bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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.

1 **Space-time dynamics in monitoring neotropical fish communities using eDNA**
2 **metabarcoding**

3 Naiara Guimarães Sales[1,2], Owen Simon Wangensteen[3], Daniel Cardoso Carvalho[4], Kristy
4 Deiner[5], Kim Præbel[3], Ilaria Coscia[1], Allan D. McDevitt[1], Stefano Mariani[1,6]
5
6 [1] Ecosystems and Environment Research Centre, School of Science, Engineering and
7 Environment, University of Salford, UK.
8 [2] CESAM – Centre for Environmental and Marine Studies, Departamento de Biologia Animal,
9 Faculdade de Ciências da Universidade de Lisboa, Lisbon, Portugal
10 [3] Norwegian College of Fishery Science, UiT - The Arctic University of Norway, Tromsø,
11 Norway.
12 [4] Programa de Pós-graduação em Biologia de Vertebrados, Pontifícia Universidade Católica
13 de Minas Gerais, Belo Horizonte, Brasil.
14 [5] Life Sciences, Natural History Museum, London, UK
15 [6] School of Natural Sciences and Psychology, Liverpool John Moores University, Liverpool,
16 UK

17 **Corresponding author:**

18 Naiara Guimarães Sales, naiarasl@gmail.com

19 CESAM – Centre for Environmental and Marine Studies, Departamento de Biologia Animal,
20 Faculdade de Ciências da Universidade de Lisboa, 1749-016 Lisbon, Portugal

21 Running title: eDNA for fish biodiversity monitoring in Brazil
22

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity.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.

23 **ABSTRACT**

24 The biodiverse Neotropical ecoregion remains insufficiently assessed, poorly managed, and
25 threatened by unregulated human activities. Novel, rapid and cost-effective DNA-based
26 approaches are valuable to improve understanding of the biological communities and for
27 biomonitoring in remote areas. Here, we evaluate the potential of environmental DNA (eDNA)
28 metabarcoding for assessing the structure and distribution of fish communities by analysing
29 sediments and water from 11 locations along the Jequitinhonha River catchment (Brazil). Each
30 site was sampled twice, before and after a major rain event in a five-week period and fish
31 diversity was estimated using high-through-put sequencing of 12S rRNA amplicons. In total,
32 252 Molecular Operational Taxonomic Units (MOTUs) and 34 fish species were recovered,
33 including endemic, introduced, and previously unrecorded species for this basin. Spatio-
34 temporal variation of fish assemblages was detected, richness during the first campaign was
35 nearly twice as high as in the second sampling round; though peaks of diversity were primarily
36 associated with only four locations. No correlation between β-diversity and longitudinal
37 distance or presence of dams was detected, but low species richness observed at sites located
38 near dams indicates that these anthropogenic barriers might have an impact on local fish
39 diversity. Unexpectedly high α-diversity levels recorded at the river mouth suggest that these
40 sections should be further evaluated as putative “eDNA reservoirs” for rapid monitoring. By
41 uncovering spatio-temporal changes, unrecorded biodiversity components, and putative
42 anthropogenic impacts on fish assemblages, we further strengthen the potential of eDNA
43 metabarcoding as a biomonitoring tool, especially in regions often neglected or difficult to
44 access.

45 **Keywords:** eDNA, biodiversity assessment, fish, freshwater, Brazil, river

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

46 ## 1 INTRODUCTION

47 Despite covering less than 1% of the Earth's surface, freshwater habitats harbour over
48 40% of global fish diversity (Nelson, 2006; Dudgeon et al., 2006). Fish from rivers, lakes, and
49 wetlands provide essential protein subsistence for a large proportion of human populations
50 worldwide (FAO, 2012; McIntyre et., 2016), and are increasingly affected by anthropogenic
51 impacts (e.g. habitat modification, fragmentation, climate change; Vörösmarty et al., 2010;
52 Grill et al., 2019). Because of the global impact to freshwater ecosystems, their associated
53 vertebrate populations are declining at alarming rates (83% decline since 1970; WWF, 2018),
54 and their conservation and management are a priority for global biodiversity (IPBES, 2019).
55 Nevertheless, despite broad agreement on the requirements to understand and monitor
56 biodiversity and ecological networks in freshwater habitats (Socolar et al., 2015), our
57 comprehension of biodiversity conservation in this realm lags behind terrestrial and marine
58 environments (Jucker et al., 2018).

59 The Neotropical region harbours one of the greatest freshwater fish diversities in the
60 world (approximately 30% of all described freshwater fish species), and is currently facing
61 unprecedented levels of anthropogenic pressure. In this region, conservation and management
62 actions in freshwater habitats are challenging due to a lack of infrastructure leading to sampling
63 constraints, as well as a shortage of taxonomic expertise to fully characterise this megadiverse
64 ichthyofauna (Reis et al., 2016). In Neotropical countries, such as Brazil, fish biodiversity
65 assessment relies on sampling using traditional survey methods (e.g. gill nets and traps)
66 followed by morphological identification, which might be selective, harmful, and have low
67 detection rates for rare and elusive species) and small life-stages (Becker et al., 2015; Sales et
68 al., 2018).

69 Use of specific fishing practices coupled with the remoteness and large geographic
70 extension of most catchments, has meant that Neotropical rivers have not been sufficiently

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

71 surveyed for baseline estimates of fish diversity. Underestimation of fish diversity resulting
72 from low sampling efficiency may provide biased metrics and hamper management and
73 conservation plans (Trimble & van Aarde, 2012),), including recovery plans for damaged
74 ecosystems (Sales et al., 2018). In addition, with a significantly reduced investment in scientific
75 research and conservation (Thomé and Haddad, 2019), there is an urge to move towards more
76 cost-effective methods to estimate biodiversity at a broad scale (i.e. detecting and monitoring
77 multiple species simultaneously in vast areas).

78 Molecular approaches offer a universal key to identify, assess and quantify biodiversity,
79 especially in biodiversity-rich and understudied ecosystems and regions (Schwartz et al.,
80 2006). One of the most effective approaches to circumvent the limitations of traditional surveys
81 in mega-diverse systems is the use of DNA barcoding and metabarcoding (Gomes et al., 2015;
82 Cilleros et al., 2019). Sequencing trace DNA present in the water (environmental DNA or
83 eDNA) can now be reliably used to detect species presence (Deiner et al., 2017) and, to some
84 extent, abundance (Doi et al. 2017; Ushio et al. 2018; Shelton et al., 2019). Recently, Cilleros
85 et al. (2019) demonstrated the efficiency of eDNA metabarcoding in providing spatially
86 extensive data on freshwater fish biodiversity in French Guyana, and a better discrimination of
87 assemblage compositions when compared to traditional sampling. We recently showed the
88 influence of sampling medium, as well as sampling preservation and time, on the
89 reconstruction of ichthyofaunal assemblages in a Brazilian catchment, inferred through eDNA
90 (Sales et al., 2019). Nevertheless, the vast majority of eDNA metabarcoding biomonitoring
91 studies remain concentrated in temperate regions, in established and fairly well-accessible
92 environments (Handley et al., 2019; McDevitt et al., 2019).

93 In this study, we use eDNA metabarcoding to unravel patterns of fish diversity in a
94 poorly studied Brazilian catchment, the Jequitinhonha River Basin (JRB). This catchment
95 belongs to the east Atlantic basin complex, characterised by a high number of species

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

96 endemism (Reis et al., 2016). Until 2010, the known ichthyofauna of this catchment included
97 63 described fish species (including 10 introduced species and a substantial number of
98 endangered species, Rosa & Lima 2008; Andrade-Neto, 2010), making this river a relatively
99 low biodiversity ecosystem when compared to its neighbouring basins. This reduced species
100 richness had been linked to historical geological and geographical features (Andrade-Neto,
101 2010). However, the geological history of the Jequitinhonha is very similar to that of adjacent
102 basins (e.g. Doce and Mucuri river), which led to the consideration that more contemporary
103 factors may explain the low biodiversity in the catchment, including the lack of adequate
104 surveys and impact from anthropogenic activities. The Jequitinhonha is known to be affected
105 by severe droughts, the impact of dams in the main river course and tributaries, and the
106 occurrence of introduced species (Sales et al., 2018;). Thus, an inadequate baseline survey of
107 the basin might still account for a great number of native and cryptic species yet to be described
108 for this catchment (Jerep et al., 2016; Dutra et al., 2016; Nielsen, Pessali & Dutra, 2017).

109 Furthermore, as other semi-arid and arid regions, the Jequitinhonha faces great
110 variation in water availability (i.e. long dry periods and sudden heavy rain periods; Leite et al.,
111 2010). However, the influence of precipitation in fish assemblages dynamics have not been
112 evaluated in this context. Here we assessed fish diversity, spatially (along the river stem and in
113 two tributaries) and temporally, (before and after heavy precipitation) using eDNA
114 metabarcoding to test whether this DNA-based method can estimate community structure
115 along the course of this anthropogenically-impacted river and thus, be used for biomonitoring
116 purposes.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

117 **2 MATERIALS AND METHODS**

118

119 ***2.1 Study Area***

120 The Jequitinhonha River basin (Figure 1), Southeast Brazil (17° S, 43° W), flows
121 between two biodiversity hotspots ('Cerrado' and the Atlantic Forest) and is characterised by
122 a tropical climate and environmental heterogeneity. The main river flows over 1,082 km, from
123 its source in Serro, at an elevation of 1200 m, to its outlet in the Atlantic Ocean at the locality
124 of Belmonte. The main river stem is interrupted by two large dams built for hydroelectric power
125 generation: the Irapé, the tallest dam in Brazil, built in 2006, and the Itapebi, established in
126 2002.

127

128 ***2.2 Historical data and local reference database construction***

129 A compiled species list was built by retrieving all papers available using a Google
130 Scholar search with the terms "fish" and "Jequitinhonha", combined with a search in
131 Portuguese language journals (applying the terms "peixe", "Jequitinhonha", "ictiofauna"), we
132 included data from research papers as well as compiling information on species occurrence
133 from unpublished environmental reports (Table S1, Supplemental information).

134 To enhance the available reference sequence database in order to obtain a better
135 taxonomic assignment, we retrieved all 12S rRNA mitochondrial gene fish sequences available
136 from GenBank and sequenced 55 additional missing species (Table S2). Information regarding
137 sample preparation and sequencing is provided in the Supplemental information.

138

139 ***2.3 eDNA sampling and processing***

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

140 Two sampling campaigns were conducted at 11 sites during a five-week interval (first
141 sampling period: 22/01 to 01/02/2017; second sampling: 19/02 to 01/03/2017). In between the
142 two sampling campaigns a major precipitation event (from 2.1-50mm in the first sampling
143 event to 100-250 mm in the second sampling event - CPTEC/INPE, 2018) occurred. Sites
144 included locations on the main river (nine) and one on each of two of the major tributaries (the
145 Itacambiruçu river and the Araçuaí river; Fig. 1). At each site, six water samples of one liter
146 each and two sediment samples (˜25 mL) each were collected. Sediments samples were
147 preserved in ethanol and kept cold during the sampling At the time of sampling proper storage
148 conditions of samples in tropical field conditions had been untested, therefore, we split half of
149 the water samples (N=3) and stored them on ice in a cooling box while for the other samples
150 (N=3) the cationic surfactant benzalkonium chloride (BAC) was added at a final concentration
151 of 0.01% as a preservation buffer to suppress the degradation of DNA by microorganisms
152 (Yamanaka et al. 2017). The effect of storage treatment (ice *vs* BAC) on MOTU diversity
153 recovery was significant only for samples obtained during the first campaign. Still, despite
154 significant ($p = 0.016$) only 2% of the variance was explained, whereas no significant
155 difference was recovered for samples obtained during the second campaign (Sales et al. 2019),
156 all replicates were used for downstream analyses in this study. In total, 132 water samples and
157 44 sediment samples were analysed.

