University of Salford
School of Environment & Life Sciences
Science Without Borders Programme

Barcoding and eDNA metabarcoding of a rapidly changing Neotropical freshwater fish community

This thesis is submitted to University of Salford, School of Environment & Life Sciences, in fulfilment of the requirements for the degree of Doctor of Philosophy (PhD).

Naiara Guimarães Sales

Supervisor:
Dr. Stefano Mariani
Co-Supervisors:
Dr. Allan D. McDevitt
Dr. Daniel Cardoso de Carvalho

March 2019

“Adote o ritmo da natureza. O segredo dela é a paciência.”

Ralph Waldo Emerson

Table of Contents

Abstract ……**10**

Chapter I - **INTRODUCTION AND BACKGROUND**……..**11**

1.1 Fish biodiversity assessment and decline in Neotropical freshwater systems……**12**

1.2 Brazilian fish biodiversity: Doce and Jequitinhonha river basins……**14**

1.3 DNA Barcoding as a tool for fish species identification……**18**

1.4 Metabarcoding using amplicon sequencing as an effective tool for biodiversity assessment……**20**

1.5 Important considerations when working with eDNA……**23**

Chapter II – **HIDDEN DIVERSITY HAMPERS CONSERVATION EFFORTS IN A HIGHLY IMPACTED NEOTROPICAL RIVER SYSTEM**……**48**

2.1 Abstract……**49**

2.2 Introduction……**50**

2.3 Material and Methods……**54**

2.4 Results……**59**

2.5 Discussion……**66**

2.6 References……**74**

2.7 Supplementary Material……**81**

Chapter III – **INFLUENCE OF PRESERVATION METHODS AND SAMPLING MEDIUM ON eDNA RECOVERY IN NEOTROPICAL LOTIC ENVIRONMENTS**……**94**

3.1 Abstract……**95**

3.2 Introduction……**96**

3.3 Material and Methods……**101**

3.4 Results……**108**

3.5 Discussion……**116**

3.6 References……**122**

3.7 Supplementary Material……**128**

Chapter IV – **SPATIO-TEMPORAL VARIATION IN NEOTROPICAL RIVERINE ICHTHYOFAUNA INFERRED BY eDNA ANALYSIS**……**134**

4.1 Abstract……**135**

4.2 Introduction……**136**

4.3 Material and Methods……**140**

4.4 Results……**150**

4.5 Discussion……**162**

4.6 References……**172**

4.7 Supplementary Material……**180**

Chapter V – **FINAL CONSIDERATIONS**……**191**

References ……**201**

APPENDIX

List of Tables

CHAPTER II

Table 1: Distance summary reports for sequence divergence between species, genus and family level including minimum, mean and maximum genetic distances (K2P)……**60**

Table 2: List of undescribed species including the nearest neighbor, BIN and genetic similarity (%)……**64**

Table 3: List of described species with high intraspecific divergence (>2%), showing the maximum and mean intraspecific genetic distance, clades and number of BIN, ABGD and bPTP clusters……**65**

Table 4: Comparison among DNA barcoding studies conducted in Brazilian basins, including the number of sequences and species analyzed, and intraspecific and intrageneric distances (Minimum and maximum. The mean is inside the parentheses)……**67**

Table S1: Sixty nine fish species barcoded from Doce River Basin and identified as a unique Barcode Index Number (BIN). LGC and RD– Laboratório de Genética da Conservação do Programa de Pós-graduação em Biologia de Vertebrados/PUC Minas, MBML – Museu de Biologia Professor Mello Leitão, MCNIP – Coleção de Ictiologia do Museu de Ciências Naturais da *PUC* Minas, ZUEC – Museu de Zoologia da Universidade Estadual de Campinas "Prof. Adão José Cardoso" ……**82**

Table S2: Species from Doce river basin including intraspecific distance and additional analyses. E=endemic, I=Invasive, N=native, NR=New Record, T=Threatened……**90**

CHAPTER III

Table 1: Treatments analysed according to sampling medium, preservation method used and sampling event……**107**

Table 2: PERMANOVA results (R^2-effect sizes and significance level) showing the effect of sampling medium on MOTU diversity recovery……**119**

Table S1: Sample sites, including GPS coordinates……**128**

Table S2: PERMANOVA results (R^2-effect sizes and significance level) showing the effect of preservation method, sampling medium, and sampling time on MOTU……**128**

Table S3: List of samples including primers and tags used……**129**

CHAPTER IV

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)……..**158**

Table S1: Taxa detected by eDNA metabarcoding and correspondent nearest neighbour species reported for Jequitinhonha river basin……**181**

Table S2: Species reported for the Jequitinhonha River Basin……**182**

Table S3: Comparison of different SWARM clustering thresholds.……**185**

Table S4: Species identified applying the minimum identity of 0.97, according to each SWARM threshold……**186**

Table S5: Species detected in each sampling event and sampling medium……**186**

Table S6: Sample sites including code, city, human population* and GPS coordinates……..**188**

Table S7: List of samples sequenced for the custom reference database……..**189**

APPENDIX

Table A1: List of all MOTUs recovered for each treatment……..**204**

List of Figures

CHAPTER I

Figure 1: Map of Doce (blue circle) and Jequitinhonha (orange circle) River basins in Brazil. Adapted from Peixe Vivo, CEMIG.……**18**

Figure 2: Maps of Jequitinhonha (A) and Doce (B) river basins, showing in detail the catchment area including main river and tributaries. ……**18**

Figure 3: Jequitinhonha river basin. Clockwise from top left: Sampling site at Jequitinhonha city, site located at Itacambiruçu, example of road and difficulties in reaching the sampling location, traditional sampling in Araçuaí river……..**19**

Figure 4: Records obtained when searching for "Environmental DNA" or "eDNA" (n=556, Web of Science)……**23**

Figure 5: Distribution of eDNA studies across countries (Web of Science), including the continents: North America and Europe (dark and light blue, respectively), Australia (brown), Asia (black), South America (green) and Africa (red)……**24**

Figure 6: Examples of sampling media (water and sediment) obtained in the Jequitinhonha river……**26**

CHAPTER II

Figure 1: Map of Doce River Basin, including sample sites distribution.……..**53**

Figure 2: Genetic divergences found for all sequences analyzed at species (A) and genus (B) levels……**59**

Figure S1: NJ Tree based on K2P distance, encompassing all analyzed species to check the occurrence of putative new MOTUs (red branches)……**80**

CHAPTER III

Figure 1: Map of Jequitinhonha river basin sampling locations……..**102**

Figure 2: Number of MOTUs recovered per sampling medium and preservation method (sediment vs water – BAC and ICE) and sampling event……**107**

Figure 3: Non-metric multidimensional scaling (nMDS) plots showing similarities of sample sites per sampling event. Analyses based on A) Sampling event 1; B) Sampling event 2; C) Sediment samples; D) Water samples preserved using BAC; and E) Water samples preserved using ICE……**111**

Figure 4 : Comparison of MOTU recovery between sampling events……**112**

Figure 5: Relative read abundance per order and family……**115**

CHAPTER IV

Figure 1: The Jequitinhonha river basin, including sampling sites used in the study, dams and respective hydrological regions……..**140**

Figure 2: Workflow illustrating the methods used in this paper and respective number of MOTUs retrieved in each dataset analysed, and the final number of species assigned with >0.97 identity……**150**

Figure 3: Species distribution in Jequitinhonha river basin, according to sampling media and campaign……..**153**

Figure 4: Heat trees displaying the fish diversity recovered for Jequitinhonha river Basin using eDNA metabarcoding, during the first (A) and second (B) campaigns. Blue = Water samples; Brown = Sediment samples……..**154**

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)……**155**

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) Sediment samples obtained in the first campaign; C) Water samples obtained in the second campaign; D) Sediment samples obtained in the second campaign……..**159**

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) Sediment samples obtained in the first campaign; C) Water samples obtained in the second campaign; D) Sediment samples obtained in the second campaign……**160**

Figure S1: Phylogenetic tree (A) and pairwise genetic distance (B) recovered for *Prochilodus* spp. ……..**180**

Acknowledgements

When I first arrived in Salford everything was new to me. I faced the unknown with curiosity, and I knew that many challenges were yet to come, and it would not be fair to start this thesis without acknowledging the people that made it all possible, from which I learnt a lot and will be forever grateful to.

I would like to start by expressing my gratitude to Stefano, who without knowing me accepted to be my supervisor and then guided me for the following 4 years. I am thankful for his continued support, ideas, patience and help. His advices surely contributed to improve me as a researcher and I hope we can continue this collaboration in the future.

I am sincerely grateful to Owen, who played a vital role in the eDNA Project. None of this would have been the same without his generosity, kindness and continued help. And I also, thank Owen and Sandra for all their support in Tromsø. ¡Gracias!

I owe thanks to my co-supervisor, Daniel Carvalho, for introducing me to this area of research and teaching me a lot during my Masters. Also, for all the support provided even before the beginning of my PhD, when I first decided to apply for the Science without Borders Scholarship.

I am thankful to Allan McDevitt for his continued help and mostly, for inviting me to do a field work in Scotland. I can definitely say, that (even after literally falling few times during the sampling) it helped me to change the way I was seeing my PhD which would not have been the same without you and Sam laughing at me.

They shared with me their knowledge, enthusiasm, and humour. They guided and believed in me even when I started to doubt it myself. I am very thankful to the MEG mentors (Allan, Ilaria and Chiara), for making me enjoy my PhD in Salford (even more) and especially, Ilaria and Chiara for being such a great inspiration to me and to whom I hold a great admiration.

Definitely, this period would not have been the same without the amazing people that shared with me joyful, and also stressful moments. Judith and Donna, for the long conversations and advices. The MEG Group (Charles, Sam, Donal, Andrew, Joe, Mari, Marine, Matt, Lydia, Matt) for the meetings and the fun moments at the uni.

I would like to express my gratitude to Mari, for being my friend along all this way, for incentivizing me, cherishing the good moments and helping in the not so good ones.

And also, for being the best travel buddy accepting to travel with me whenever I needed a break, without even asking the destination.

Professor Rob Young, who I am grateful to for inviting me to come to Salford and for the help during the initial process. The Little Brazil (Lu Barçante, Vinícius, Luiza, Marina, Isabela, Ivana, Luísa, Felipe) for sharing the first experiences in this new country.

I am grateful to the helpful staff from the University of Salford, for the constant help (specially Kelly Milward) and the assessors who provided great insights and feedbacks during the project development (Jean Boubli, Robert Jehle and Katherine Yates).

I owe special thanks the Science without Borders programme, from the National Council for Scientific and Technological Development (CNPq) for the scholarship, which has provided me with the unique opportunity to conduct my studies in the UK and have an excellent formation. I also owe thanks to Santander Universities and Association of the Tropical Biology (ATBC) for the grants provided.

I am also very grateful to many people, that even being distant played a significant role during the entire process.

Professor Enemir, who was my first supervisor and my mentor since the beginning of my academic life, and, the first to show me how to do research and encouraged me to keep my studies.

Gilberto and Yoshi, for being examples of committed and great Biologists, for sharing with me their knowledge and for their vital contribution in the DNA barcoding project.

I owe thanks to all professors from PUC Minas, for providing me an excellent formation without which I would not have been able to do a PhD.

I am glad to also have biologist friends that understand the ups and downs of this journey and were not distant even with all the distance between us. Nay, Vitor and Bela, thank you for being so patient! Specially, Nay for being the kindest person I've ever met.

Special thanks to the communities in the Jequitinhonha Valley, who welcomed us during the field work, for their generosity, happiness and valuable help. I am glad I had this incredible opportunity to get to know this part of Brazil.

I am very grateful to Filippe, who really experienced the entire process with me, celebrating successful moments and cherishing me in the tough ones. For being so understanding, calm, kind, and patient.

"He who has a why to live for can bear almost any how" (Nietzsche). I am grateful for having a big and loving family. Pri, Lé and Doia, thank you for being always present in my life! Filipe and Benício, the reason why I always want to improve myself and be a better person. And I am grateful for the support and comprehension provided when I decided to move and life so far away from them.

And the most beloved ones, (meus amados!) Virginia, Emir, Mah, Iaia and Pê. This is all for you.

