.
- 639 61. Herlemann DPR, Labrenz M, Jürgens K, Bertilsson S, Waniek JJ, Andersson AF. Transitions
640 in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J* 2011;
641 **5**: 1571–1579.
- 642 62. Caporaso JG, Lauber CL, Walters W a, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-
643 throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME*
644 *J* 2012; **6**: 1621–4.
- 645 63. Logares R. Workflow for Analysing MiSeq Amplicons based on Uparse v1.5.
646 <https://doi.org/105281/zenodo259579> 2017.
- 647 64. Edgar RC. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon

sequencing. *bioRxiv* 2016; 081257.

65. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. *Vegan: Community Ecology Package*. R package version 2.0-9. 2013.

66. R Development Core Team 3.0.1. A Language and Environment for Statistical Computing. *R Found Stat Comput* . 2013.

67. Ranjard L, Dequiedt S, Chemidlin Prévost-Bouré N, Thioulouse J, Saby NPA, Lelievre M, et al. Turnover of soil bacterial diversity driven by wide-scale environmental heterogeneity. *Nat Commun* 2013; **4**: 1–10.

68. Anderson MJ. A new method for non-parametric multivariate analysis of variance. *Austral Ecol* 2001; **26**: 32–46.

69. Hammer Ø, Harper DAT, Ryan PD. Past: Paleontological statistics software package for education and data analysis. *Palaeontol Electron* 2001; **4**: 9.

70. Pedrós-Alió C. The Rare Bacterial Biosphere. *Ann Rev Mar Sci* 2012; **4**: 449–466.

71. Zhou J, Ning D. Stochastic Community Assembly: Does It Matter in Microbial Ecology? *Microbiol Mol Biol Rev* 2017; **81**: 1–32.

72. Fine PVA, Kembel SW. Phylogenetic community structure and phylogenetic turnover across space and edaphic gradients in western Amazonian tree communities. *Ecography (Cop)* 2011; **34**: 552–565.

73. Webb CO, Ackerly DD, McPeck M a., Donoghue MJ. Phylogenies and Community Ecology. *Annu Rev Ecol Syst* 2002; **33**: 475–505.

74. Webb CO, Ackerly DD, Kembel SW. Phylocom: Software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics* 2008; **24**: 2098–2100.

75. Kembel SW. Disentangling niche and neutral influences on community assembly: Assessing the performance of community phylogenetic structure tests. *Ecol Lett* 2009; **12**: 949–960.

- 672 76. Oden NL, Sokal RR. Directional autocorrelation: An extension of spatial correlograms to two
673 dimensions. *Syst Zool* 1986; **35**: 608–617.
- 674 77. Diniz-Filho JAF, Terribile LC, Da Cruz MJR, Vieira LCG. Hidden patterns of phylogenetic
675 non-stationarity overwhelm comparative analyses of niche conservatism and divergence.
676 *Glob Ecol Biogeogr* 2010; **19**: 916–926.
- 677 78. Dini-Andreote F, Stegen JC, Van Elsas JD, Salles JF. Disentangling mechanisms that
678 mediate the balance between stochastic and deterministic processes in microbial succession.
679 *Proc Natl Acad Sci U S A* 2015; **112**: E1326–E1332.
- 680 79. Andersson AF, Riemann L, Bertilsson S. Pyrosequencing reveals contrasting seasonal
681 dynamics of taxa within Baltic Sea bacterioplankton communities. *ISME J* 2010; **4**: 171–81.
- 682 80. Chase JM, Myers JA. Disentangling the importance of ecological niches from stochastic
683 processes across scales. *Philos Trans R Soc B Biol Sci* 2011; **366**: 2351–2363.
- 684 81. Friedman J, Alm EJ. Inferring Correlation Networks from Genomic Survey Data. *PLoS*
685 *Comput Biol* 2012; **8**: e1002687.
- 686 82. Watts SC, Ritchie SC, Inouye M, Holt KE. FastSpar: Rapid and scalable correlation
687 estimation for compositional data. *Bioinformatics* 2019; **35**: 1064–1066.
- 688 83. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A
689 software Environment for integrated models of biomolecular interaction networks. *Genome*
690 *Res* 2003; **13**: 2498–2504.
- 691 84. Assenov Y, Ramírez F, Schelhorn SESE, Lengauer T, Albrecht M. Computing topological
692 parameters of biological networks. *Bioinformatics* 2008; **24**: 282–284.
- 693 85. Röttgers L, Faust K. From hairballs to hypotheses—biological insights from microbial
694 networks. *FEMS Microbiol Rev* 2018; **42**: 761–780.
- 695 86. Cao X, Zhao D, Xu H, Huang R, Zeng J, Yu Z. Heterogeneity of interactions of microbial