158 Environmental DNA sample filtration, DNA extraction from filtered water and
159 sediment samples, amplification of the 12S rRNA fragment using the MiFish primer set (Miya
160 et al., 2015), multiplexed library preparation, and sequencing of two separate libraries (Library
161 1/LIB1 – first sampling event; Library 2/LIB2 – second sampling event) in one Illumina MiSeq
162 platform run were conducted as described in Sales et al. (2019), and detailed in the Appendix
163 included in the Supplemental information. Detailed procedures to control for contamination
164 are also described in Supplemental information.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

165

166 *2.4 Bioinformatic analyses and taxonomic assignment*

167 The metabarcoding bioinformatics pipeline used for data analysis was based on the
168 OBITools software suite (Boyer et al., 2016), following the protocol described in Sales et al.
169 (2019). Clustering was conducted using a step-by-step aggregation method (SWARM, Mahé
170 et al., 2014) applying a clustering value of d=1 (detailed information on evaluation of different
171 clustering values can be found on Supplemental information). Molecular operational
172 taxonomic units (MOTUs) and the inferred species (based on at at least 97% of similarity with
173 reference sequences; Sales et al., 2020) richness were compared among the three obtained
174 datasets.

175 For the diversity analyses (species richness and β-diversity), we applied a conservative
176 approach and treated our results as presence/absence-based as suggested by Li et al (2018).
177 Often MOTUs are used as a proxy for species, however, the correlation between these two
178 classifications of diversity are not straightforward. Richness in MOTUs is highly influenced
179 by the occurrence of cryptic species and by the thresholds applied during the bioinformatic
180 analyses (Pawlowski et al., 2018), which may cause an overestimation of true richness (e.g.
181 inflation of different MOTUs belonging to the same species due to natural intraspecific
182 variability, PCR amplification and/or sequencing errors). On the other hand, richness based on
183 MOTUs being assigned to a species may be an underestimate due to the lack of a complete
184 reference database or due to a low taxonomic resolution of the target gene fragment analysed.

185 To verify whether the inferred community diversity patterns significantly varied
186 because of the species assignment process, two datasets were used for estimating community
187 metrics of alpha and beta. Specifically, the filtered dataset included only MOTUs that could be
188 identified to the rank of species, whereas the non-filtered dataset included all MOTUs retrieved

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

189 after quality filtering steps. The filtered dataset is a subset of the total MOTU diversity
190 recovered, and thus it provides a more conservative overview for known fish diversity (Li et
191 al., 2018).

192 A species name assigned to each MOTU might not correspond exactly to the species
193 occurring in the Jequitinhonha River Basin (based on the compiled species list; Table S1)
194 because when the correct species is not present in the reference database, the taxonomic
195 assignment is based on the closest congeneric species. In this case, species not previously
196 reported for this basin are marked with an asterisk in order to highlight that the species herein
197 included might be an indicative of occurrence of the genus and not the exact species present in
198 this river basin.

199 Statistical analyses were performed in R v3.5.1 (R Core Team 2019). Replicates were
200 pooled (water=6 samples per site, sediment=2 samples) before the following statistical
201 analyses. Alpha-diversity (species richness) was estimated as the total number of MOTUs
202 (unfiltered dataset), or number of MOTUs assigned to species level (filtered dataset), at each
203 sample site. β-diversity was obtained by generating a distance matrix based on the Jaccard
204 coefficient, using the *vegdist* function implemented in vegan 2.5-2 (Oksanen et al. 2013). The
205 Jaccard distance is based on presence or absence of species (value of 0 means both samples
206 share the same species whereas 1 means samples have no species in common). Principal
207 Coordinates Analysis (PCoA) was used to determine the relationship between distance and
208 sites in the β-diversity matrix (cmdscale function) and the correlation between β-diversity and
209 longitudinal distance and the β-diversity and presence of physical barriers (dams) was tested
210 using a Mantel test (Li et al., 2018). The geographic distance matrix between sites was
211 estimated using the road route because the road follows the river course and thus, this distance
212 would provide a better estimate when compared to linear distance between two sample
213 locations. The matrix used for testing the influence of physical barriers was constructed by

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

214 weighting distance values between sites according to the existence of barriers (e.g. 0 – no
215 physical barrier between sites, 1- one barrier between sites and 2 – two barriers).

216 Even after our extensive effort to supplement the reference database for taxonomic
217 assignment improvement, most of the MOTUs recovered were not identified to species level
218 (see above) and, thus, a great portion of biodiversity information that could be used for diversity
219 assessments is not included in the reduced filtered dataset. To verify the total diversity
220 recovered and to visualize the community data, we used a hierarchical structure of taxonomic
221 classifications, in the R package Metacoder (Foster et al., 2017). This package, designed for
222 metabarcoding data, provides "heat tree" plots using statistics associated with taxa (e.g. read
223 abundances) and allows for a visual comparison between samples that takes into account their
224 taxonomic/phylogenetic diversity. Venn diagrams were obtained by comparing the orders and
225 families included in the compiled species list, and orders and families detected in each of the
226 eDNA datasets (filtered and non-filtered) using BioVenn (Hulsen, Vlieg, & Alkema, 2008).

227

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

228 ## 3 RESULTS

229 Our extensive review of both published and non-published literature sources resulted
230 in 111 species records for the Jequitinhonha River Basin (Table S1).

231 We obtained 16.1 million raw reads (LIB1 - 6,399,823; LIB2 - 9,704,699) in one
232 Illumina MiSeq run (See Supplemental information for details). After quality control,
233 clustering and all initial filtering steps, 2056 (LIB1) and 967 (LIB2) MOTUs were kept, with
234 154 and 59 MOTUs being assigned to species with >0.97 -min-identity, respectively. The
235 number of retained MOTUs varied considerably between filtered and unfiltered datasets and
236 for several species, more than one MOTU was also recovered (Figure 2, Table S3 and Table
237 S4).

238

239 ### *3.1 Taxonomic assignment*

240 Based on the combined data (including all filtered datasets - species >0.97 identity)
241 detected fish diversity included six orders, 20 families, 28 genera and at least 34 fish species
242 (Figure 2, Table S4). Characiformes (n=12) and Siluriformes (n=12) were the two orders
243 represented by the largest number of species identified and all the remaining orders were
244 comprised by less than five species.

245 A comparison between species identified by eDNA and closely related species reported
246 for the JRB suggests that several congeneric species (e.g. *Leporinus, Prochilodus,*
247 *Trichomycterus*) are not discernible using our generally applied bioinformatic threshold of
248 3.0% due to a lack of taxonomically informative variation in the ~170 bp fragment of the 12
249 rRNA gene, for these groups (Table S5).

250 Comparing the data obtained for both sampling times (Figure 3, Table S6), four species
251 were detected only during the first sampling (*Australoheros facetus*, *Cyprinus carpio**,*

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

252 *Hypostomus* sp., *Trichomycterus* sp.), whilst *Coptodon zilli* and Hoplias intermedius* were
253 detected only in the second sampling.

254 Sediment samples failed to detect five species (*Australoheros facetus*, *Cyprinus*
255 *carpio**, *Hypostomus gymnorhyncus**, *Poecilia reticulata*, *Trichomycterus* sp.), whilst water
256 samples detected all species present in the sediments. Analyses of water and sediment samples
257 demonstrated the occurrence of both widely distributed as well as less abundant species.
258 Several taxa (e.g. *Leporinus* sp., *Prochilodus* sp., *Rhamdia quelen*) were detected in both water
259 and sediment samples in most of sampling sites, in at least one sampling campaign, and
260 therefore seem to have a broad geographic distribution in the Jequitinhonha river basin.

261 A remarkable result obtained by eDNA included the detection in all analysed sites of
262 species rarely reported in traditional sampling studies (e.g. *Crenicichla* sp., Figure 3). Also we
263 may highlight, the occurrence of putative new records for this basin including invasive species
264 such as the dourado - *Salminus brasiliensis** and pacamã - *Lophiosilurus alexandri**.
265 Furthermore, some species, including native and non-indigenous species, were restricted to a
266 few locations (e.g. native: roncador *Wertheimeria maculata* (sample sites 1, 3, 8 and 10); non-
267 indigenous: oscar *Astronotus ocellatus* (sample site 7); chameleon cichlid *Australoheros*
268 *facetus* (sample site 11); tilapias *Coptodon* sp.* (sample sites 1 and 2); or were detected in only
269 one campaign (e.g. *Australoheros facetus*, *Coptodon* sp.*, carp *Cyprinus carpio**, wolf fish
270 *Hoplias intermedius*, pleco *Hypostomus gymnorhyncus**, pencil catfish *Trichomycterus* sp.).

271 The filtered dataset provides a potentially more conservative estimate of fish diversity
272 at the rank of species because many MOTUs could not be assigned a name using the 97%
273 similarity threshold. Fish diversity depicted by the heat trees based on all detected MOTUs (i.e.
274 the unfiltered dataset) shows that diversity remains especially high for the Order
275 Characiformes, as many families appear to be comprised of several MOTUs (e.g. Anostomidae,

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

276 Prochilodontidae; Figure 4). Comparisons between the filtered and unfiltered datasets
277 demonstrated that a conservative approach (i.e. using filtered data) might lead to a biodiversity
278 information loss since it greatly reduces the diversity in MOTUs recovered and fails in
279 detecting orders and families known to occur in this catchment but that were not identified up
280 to the species level (Figure 5).

281

282 ***3.2 Species richness and Beta diversity***

283 During the first campaign, highest MOTU richness was found in water samples from
284 the most upstream (site 1) and downstream (site 11) sampling sites, followed by sampling sites
285 4 and 8 (Figure 6A). The lowest number of MOTUs was recovered for sample site 7. Beta
286 diversity patterns showed similarities between sample sites 4 and 11, and sample sites 1 and 8,
287 whereas sample site 7 showed the most distinct fish assemblage when compared to all
288 locations. Environmental DNA recovered from water samples collected three weeks later,
289 demonstrated that species richness among sites fluctuate in time in this catchment (Figure 6B),
290 with generally greater homogeneity in the species richness amongst all sample sites in the late
291 sampling event. Still, the most upstream and downstream locations (1, 2, 10, 11), alongside
292 sample site 8, still harboured the highest number of species.

293 Data recovered from sediment samples provided a different overview of species
294 richness and beta diversity. Overall, the number of species recorded for sediment samples was
295 lower compared to water samples in the first campaign (Figure 6C). Sample site 1 had a much
296 lower species richness compared to water samples along with sampling sites 2, 4, 8, 9, 10. An
297 increase in the species richness was detected for sampling sites 3, 5 and 7, while sample sites
298 11 and 8 were confirmed as highly species-rich locations. In the second campaign (Figure 6D),

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

299 when compared to data recovered from water samples, six sample sites (1, 2, 6, 8, 9, 10) had a
300 lower species richness, while higher values were obtained for sample sites 3, 4, 7.

301 Over time, the pattern of harbouring the highest species richness appeared relatively
302 constant in sites 1 and 11 for both sampling media, except in the first campaign where fewer
303 species were detected in location 1 for sediment. Yet, the most downstream location kept an
304 almost stable species richness in both sampling media for both sampling campaigns.