Abstract

The Neotropical region comprises one of the greatest freshwater fish diversities in the world. Conservation and management actions in freshwater realms face great challenges in this region due to an insufficient knowledge base (e.g. shortage of taxonomic expertise, lack of robust, routine, standardised monitoring programmes), infrastructure limitations and logistic constraints (e.g. access to remote areas, insufficient funds to cover surveys). Biodiversity assessment depends on reliable detection and accurate identification of species; thus, additional methods (e.g. integrative taxonomy, DNA barcoding and other molecular diagnostic methods), associated with traditional taxonomic identification, are being increasingly implemented worldwide. Surprisingly, despite being a hugely biodiverse country, Brazil has not yet embraced these novel DNA-assisted approaches to biodiversity monitoring. Therefore, with this thesis, I aim to bolster the implementation of DNA barcoding and metabarcoding in Brazilian riverine ecosystems. With the cooperation of collaborators in both Brazilian and British institutions, I built a barcode library and provided a more robust biodiversity record for the ichthyofauna of the Doce river, reflecting communities as they were prior to a major chemical pollution disaster in that catchment. Furthermore, I evaluated the application of eDNA metabarcoding as a fish biodiversity assessment tool, along the course of the Jequitinhonha river. Results for the Doce suggested the occurrence of potentially cryptic species, species complex, or historical errors in morphological identification. Metabarcoding of the environmental samples in the Jequitinhonha allowed the detection of native and introduced species and provided data from localities often neglected due to the difficulties of traditional sampling. Collectively, my studies indicate that a range of powerful and cost-effective molecular approaches are now available to biologists and conservationists, which will empower and fast-track the process of characterising biodiversity, and ultimately ecosystem function, in Brazilian freshwater habitats.

Keywords: DNA barcoding, environmental DNA, fish, metabarcoding, Neotropical

Chapter I

INTRODUCTION AND BACKGROUND

Chapter I

INTRODUCTION AND BACKGROUND

1.1 *Fish biodiversity assessment and decline in Neotropical freshwater systems*

Anthropogenic alterations of the global environment are one of the key factors leading to biodiversity loss, leading to local extinction events, changes in community structure and loss of important ecosystem services. Human impacts on natural populations are increasing at a fast pace, and many habitats and biological communities are likely to be currently undergoing severe ecological shifts (Krausmann et al., 2013; Steffen et al., 2015). Freshwater ecosystems are among the most vulnerable and susceptible to such rapid changes.

Many anthropogenic impacts (e.g. pollution, deforestation, habitat loss) have been attributed as a main cause of species decline. Water pollution is a well known factor imposing threats to natural aquatic species. Regarding freshwater fishes, increasing levels of toxicant residues originated from industrial, agricultural or other human activities and discharged in rivers have demonstrated significant impacts on fish health (e.g. alteration of migration patterning, increase of mutation rates) and ecological integrity of habitats (Van Straalen & Timmermans, 2002; Viana et al., 2018). Additionaly, the construction of dams is well known as a fish diversity reduction factor, due to the modification of physical and ecological characteristics of the habitats (e.g. water flow, impoundments and flood control, nutrient dynamics, changes in water quality and temperature, increased predation pressure, and

habitat loss). These barriers also block movements responsible for connecting populations and enabling migratory fish species to complete their reproductive cycle (usually correlated to seasonal flood pulses) (FAO, 2002; Pelicice & Agostinho, 2007; Ziv et al., 2012; Pelicice et al., 2015; Pompeu et al., 2012).

In addition to habitat modifications, anthropogenic actions cause changes in species distribution, enhancing the mobility of organisms through biological introductions. The accidental or intentional spread of species beyond their native ranges represents a great threat to biodiversity, being second only to habitat loss as the leading cause of extinctions in vertebrates (Bellard et al., 2016; Vitousek, 1997; Schmitz & Simberloff, 1997, Chapin III et al., 2000). When alien species become invasive, ecological change can be irreversible, or can only be curbed through costly, complex and challenging management or eradication measures. Thus, great efforts are necessary to detect the introduction quickly to maximize response success (Vilà et al., 2011; Rejmánek & Pitcairn, 2002; Simberloff et al., 2013). These factors combined are leading to the decline of freshwater species/populations (83% decline since 1970) and management of freshwaters are deemed as a priority target for global biodiversity conservation (WWF, 2018).

Species description and accurate identification are crucial factors for biodiversity conservation. Described neotropical ichthyofauna comprises more than 5,000 species and it is estimated that the true number may exceed 7,000 in this region (Albert & Reis, 1999; Reis et al., 2016). Still, fish biodiversity is often underestimated due to the great challenge imposed by infrastructure problems and difficulties of sampling (e.g. access to remote areas,

insufficient funds to cover expenses), a shortage in taxonomic expertise and the absence of morphological characters available to identify and distinguish between specimens at early development stages of life (Reis et al., 2016; Ely et al., 2017).

The threats faced by fish species in freshwater realms require urgent action for management and conservation plans. A number of novel methods (e.g. integrative taxonomy, DNA barcoding, eDNA – associated with traditional taxonomic identification) are being employed worldwide to contribute to biodiversity assessment (Dayrat, 2005; Carvalho et al., 2011; Gomes et al., 2015; Hänfling et al., 2015). An integrative approach, through the association of distinct methodologies (e.g. morphological and molecular identifications) has yet to contribute to biodiversity assessment and conservation in neotropical rivers.

1.2 *Brazilian fish biodiversity: Doce and Jequitinhonha river basins*

The Doce and Jequitinhonha river basins are located in Southeastern Brazil (Fig.1 and Fig.2), belong to a group of eastern coastal catchments and are characterized by a high level of endemic and endangered species (Rosa & Lima, 2008). Both catchments are facing unprecedent levels of threats, such as the construction of dams, siltation, pollution, water contamination and introduction of non-native fishes.

The Doce river basin comprises approximately 83.400 km², extending by 853 km and draining 230 cities in two states. This drainage is inserted in two Brazilian hotspots (Atlantic forest and Brazilian Savanna) and recognized as an area rich in fish biodiversity, with many

species not yet known to science (i.e. estimates indicate that the number of species reaches as twice the number found in Great Britain). Seventy-two native freshwater fish species have already been described for this basin and are considered currently threatened by environmental impacts such as dams, species introductions and siltation (Vieira, 2009). Furthermore, in 2015 a mining collapse lead to an important environmental modification in this already impacted catchment by greatly increasing the levels of heavy metals in the water (GFT, 2015). The accidental discharge of approximately 62 million m^3 of a toxic mud along the main course of the Doce river affected the biodiversity and human communities of this basin, and thus was considered as the worst environmental accident reported for any South American catchment. After the mine burst, the metal-rich tailings were released into the watershed and caused a massive loss of vegetation and die-off of fishes (GFT, 2015; IBAMA, 2015; Fernandes et al., 2016). Understanding the impacts of this disaster on the fish biodiversity is crucial, as many riverine communities rely on fisheries for their livelihood (e.g. ecoturism, source of income and subsistence -Ecoplan-Lume, 2010; GFT, 2015; Neves et al., 2016). However, the effects of this ecological disaster on fish populations still remain to be completely elucidated yet due to the high frequency of cryptic species and the occurrence of many putative endemic and undescribed species (Ramirez et al., 2016). The recovery of fish populations in the Doce river basin, after the ecological disaster, depends on the recolonization of the main course of this river and on the diversity, size and conservation status of the remnant fish populations in the tributaries (Fernandes et al., 2016). Until this date, the only genetic studies conducted in this river basin focused on the genetic diversity

and cytogenetic data of few species, or on phylogenetic/phylogeographic analyses comprising species from distinct Brazilian catchments (Santos et al., 2010; Barros et al., 2015; Ramirez et al., 2016; Swarça et al., 2018) and none study aiming to describe the entire fish biodiversity was published so far.

The Jequitinhonha River basin (Fig.3), located in Southeast Brazil (17°, 43° W) is inserted in two biodiversity hotspots (Cerrado and Atlantic Forest) and characterized by a tropical climate and environmental heterogeneity. The main river flows over 1,082 km, from its source in Serro (Minas Gerais) at an elevation of 1200m, to drain its water in the ocean at the locality of Belmonte (Bahia). The known ichthyofauna of Jequitinhonha river basin comprises 52 species and is already composed by 16% of non-native species (Andrade-Neto, 2009). Also, this catchment has two large dams constructed: the hydroelectric power plant of Irapé, the tallest dam in Brazil and implemented in 2006 and the hydroelectric power plant of Itapebi which was implemented in 2002. This basin is still poorly studied, and the only genetic studies conducted in this area were a DNA Barcoding of its ichthyofauna (Pugedo et al., 2016) and a population genetic study on hybridization between a native and introduced species (Sales et al., 2017). Interestingly, the species richness of the Jequitinhonha river basin (JRB) remains as a conundrum because, despite being considered as a low biodiversity catchment, it has the same geological formation as neighboring basins with higher biodiversity, thus leading to the hypothesis that the low biodiversity of this basin might be due to the lack of efficient sampling and knowledge regarding its ichthyofauna. Additionally, one of the most

abundant endemic fish species (*Prochilodus hartii*) was found to be threatened by hybridization with introduced congeneric species in this river system (Sales et al., 2017).

FIGURE 1 | Map of Doce (blue circle) and Jequitinhonha (orange circle) River basins in Brazil. Adapted from Peixe Vivo, CEMIG.

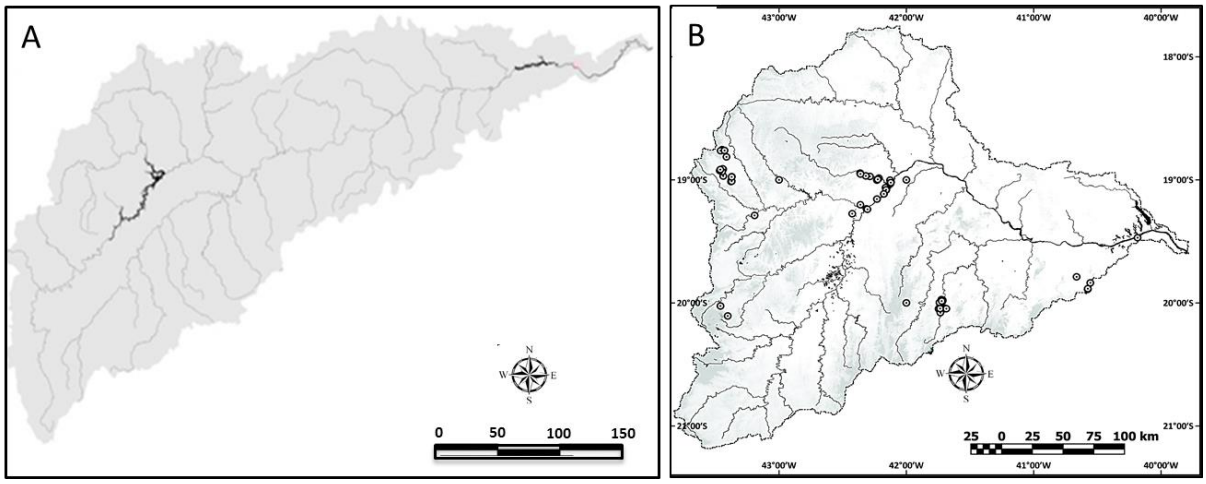


FIGURE 2 | Maps of Jequitinhonha (A) and Doce (B) river basins, showing in detail the catchment area including main river and tributaries.

FIGURE 3 | Jequitinhonha river basin. Clockwise from top left: Sampling site at Jequitinhonha city, site located at Itacambiruçu, example of road and difficulties in reaching the sampling location, traditional sampling in Araçuaí river.