- 696 communities in regions of Taihu Lake with different nutrient loadings: A network analysis.
697 *Sci Rep* 2018; **8**: 1–11.
- 698 87. Newman MEJ. The structure and function of networks. *Comput Phys Commun* 2002; **147**:
699 40–47.
- 700 88. Newman MEJ. The structure and function of complex networks. *SIAM Rev* . 2003. Society
701 for Industrial and Applied Mathematics Publications. , **45**: 167–256
- 702 89. Barabási AL, Oltvai ZN. Network biology: Understanding the cell’s functional organization.
703 *Nat Rev Genet* . 2004. , **5**: 101–113
- 704 90. Banerjee S, Schlaeppi K, van der Heijden MGA. Keystone taxa as drivers of microbiome
705 structure and functioning. *Nat Rev Microbiol* 2018; **16**: 567–576.
- 706 91. Herren CM, McMahon KD. Keystone taxa predict compositional change in microbial
707 communities. *Environ Microbiol* 2018; **20**: 2207–2217.
- 708 92. Berry D, Widder S. Deciphering microbial interactions and detecting keystone species with
709 co-occurrence networks. *Front Microbiol* 2014; **5**: 1–14.
- 710 93. Newman MEJ. A measure of betweenness centrality based on random walks. *Soc Networks*
711 2005; **27**: 39–54.
- 712 94. Brandes U. A faster algorithm for betweenness centrality. *J Math Sociol* 2001; **25**: 163–177.
- 713 95. ter Braak CJF, Šmilauer P. CANOCO Reference Manual and CanoDraw for Windows User’s
714 Guide: Software for Canonical Community Ordination (Version 5). *Sect Permut Methods*
715 *Microcomput Power, Ithaca, New York*, 2012.
- 716 96. Powell JR, Karunaratne S, Campbell CD, Yao H, Robinson L, Singh BK. Deterministic
717 processes vary during community assembly for ecologically dissimilar taxa. *Nat Commun*
718 2015; **6**: 1–10.
- 719 97. Weilhoefer CL, Pan Y, Eppard S. The effects of river floodwaters on floodplain wetland

720 water quality and diatom assemblages. *Wetlands* 2008; **28**: 473–486.

721 98. Faust K, Sathirapongsasuti JF, Izard J, Segata N, Gevers D, Raes J, et al. Microbial co-
722 occurrence relationships in the Human Microbiome. *PLoS Comput Biol* 2012; **8**: e1002606.

723 99. Lupatini M, Suleiman AKA, Jacques RJS, Antoniolli ZI, de Siqueira Ferreira A, Kuramae
724 EE, et al. Network topology reveals high connectance levels and few key microbial genera
725 within soils. *Front Environ Sci* 2014; **2**: 1–11.

726 100. Ma B, Wang H, Dsouza M, Lou J, He Y, Dai Z, et al. Geographic patterns of co-occurrence
727 network topological features for soil microbiota at continental scale in eastern China. *ISME J*
728 2016; **10**: 1891–1901.

729

730

Figures

Fig. 1: a) The Paraná River in South America showing the location of the study area and b) the environments sampled. c) Daily water level from 2013 to 2016 in the Paraná River. The sampled periods are indicated with arrows.

ID	Env. Type	Env. Name	ID	Env. Type	Env. Name	ID	Env. Type	Env. Name	ID	Env. Type	Env. Name
1	MC	Paraná River	7	SC	Colastinecito de las Cruces Stream	13	MC	Coronda River	19	IL	Chancha Lake
2	CL	Miní Lake	8	IL	Chajá Lake	14	IL	Las Garzas Lake	20	IL	Escondido Swamp
3	MC	Colastiné River	9	SC	Correntoso Swamp	15	IL	Chicana Lake	21	SC	Paraná River (Viejo)
4	CL	Moreira Lake	10	IL	Chajacito Lake	16	SC	Cataratas Stream	22	IL	Sauces Lake
5	CL	Ferranda Lake	11	SC	Colastinecito de las Cruces Stream	17	IL	La Curva Lake			
6	CL	Perla Lake	12	IL	Del Medio Lake	18	IL	Los Gansos Lake			

Fig. 2: Environmental heterogeneity in the Paraná fluvial system at each hydrological phase expressed as average environmental dissimilarity between sites ($\overline{Ed}$, grey bars) and the variability (coefficient of variation, CV%) of nine selected abiotic variables among environments. **LW:** low water, **LWs:** low water with sedimentological pulse, **HW:** high water, **eHW:** extraordinary high water.