305 Longitudinal distance had a negligible effect on beta diversity amongst sample sites (*p*-
306 value > 0.05, Table 1) and the presence of physical barriers (e.g. dams) also did not show a
307 significant influence on beta diversity of different sample types (water and sediment, Table 1).
308 A positive significant correlation was found between filtered and unfiltered datasets, for both
309 water and sediment (Table 1.)

310 For both sampling media, despite the variation in taxa richness showed by both datasets,
311 the pattern of alpha diversity variation among sample sites obtained for filtered (species) and
312 unfiltered (MOTUs) datasets were still quite congruent (Figure 7). However, for sediment
313 samples collected in the first campaign, sites 3 and 11 had a greater MOTU diversity when
314 compared to all nine remaining locations (Figure 7C). Despite also being the most species rich
315 sites, the great amount of MOTUs obtained and not assigned indicates that a great diversity
316 remains hidden in this sampling medium. Also, as demonstrated by the PCoA (Figure 7C), in
317 the first campaign these sites had a more distinct fish assemblage when compared to the others.
318 Furthermore, a higher resolution was obtained for the unfiltered dataset as a more segregated
319 sample clustering is evident in the PCoA ordination. Sediment samples from the first campaign
320 exhibited a peculiar clustering, with highly diverse samples in 3 and 11 strongly separated from
321 all other sites.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

322 **4 DISCUSSION**

323 The understanding of species distribution and the processes shaping spatial variation
324 and community composition are crucial for applying sustainable management schemes and
325 ensure timely conservation of biodiversity, especially for endemic and threatened species. Such
326 actions also require methods that allow for rapid and robust detection of biodiversity at
327 different spatial scales (Kelly et al., 2014). Here, we used eDNA metabarcoding of water and
328 sediment samples to investigate fish community variation over time along the course of a
329 Neotropical river.

330 We found that eDNA metabarcoding applied to understanding fish distributions in a
331 neotropical setting greatly enhanced our ability to not only measure richness along the course
332 of a large river, but also to reveal hidden diversity and putative unrecorded species invasions.
333 The compiled list of species (N=111) reported for the Jequitinhonha river basin herein was
334 higher than previously recorded (N=63) in 2010 (Andrade-Neto, 2010). Hence, our thorough
335 evaluation of all possible taxonomic information available at the time of our study estimates
336 the occurrence of more than 80 species in this catchment (Andrade-Neto, 2010; Godinho et al.,
337 1999). Our molecular assessment based on eDNA metabarcoding demonstrates that, as of yet,
338 there may be even more species yet to be recorded and putting the richness of this basin on par
339 with other closely adjacent basing thought to harbour higher diversity. These results
340 demonstrate our current lack of understanding of tropic diversity in many systems and
341 corroborates that new DNA based methods are ideal in generating new baselines for
342 biodiversity monitoring.

343

344 **4.1 Introduced and native species**

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

345 Environmental DNA metabarcoding allows the detection of multiple species
346 simultaneously, including species not expected to occur in an area (Deiner et al., 2017), helping
347 to track biological invasions and providing an early warning of species introduction. Here,
348 almost 30% of the taxa detected by eDNA were non-indigenous species, including species not
349 reported yet for this catchment. To our knowledge, previous records of *Salminus brasiliensis*
350 and *Lophiosilurus alexandri* occurrence in the JRB are absent from the literature. These are
351 commercially important species, already introduced for fishery purposes in several Brazilian
352 basins (Vitule et al., 2014). Hence, their occurrence in the Jequitinhonha is not necessarily a
353 surprise. However, it raises concerns about the ecological consequences of such unmanaged
354 introductions. Biodiversity loss is not only restricted by species disappearance, but also by a
355 reduction in ecosystem services due to an increase of biological similarity between areas (i.e.
356 species loss or increase through biological introductions leading to biotic homogenization;
357 Rahel, 2000).

358 It has been widely documented that analysis of eDNA surpasses traditional methods for
359 assessment of biodiversity and detection of invasive species (Schmelzle & Kinziger, 2016;
360 McDevitt et al., 2019). The only cyprinid previously documented in this basin was
361 *Hypophthalmichthys molitrix*. Herein, we registered the presence of *Cyprinus carpio*, another
362 species that has been widely introduced to Brazilian waters (Alves et al., 2007). Environmental
363 DNA metabarcoding also detected various species of tilapia (*Oreochromis* sp. and *Coptodon*
364 *zilli*). The impacts of tilapia invasion are well known worldwide, and all species show high
365 invasive potential, including in Neotropical countries (Cassemiro et al., 2017).

366 Our study also detected remarkable cases, such as the native species *Crenicichla* sp.
367 The genus *Crenicichla* is one of the most species rich among the South American Cichlids,
368 where it is known to widely occur. However, the genus is still lacking an improved taxonomic
369 resolution and conservation status evaluation (Kullander & de Lucena, 2006). In 2006, an

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

370 expedition applied extensive sampling efforts to collect *Crenicichla* sp. in the Jequitinhonha,
371 without any success, and this species was only documented in 2009 by an environmental report
372 based on traditional sampling and morphological identification (Kullander & Lucena, 2006;
373 Intertechne, 2009). An issue reported worldwide, is that even when monitoring programmes
374 are conducted, most of the data obtained are often not published or made available and thus
375 remain inaccessible to further scientific studies (Lindenmayer & Likens, 2009; Revenga et al.,
376 2005). Here, eDNA metabarcoding data revealed that this species might be present at several
377 locations in the Jequitinhonha, indicating a possible large geographical distribution.

378 Taxonomic issues are often present in monitoring programs and the risk of
379 misidentification exists, regardless of the method applied (i.e. traditional sampling,
380 morphological identification, eDNA; Radinger et al., 2018; Jerde, 2019). Erroneous
381 identifications might also be present in the reference databases, especially in highly biodiverse
382 regions such as the Neotropics, where the amount of unknown and undescribed taxa and the
383 occurrence of cryptic species represent substantial issues. As demonstrated in previous studies,
384 identification of some species might be problematic when using eDNA metabarcoding based
385 on the 12S fragment employed here, due to its lack of taxonomic resolution and the
386 incompleteness of the reference databases (Yu et al., 2012; Eiler et al., 2013). Because a gene
387 tree is not necessarily related to a species tree, the phylogenetic resolution it provides can by
388 obscured for groups of taxa. The imperfect taxonomic resolution might allow the multiple
389 assignment of congeneric species (i.e. one species being concomitantly assigned to its multiple
390 congeners) when several reference sequences are available (please see example of *Prochilodus*
391 sp. below). In contrast, when the reference database is not complete for all species occurring
392 in the area, several MOTUs belonging to distinct species might be assigned to and errouneously
393 identified as the single closely related species available in the database (Sales et al., 2020) For
394 instance, most MOTUs belonging to *Prochilodus* sp. could not be assigned to species level due

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

395 to a high similarity among orthologous sequences from congeneric species. This poses a
396 conservation issue, since *Prochilodus argenteus* is an invasive species in the Jequitinhonha,
397 and is believed to have recently diverged from the endemic species *P. hartii* (Melo et al., 2018).
398 Henceforth, due to the conservative criteria applied to analyse the data, the number of species
399 detected is surely underestimated.

400 Six anostomids are described for the Jequitinhonha, and here we identified one of these
401 species (*Megaleporinus garmanii*), but also identified two species not previously reported
402 (*Leporinus copelandii* and *Hypomasticus mormyrops*). The only previous record of *Leporinus*
403 *copelandii* was deemed as an historical error (Andrade-Neto, 2010). Cilleros et al. (2019),
404 despite using a different 12S fragment, also reported the limitations in the taxonomic
405 assignment of species belonging to the genus *Leporinus*, therefore our data set is unable to
406 clarify the nuances within this group.

407

408 **4.2 Anthropogenic impacts and species richness**

409 Ecological communities vary in time and space, and the monitoring of these dynamics
410 is essential for conservation purposes (Bálint et al., 2018). In the Jequitinhonha River basin,
411 significant spatial and temporal fluctuations in fish assemblages inferred from eDNA were
412 detected. The longitudinal distance and presence of barriers did not explain community
413 variation ($p>0.05$); however, anthropogenic impacts might still have an influence on fish
414 diversity distribution in this river basin. Regarding data recovered from water samples, low
415 species richness were recovered from the reservoirs (3 – José Gonçalves, 9 – Salto da Divisa)
416 and the first sites located downstream the dams (5 – Coronel Murta and 10 – Itapebi). The
417 presence of dams is a well known fish diversity reduction factor since these barriers greatly
418 impact the environment (i.e. modification of physical and ecological characteristics of the

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

419 habitats, such as modifications in water flow, nutrient dynamics, water quality and temperature;
420 Pelicice & Agostinho, 2007; Pompeu et al., 2012). However, changes in fish distribution and
421 communities composition may also arise from plenty of distinct alterations and complex
422 interactions in the impounded environment (Agostinho, Pelicice & Gomes, 2008). Therefore,
423 despite no significant correlation between dams and fish diversity was herein found, the use of
424 eDNA metabarcoding offers a promising tool for evaluating the impoudment's impact on fish
425 distribution and thus, should be further investigated.

426 The sites comprising the highest fish diversity in this basin were represented by
427 locations characterized by different anthropogenic influences. The most upstream site
428 (Mendanha) is located in a less populated and more pristine region (Table S7, Supplementary
429 Material), near two areas of natural preservation (State Parks Biribiri and Rio Preto). The other
430 two sampling sites (Almenara, 8, and Belmonte, 11) are located near more densely populated
431 cities and impacted areas (i.e. due to the deforestation and mining activities, siltation increases
432 towards the river mouth and represents one of the greatest impacts in the Jequitinhonha river -
433 IBGE, 1997). Almenara, is a particularly impacted area, and during the sampling had a low
434 water level and accumulation of sediments, which might have contributed to increase the eDNA
435 concentration and accumulation, increasing the species diversity, despite the low
436 environmental quality.

437 The putative effect of anthropogenic activities on fish eDNA recovery herein described
438 corroborates the well-known impacts of human actions (e.g. construction of dams, species
439 introduction, pollution) leading to biotic homogenization (Agostinho et al., 2008, Ribeiro et
440 al., 2017).

441

442 **4.3 Seasonal changes in fish assemblages**

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

443 Seasonal changes driven by natural factors (e.g. water flow, rainfall) could also
444 contribute to explain assemblage variation even over a short time frame (i.e. weeks) as mobile
445 species, such as fish, can rapidly disperse and vary their distribution in response to changing
446 abiotic conditions (Arrington & Winemiller, 2006; Fitzgerald et al., 2017).

447 Water availability shows a great temporal variability in semi-arid and arid regions, with
448 short, but intense, rainfall episodes followed by long dry periods (Leite et al., 2010). The
449 Jequitinhonha river basin is inserted in a semi-arid region and in the first sampling campaign
450 it was facing a severe drought. Before the second sampling campaign, an increase in the
451 average accumulated rainfall (from 2.1-50mm in the first sampling event to 100-250 mm in the
452 second sampling event; CPTEC/INPE, 2018) might have contributed to a higher evenness in
453 MOTU richness/fish diversity amongst sample sites (regarding the contemporary species
454 richness inferred through water samples). The climatic and hydrological changes followed by
455 the onset of the rainy season usually triggers the start of fish migration in the semi-arid regions
456 (Chellappa et al., 2003; Chellappa et al., 2009). An increased water volume and subsequently
457 higher connectivity of aquatic habitats might stimulate the dispersal and result in reduced
458 densities of organisms (Fitzgerald et al., 2017). Therefore, the result here presented might
459 suggest that freshwater fish assemblages in tropical habitats may vary significantly between
460 dry and wet seasons. Besides the apparent homogenization found after the rainfall event, an
461 important factor to take into consideration is the reduction of diversity recovered in the second
462 campaign when compared to the first. The ecology of DNA might play an important role
463 regarding this matter, as eDNA molecules could be more diluted in the water column
464 decreasing the detectability of some species (e.g. rare or less abundant species).