1.3 DNA barcoding as a tool for fish species identification

Molecular approaches offer a universal key to identify, assess and quantify biodiversity, especially in biodiversity-rich and understudied ecosystems and regions (Schwartz et al., 2006). A new system for species identification and discovery through DNA barcodes, based on the use of a fragment of the mitochondrial gene cytochrome oxidase subunit I (COI) (˜650bp), was proposed by Hebert et al. (2003) as the standard molecular method. The Barcode of Life Project aims to describe the Earth biodiversity by sequencing and obtaining

barcodes for all living species and making it available as a reference library of DNA barcodes in the Barcode of Life Data Systems database (BOLD, www.bold.org). The database includes now over 6 million barcodes belonging to almost 200,000species, and more than 12,000 studies using DNA barcoding have been published since the concept was introduced in 2003. Since the Barcode of Life project launch, the DNA barcoding system has been widely applied to different taxa worldwide. Identification of fish fauna using DNA barcoding was successfully conducted globally (e.g. mislabelling detection (Cawthorn et al., 2018; Carvalho et al. 2017), forensics analyses (Dawnay et al. 2007; Kumar at al., 2018), management of long term fisheries (Ardura et al., 2010; Metcalf et al. 2007), data regarding spawning and recruitment areas (Almeida et al., 2018; Becker et al., 2015; Frantine-Silva et al., 2015), description of cryptic and putative new species (Iyiola et al., 2018; Hou et al., 2018), and has proven to be a valuable tool for the identification of specimens at early life development stage, flag potential overlooked species and potential new candidate species (Nwani et al., 2011; Ward et al., 2005; Hubert et al. 2008; Carvalho et al., 2011; Pereira et al., 2011; Becker et at., 2015).

In Brazil, this approach was tested in distinct catchments and has already been applied as a biodiversity identification method and contributed to highlight species introductions, and the occurrence of putative cryptic and new species (Carvalho et al., 2011; Pereira et al., 2011; Pugedo et al., 2016; Gomes et al., 2015) which has contributed to improve the fish community composition knowledge, leading to the description of several new species (Jerep, Camelier & Zanata, 2016; Dutra et al., 2016, Nielsen et. al, 2017; Zawadzki et al., 2016, Pereira et al., 2017). However, the effective performance of DNA barcoding relies in the so-called barcode

gap which is respresented by the existence of a gap between intra and interspecific genetic divergences, meaning that DNA sequences obtained for different individuals of the same species (intraspecific variability) need to present a greater similarity than between different species. For fish species, a 2% threshold of genetic distance was suggested after analysing 1088 species (Ward et al., 2005), however, taking into account the different life histories of species, despite been reported the occurrence of a barcode gap for most fish species analysed, there are still exceptions (Hurst & Jiggins, 2005), especially in the case of recently diverged species (Prosdocimi et al,. 2012). Thus, when using the DNA barcoding data for infering the occurrence of cryptic and new candidate species, a robust and careful analyses is required and can greatly be improved when including an integrative approach in order to provide reliable information regarding the biodiversity studied.

1.4 Metabarcoding using amplicon sequencing as an effective tool for biodiversity assessment

Recent advancements have now opened new opportunities for studying biodiversity by sequencing trace DNA present in the environment – the so-called "environmental DNA" (eDNA), to identify species presence and, to some extent abundance, in aquatic environments (Thomsen et al., 2012). Since animals release DNA continuously into the surrounding environment (e.g. shed cells, faeces, gametes), this method consists in extracting such DNA remnants from environmental samples (e.g. water, sedimentsand ice cores) and using specific molecular markers to target taxa of interest. Single-taxon approaches use specific narrow-

target markers using PCR or qPCR amplification (Ficetola et al., 2008;, Takahara et al., 2013; Piaggio et al., 2014) while community-level approaches use universal markers and parallel sequencing to detect a broad range of taxa (i.e. "metabarcoding" - Rees et al., 2014; Thomsen & Willerslev, 2015). Environmental DNA metabarcoding is a more efficient and cost-effective option when a broad characterisation of the ecosystem is required, also allowing the detection of unexpected species (Gillet et al., 2018).

The term "environmental DNA" was used for the first time in 1987 to describe a DNA extraction protocol to obtain DNA from sediments (Ogram et al., 1987) and since then studies started to be conducted focusing primarly on microorganisms. Willerslev et al. (2003) conducted the first metagenomics/metabarcoding study on macroorganisms, demonstrating the feasibility of applying this method to reconstruct paleocommunities. However, the method applied (e.g. using cloning vectors) was expensive and time consuming. The advent of Next Generation sequencing in 2005 contributed to stimulate studies in this field. Ficetola et al. (2008) published the first study detecting species from freshwater samples, and since then a rapid and accelerated increase in publications occurred in the past decade (Figure 4).

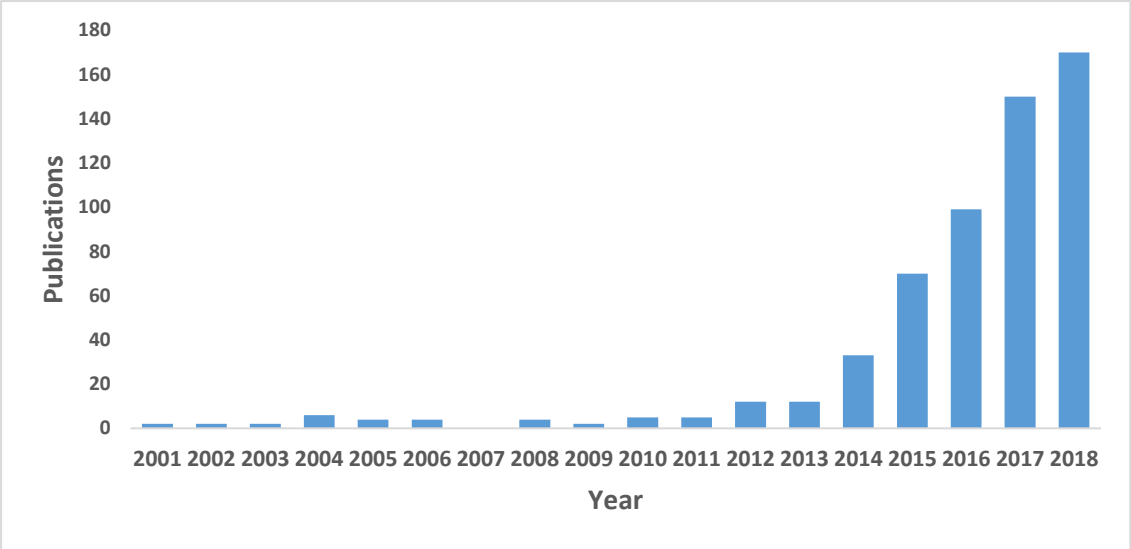


FIGURE 4 | Records obtained when searching for "Environmental DNA" or "eDNA" (Period= 2001-2019, n=556, Web of Science).

However, the efforts are not distributed homogeneously worldwide. Most of the studies have been conducted in temperate regions and the use of eDNA in neotropical environments is still scarce (Figure 5). When looking for studies containing the word "Brazil", the search returned nine records from which only five records represent studies conducted in this country, including characterization of amphibian communities (Sasso et al., 2017, Lopes et al., 2017), qPCR amplification for detecting the golden mussel (*Limnoperna fortunei*) (Pie et al., 2017), detection of freshwater bacterioplankton (Tessler et al., 2017) and development of primers targeting the 16S gene for freshwater vertebrates (Vences et al., 2016). The remaining four records belonged to authors based in Brazil but with studies focused on a different geographical region.

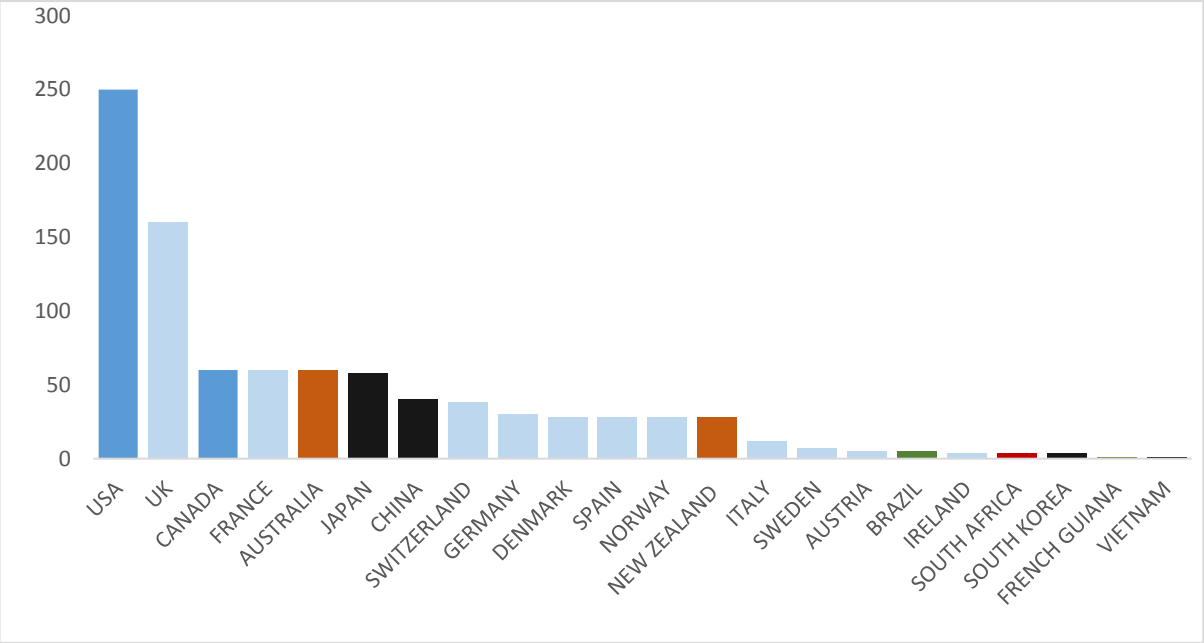


FIGURE 5 | Distribution of eDNA studies across countries (based on authors address, Web of Science), including the continents: North America and Europe (dark and light blue, respectively), Oceania (brown), Asia (black), South America (green) and Africa (red).

Since its advent, this approach has been tested in different contexts and areas and despite being considered as an attractive option for environmental monitoring, there remains much to be explored and ground-truthed before it can be routinely applied as a biomonitoring (i.e. continuous assess of current state and ongoing changes in environments, including habitats, species and populations)tool (Jeunen et al., 2019; Salter et al., 2018; Taberlet et al., 2018). Some aspects, including laboratory procedures and bioinformatics, remain challenging, but an ever increasing number of studies worldwide are contributing to making eDNA metabarcoding an established tool in ecological analysis (Taberlet et al., 2018).

Below, I illustrate and discuss each step of the eDNA metabarcoding workflow.

1.5 Important considerations when working with eDNA

- Sampling medium

Persistence of eDNA molecules in the environment can vary significantly depending on the species shedding rates and DNA degradation in diverse habitats. Sampling media (Fig.6) can harbor eDNA for distinct temporal periods, from days in water samples (Dejean et al., 2011; Pilliod et al., 2014) to thousands of years in ice cores (Thomsen and Willerslev, 2015). Furthermore, studies demonstrated that the choice of sampling medium can critically affect the results obtained. For example, sediment samples can harbour up to 1800 times more eDNA when compared to water samples and even provide a higher number of OTUs and a different community composition when compared to water samples (Turner et al., 2015). On the other hand, for specific taxa (e.g. fish) this sampling medium might yield reduced detection of species compared to water (Holman et al., 2018, Koziol et al., 2018, Shaw et al., 2016). When using a single substrate, the data obtained through metabarcoding are likely to be an underestimation of local biodiversity, and thus some authors recommend experimental designs that include multiple substrates (Koziol et al., 2018). However, despite the great difference found among sampling media and the importance of testing the suitability of each substrate before conducting monitoring surveys, few eDNA studies have incorporated this aspect in their analyses (Shaw et al., 2016).

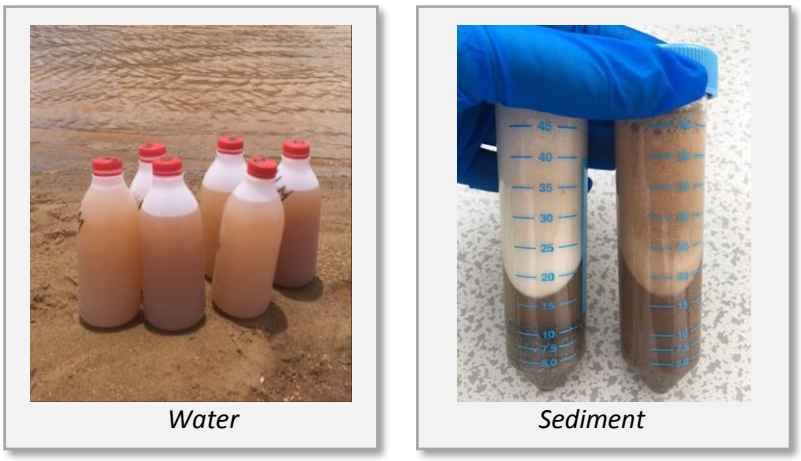


FIGURE 6 | Examples of sampling media (water and sediment) obtained in the Jequitinhonha river.