Fig. 3: a) Samples ordination in a non-metric multidimensional scaling (NMDS) according to the similarity in bacterial communities structure (zOTU relative abundance) among four hydrological phases (indicated with color) at the different type of environments (indicated with symbols) of the Paraná fluvial system. Stress value: 0.112. **b-c)** Richness and Shannon-Weaver diversity index H' of local bacterial communities at the Paraná fluvial system in four hydrological phases. **LW:** low water, **LWs:** low water with sedimentological pulse, **HW:** high water, **eHW:** extraordinary high water. **IL:** isolated lake, **CL:** connected lake, **SC:** secondary channel, **MC:** main channel.

Fig. 4: Relative importance of high-level processes structuring the bacterial metacommunity of the Paraná fluvial system in four hydrological phases. Values indicate percentage of community turnover associated to each process: homogeneous selection, heterogeneous selection, homogeneous dispersal, dispersal limitation combined with drift, and drift. **LW:** low water, **LWs:** low water with sedimentological pulse, **HW:** high water, **eHW:** extraordinary high water.

Fig. 5: Bacterial associations networks for each hydrological phase of the Paraná fluvial system, arranged according to bacterial taxonomy (classes or orders). The size of each node is proportional to the number of connections (degree). Only significant ($P < 0.01$) and strong correlations ($|r| > 0.7$) were considered. The number of nodes, number of edges, average number of neighbors, and clustering coefficient of the overall network was indicated below each network.

● Acidobacteria ● Bacteroidetes ● Cyanobacteria ● Planctomycetes ● Betaproteobacteria ● Epsilonproteobacteria ● Verrucomicrobia
 ● Actinobacteria ● Chloroflexi ● Firmicutes ● Alphaproteobacteria ● Deltaproteobacteria ● Gammaproteobacteria ● Others

Fig. 6: Complexity of bacterial associations networks for each hydrological phase of the Paraná fluvial system expressed as the distribution of degrees. Degree indicates the number of associations shared by each node in a network.

Fig. 7: Redundancy analysis based on the abundance of keystone taxa per site versus environmental variables. Only the significant variables ($P < 0.05$, forward selection) are shown (red arrows). CDOM (A_{440}) is the light absorption coefficients at 440 nm that is proportional to the CDOM concentration. CDOM ($S_{275-295}$) is spectral slope from 275 to 295 nm and indicates the CDOM molecular weight. Circles represent the position of each zOTU in the biplot and the color indicates the hydrological phase in which they were selected as keystone: low water (LW), low water with sedimentological pulse (LWs), high water (HW), extraordinary high water (eHW). Those zOTUs selected in more than one hydrological phase are indicated with different circles.

787 **Fig. 8:** Empirical evidence (a-b) that support a new conceptual model (c) on how environmental
788 heterogeneity regulates the four ecological processes that shape bacterial metacommunity
789 assemblage (*sensu* Vellend [3,11]). At high environmental heterogeneity, community structure is
790 mainly determinate by heterogeneous selection. The heterogeneous selection promotes a greater β -
791 diversity increasing divergence in local communities, and a greater complexity bacterial association
792 networks. At intermediate values of environmental heterogeneity both heterogeneous and
793 homogeneous selections act with similar strength, and stochastic processes reaches more
794 importance. In this scenario, the selection acting in opposite ways does not allow homogenization or
795 differentiation of local community structure, leading to an overall reduction of turnover and β -
796 diversity. Simultaneously, the metacommunity would have a relatively high randomness due to
797 intermediate dispersal rates. As consequence, the association network would tend to be less
798 complex and more random. At extremely homogenous environmental conditions, the low diversity
799 of niche habitat leads to an increase of homogeneous selection, and local communities are strongly
800 filtered by common environmental factors. The homogeneous selection leads to a low turnover and
801 β -diversity, promoting less complex associations.