465

466 **4.4 eDNA transport and species richness**

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

467 Another factor we need to take into account is eDNA transport from locations upstream
468 from our sample sites. This transport could lead to an overestimation of species richness
469 recovered for each sample site, and, the species identification per site therefore does not mean
470 that the species themselves are present there at the time of collection (Barnes & Turner, 2016;
471 Deiner et al., 2014). Still, eDNA transport distances may vary between river systems due to
472 abiotic and biotic factors (e.g. temperature, pH, bacterial load, or seasonal changes such as
473 drought or intense rainfall periods; Deiner et al., 2016). Most of the studies evaluating the effect
474 of eDNA upstream transportation reported travel distances of few kilometers, whereas, a travel
475 distance higher than 100km was demonstrated by Pont et al. (2018) for a high discharge (m^3/s)
476 river system. Still, despite the eDNA downstream transportation, the latter study demonstrated
477 the capability of eDNA in providing an accurate snapshot of fish assemblage composition in a
478 large river and finally, suggested that a distance of around 70 km would be enough to limit the
479 potential noise of eDNA transport. Therefore, despite having a high discharge rate (average of
480 409 m^3/s), the approximate distance between sites was 100 km and thus, the influence of eDNA
481 transport on species detected at each site might not be considered as a great concern here.
482 However, as no study has been conducted in Brazilian lotic environments focusing on
483 understanding eDNA transport and diffusion, it is difficult to draw sound conclusions regarding
484 this matter and so, additional studies focusing on the information recovered from eDNA in
485 large neotropical rivers might contribute to expand the knowledge of its complex
486 spatiotemporal dynamics.

487 The high alpha diversity values found for the site located at the river mouth (site 11,
488 Belmonte) deserves some consideration since this region has marine influence (including the
489 detection of one marine family, Engraulidae, by sediment samples in this sample site, Figure
490 4) and its abiotic characteristics (e.g. increased salinity) would be expected to restrict the
491 occurrence of some freshwater species. A hypothesis that could explain the detection of species

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

492 not expected to occur in this area includes eDNA transport and accumulation. Species shed
493 DNA constantly, which can be available in the water column or bound to superficial sediment.
494 A higher concentration and longer persistence of fish eDNA in the sediments might contribute
495 to eDNA molecule resuspension which might affect inferences from aqueous DNA in both
496 spatial and temporal scales (Turner et al., 2015; Graf & Rosenberg, 1997; Bloesch, 1995;).

497 Due to the fragmentation of the Jequitinhonha River, this site (site 11, Belmonte) is
498 located in a region characterized by a high level of sediment trapping
499 (freeflowingriver.org/maptool/) and possibly, this segment can act as an "eDNA reservoir" due
500 to the accumulation of molecules transported throughout the river. In addition to that, an
501 increase in water flow and tidal movements can also cause eDNA particle resuspension
502 (increasing the probability of retrieving old eDNA from the sediment beds – Jamieson et al.,
503 2005), which, associated with the resistance applied by the incursion of the marine waters into
504 the river, can contribute to retain and resuspend the eDNA accumulated in this area, making it
505 available in the water column. Considering this, river mouths should then be further
506 investigated as putative eDNA reservoirs since it could contribute in future sampling strategies
507 focusing on obtaining a snapshot of the entire fish community at a large scale.

508 Bioinformatics and technical aspects also play an important role in diversity recovery
509 from eDNA samples, and the existing trade-off between uncertainty and stringency may be
510 carefully considered when interpreting eDNA results as it might lead to false negative or false
511 positive detections (Evans et al., 2017; Grey et al., 2018). Regarding the analysed datasets, the
512 filtered data is considered as a subset of the total diversity recovered and showed a lower
513 diversity at the order and family levels. However, the significant positive correlation between
514 datasets demonstrated that beta-diversity is not influenced by the filtering criteria applied as
515 much as the effect of sampling medium or sampling time. As suggested by Li et al. (2018), the
516 filtered dataset provided a more conservative overview of fish diversity, compared to the

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

517 unfiltered dataset and thus did not detect several families and orders known to be present in
518 this catchment.

519 Fish diversity depicted by the heat trees based on the unfiltered data shows that a hidden
520 diversity might be present, especially for the Order Characiformes, as many families appear to
521 comprise several MOTUs (e.g. Anostomidae, Prochilodontidae). This likely reflects the
522 presence of multiple genera/species such as in the Anostomidae, known to harbour at least
523 seven species in this basin, which are absent from the reference sequence databases. Therefore,
524 to avoid underestimating the biodiversity, and reduce ambiguity in eDNA-based species
525 detection, we stress the importance of coordinating morphological surveys alongside DNA
526 assessments. Most importantly, there is also a need of increasing efforts towards building more
527 complete genetic reference databases, ideally composed of whole mitochondrial genomes, as
528 the lack of reference sequences has been considered as a great hindrance to fullfill the potential
529 of eDNA metabarcoding in assessing biodiversity rich ecosystems (Cilleros et al., 2019; Sales
530 et al., 2020).

531 Given the unprecedented rates of population and species decline and the increasing
532 anthropogenic impacts on freshwater communities, the importance of a rapid, robust and
533 efficient monitoring program has never been more in need for this ecosystem. Here we
534 illustrated eDNA ecology when analysing an entire river basin from the headwater to the river
535 mouth, and highlighted some of the challenges of applying eDNA metabarcoding in spatio-
536 temporal ecological studies, including recommendations for future work. Understanding eDNA
537 metabarcoding dynamics is an important step to make it a complementary monitoring tool to
538 traditional methods. This enhancement can improve the applicability of eDNA metabarcoding
539 for biomonitoring purposes in Brazilian freshwaters and therefore, allow the detection of
540 elusive, rare or patchily distributed species and provide data for neglected and difficult to
541 access localities.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

542

543 **ACKNOWLEDGEMENTS**

544 This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico
545 (CNPq) and Science without Borders Program (Grant #204620/2014-7). NGS also thanks
546 FCT/MCTES for the financial support to CESAM (UID/AMB/50017/2019), through national
547 funds. DCC is grateful to CNPq for the productivity fellowship (CNPq 306155/2018-4). We
548 are grateful to Letícia Sales for the invaluable fieldwork assistance and Gilberto Nepomuceno
549 Salvador for assistance with the map.

550

551

552

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

553 **REFERENCES**

554 Agostinho, A. A., Pelicice, F. M., & Gomes, L. C. (2008). Dams and the fish fauna of the
555 Neotropical region: impacts and management related to diversity and fisheries. *Brazilian*
556 *Journal of Biology*, 68(4, Suppl.), 1119-1132. doi: 10.1590/S1519-69842008000500019

557 Alves, C. B. M., Vieira, F., Magalhães, A. L. B. & Brito, M. F. G. 2007b. Impacts of non-native
558 fish species in Minas Gerais, Brazil: Present situation and prospects. Pp. 291-314. In: Bert, T.
559 M. (Ed). Ecological and genetic implications of aquaculture activities. Florida, Springer, 548p.

560 Andrade-Neto, F. R. (2010). Estado atual do conhecimento sobre a fauna de peixes da bacia do
561 Jequitinhonha. MG Biota, 2, 23–35.

562 Arrington, D. A., and K. O. Winemiller. (2006). Habitat affinity, the seasonal flood pulse, and
563 community assembly in the littoral zone of a Neotropical floodplain river. *Journal of the North*
564 *American Benthological Society*, 25, 126–141.

565 Bálint, M., Pfenninger, M., Grossart, H. P., Taberlet, P., Vellend, M., Leibold, M. A., … Jansen,
566 M. (2018). Environmental DNA time series in ecology. *Trends in Ecology and Evolution*,
567 33(12), 945–957. doi: 10.1016/j.tree.2018.09.003

568 Barnes, M. A., Turner, C. R., Jerde, C. L., Renshaw, M. A., Chadderton, W. L., & Lodge, D.
569 M. (2014). Environmental conditions influence eDNA persistence in aquatic systems.
570 *Environmental Science and Technology*, 48(3), 1819–1827. doi: 10.1021/es404734p

571 Becker, R. A., Sales, N. G., Santos, G. M., Santos, G. B.,& Carvalho, D. C. (2015). DNA
572 barcoding and morphological identification of neotropical ichthyoplankton from the Upper
573 Paraná and São Francisco. *Journal of Fish Biology*, 87, 159–168. doi: 10.1111/jfb.12707

574 Bloesch, J. (1995). Mechanisms, measurement and importance of sediment resuspension in
575 lakes. *Marine and Freshwater Research*, 46, 295–304.

576 Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac., E. (2016). obitools: A
577 unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, 16(1),
578 176–182. doi: 10.1111/1755-0998

579 Cassemiro, F. A. S., Bailly, D., da Graça, W. J., & Agostinho, A. A. (2018). The invasive
580 potential of tilapias (Osteichthyes, Cichlidae) in the Americas. *Hydrobiologia*, 817(1), 133–
581 154. doi: 10.1007/s10750-017-3471-1

582 Cilleros, K., Valentini, A., Allard, L., Dejean, T., Etienne, R., Grenouillet, G., … Brosse, S.
583 (2019). Unlocking biodiversity and conservation studies in high diversity environments using
584 environmental DNA (eDNA): a test with Guianese freshwater fishes. *Molecular Ecology*
585 *Resources*, 7, 1–20. doi: 10.1111/1755-0998.12900

586 Chellappa, S., Câmara, M. R., Chellappa, N. T., Beveridge, M. C. M., & Huntingford, F. A..
587 (2003). Reproductive ecology of a neotropical cichlid fish, *Cichla monoculus* (Osteichthyes:
588 Cichlidae). *Brazilian Journal of Biology*, 63(1), 17-26. doi: 10.1590/S1519-
589 69842003000100004

590 Chellappa, S., Bueno, R. M. X., Chellappa, T., Chellappa, N. T., Almeida e Val, V. M. F. (2009).
591 Reproductive seasonality of the fish fauna and limnoecology of semi-arid Brazilian reservoirs.
592 *Limnologica*, 39, 4, 325-329.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

593 CPTEC (Centro de Previsão de Tempo e Estudos Climáticos), INPE (Instituto Nacional de
594 Pesquisas Espaciais). (2018). Available at: http://clima1.cptec.inpe.br/monitoramentobrasil/pt

595 Deiner, K., & Altermatt, F. (2014). Transport distance of invertebrate environmental DNA in a
596 natural river. PLoS ONE, 9(2). doi: 10.1371/journal.pone.0088786

597 Deiner, K., Fronhofer, E. A., Mächler, E., Walser, J. C., & Altermatt, F. (2016). Environmental
598 DNA reveals that rivers are conveyer belts of biodiversity information. *Nature*
599 *Communications*, 7, 1-8. doi:10.1038/ncomms12544.