- Preservation of samples

After being released in the environment, eDNA can degrade at a fast rate or settle and persist for longer times when bound to the sediment. Still, to overcome the degradation of environmental samples, the most recommended approach is to extract the DNA as quickly as possible after sampling. In some cases, when filtration and DNA extraction are not feasible to be conducted immediately after sampling (e.g. field work conducted in remote sites), a preservation method must be employed in order to stop or decrease the microbial activities and minimize DNA degradation in the samples. The most widely employed method is storing the samples at low temperatures (i.e. freezing the samples or cooling using a cool box) which requires a substantial equipment increase (Eichmiller et al., 2016; Pilliod et al., 2014). The inclusion of buffers, such as EtOH–NaAc (ethanol-sodium acetate) solution have been shown

to keep an eDNA persistence rate similar to samples stored in ice (Ladell et al., 2018), however, when sampling larger volumes of water the increased final volume obtained (i.e. addition of over 2x of solution) might be considered problematic during long sampling campaigns. Recently, Yamanaka et al. (2017) tested the addition of cationic surfactants as preservatives to suppress DNA degradation at ambient temperatures, demonstrating the efficiency of Benzalkonium chloride (0.01%) in retaining eDNA concentration even after 10-day incubation at 21ºC. However, as these methods have been tested individually in different contexts, a further evaluation should be conducted in order to provide a more reliable information regarding the best preservation method to be applied according to the environmental samples obtained.

- DNA capture

A critical step in eDNA analyses is DNA capture and the most used methods to obtain eDNA from environmental samples are filtration and precipitation. In the former, filters are employed, which permit the passage of water and retention of eDNA sources (e.g. tissue, cells, organelles)within the fine mesh of the filter; in the latter, a chemical process of precipitation of nucleic acids using ethanol is followed. Despite both being commonly used, filtration has been shown to be a better option for recovering eDNA from water samples (Jerde et al., 2011, Deiner et al., 2015, Eichmiller et al., 2016). A broad range of filters and DNA extraction techniques have been employed in eDNA surveys, but there is still there is no consensus about the best filter pore size and material and DNA extraction protocol to be

followed as the results varied accordingly to the environment sampled (e.g. marine, freshwater) and type of samples analyzed. Filtering using a syringe method (e.g. Sterivex filter unit) allows processing multiple samples independently, reducing cross contamination and allowing sampling multiple sites at the same time (, Lacoursière-Roussel et al., 2018). However, the filtering capacity of enclosed capsules filters might be shorter when compared to cellulose nitrate filters mounted to a funnel adapted to a peristaltic automatic pump, which were shown to outperform Sterivex capsule filters (Spens et al., 2017).

Filter pore size should also be analyzed to obtain a tradeoff between DNA recovery and filtering time, as for turbid waters a larger pore size filters can greatly decrease filtration time but might also reduce eDNA recovery (Eichmiller et al., 2016). Regarding filter composition and pore size, recently, Majaneva et al. (2018) demonstrated that cellulose nitrate and mixed cellulose ester (MCE) filters yielded more eDNA than polyethersulfone filters, and filters with pore size of 0.45 µm despite yielding less DNA than small pore size filters (0.2 µm). This may be useful for sampling in turbid waters as the latter may clog easily. Additionally, 0.45 µm MCE filters have been considered as a better option when compared to larger pore size filters and the 0.45 µm Sterivex filter (Li et al., 2018). Li et al. (2018) also stress that after filtering, the filters should be kept either dry or in lysis buffer to avoid degradation.

- Metabarcoding: markers choice

In metabarcoding studies the choice of metabarcode is crucial as it greatly impacts the end results. Thus, several factors may be taken into account before making a choice: clear

definition of the target taxonomic group including the level of taxonomic resolution required, size of selected barcode as shorter fragments might be easier to detect in the case of degraded DNA; and the existence of reference sequences in the databases (Taberlet et al., 2018). Therefore, an ideal primer for eDNA metabarcoding surveys should be specific to the target taxonomic group, amplify a short fragment (~150-200bp) containing sufficient taxonomic resolution to allow species assignment at high confidence rates, and amplify DNA from all species of the target group without favoring any of those (Coissac et al., 2012; Elbrecht & Leese, 2017). A trade-off between fragment size and taxonomic resolution represents a significant challenge when choosing the marker, but it should also be noted that fragment size is additionally constrained by the sequencing platform used. The Illumina MiSeq generates large amout of data through a fast and high-quality sequencing approach (i.e. low substitution and indel errors) when compared to other platforms (e.g. 454 GS, Ion Torrent PGM - Loman et al., 2012). However, for achivieng this high-quality standard and reduce errors during the sequencing proccess, the chemistry available for this platform allows the sequencing of fragments up to 500bp, limiting the fragment size analysed in metabarcoding studies (Slatko et al., 2018).

The cytochrome c oxidase subunit I (COI) is the mitochondrial DNA marker of choice in DNA barcoding studies, having an associated robust reference database (BOLD Systems, www.boldsystems.org); however, COI is a coding gene (i.e. all third codons positions are variable) and is considered challenging for most of eDNA metabarcoding studies due the absence of highly conserved regions required for robust primer design (Deagle et al., 2014).

Mitochondrial ribosomal RNA (rRNA) gene regions generally show relatively similar taxonomic resolution to COI and are currently being widely used in metabarcoding studies (Riaz et al., 2011; Valentini et al., 2009). In regards to assessing fish biodiversity, several primers targeting fragments of the mitochondrial 12S rRNA gene have been described for Actinopterygii species. Kelly et al. (2014) tested a set of primers, amplifying a 106bp fragment, which allowed the detection of bony fishes in a mesocosm, however those were only identified up to the genus level due to the limited variability of the amplicons analyzed. The same constraint (low taxonomic resolution) was also reported for the primers described by Valentini et al. (2016, ˜70bp), whereas the MiFish primer set (172bp, Miya et al., 2015) allowed higher taxonomic assignments. Howewer, despite the increase in the taxonomic assignment, due to the still low phylogenetic resolution of the fragments been currently analysed many species migh not be realiably identified. Thus, the lack of appropriate genetic regions targeted hampers the identification at species level and then, the biodiversity detected remains surely underestimated. Optimization of new primer sets are still ongoing and this will allow a great improvement in the taxonomic assignment and monitoring using eDNA.

Although eDNA studies have proved that this method can be very efficient in detecting species, most of the data obtained still represent an underestimation of true biodiversity, as a large part of the information recovered is lost due to the incompleteness of the reference databases and a great effort should be made in order to overcome this problem and improve the application of eDNA metabarcoding as a biomonitoring tool. While rapid improvements of the databases based on sequencing short fragments will offer increasingly accurate

metabarcoding results, an attractive perspective is provided by the possibility of generating whole mitogenome data. In addition to significantly improve eDNA studies by alloowing species detection after retrieving whole mitogenomes from the environment (Deiner et al., 2017), it would also contribute to overcome the problems caused by the constant optimization of markers.

- Potential use of eDNA in Brazilian rivers

Environmental DNA metabarcoding offers a great opportunity to take biomonitoring to a higher level, improving species conservation and management for freshwater fishes in understudied regions. In Neotropical freshwater ecosystems, teleost populations are often monitored by live capture followed by morphological identification, and due to the water features (e.g. low conductivity, high water turbity and current velocity) observation and use of electrofishising are not feasible and traditional sampling is restricted to the use of invasive and selective methods (e.g. nets, toxicants). These methods are destructive to the ichthyofauna, and also provide biased or incomplete representation of the monitored community by selecting few species (e.g. size selection of nets) (Dalu et al., 2015; Gunzburger, 2007).

Traditional field monitoring techniques might also fail in detecting species when they are elusive, rare or occurring at low abundances or densities, and new non-invasive methods have been recently proposed to improve biodiversity assessment of fish by obtaining the DNA present in the environment. The effectiveness of non-invasive methods such as eDNA

metabarcoding has been proven for different environments and taxa, and when used as a complementary tool (i.e. associated with traditional sampling) has shown to increase the fish diversity assessment proving more accurate and reliable estimates. This method was shown to be more sensitive than traditional sampling approaches such as electrofishing (Shaw et al., 2016; McDevitt et al., 2018) and BRUVs (Boussarie et al., 2018; Stat et al., 2018). As a notable example, when compared to gillnets surveys conducted in the Mekong basin, eDNA metabarcoding detected 30 more species that were not caught by the traditional methods in a 3 year-period (Gillet et al., 2018). Furthermore, eDNA is revolutionizing the monitoring of non-native species and has been used extensively to detect and track species invasion (Jerde et al., 2011; Dejean et al., 2012; Goldberg et al., 2013; Takahara et al., 2013; Fukumoto et al., 2015; Hänfling et al., 2015; Hunter et al., 2015; Uchii & Minamoto, 2016).

Biological invasions represent a major challenge in conservation biology and the management of natural population (Rhymer & Simberloff, 1996; Magalhães & Jacobi, 2013; Metcalf et al., 2007). Introduced species can spread pathogens, alter ecosystem structure, change ecological interactions, outcompete native species and promote biotic homogenization (Latini & Petrere, 2004; Catford et al., 2012). Introduced species can remain undiscovered for a long period before spreading into new habitat (Crooks, 2011; Essl et al., 2011). Preventing introduction is considered the ideal scenario, but an early detection is particularly important to reduce the impacts invasive species may have further along the invasion process (Simberloff et al., 2013). Detecting alien species and mitigating against their negative impact is inherently difficult, and especially so in mega-diverse countries. In Brazil,

non-native fish introduction has become commonplace due to the absence of effective methods to deter such introductions (Agostinho et al., 2007; Pelicice et al., 2014; Azevedo-Santos et al., 2015). Although different laws prohibit the release of non-native specimens, poor resources for enforcement, in a continent-sized country, hamper the effectiveness of the normative approach (Azevedo-Santos et al., 2015). In this context the use of innovative molecular tools such as the environmental DNA (eDNA) technique could prove to be an effective tool for general biodiversity assessment, and for offering an "early warning" system for the detection of non-native species and quickly enable swift eradication or mitigating measures (Takahara et al., 2013; Simberloff et al., 2013; Ficetola et al., 2008; Chown et al., 2008; Rees et al., 2014).

DNA barcoding and eDNA metabarcoding data can provide valuable information about freshwater fish community dynamics, and evaluate the impact of anthropogenic actions, such as pollution, species introduction, and the construction of dams and other forms of habitat modification. Biomonitoring of fish species using molecular tools could also contribute to establishing long-term monitoring schemes and obtain information from areas that are often negletected due to poor accessibility.

DNA barcoding has already proven to be a valuable tool in describing fish biodiversity in Neotropical rivers, by contributing to flag taxonomic problems and uncover the occurrence of undescribed and/or cryptic species (Carvalho et al., 2011; Pugedo et al., 2015). Since the onset of eDNAstudies, the approach has been favourably received by monitoring agencies; however, the efficacy of eDNA metabarcoding in a neotropical context remains to be

explored in-depth, especially due to a range of factors (e.g. high biodiversity, higher temperatures and solar radiation, water acidity and turbidity,high discharges), which may have strong influence on eDNA recovery and species detection in these habitats.

Aims of the PhD thesis

My thesis work aims to:

I) Apply different molecular tools (DNA barcoding and eDNA metabarcoding);

II) Test their potential to improve knowledge of Brazilian freshwater communities;

III) Establish these techniques as regular tools for ecological analysis in the Neotropics.

Here, I enclose three research chapters that focus on three main issues:

I) The DNA barcoding section (Chapter 2) details on the generation of a DNA barcode library for the Doce River Basin (DRB) ichthyofauna, using data collected prior to the chemical spillage disaster of 2015. I contribute to an improved biodiversity baseline record for this recently impacted ecosystem, including the detection of invasive species and cryptic, likely undescribed, species.

II) In Chapter 3, I investigate the influence of methodological aspects on eDNA metabarcoding inference, namely: i) the effect of two different sampling media (sediment and water); ii) the effect of preservation methods for water samples (low temperatures vs cationic surfactant solution); iii) the effect of sampling time, by conducting two sampling campaigns conducted at a three-week interval;

III) In Chapter 4, I examine the fish communities along the Jequitinhonha river basin, based on environmental DNA metabarcoding of water and sediment samples.