600 Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., …
601 Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey
602 animal and plant communities. *Molecular Ecology*, 26(21), 5872–5895.
603 doi:10.1111/mec.14350

604 Doi, H., Inui, R., Akamatsu, Y., Kanno, K., Yamanaka, H., Takahara, T. & Minamoto, T. (2017).
605 Environmental DNA analysis for estimating the abundance and biomass of stream fish.
606 *Freshwater Biology*, 62(1), 30-39. doi:10.1111/fwb.12846

607 Dudgeon, D., Arthington, A. H., Gessner, M. O., Kawabata, Z. I., Knowler, D. J., Leveque, C.,
608 … Sullivan, C. A. (2006). Freshwater biodiversity: Importance, threats, status and conservation
609 challenges. *Biological Reviews*, 81, 163–182. doi:10.1017/S1464793105006950

610 Dutra, G., Penido, I., Mello, G., & Pessali, T. C. (2016). Two new species of Cyphocharax
611 (Teleostei: Characiformes: Curimatidae) from headwaters of the Jequitinhonha and São
612 Francisco river basins, Minas Gerais, Brazil. *Zootaxa*, 4103(2), 154-158. doi:
613 10.11646/zootaxa.4103.2.5

614 Eiler, A., Drakare, S., Bertilsson, S., Pernthaler, J., Peura, S., Rofner, C., ... Lindström, E. S.
615 (2013). Unveiling distribution patterns of freshwater phytoplankton by a next generation
616 sequencing based approach. *PLoS ONE*, 8. doi: 10.1371/journal.pone.0053516

617 Evans, N. T., Li, Y., Renshaw, M. A., Olds, B. P., Deiner, K., Turner, C. R., … Pfrender, M. E.
618 (2017). Fish community assessment with eDNA metabarcoding: effects of sampling design and
619 bioinformatic filtering. *Canadian Journal of Fisheries and Aquatic Sciences*, 1374(January),
620 1–13. doi: 10.1139/cjfas-2016-0306

621 FAO. (2012). Cold water fisheries in the trans-Himalayan countries. Available in:
622 ftp://ftp.fao.org/docrep/fao/005/y3994e/y3994e00.pdf

623 Fitzgerald, D. B., Winemiller, K. O., Sabaj Pérez, M. H., & Sousa, L. M. (2017). Seasonal
624 changes in the assembly mechanisms structuring tropical fish communities. *Ecology*, 98(1),
625 21–31. doi: 10.1002/ecy.1616

626 Foster, Z., Sharpton, T., & Grünwald, N. (2017). Metacoder: An R package for visualization
627 and manipulation of community taxonomic diversity data. *PLoS Computational Biology*, 13(2),
628 1-15. doi: 10.1371/journal.pcbi.1005404

629 Godinho, H. P., Godinho, A. L., & Vono, V. (1999). Peixes da bacia do rio Jequitinhonha. In:
630 Lowe Mcconnel (ed) Estudos ecológicos de comunidades de peixes tropicais. EDUSP, São
631 Paulo, 414-423.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

632 Gomes, L. C., Pessali, T. C., Sales, N. G., Pompeu, P. S., & Carvalho, D. C. (2015). Integrative
633 taxonomy detects cryptic and overlooked fish species in a neotropical river basin. *Genetica*,
634 143(581). doi: 10.1007/s10709-015-9856-z

635 Graf, G. & Rosenberg, R. (1997). Bioresuspension and biodeposition: a review. *Journal of*
636 *Marine Systems*, 3-4, 269-278.

637 Grey, E. K., Bernatchez, L., Cassey, P., Deiner, K., Deveney, M., Howland, K. L., … Lodge,
638 D. M. (2018). Effects of sampling effort on biodiversity patterns estimated from environmental
639 DNA metabarcoding surveys. *Scientific Reports*, 8(1), 2–11. doi: 10.1038/s41598-018-27048-
640 2

641 Grill, G., Lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., … Zarfl, C. (2019).
642 Mapping the world's free-flowing rivers. *Nature*, 569(7755), 215–221. doi: 10.1038/s41586-
643 019-1111-9

644 Handley, L. L., Read, D. S., Winfield, I. J., Kimbell, H., Li, J., Hahn, C., … Hänfling, B. (2019).
645 Temporal and spatial variation in distribution of fish environmental DNA in England's largest
646 lake. *Environmental DNA,* 1, 26-39.

647 Holman, L. E., de Bruyn, M., Creer, S., Carvalho, G., Robidart, J., & Rius, M. (2018). The
648 detection of introduced and resident marine species using environmental DNA metabarcoding
649 of seawater and sediment. *Scientific Reports*, 9, 11559.

650 Hulsen, T., de Vlieg, J., & Alkema, W. (2008). BioVenn—A web application for the comparison
651 and visualization of biological lists using area- proportional Venn diagrams. *BMC Genomics*,
652 9(1),), 488. doi: 10.1186/1471-2164-9-488

653 IBGE (Fundação Instituto Brasileiro de Geografia e Estatística). (1997). Diagnóstico ambiental
654 da Bacia do Rio Jequitinhonha. Ronaldo do Nascimento Gonçalves, Ministério do
655 Planejamento e Orçamento, Salvador. Available at:
656 https://biblioteca.ibge.gov.br/visualizacao/livros/liv95902.pdf

657 Intertechne. (2010). Inventário hidrelétrico dos rios Jequitinhonha e Araçuaí – Relatório final
658 dos estudos de inventário. Apêndice D, estudos ambientais, 259 p.

659 IPBES (2019): Global assessment report on biodiversity and ecosystem services of the
660 Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. E. S.
661 Brondizio, J. Settele, S. Díaz, and H. T. Ngo (editors). IPBES secretariat, Bonn, Germany.

662 Jamieson, R. C., Joy, D. M., Lee, H., Kostaschuk, R., & Gordon, R. J. (2005). Resuspension of
663 sediment-associated *Escherichia coli* in a natural stream. *Journal of Environmental Quality*.
664 34, 581–589.

665 Jerde, C.L. (2019). Can we manage fisheries with the inherent uncertainty from eDNA?.
666 *Journal Fish Biology*. doi:10.1111/jfb.14218

667 Jerep, F. C., Camelier, P., & Malabarba, L. R. (2016). *Serrapinnus zanatae*, a new species from
668 the rio Jequitinhonha basin, Minas Gerais State, Brazil (Teleostei: Characidae:
669 Cheirodontinae). *Ichthyological Exploration of Freshwaters*, 26(4), 289-298.

670 Jucker, T., Wintle, B., Shackelford, G., Bocquillon, P., Geffert, J. L., Kasoar, T., … Mukherjee,
671 N. (2018). Ten-year assessment of the 100 priority questions for global biodiversity
672 conservation. *Conservation Biology*, 32, 1457-1463. doi: 10.1111/cobi.13159

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

673 Kelly, R. P., Port, J. A., Yamahara, K. M., & Crowder, L. B. (2014). Using environmental DNA
674 to census marine fishes in a large mesocosm. *PLoS ONE*, 9(1), e86175. doi:
675 10.1371/journal.pone.0086175

676 Kullander, S. O., & Lucena, C. A. S. (2006). A review of the species of Crenicichla (Teleostei:
677 Cichlidae) from the Atlantic coastal rivers of southeastern Brazil from Bahia to Rio Grande do
678 Sul States, with descriptions of three new species. *Neotropical Ichthyology*, 4(2), 127-146. doi:
679 10.1590/S1679-62252006000200001

680 Leite, M. G. P., & Fujaco, M. A. G. (2010). A long-term annual water balance analysis of the
681 Araçuaí River Basin, Brazil. *Journal of Geographical Sciences*, 20(6), 938–946. doi:
682 10.1007/s11442-010-0822-5

683 Li, Y., Evans, N. T., Renshaw, M. A., Jerde, C. L., Olds, B. P., Shogren, A. J., … Pfrender, M.
684 E. (2018). Estimating fish alpha and beta-diversity along a small stream with environmental
685 DNA metabarcoding. *Metabarcoding and Metagenomics*, 2, e24262. doi:
686 10.3897/mbmg.2.24262

687 Lindenmayer, D.B., & Likens, G.E. (2009). Adaptive monitoring: a new paradigm for long-
688 term research and monitoring. *Trends in Ecology and Evolution*, 24, 482–486. doi:
689 10.1016/j.tree.2009.03.005

690 Mahé, F., Rognes, T., Quince, C., de Vargas, C., & Dunthorn, M. (2014). Swarm: robust and
691 fast clustering method for amplicon-based studies. *PeerJ*, 2, e593. doi: 10.7717/peerj.593

692 McDevitt, A. D., Sales, N. G., Browett, S., Sparnenn, A., Mariani, S., Wangensteen, O. S., …
693 Benvenuto, C. (2019). Environmental DNA metabarcoding as an effective and rapid tool for
694 fish monitoring in canals. *Journal Fish Biology*, 95, 179-182.

695 McIntyre, P. B., Liermann, C. A. R., & Revenga, C. (2016). Linking freshwater fishery
696 management to global food security and biodiversity conservation. *Proceedings of the National*
697 *Academy of Sciences*, 113(45), 12880-12885.

698 Melo, B. F., Dorini, B. F., Foresti, F., & Oliveira, C. (2018). Little divergence among
699 mitochondrial lineages of Prochilodus (Teleostei, Characiformes). *Frontiers in Genetics*, 9, 1-
700 11. doi: 10.3389/fgene.2018.00107

701 Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., Iwasaki, W. (2015). MiFish,
702 a set of universal PCR primers for metabarcoding environmental DNA from fishes. *Royal*
703 *Society Open Science*, 2, 150088. doi: 10.1098/rsos.150088

704 Nelson, J. S. (2006). Fishes of the World, 4th edn. John Wiley and Sons, Hoboken, New Jersey,
705 601 p.

706 Nielsen, D. T. B., Pessali, T. C., & Dutra, G. (2017). A new annual fish of the genus
707 Simpsonichthys (Cyprinodontiformes: Cynolebiidae) from the upper Rio Jequitinhonha basin,
708 Brazil. *Zootaxa,* 4263(1), 165-172. doi: 10.11646/zootaxa.4263.1.8

709 Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., Ohara, R., … Wagner, R.
710 (2013). Package "vegan." Community ecology package.
711

712 Pawlowski, J., Kelly-Quinn, M., Altermatt, F., Apothéloz-Perret-Gentil, L., Beja, P., Boggero,
713 A., … Kahlert, M. (2018). The future of biotic indices in the ecogenomic era: Integrating

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

714 (e)DNA metabarcoding in biological assessment of aquatic ecosystems. *Science of the Total*
715 *Environment*, 637, 1295–1310. doi: 10.1016/j.scitotenv.2018.05.002

716 Pelicice, F. M. & Agostinho A. A. (2007). Fish-passage facilities as ecological traps in large
717 neotropical rivers. *Conservation Biology*, 22, 180-188. doi: 10.1111/j.1523-1739.2007.00849.x

718 Pompeu, P. S., Agostinho, A. A., & Pelicice, F. M. (2012). Existing and future challenges: the
719 concept of successful fish passage in South America. *River Research and Applications*, 28,
720 504-512. doi: 10.1002/rra.1557

721 Pont, D., Rocle, M., Valentini, A., Civade, R., Jean, P., Maire, A., … Dejean, T. (2018).
722 Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its
723 downstream transportation. *Scientific Reports*, 8(1), 10361. https://doi.org/10.1038/s41598-
724 018-28424-8

725 Radinger, J., Britton, J. R., Carlson, S. M., Magurran, A. E., Alcaraz-Hernández, J. D.,
726 Almodóvar, A., … García-Berthou, E. (2018). Effective monitoring of freshwater fish. *Zenodo.*
727 doi: 10.5281/zenodo.1480407

728 Rahel, F. J. (2000). Homogenization of fish faunas across the United States. *Science,*
729 288:854–856

730 Reis, R. E., Albert, J. S., Di Dario, F., Mincarone, M. M., Petry, P., ... Rocha, L. A. (2016). Fish
731 biodiversity and conservation in South America. *Journal of Fish Biology*, 89(1), 12-47. doi:
732 10.1111/jfb.13016

733 Revenga, C., Campbell, I., Abell, R., de Villiers, P., & Bryer, M. (2005). Prospects for
734 monitoring freshwater ecosystems towards the 2010 targets. *Philosophical Transactions of the*
735 *Royal Society B: Biological Sciences*, 360, 397–413. doi: 10.1098/rstb.2004.1595

736 Ribeiro, V. R., Gubiani, E., A., & Cunicu, A. M. (2017). Occurrence of non-native fish species
737 in neotropical river under the influence of aquaculture activities. *Boletim do Instituto de Pesca*,
738 44(1), 80-90. doi:10.20950/1678-2305.2018.288

739 Rosa, R. S., & Lima, F. C. T. (2008). Livro Vermelho da Fauna Brasileira Ameaçada de extinção
740 In: Machado ABM, Drummond GM, Paglia AP (eds). Os peixes brasileiros ameaçados de
741 extinção Ministério do Meio Ambiente, Brasília.

742 Sales, N. G., Pessali, T. C., Andrade-Neto, F., & Carvalho, D. C. (2018). Introgression from
743 non-native species unveils a hidden threat to the neotropical fish *Prochilodus hartii*. *Biological*
744 *Invasions*, 20(3), 555-566. doi: 10.1007/s10530-017-1556-4

745 Sales, N. G., Wangensteen, O. S., Carvalho, D. C., & Mariani, S. (2019). Influence of
746 preservation method, sampling medium and time on eDNA recovery in a Neotropical river.
747 *Environmental DNA,* 1, 119-130.

748 Sales, N. G., Kaizer, M. C., Coscia, I., Perkins, J. C, Highlands, A., Boubli, J. P., … McDevitt,
749 A. D. (2020). Assessing the potential of environmental DNA metabarcoding for monitoring
750 Neotropical mammals: a case study in the Amazon and Atlantic Forest, Brazil. *Mammal*
751 *Review*, doi: 10.1111/mam.12183.

752 Schmelzle, M. C., & Kinziger, A. P. (2016). Using occupancy modelling to compare
753 environmental DNA to traditional field methods for regional-scale monitoring of an

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

754 endangered aquatic species. *Molecular Ecology Resources*, 16(4), 895–908. doi: 10.1111/1755-
755 0998.12501

756 Schwartz, M. K., Luikart, G. & Waples, R.S. (2006). Genetic monitoring as a promising tool
757 for conservation and management. *Trends in Ecology and Evolution*, 22 (1), 25-33. doi:
758 10.1016/j.tree.2006.08.009

759 Shelton, A. O., Kelly, R. P., O'Donnell, J. L., Park, L., Schwenke, P., Greene, C., … Beamer,
760 E. M. (2019). Environmental DNA provides quantitative estimates of a threatened salmon
761 species. *Biological Conservation*, 237, 383–391. doi: 10.1016/j.biocon.2019.07.003

762 Socolar, J. B., Gilroy, J. J., Kunin, W. E., & Edwards, D. P. (2015). How should beta-diversity
763 inform biodiversity conservation? *Trends in Ecology and Evolution*, 1, 1–14. doi:
764 10.1016/j.tree.2015.11.005

765 Thomé, M. T. C., & Haddad, C. F. B. (2019). Brazil's biodiversity researchers need help.
766 *Science*, *364*(6446), 1144-1145.

767 Trimble, M. J., & R. J. van Aarde. (2012). Geographical and taxonomic biases in research on
768 biodiversity in human-modified landscapes. *Ecosphere*, 3(12), 119-129. doi: 10.1890/ES12-
769 00299.1

770 Turner, C. R., Uy, K. L., & Everhart, R. C. (2015). Fish environmental DNA is more
771 concentrated in aquatic sediments than surface water. Biological Conservation, 183, 93–102.
772 doi: 10.1016/j.biocon.2014.11.017

773 Ushio, M., Murakami, H., Masuda, R., Sado, T., Miya, M., Sakurai, S., Yamanaka, H.,
774 Minamoto, T., Kondoh, M. (2018). Quantitative monitoring of multispecies fish environmental
775 DNA using high-throughput sequencing. *Metabarcoding and Metagenomics*. 2, e23297.
776 doi:10.3897/mbmg.2.23297

777 Vitule, J. R. S., Bornatowski, H., Freire, C. A., & Abilhoa, V. (2014). Extralimital introductions
778 of *Salminus brasiliensis* (Cuvier, 1816) (Teleostei, Characidae) for sport fishing purposes: a
779 growing challenge for the conservation of biodiversity in neotropical aquatic ecosystems.
780 *BioInvasions Records*, 3(4), 291–296. doi: 10.3391/bir.2014.3.4.11

781 Vorosmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Glidden, S.
782 (2010). Global threats to human water security and river biodiversity. *Nature*, 467, 555-561.
783 doi: 10.0.4.14/nature09440

784 WWF. 2018. Living Planet Report – 2018: Aiming Higher. Grooten, M. and Almond, R. E. A.
785 (Eds.) WWF, Gland, Switzerland.

786 Yamanaka, H., Minamoto, T., Matsuura, J., Sakurai, S., Tsuji, S., Motozawa, H., ... Kondo, A.
787 (2017). A simple method for preserving environmen- tal DNA in water samples at ambient
788 temperature by addition of cat- ionic surfactant. Limnology, 18(2), 233–324. doi: 10.1007/
789 s10201-016-0508-5

790 Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C. & Ding, Z. (2012). Biodiversity
791 soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring.
792 *Methods in Ecology and Evolution*, 3, 613-623. doi: 10.1111/j.2041-210X.2012.00198.x

793

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

794

795 **DATA ACESSIBILITY**

796 Raw data will be made available on DRYAD upon acceptance.

797

798 **AUTHOR CONTRIBUTIONS**

799 NGS, OSW and SM designed the study. NGS carried out the fieldwork. NGS and OSW
800 performed the laboratory work and the bioinformatics. NGS analysed the data primarily, with
801 contributions from ADM, IC, KD and KP. All authors discussed the results and implications.
802 NGS drafted the manuscript, all authors provided manuscript input and contributed in
803 discussion that developed the study.

804

805

806

807

808

809

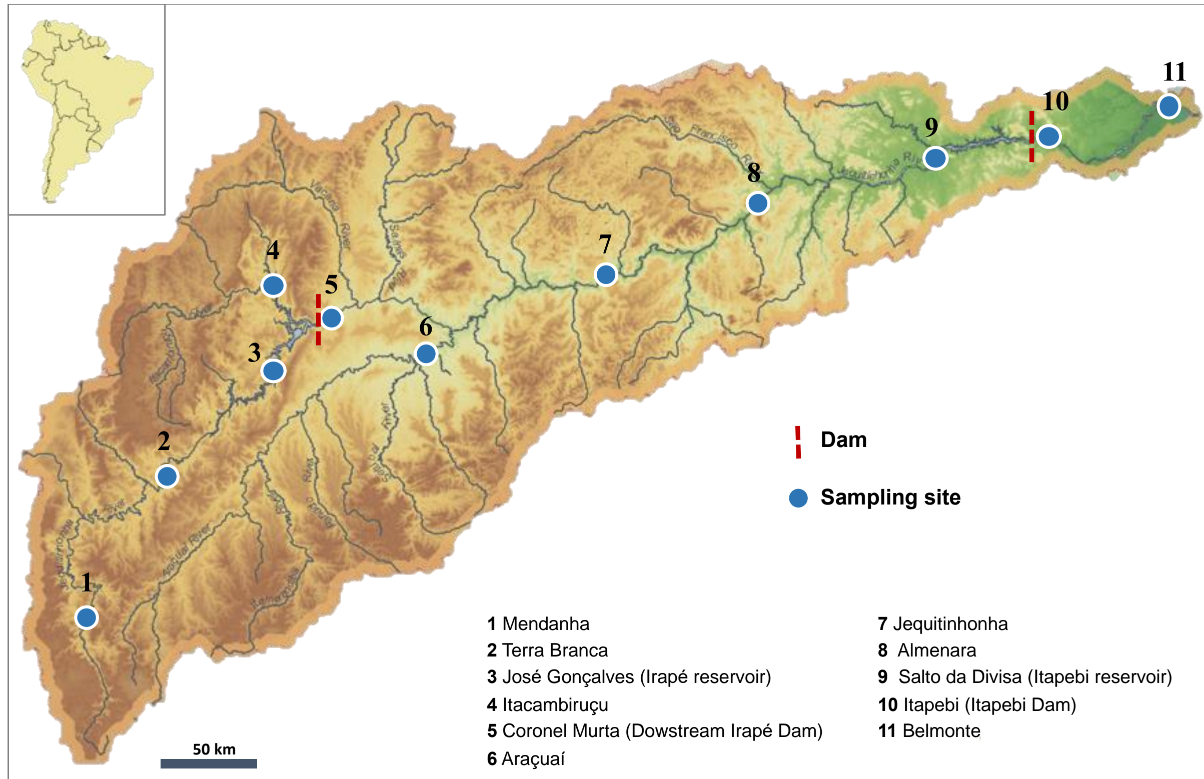
810

811

812

813

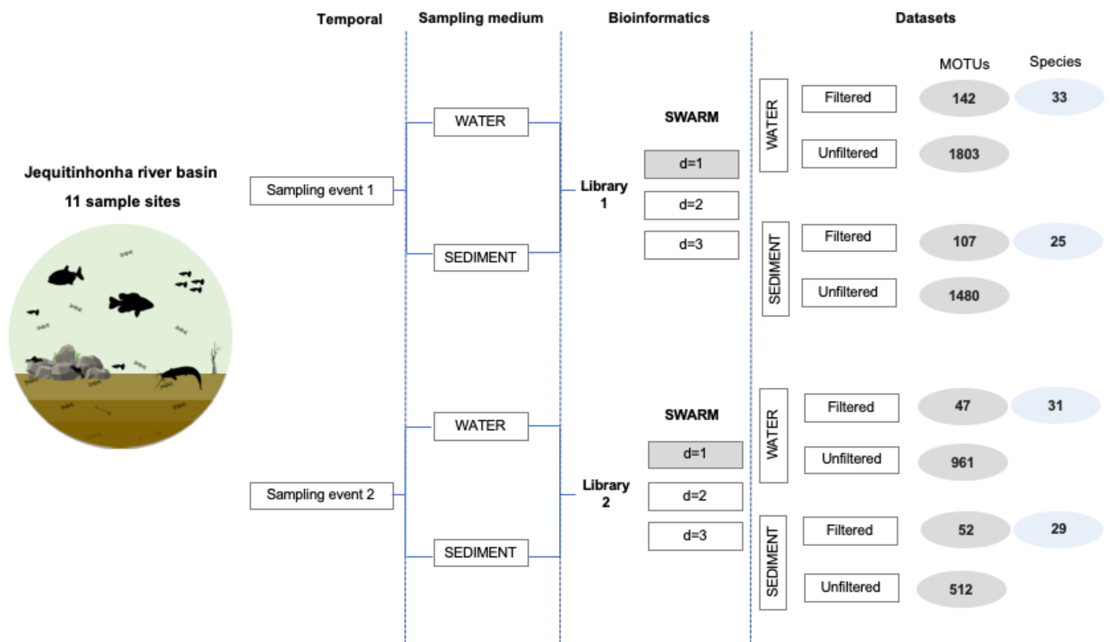
bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



814
815

816 FIGURE 1 | The Jequitinhonha river basin, including sampling sites used in the study, dams
817 and respective hydrological regions.