REFERENCES

Agostinho, A. A., Gomes, L. C. and Pelicice, F. M. (2007). Ecologia e Manejo de Recursos Pesqueiros em Reservatórios do Brasil. EDUEM, Maringá. 208p.

Albert, J. S., Fernandes-Matioli, F. M. C., and Almeida-Toledo, L. F. (1999). New species of *Gymnotus* (Gymnotiformes, Teleostei) from south eastern Brazil: toward the deconstruction of *Gymnotus carapo*. *Copeia*, 1999, 410-42. doi: 10.2307/1447486

Almeida, F. S., Frantine-Silva, W., Lima, S. C., Garcia, D. A. Z., and Orsi, M. L. (2018). DNA barcoding as a useful tool for identifying non-native species of freshwater ichthyoplankton in the neotropics. *Hydrobiologia*, 817, 111-119. doi: 10.1007/s10750-017-3443-5

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

Azevedo-Santos, V. M., Pelicice, F. M., Lima-Junior, D. P., Magalhães, A. L. B., Orsi, M. L., Vitule, J. R. S., and Agostinho, A. A. (2015). How to avoid fish introductions in Brazil: education and information as alternatives. *Natureza & Conservação*, 13, 123-132. doi: 10.1016/j.ncon.2015.06.002

Barros, L. C., Santos, U., Cioffi, M. de B., and Dergam, J. A. (2015). Evolutionary divergence among *Oligosarcus* spp. (Ostariophysi, Characidae) from the São Francisco and Doce River Basins: *Oligosarcus solitarius* Menezes, 1987 shows the highest rates of chromosomal evolution in the neotropical region. *Zebrafish*, 12, 102–110. doi:10.1089/zeb.2014.1030.

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

Bellard, C., Cassey, P., and Blackburn, T. M. (2016). Alien species as a driver of recent extinctions. *Biology Letters*, 12, 24–27. doi: 10.1098/rsbl.2015.0623

Boussarie, G., Bakker, J., Wangensteen, O. S., Mariani, S., Bonnin, L., Juhel, J. B., et al. (2018). Environmental DNA illuminates the dark diversity of sharks. *Science Advances*, 4(5), 1-8. doi: 10.1126/sciadv.aap9661

Carvalho, D. C., Oliveira, D. A. A., Pompeu, P. S., Leal, C. G., Oliveira, C., and Hanner, R. (2011). Deep barcode divergence in Brazilian freshwater fishes: the case of the São Francisco River basin. *Mitochondrial DNA*, 22, 80-86. doi: 10.3109/19401736.2011.588214

Carvalho, D. C., Guedes, D., da Gloria Trindade, M., Coelho, R. M. S., and Araujo, P. H. (2017). Nationwide Brazilian governmental forensic programme reveals seafood mislabelling trends

and rates using DNA barcoding. *Fisheries Research*, 191, 30–35. doi: 10.1016/j.fishres.2017.02.021

Catford, J. A., Vesk, P. A., Richardson, D. M., and Pysek, P. (2012). Quantifying levels of biological invasion: towards the objective classification of invaded and invasible ecosystems. *Global Change Biology*, 18, 44-62. doi: 10.1111/j.1365-2486.2011.02549.x

Cawthorn, D. M., Baillie, C. and Mariani, S. (2018). Generic names and mislabeling conceal high species diversity in global fisheries markets. *Conservation Letters*, 11:e12573. doi: 10.1111/conl.12573

Chapin III, F. S., Zavaleta, E. S., Eviner, V. T., Naylor, R. L., Vitousek, P. M., Reynolds, H. L., et al. (2000). Consequences of changing biodiversity. *Nature*, 405, 234-242. doi: 10.1038/35012241

Chown, S. L., ·Sinclair, B. J., and ·van Vuuren, B. J. (2008). DNA barcoding and the documentation of alien species establishment on sub-Antarctic Marion Island. *Polar Biology*, 31, 651–655. doi: 10019.1/11628

Coissac, E., Riaz, T., and Puillandre, N. (2012). Bioinformatic challenges for DNA metabarcoding of plants and animals. *Molecular Ecology*, 21(8), 1834-47. doi: 10.1111/j.1365-294X.2012.05550.x.

Crooks, J. A. (2011). Lag times. In Encyclopedia of Biological Invasions (Simberloff, D. and Rejmánek, M., eds). University of California Press

Dalu, T., Wasserman, R. J., Jordaan, M., Froneman, W. P., and Weyl, O. L. F. (2015). An assessment of the effect of rotenone on selected non- target aquatic fauna. PLoS ONE, 10(11), e0142140. doi: 10.1371/journal.pone.0142140

Dayrat, B. (2005). Towards integrative taxonomy. *Biological Journal of the Linnean Society*, 85: 407-415. doi:10.1111/j.1095-8312.2005.00503.x

Dawnay, N., Ogden, R., Mcewing, R., Carvalho, G., and Thorpe, R. (2007). Validation of the barcoding gene COI for use in forensic genetic species identification. Forensic Science International, 173, 1-6. doi: 10.1016/j.forsciint.2006.09.013

Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F., and Taberlet, P. (2014). DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. *Biology Letters*, 10(9). doi: 10.1098/rsbl.2014.0562

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

Deiner, K., Renshaw, M. A., Li, Y., Olds, B. P., Lodge, D. M., and Pfrender, M. E. (2017). Long-range PCR allows sequencing of mitochondrial genomes from environmental DNA. *Methods in Ecology and Evolution*, 2017, 1–11. doi: 10.1111/2041-210X.12836

Dejean, T., Valentini, A., Duparc, A., Pellier-Cuit, S., Pompanon, F., Taberlet, P., et al. (2011). Persistence of Environmental DNA in Freshwater Ecosystems. *PLoS ONE*, 6: e23398. doi: 10.1371/journal.pone.0023398

Dutra, G., Penido, I., Mello, G., and Pessali, T. C. (2016). Two new species of *Cyphocharax* (Teleostei: Characiformes: Curimatidae) from headwaters of the Jequitinhonha and São Francisco river basins, Minas Gerais, Brazil. *Zootaxa*, 4103(2), 154-156. doi: 10.11646/zootaxa.4103.2.5.

Ecoplan-Lume. (2010). Plano integrado de recursos hídricos da bacia hidrográfica do rio Doce. Belo Horizonte: IGAM.

Eichmiller, J. J., Best, S. E., and Sorensen, P. W. (2016). Effects of temperature and trophic state on degradation of environmental DNA in lake water. *Environmental Science and Technology*, 50(4), 1859–1867. doi: 10.1021/acs.est.5b05672

Elbrecht, V., and Leese, F. (2017). Validation and Development of COI Metabarcoding Primers for Freshwater Macroinvertebrate Bioassessment. *Frontiers in Environmental Science*, 5, 1-11. doi: 10.3389/fenvs.2017.00011

Ely, C. V., Bordignon, S. A. L., Trevisan, R., and Boldrini, I. I. (2017). Implications of poor taxonomy in conservation. *Journal of Nature Conservation*, 36, 10-13. doi: 10.1016/j.jnc.2017.01.003

Ermakov, O. A., Simonov, E., Surin, V. L., Titov, S. V., Brandler, O. V., and Borisenko, A. V. (2015). Implications of hybridization, NUMTs, and overlooked diversity for DNA barcoding of Eurasian ground squirrels. *PLoS ONE*, 10(1): e0117201. doi: 10.1371/journal.pone.0117201

Essl, F., Dullinger, S., Rabitsch, W., Hulme, P. E., Hulber, K., Jarosik, V., et al. (2011). Socioeconomic legacy yields an invasion debt. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 203–207. doi: 10.1073/pnas.1011728108

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

Fernandes, G. W., Goulart, F. F., Ranieri, B. D., Coelho, M. S., Dales, K., Boesche, N., et al. (2016). Deep into the mud: ecological and socio-economic impacts of the dam breach in Mariana, Brazil. *Natureza & Conservação*, 14, 35-45. doi: 10.1016/j.ncon.2016.10.003

Ficetola, G. F., Miaud, C., Pompanon, P. and Taberlet, P. (2008). Species detection using environmental DNA from water samples. *Biology Letters*, 4, 423–425. doi: 10.1098/rsbl.2008.0118

Fukumoto, S., Ushimaru, A., and Minamoto, T. (2015). A basin-scale application of environmental DNA assessment for rare endemic species and closely related exotic species in rivers: a case study of giant salamanders in Japan. *Journal of Applied Ecology*, 52(2), 358-365. doi: 10.1111/1365-2664.12392

Gillet, B., Cottet, M., Id, T. D., Kue, K., Chanudet, V., and Id, S. H. (2018). Direct fishing and eDNA metabarcoding for biomonitoring during a 3-year survey significantly improves number of fish detected around a South East Asian reservoir. *PLoS ONE*, 3(12): e0208592. doi: 10.1371/journal.pone.0208592

Gomes, L. C., Pessali, T. C., Sales, N. G., Pompeu, P. S., and Carvalho, D. C. (2015). Integrative taxonomy detects cryptic and overlooked fish species in a neotropical river basin. *Genetica*, 143(5), 581-588. doi: 10.1007/s10709-015-9856-z

GFT. (2015). Avaliação dos efeitos e desdobramentos do rompimento da Barragem de Fundão em Mariana-MG. Available at: www.agenciaminas.mg.gov.br/ckeditor_assets/attachments/770/relatorio_final_ft_03_02_2016_15h5min.pdf

Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., et al. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution*, 7(11), 1299–1307. doi: 10.1111/2041-210X.12595

Gunzburger, M. S. (2007). Evaluation of seven aquatic sampling methods for amphibians and other aquatic fauna. *Applied Herpetology*, 4(1), 47–63. doi: 10.1163/157075407779766750

Holman, L. E., de Bruyn, M., Creer, S., Carvalho, G., Robidart, J., and Rius, M. (2018). The detection of novel and resident marine non-indigenous species using environmental iDNA metabarcoding of seawater and sediment. *bioRxiv*, 440768. doi: 10.1101/440768

Hou, G., Chen, W. T., Lu, H. S., Cheng, F., and Xie, S. G. (2018). Developing a DNA barcode library for perciform fishes in the South China Sea: Species identification, accuracy and cryptic diversity. *Molecular Ecology Resources*, 18, 137–146. doi: 10.1111/1755-0998.12718

Hubert, N., Hanner, R., Holm, E., Mandrak, N. E., Taylor, E., Burridge, M., et al. (2008). Identifying Canadian freshwater fishes through DNA barcodes. *PLoS ONE*, 3, e2490. doi: 10.1371/journal.pone.0002490

Hunter, M. E., Oyler-McCance, S. J., Dorazio, R. M., Fike, J. A., Smith, B. J., Hunter, C. T., et al. (2015). Environmental DNA (eDNA) sampling improves occurrence and detection estimates of invasive Burmese pythons. *PLoS ONE*, 10(4), e0121655. doi: 10.1371/journal.pone.0121655

Hurst, G., and Jiggins, F. M. (2005). Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society B: Biological Sciences*, 272, 1525-1534. doi: 10.1098/rspb.2005.3056

Iyiola, O. A., Nneji, L. M., and Mustapha, M. K. (2018). DNA barcoding of economically important freshwater fish species from north-central Nigeria uncovers cryptic diversity. *Ecology and Evolution*, 8, 6932–6951. doi: 10.1002/ece3.4210

Jerde, C. L., Mahon, A. R., Chadderton, W. L., and Lodge, D. M. (2011). "Sight-unseen" detection of rare aquatic species using environmental DNA. *Conservation Letters*, 4(2), 150-157. doi: 10.1111/j.1755-263X.2010.00158.x

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

Jeunen, G.-J., Knapp, M., Spencer, H. G., Taylor, H. R., Lamare, M. D., Stat, M., et al. (2019). Species-level biodiversity assessment using marine environmental DNA metabarcoding requires protocol optimization and standardization. *Ecology and Evolution*, 1–13. doi: 10.1002/ece3.4843

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

Knudsen, S. W., Ebert, R. B., Hesselsøe, M., Kuntke, F., Hassingboe, J., Mortensen, P. B., et al. (2019). Species-specific detection and quantification of environmental DNA from marine fishes in the Baltic Sea. *Journal of Experimental Marine Biology and Ecology*, 51031–45. doi: 10.1016/j.jembe.2018.09.004

Koziol, A., Stat, M., Simpson, T., Jarman, S., DiBattista, J. D., Harvey, E. S., et al. (2018). Environmental DNA metabarcoding studies are critically affected by substrate selection. *Molecular Ecology Resources*, 0–2. doi: 10.1111/1755-0998.12971

Krausmann, F., Erb, K.-H., Gingrich, S., Haberl, H., Bondeau, A., Gaube, V., and Searchinger, T. D. (2013). Global human appropriation of net primary production doubled in the 20th century. *Proceedings of the National Academy of Sciences*, 110(25), 10324. doi: 10.1073/pnas.1211349110

Kumar, V. P., Shukla, M., Rajpoot, A., Thakur, M., Nigam, P., Kumar, D., Mehta, A. K., and Goyal, S. P. (2018). DNA barcoding as a tool for robust identification of cervids of India and its utility in wildlife forensics, *Mitochondrial DNA Part B*, 3(1), 250-255. doi: 10.1080/23802359.2018.1438858

Lacoursière-Roussel, A., Howland, K., Normandeau, E., Grey, E. K., Archambault, P., Deiner, K., et al. (2018). eDNA metabarcoding as a new surveillance approach for coastal Arctic biodiversity. *International Journal of Business Innovation and Research*, 17(3), 7763–7777. doi: 10.1002/ece3.4213

Latini, A. O., and Petrere, M. (2004). Reduction of a native fish fauna by alien species: an example from Brazilian freshwater tropical lakes. *Fish Management Ecology*, 11:71-79. doi: 10.1046/j.1365-2400.2003.00372.x

Li, J., Lawson Handley, L. J., Read, D. S., and Hänfling, B. (2018). The effect of filtration method on the efficiency of environmental DNA capture and quantification via metabarcoding. *Molecular Ecology Resources*, 18(5), 1102–1114. doi: 10.1111/1755-0998.12899

Loman, N. J., Misra, R. V., Dallman, T. J., Constantinidou, C., Gharbia, S. E., Wain, J., and Pallen, M. J. (2012). Performance comparison of benchtop high-throughput sequencing platforms. *Nature Biotechnology*, 30(5), 434–439. doi:10.1038/nbt.2198

Lopes, C. M., Sasso, T., Valentini, A., Dejean, T., Martins, M., Zamudio, K. R., and Haddad, C. F. (2017). eDNA metabarcoding: a promising method for anuran surveys in highly diverse tropical forests. *Molecular Ecology Resources*, 17, 904-914. doi:10.1111/1755-0998.12643

Magalhães, A. L. B., and Jacobi, C. M. (2013). Invasion risks posed by ornamental freshwater fish trade to southeastern Brazilian rivers. *Neotropical Ichthyology*, 11(2), 433-441. doi: 10.1590/S1679-62252013005000003.

Majaneva, M., Diserud, O. H., Eagle, S. H. C., Boström, E., Hajibabaei, M. and Ekrem, T. (2018). Environmental DNA filtration techniques affect recovered biodiversity. *Scientific Reports*, 8(1), 1-11. doi: 10.1038/s41598-018-23052-8

McDevitt, A. D., Sales, N. G., Browett, S. S., Sparnenn, A. O., Mariani, S., Wangensteen, O. S., Coscia, I., and Benvenuto, C. (2018). Environmental DNA metabarcoding as an effective and rapid tool for fish monitoring in canals. *bioRxiv*. doi: 10.1101/498451.

Metcalf, J. L, Pritchard, V. L, Silvestri, S. M, Jenkins, J. B, Wood, J. S, and Cowley, D. E. (2007). Across the great divide: genetic forensics reveals misidentification of endangered cutthroat trout populations. *Molecular Ecology*, 16, 4445-4454. doi: 1 0.1111/j.1365-294X.2007.03472.x

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., et al. (2015). MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes. *Royal Society Open Science*, 2, 1-8. doi: 10.1098/rsos.150088

Neves, A. C. O., Nunes, F. P., Carvalho, F.A., and Fernandes, G. W. (2016). Neglect of ecosystems services by mining, and the worst environmental disaster in Brazil. *Natureza & Conservação*, 14, 24-27. doi: 10.1016/j.ncon.2016.03.002

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

Nwani, C. D., Becker, S., Braid, H. E., Ude, E. F., Okogwu, O. I., and Hanner, R. (2011). DNA barcoding discriminates freshwater fishes from southeastern Nigeria and provides river system-level phylogeographic resolution within some species. *Mitochondrial DNA*, 22, 43–51. doi: 10.3109/19401736.2010.536537

Ogram, A., Sayler, G. S., and Barkay, T. (1987). The extraction and purification of microbial DNA from sediments. *Journal of Microbial Methods*, 7, 57-66. doi: 10.1016/0167-7012(87)90025-X

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

Pelicice, F. M., Vitule, J. R. S., Lima Junior, D. P., Orsi, M. L., and Agostinho, A. A. (2014). A serious new threat to Brazilian freshwater ecosystems: the naturalization of nonnative fish by decree. *Conservation Letters*, 7, 55-60. doi: 10.1111/conl.12029

Pereira, L. H., Pazian, M. F., Hanner, R., Foresti, F., and Oliveira, C. (2011). DNA barcoding reveals hidden diversity in the Neotropical freshwater fish *Piabina argentea* (Characiformes: Characidae) from the Upper Parana Basin of Brazil. *Mitochondrial DNA*, 22, 87-96. Doi: 10.3109/19401736.2011.588213

Pereira, E. H. L., Pessali, T. C., Andrade, F., and Reis, R. E. (2017). Description of a new species of Pareiorhaphis (Loricariidae: Neoplecostominae) from the rio Jequitinhonha basin, Minas Gerais, eastern Brazil. *Neotropical Ichthyology*, 15(3). doi: 10.1590/1982-0224-20170007

Piaggio, A. J., Engeman, R. M., Hopken, M. W., Humphrey, J. S., Keacher, K. L., Bruce, W. E., et al. (2013). Detecting an elusive invasive species: a diagnostic PCR to detect Burmese python in Florida waters and an assessment of persistence of environmental DNA. *Molecular Ecology Resources*, 14 (2), 374-380. doi: 10.1111/1755-0998.12180.

Pie, M. R., Ströher, P. R., Agostinis, A. O., Belmonte-Lopes, R., Tadra-Sfeir, M. Z. (2017). Development of a real-time PCR assay for the detection of the golden mussel (*Limnoperna fortunei*, Mytilidae) in environmental samples. *Anais da Academia Brasileira de Ciências*, 89(2), 1041-1045. doi: 10.1590/0001-3765201720160723

Pilliod, D. S., Goldberg, C. S., Arkle, R. S., and Waits, L. P. (2014). Factors influencing detection of eDNA from a stream-dwelling amphibian. *Molecular Ecology Resources,* 14(1), 109–116. doi: 10.1111/1755-0998.12159

Pompeu, P. S., Agostinho, A. A., and Pelicice, F. M. (2012). Existing and future challenges: the concept of successful fish passage in South America. *River Research and Applications*, 28, 504-512. doi: 1 0.1002/rra.

Prosdocimi, F., Carvalho, D., De Almeida, R., and Beheregaray, L. (2012). The complete mitochondrial genome of two recently derived species of the fish genus Nannoperca (Perciformes, Percichthyidae). *Molecular Biology Reports*, 39, 2767-2772. doi: 10.1007/s11033-011-1034-5.

Pugedo, M. L., Andrade-Neto, F. R., Pessali, T. C., Birindelli, J. L. O., and Carvalho, D. C. (2016). Integrative taxonomy supports new candidate fish species in a poorly studied neotropical region: the Jequitinhonha River Basin. *Genetica*, 144, 341-349. doi: 10.1007/s10709-016-9903-4

Ramirez, J. L., Carvalho-Costa, L. F., Venere, P. C., Carvalho, D. C., Troy, W. P. and Galetti, P. M. (2016), Testing monophyly of the freshwater fish Leporinus (Characiformes, Anostomidae) through molecular analysis. *Journal Fish Biology*, 88: 1204-1214. doi:10.1111/jfb.12906

Ramirez, J. L., Birindelli, J. L., Carvalho, D. C., Affonso, P. R. A. M., Venere, P. C., Ortega, H., et al. (2017). Revealing hidden diversity of the underestimated Neotropical ichthyofauna: DNA Barcoding in the recently described genus Megaleporinus (Characiformes: Anostomidae). *Frontiers in Genetics*, 8, 149. doi: 10.3389/fgene.2017.00149

Rees, H. C., Maddison, B. C., Middleditch, D. J. Patmore, J. R. M., and Gough, K. C. (2014). The detection of aquatic animal species using environmental DNA – a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology*, 5, 1-11. doi: 10.1111/1365-2664.12306

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

Rejmánek, M. and Pitcairn, M.J. (2002). When is eradication of exotic pest plants a realistic goal? In Turning the Tide: the Eradication of Invasive Species (Veitch, C.R. and Clout, M.N., eds), pp. 249–253, IUCN SSC Invasive Species Specialist Group

Rhymer, J. M, and Simberloff, D. (1996). Extinction by hybridisation and introgression. *Annual Review of Ecology, Evolution, and Systematics*, 27, 83-109. doi: 10.1146/annurev.ecolsys.27.1.83

Riaz, T., Shehzad, W., Viari, A., Pompanon, F., Taberlet, P., and Coissac, E. (2011). ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, 39(21), e145. doi: 10.1093/nar/gkr732

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

Sales, N. G., Pessali, T. C., Andrade-Neto, F., and Carvalho, D. C. (2017). Introgression from non-native species unveils a hidden threat to the neotropical fish *Prochilodus hartii*. *Biological Invasions*, . doi: 10.1007/s10530-017-1556-4. doi: 10.1007/s10530-017-1556-4

Santos, U. , Völcker, C. M., Belei, F. A., Cioffi, M. B., Bertollo, L. A., Paiva, S. R., et al. (2009). Molecular and karyotypic phylogeography in the Neotropical *Hoplias malabaricus* (Erythrinidae) fish in eastern Brazil. *Journal of Fish Biology*, 75, 2326-2343. doi:10.1111/j.1095-8649.2009.02489.x

Sasso, T., Lopes, C. M., Valentini, A., Dejean, T., Zamudio, K. R., Haddad, C. F. B., et al. (2017). Environmental DNA characterization of amphibian communities in the Brazilian Atlantic forest: Potential application for conservation of a rich and threatened fauna. *Biological Conservation*, 215, 225–232. doi: 10.1016/j.biocon.2017.09.015

Schmitz and Simberloff, D. (1997). Biological invasions: a growing threat. *Issues in Science and Technology*, 33-40.

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

Shaw, J. L. A., Clarke, L. J., Wedderburn, S. D., Barnes, T. C., Weyrich, L. S., and Cooper, A. (2016). Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. *Biological Conservation*, 197, 131–138. doi: 10.1016/j.biocon.2016.03.010

Simberloff, D., Martin, J. L., Genovesi, P., Maris, V., Wardle, D. A., Aronson, J., et al. (2013). Impacts of biological invasions: what's what and the way forward. *Trends in Ecology and Evolution*, 28. doi: 10.1016/j.tree.2012.07.013

Slatko, B. E., Gardner, A. F., and Ausubel, F. M. (2018). Overview of Next-Generation Sequencing Technologies. *Current Protocols in Molecular Biology*, 122(1), e59. doi:10.1002/cpmb.59

Spens, J., Evans, A. R., Halfmaerten, D., Knudsen, S. W., Sengupta, M. E., Mak, S. S. T., et al. (2016). Comparison of capture and storage methods for aqueous macrobial eDNA using an optimized extraction protocol: advantage of enclosed filter. *Methods in Ecology and Evolution*, 8, 635-645. doi: 10.1111/2041-210X.12683

Stat, M., John, J., DiBattista, J. D., Newman, S. J., Bunce, M., and Harvey, E. S. (2018). Combined use of eDNA metabarcoding and video surveillance for the assessment of fish biodiversity. *Conservation Biology*, 33(1), 196–205. doi: 10.1111/cobi.13183

Steffen, W., Broadgate, W, Deutsch, L., Gaffney, O., and Ludwig, C. (2015). The trajectory of the Anthropocene: The Great Acceleration. *The Anthropocene Review*, 2, 81–98. doi: 10.1177/2053019614564785

Swarça, A. C., Orsi, M. L., Takagui, F. H., Dias, A. L., Dergam, J. A., and Fenocchio, A. S. (2018). First chromosome data on *Steindachneridion doceanum* (Siluriformes: Pimelodidae): a critically endangered catfish endemic of the Doce River basin, Brazil. *Neotropical Ichthyology*, 16(4). doi: 10.1590/1982-0224-20180066

Taberlet, P., Bonin, A., Zinger, L., and Coissac, E. (2018). Environmental DNA for Biodiversity Research and Monitoring. Oxford University Press. 253p.