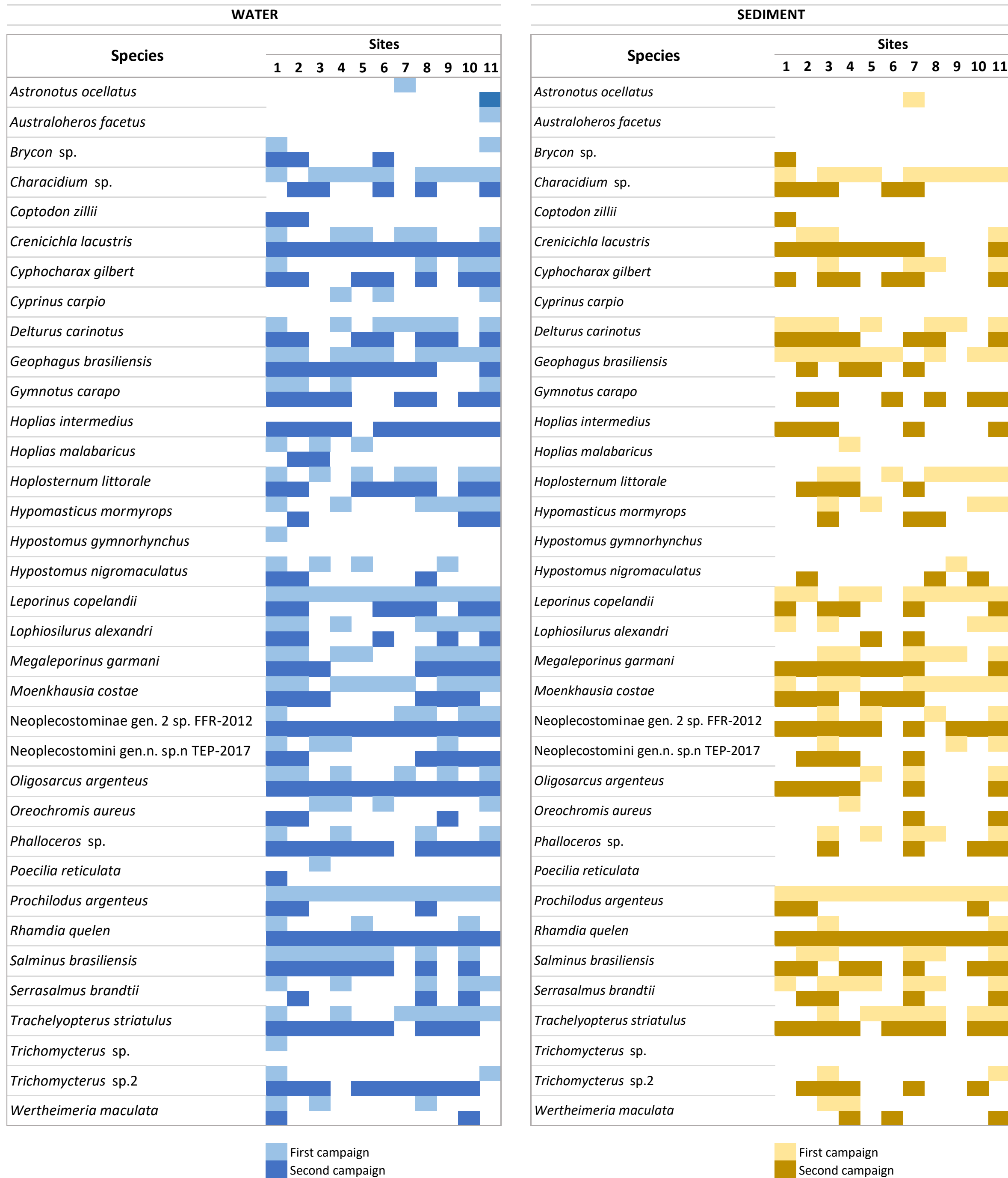
818



819

820 FIGURE 2 | Workflow illustrating the methods used in this paper and respective number of
821 MOTUs retrieved in each dataset analysed, and the final number of species assigned with >0.97
822 identity.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.



823

824 FIGURE 3 | Species distribution in the Jequitinhonha River Basin, according to sampling
825 media and campaign.

826

827

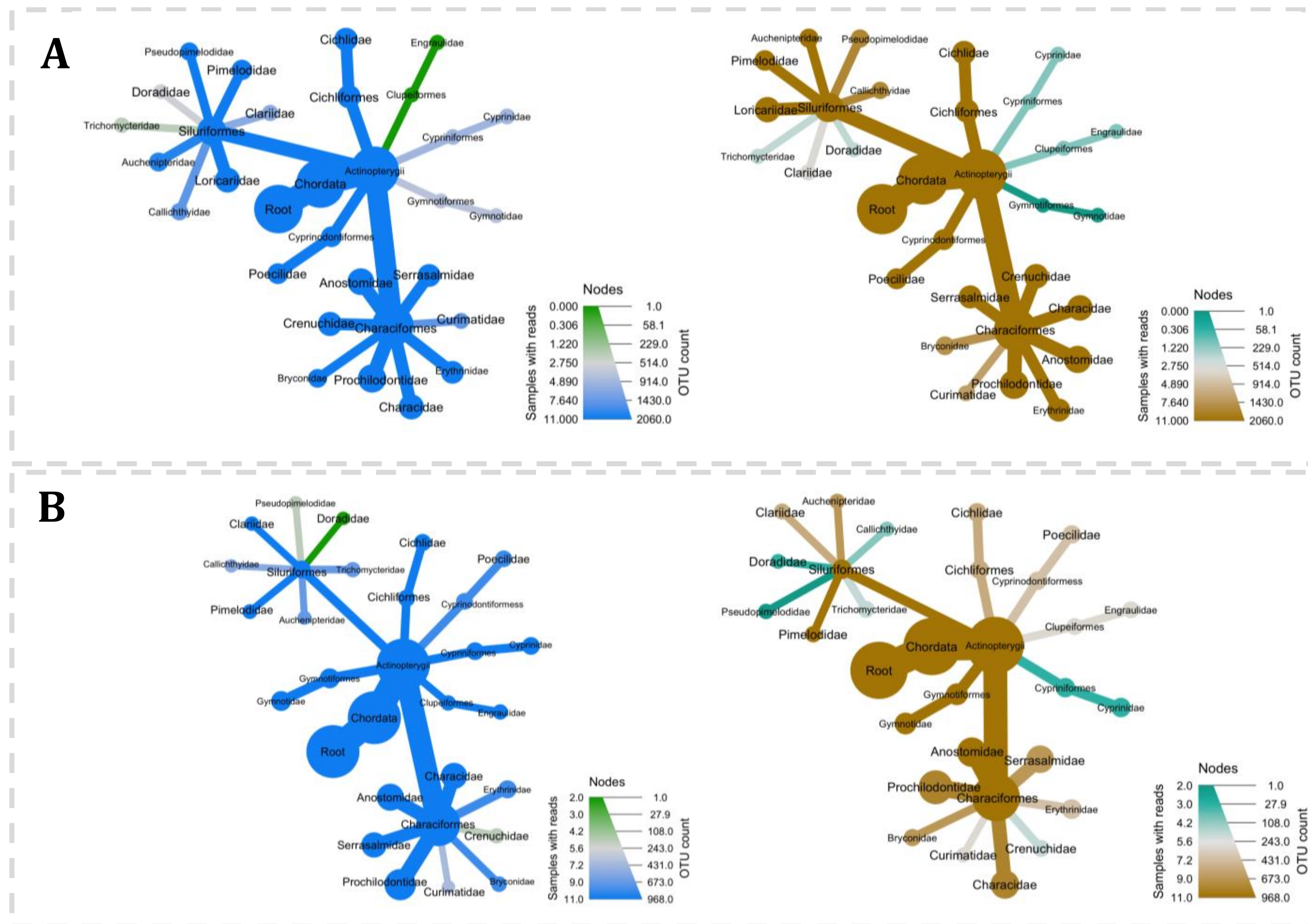


FIGURE 4 | Heat trees displaying the fish diversity recovered for Jequitinhonha River Basin using eDNA metabarcoding unfiltered datasets, during the first (A) and second (B) campaigns. Blue = Water samples; Brown = Sediment samples.

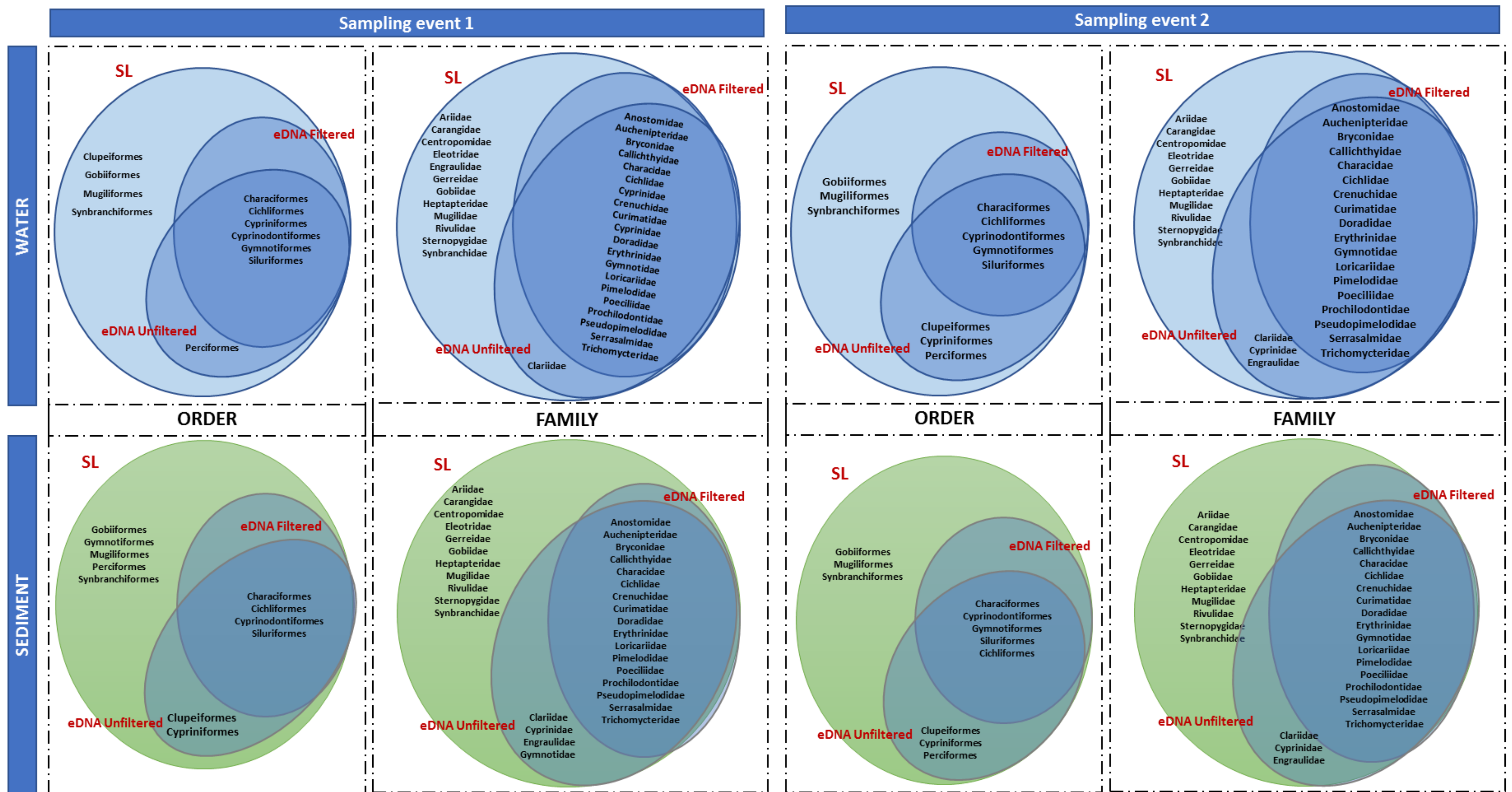


FIGURE 5 | Venn diagram of fish orders and families comparing the data included in the species list based on traditional sampling (SL) to eDNA detected in distinct sampling media (water vs sediment); sampling campaign; and datasets analysed (unfiltered vs filtered).