Takahara, T., Minamoto, T., and Doi, H. (2013). Using Environmental DNA to Estimate the Distribution of an Invasive Fish Species in Ponds. *PLoS ONE,* 8(2): e56584. doi:10.1371/journal.pone.0056584

Tessler, M., Brugler, M. R., DeSalle, R., Hersch, R., Velho, L. F. M., Segovia, B. T., et al. (2017). A Global eDNA Comparison of Freshwater Bacterioplankton Assemblages Focusing on Large-River Floodplain Lakes of Brazil. *Microbial Ecology*, 73(1), 61–74. doi: 10.1007/s00248-016-0834-5

Thomsen, F. P. and Willerslev, E. (2015). Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, 183, 4-18. doi: 10.1016/j.biocon.2014.11.019

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

Uchii, K., and Minamoto, T. (2016). A novel environmental DNA approach to quantify the cryptic invasion of non-native genotypes. *Molecular Ecology Resources*, 16(2), 415-422. doi: 10.1111/1755-0998.12460

Valentini, A., Pompanon, F., and Taberlet, P. (2009). DNA barcoding for ecologists. *Trends in Ecology and Evolution*, 24(2), 110–117. doi: 10.1016/j.tree.2008.09.011

Valentini, A., Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., et al. (2016). Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology*, 25(4), 929–942. doi: 10.1111/mec.13428

van Straalen, N. M., and Timmermans, M. J. T. N. (2002). Genetic Variation in Toxicant-Stressed Populations: An Evaluation of the "Genetic Erosion" Hypothesis, Human and Ecological Risk Assessment: *An International Journal*, 8:5, 983-1002, doi: 10.1080/1080-700291905783

Vences, M., Lyra, M. L., Perl, R. G. B., Bletz, M. C., Stankovic, D., Lopes, C.M., et al. (2016). Freshwater vertebrate metabarcoding on Illumina platforms using double-indexed primers of the mitochondrial 16S rRNA gene. *Conservation Genetics Resources*, 8, 323-327. doi: 10.1007/s12686-016-0550-y

Viana, L. F., Súarez, Y. R., Cardoso, C. A. L., Crispim, B. do A., Cavalcante, D. N. de C., Grisolia, A. B., et al. (2018). The response of neotropical fish species (Brazil) on the water pollution: metal bioaccumulation and genotoxicity. *Archives of Environmental Contamination and Toxicology*, 5, 43-49. doi:10.1007/s00244-018-0551-9

Vieira, F. (2009). Distribuição, impactos ambientais e conservação da fauna de peixes da bacia do rio Doce. MG BIOTA 2, 5-22.

Vilà, M., Espinar, J. L., Hejda, M., Hulme, P. E., Jarosik, V., Maron, J. Let al. (2011). Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. *Ecology Letters*, 14(7), 702-708. doi: 10.1111/j.1461-0248.2011.01628.x.

Vitousek, P. M., D'Antonio, M. D., and Loope, L. L. (1997). Introduced species: a significant component of human caused global change. *New Zealand Journal of Ecology*, 21, 11-16.

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., and Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philosofical Transactions of the Royal Society of London*, 360(1462), 1847–1857. doi: 10.1098/rstb.2005.1716

Wilcox, T. M., McKelvey, K. S., Young, M. K., Jane, S. F., Lowe, W. H., Whiteley, A. R. et al. (2013). Robust detection of rare species using Environmental DNA: the importance of primer specificity. *PLoS ONE*, 8 (3). doi: 10.1371/journal.pone.0059520

Willerslev, E., Hansen, A. J., Binlanden, J., Brand, R. B., Gilbert, T. P., Shapiro, B., et al. (2003). Diverse plant and animal genetic records from Holocene and Pleistocene sediments. *Science*, 300, 791-795. doi: 10.1126/science.1084114

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

Zawadzki, C. H., Carvalho, P. H., Birindelli, J. O., and Azevedo, F. M. (2016). *Hypostomus nigrolineatus*, a new dark-striped species from the rio Jequitinhonha and rio Pardo basins, Brazil (Siluriformes, Loricariidae). *Ichthyological Exploration of Freshwaters*, 27(3), 263-274.

Ziv, G., Baran, E., Nam, S., Rodriguez, Iturbe, I., and Levin, S. A. (2012). Trading-off fish biodiversity, food security, and hydropower in the Mekong River basin. *PNAS*, 109, 5609-5614. doi: 10.1073/pnas.1201423109

Chapter II

HIDDEN DIVERSITY HAMPERS CONSERVATION EFFORTS IN A HIGHLY IMPACTED NEOTROPICAL RIVER SYSTEM

Chapter II

HIDDEN DIVERSITY HAMPERS CONSERVATION EFFORTS IN A HIGHLY IMPACTED NEOTROPICAL RIVER SYSTEM

Naiara Guimarães Sales[1*], Stefano Mariani[1], Gilberto Nepomuceno Salvador[2], Tiago Casarim Pessali[3], Daniel Cardoso Carvalho[4]

NS performed the molecular genetic analyses and drafted the manuscript. GS and TP collected the samples, conducted the morphological analyses and contributed to the correction of the text. DC designed and coordinates the study. DC and SM conceived the study, participated in its elaboration and helped to draft the manuscript. All authors read and approved the final manuscript.

Journal: Frontiers in Genetics
Published: 24 July 2018
doi: 10.3389/fgene.2018.00271

2.1 Abstract

Neotropical Rivers host a highly diverse ichthyofauna, but taxonomic uncertainty prevents appropriate conservation measures. The Doce River Basin (DRB), lying within two Brazilian threatened hotspots (Atlantic Forest and Brazilian Savanna) in south-east Brazil, faced the worst ever environmental accident reported for South American catchments, due to a dam collapse that spread a toxic mining tailing along the course of its main river. Its ichthyofauna was known to comprise 72 native freshwater fish species, of which 13 are endemic. Here, we build a DNA barcode library for the DRB ichthyofauna, using samples obtained before the 2015 mining disaster, in order to provide a more robust biodiversity record for this basin, as a baseline for future management actions. Throughout the whole DRB, we obtained a total of 306 barcodes, assigned to 69 putative species (with a mean of 4.54 barcodes per species), belonging to 45 genera, 18 families and 5 orders. Average genetic distances within species, genus and families were 2.59%, 11.4% and 20.5% respectively. The 69 species identified represent over 76% of the known DRB ichthyofauna, comprising 43 native (five endemic, of which three threatened by extinction), 13 already known introduced species, and 13 unknown species (such as *Characidium* sp., *Neoplecostomus* sp. and specimens identified only at the sub-family level Neoplecostominae, according to morphological identification provided by the museum collections). Over one fifth of all analyzed species (N=16) had a mean intraspecific genetic divergence higher than 2%. An integrative approach, combining NND (nearest neighbor distance), BIN (barcode index number), ABGD (automatic barcode gap discovery) and bPTP (Bayesian Poisson Tree Processes model) analyses, suggested the occurrence of potential cryptic species, species complex, or historical errors in morphological identification. The evidence presented calls for a more robust, DNA-assisted cataloguing of biodiversity-rich ecosystems, in order to enable effective monitoring and informed actions to preserve and restore these delicate habitats.

Keywords: barcode, biodiversity, cryptic diversity, Doce River, ichthyofauna, molecular identification

2.2 Introduction

Neotropical rivers host an extremely diverse ichthyofauna, but anthropogenic impact associated with the occurrence of many still undescribed or unknown species may hamper conservation effort (Reis et al., 2016; Ely et al., 2017). Due to increasing, rapid anthropogenic environmental impacts (e.g. pollution, siltation, mining, damming), biodiversity in Neotropical rivers may be lost before scientists can fully describe and comprehend it (Agostinho et al., 2005).

Effective biodiversity conservation relies on unequivocal and precise species identification, especially in the case of ecosystems that underwent degradation and require restoration. However, high biodiversity regions, such as the Neotropics, and the increasingly reduced budget for basic taxonomical research, have led to the so-called "taxonomic impediment" or "poor taxonomy", in which the shortage of funding and trained taxonomists, and the gaps in taxonomic knowledge, have delayed advances in assessment and description of biodiversity or even contributed to overestimate or underestimate species richness due to species misidentification or taxonomic confusions (Taylor, 1983; Ely et al., 2017).

The DNA barcoding initiative offers a powerful and cost-effective tool to assist with the detection of cryptic species and flag potentially problematic taxa, with the standard universal COI marker having proven particularly successful in invertebrates (Hebert et al., 2004a), birds (Hebert et al., 2004b), and fish (Ward et al., 2005; Hubert et al., 2008; Valdez-Moreno et al., 2009; Rosso et al., 2012; Carvalho et al., 2011). For effective DNA barcode performance, intraspecific variability must be lower than variability among congeneric species, the so-called

'Barcode Gap' (Meyer and Paulay, 2005). While the barcode gap tends to be around <1-2% sequence variability within species in most fish, there are exceptions (Hurst and Jiggins, 2005), especially in the case of recently diverged species (Vinas and Tudela, 2009; Shum et al., 2017). Moreover, the unambiguous identification of species from early larval stage to adulthood can aid a variety of conservation management actions. Accurate molecular identification may contribute to improving management and sustainability of long term fisheries (Metcalf et al., 2007), tracking invasive species (Corin et al., 2007; Carvalho et al., 2009), offer insights into community ecology (Pfenninger et al., 2007) and genetic certification of species used in restocking programs (Metcalf et al., 2007), as well as improving fundamental knowledge on cryptic and putatively new species (Pereira et al., 2011). Furthermore, molecular identification of eggs and larvae can provide data regarding spawning and recruitment areas, supporting a definition of priority areas for conservation (Becker et al., 2015; Frantine-Silva et al., 2015).

DNA barcode libraries have been developed for several Neotropical river systems as a biodiversity identification tool and have contributed to reveal the existence of putatively cryptic/new fish species (Carvalho et al., 2011; Pereira et al., 2011; Gomes et al., 2015; Pugedo et al., 2016; Nascimento et al., 2016). However, the biodiversity complexity remains unknown in many already impacted catchments in Brazil. One emblematic case is that of the Doce River Basin (DRB), which faced the worst environmental accident reported for any South American catchment, in the form of the largest tailings dam burst in modern history; as a result, a toxic mud (i.e. extreme high concentration of iron) spread along its main river course, affecting wild

communities, as well as the local human populations (Fernandes et al., 2016; Neves et al., 2016). As the local riverine human communities rely on fisheries for their livelihood (e.g. source of income and subsistence, resource for ecotourism), understanding the impacts of this disaster on the ichthyofauna is crucial for effective management actions (Ecoplan-Lume, 2010; GFT, 2015; Neves et al., 2016). Moreover, the recovery of fish populations in DRB, after the ecological disaster, relies on the recolonization of the main course of this river and on the diversity, size, and conservation status of the remnant fish populations in the tributaries (Olds et al., 2012).