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

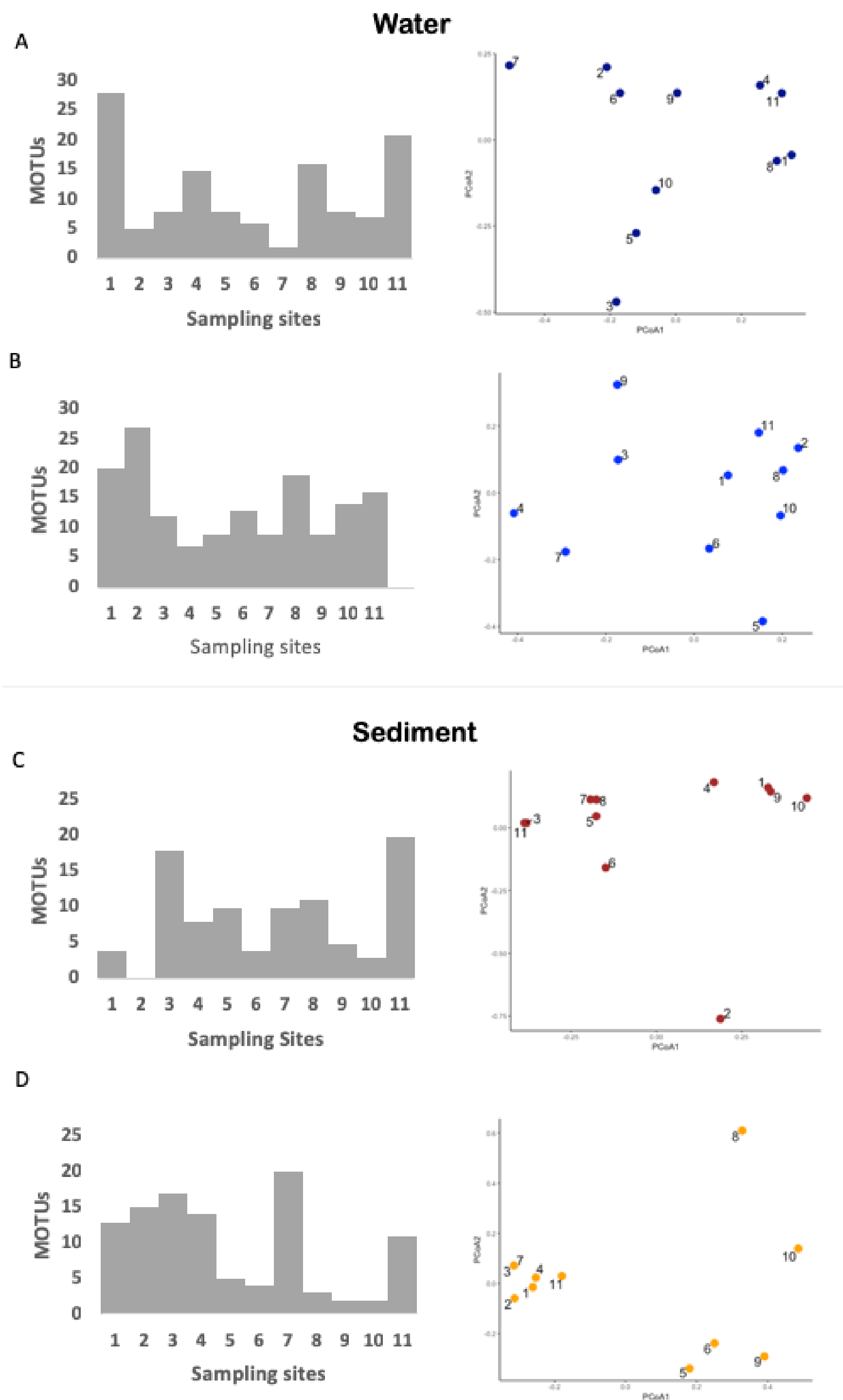


FIGURE 6 | Filtered dataset, showing the species richness distribution along the Jequitinhonha River Basin and Principal Coordinates Analysis (PCoA) of β-diversity of sampling locations (Jaccard distance). A) Water samples obtained in the first campaign; B) Water samples obtained in the second campaign; C) Sediment samples obtained in the first campaign; D) Sediment samples obtained in the second campaign.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

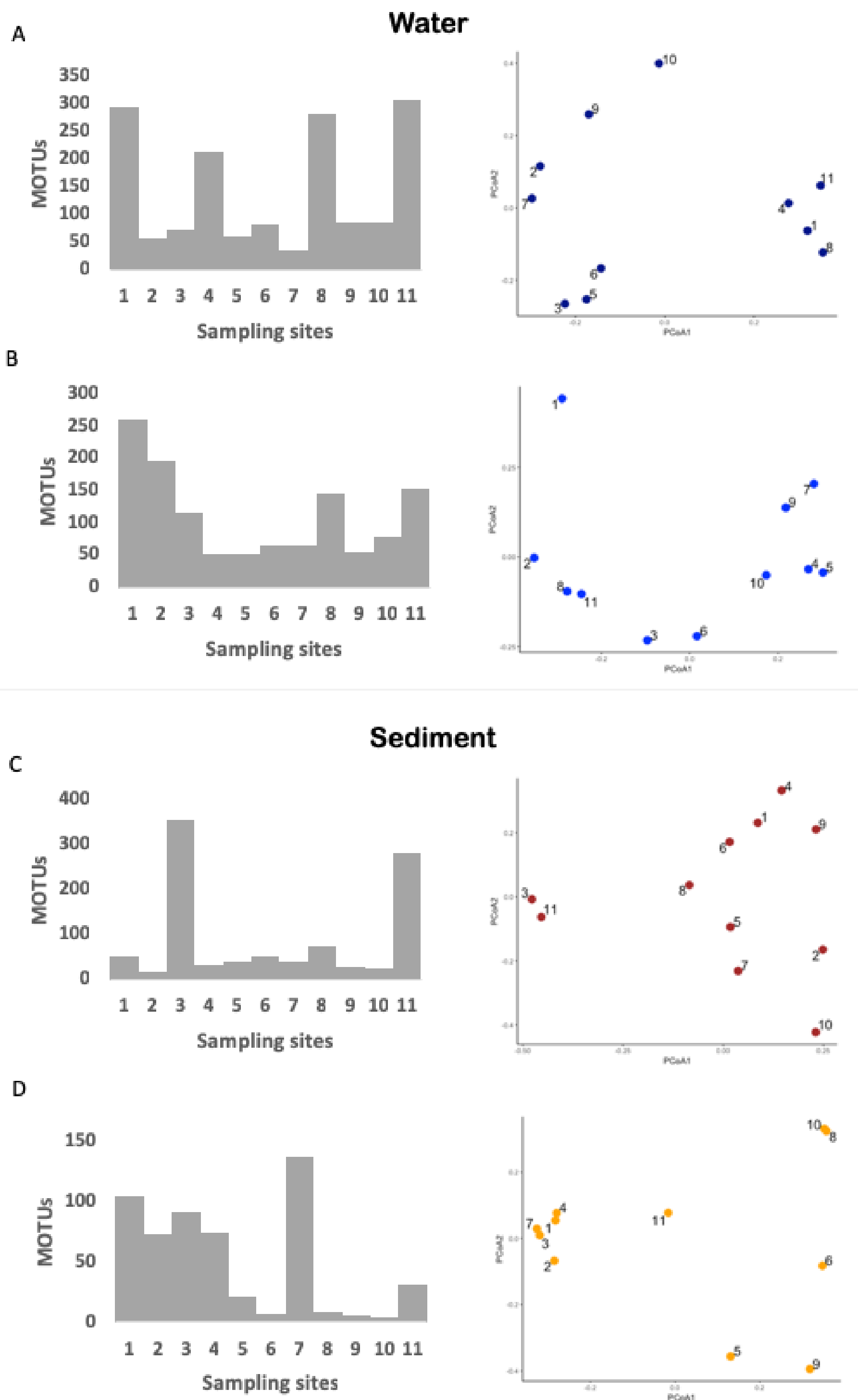


FIGURE 7 | Unfiltered dataset, showing the species richness distribution along the Jequitinhonha River Basin and Principal Coordinates Analysis (PCoA) of β-diversity of sampling locations (Jaccard distance). A) Water samples obtained in the first campaign; B) Water samples obtained in the second campaign; C) Sediment samples obtained in the first campaign; D) Sediment samples obtained in the second campaign.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.04.933366; this version posted February 4, 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.

TABLE 1 | Mantel *r* and *p*-values (in parentheses) for all the pairwise comparisons between datasets, sampling media, geographic distance and presence of barriers (dams).

			First campaign				**Second campaign**			
			Water		**Sediment**		**Water**		**Sediment**	
			Unfiltered	**Filtered**	**Unfiltered**	**Filtered**	**Unfiltered**	**Filtered**	**Unfiltered**	**Filtered**
1	W	**Unfiltered**	1							
		Filtered	0.689 (p=0.001)	1						
	S	**Unfiltered**	0.050 (p=0.359)	-0.268 (p=0.939)	1					
		Filtered	0.219 (p=0.162)	0.134 (p=0.250)	0.534 (p=0.005)	1				
2	W	**Unfiltered**	0.193 (p=0.445)	-0.142 (p=0.815)	0.110(p=0.221)	0.029 (p=0.386)	1			
		Filtered	0.011 (p=0.444)	-0.017 (p=0.491)	0.055(p=0.309)	-0.034 (p=0.555)	0.572 (p=0.001)	1		
	S	**Unfiltered**	-0.100 (p=0.656)	-0.235 (p=0.914)	0.017(p=0.389)	-0.047 (p=0.548)	-0.025 (p=0.544)	-0.174 (p=0.870)	1	
		Filtered	-0.121 (p=0.691)	-0.278 (p=0.929)	0.109(p=0.269)	-0.104 (p=0.645)	0.075 (p=0.309)	-0.040 (p=0.528)	0.822 (p=0.001)	1
		Longitudinal distance	-0.213 (p=0.897)	-0.258 (p=0.947)	-0.041(p=599)	-0.028 (p=0.561)	0.137 (p=0.154)	-0.043 (p=0.597)	0.189 (p=0.114)	0.290 (p=0.052)
		Presence of dam	-0.102 (p=0.690)	-0.172 (p=0.859)	0.028 (p=0.416)	-0.004 (p=0.514)	-0.018 (p=0.488)	-0.181 (0.876)	0.178 (p=0.161)	0.108 (p=0.26)