The Doce River Basin runs through two Brazilian biodiversity hotspots (Atlantic forest and Brazilian Savanna) located in south-east Brazil (Myers et al., 2000). The river is 853 km long and the catchment covers a total drainage area of 83.400 km² in south-eastern Brazil, between the states of Minas Gerais (86%) and Espírito Santo (14%), an area inhabited by 3 million people. DRB harbors a rich ichthyofauna, including several undescribed species, with the number of presently recognized native species summing up to 72 (Vieira, 2009). The Santo Antônio River, the second largest tributary of the Doce, was selected as a conservation priority area, since it hosts a great number of species considered endemic and threatened by extinction (Vieira et al., 2000; Vieira and Alves, 2001; Rosa and Lima, 2005). Historically, DRB is affected by human impacts by many ways. Native forest cover only 27% of DRB area (ANA, 2016), and the remained area is used to cattle, forestry, agriculture, and mining (Vieira, 2009), resulting in high rate of siltation (da Silva et al., 2011). Habitat fragmentation lead by hydroelectric construction is also affect DRB, where there are 40 hydroelectric built along

main channel of Doce River and its principal tributaries (ANEEL, 2010). However, without accurate biodiversity knowledge, species conservation may be hindered in this river system, and it had already been suggested that the environmental disaster involving the mining collapse could have led to the depletion/extinction of many still unknown endemic species (Fernandes et al., 2016). Here, we develop a DNA barcode library for the DRB ichthyofauna, using data obtained prior to the dam burst environmental disaster, contributing to an improved biodiversity baseline record for this recently impacted ecosystem.

2.3 Material and Methods

2.3.1 Sampling

We obtained fish tissue samples from 306 specimens collected between 2011 and 2015 along the main river channel and tributaries (Figure 1), identified and deposited by taxonomists in four Brazilian ichthyological collections: PUC Minas Natural History Museum (MCNIP), Museu de Biologia Professor Mello Leitão (MBML), Museu de Zoologia da Universidade Estadual de Campinas (ZUEC), and Núcleo de Pesquisas em Limnologia, Ictiologia e Aquicultura (NUPELIA). All analyzed specimens were photographed, geo-referenced, and identified to the lowest taxonomic level from identification keys or previously published works (Vari, 1992; Albert et al., 1999; Castro and Vari, 2004; Zanata and Camelier, 2009).

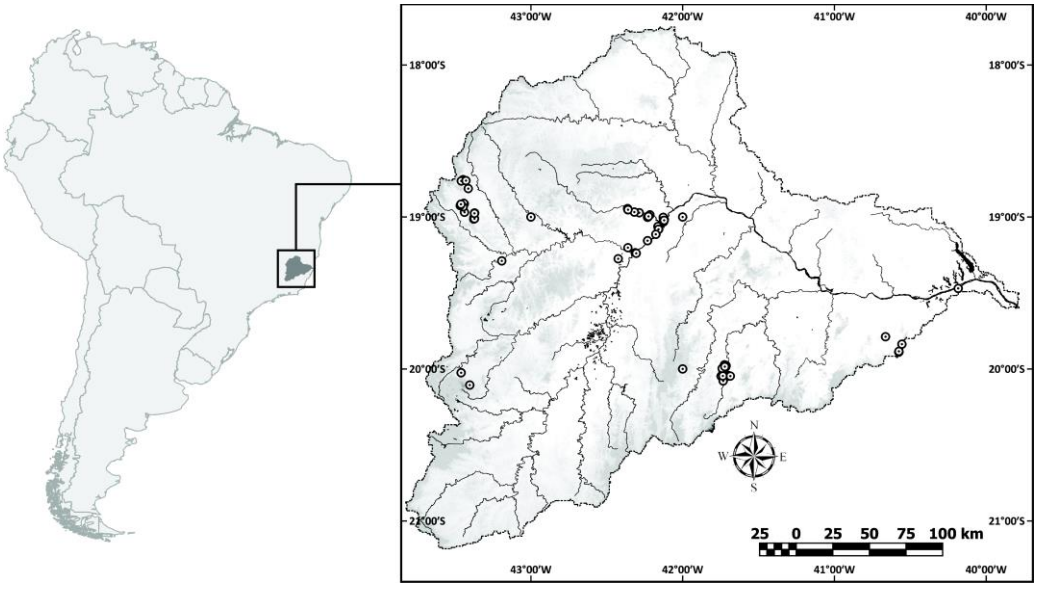


FIGURE 1 | Map of Doce River Basin, including sample sites distribution.

2.3.2 Ethics Statement

All fish analyzed in this study were collected in accordance with Brazilian legislation (Collection license 6421-1, number 5498740) or obtained from Ichthyological collections. Fish were collected, and euthanized, samples of fins were clipped from each individual and stored in absolute ethanol for subsequent molecular analysis. Specimens were fixed in 10% formaldehyde and then stored in 70% ethanol.

2.3.3 DNA extraction, amplification and sequencing

Genetic analyses were conducted, whenever possible, on a minimum of five specimens from different sample sites per species. DNA extraction followed the salting out protocol (adapted from Aljanabi and Martinez (1997)). The cytochrome c oxidase I (COI) gene (~650bp) was amplified by polymerase chain reaction (PCR) using the primers FishF1/FishR1 described by Ward et al. (2005) and the Cocktail COI-3/C_FishF1t1-C_FishR1t1 described by Ivanova et al. (2007), and following the PCR protocol described in Gomes et al. (2015). The PCR products were visualized on 1% agarose gel, alongside negative controls and a size ladder, and positive amplifications were selected for DNA sequencing. DNA sequencing was conducted in both directions in an automated DNA analyzer ABI 3500 (Life Technologies).

2.3.4 Data Analysis

Barcode sequences were edited using DNA Baser® v.3.5.4 (DNA Sequence Assembler v4 (2013), Heracle BioSoft, www.DnaBaser.com) and SeqScape v.2.1.1 (Applied Biosystems,

Foster City, CA, USA) (Díaz et al., 2016) softwares. DNA alignment was conducted using the CLUSTAL W alignment tool (Thompson et al., 1997). The neighbor-joining (NJ) trees (Saitou and Nei, 1987) and genetic distances estimations, using the K2P (Kimura-2-parameter) nucleotide evolution model (Kimura, 1980) were generated using MEGA 7 software (Kumar et al., 2016).

Intra- and inter-specific genetic distances, nearest neighbor distance (NND), and the barcode gap were calculated in the on-line BOLD Workbench (http://www.boldsystems.org) (Ratnasingham and Hebert, 2007). The nearest neighbor distance (NND) was used to estimate the minimum genetic distance between pairs of species. Different approaches were used to delimitate the Molecular Operational Taxonomic Units (MOTUs), two clustering algorithms (BIN and ABGD) and one phylogenetic-coalescent methods (bPTP). The Barcode Index Number (BIN) (Ratnasingham and Hebert, 2013) was estimated automatically in BOLD Workbench and allowed comparing DNA barcodes obtained here with other river basins that have a comprehensive DNA Barcode library such as the São Francisco, the Mucuri, the Jequitinhonha, the Paraná, and the Paranaíba River Basins (Carvalho et al., 2011; Pereira et al., 2011; Gomes et al., 2015; Díaz et al., 2016; Pugedo et al., 2016). Using this approach, it is possible to identify endemic lineages and shared ichthyofauna. Automatic Barcode Gap Discovery (ABGD) analyses (Puillandre et al., 2012) were performed using the web interface (http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html, web version 'May 31 2017') with a relative gap width value of X=1.0 and two available distance metrics [JC69 (Jukes and Cantor, 1969) and K2P (Kimura, 1980)], while the other parameter values employed default settings.

The Bayesian Poisson Tree Processes model (bPTP) was conducted using both ML (maximum likelihood) and Bayesian approaches (Zhang et al., 2013). The PTP file input consisted in a nexus tree generated in MrBayes (Ronquist and Huelsenbeck, 2003) using six random parsimony trees, with the GTRGAMMA substitution model (obtained by MEGA 7 under BIC criteria), without rooting and applying the parameters of 20 million MCMC generations and a burn-in of 10%. Analysis was conducted applying default values through the bPTP server (500,000 generations, thinning = 100, burn-in= 10%).

All data including fish photos, GPS coordinates of each sample site, vouchers numbers, detailed taxonomic identifications, and the corresponding sequence data and trace files were submitted to the Barcode of Life Data System (BOLD, http://www.boldsystems.org, see Ratnasingham and Herbert, 2007) within the project file 'DNA Barcoding of Doce River Basin'.

2.3.5 Species delimitation and hidden biodiversity

Species delimitation based on integrative approaches that combine a diverse range of statistical methods has been extensively used to identify hidden biodiversity (i.e. Padial et al. 2010; Costa-Silva et al., 2015, Gomes et al., 2015; Rossini et al., 2016, Ramirez et al., 2017). Here, species with >2% of intraspecific genetic divergences, still undescribed or unknown and identified only at genus or family level were investigated individually to detect the occurrence of new molecular operational taxonomic units (MOTUs) according to the congruence among BIN, ABGD, bPTP outputs.

Undescribed species or those only identified at genus or family level were checked

using the BIN and NND analyses in order to verify their occurrence in clusters composed by other nominal species, and their genetic divergence from the nearest neighbor (including species from DRB and/or distinct Brazilian basins). Were considered as new MOTUs when intraspecific genetic divergence was higher than 2% for described species and distinguished clusters identified by BIN, ABGD and bPTP outputs.

2.4 Results

Morphological identification on the 306 specimens yielded 69 species (see Table S1 in of which 43 are native species (five endemic, three threatened by extinction and one endemic and threatened), 13 non-native species and 13 new records to the DRB (see Table S2 in the Supplementary material), representing over 76% of its known freshwater ichthyofauna (Vieira, 2009). We then obtained 306 partial sequences of the COI gene, consisting of 665 bp on average, and no insertions, deletions, or stop codons were detected, indicating that there was no case of NUMTS (Nuclear mitochondrial DNA sequences) (Bensasson et al., 2001).

A mean of 4.54 individuals per species were sequenced, comprising 45 genera, 18 families, and 5 orders (Characiformes (41.9%), Siluriformes (40.6%), Perciformes (9.4%), Gymnotiformes (4.7%), Cyprinodontiformes (3.4%)). Species represented by one or two specimens (N=19) were not included in the estimation of intraspecific divergences (*Callichthys callichthys*, *Cichla kelberi*, *Clarias gariepinus*, *Hoplosternum littorale*, *Hyphessobrycon bifasciatus*, *H. eques*, *Hypostomus* sp., *Lophiosilurus alexandri*, *Metynnis maculatus*, *Parotocinclus maculicauda*, *Pimelodus maculatus*, *Poecilia vivipara*, *Prochilodus vimboides*, *Pygocentrus nattereri*, *Salminus brasiliensis*, *Steindachneridion doceanum*, *Trichomycterus* aff. *Auroguttatus*, *T*. cf. *brasiliensis* and *T*. *longibarbatus*). The NJ tree identified species-specific clades for 80.9% of all species. The mean genetic distances found within species, genera and families were: 2.59%, 11.4% and 20.5% (Table 1), respectively. Over 65% of the analyzed species showed genetic distances lower than 1% and for 70% of the species the divergence value was below 2% (Figure 2A). When considering intra-generic distance, 19% of

the species had a divergence higher than 20% (Figure 2B), suggesting the possibility of taxonomic errors or cryptic species.

TABLE 1 | Distance summary reports for sequence divergence between species, genus and family level including minimum, mean and maximum genetic distances (K2P).

	Minimum distance (%)	Mean distance (%)	Maximum distance (%)
Within species	0	2.59	21.82
Within genera	0	11.4	24.2
Within families	0	20.5	30.99

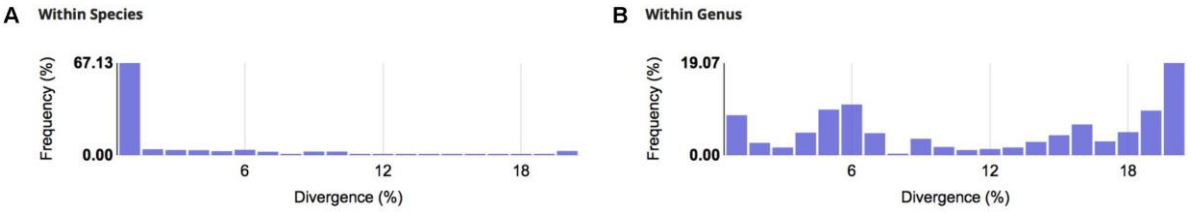


FIGURE 2 | Genetic divergences found for all sequences analyzed at species **(A)** and genus **(B)** levels.

2.4.1 Intra- and inter-specific divergence

Intraspecific distance varied from 0% to 21.82%. Particularly high genetic distances (>10%) were recovered among specimens of *Astyanax fasciatus* (20.69%), *Astyanax scabripinnis* (21.82%), *Astyanax* sp. (20.5%), *Characidium* sp. (10.17%), *Crenicichla lacustris* (21.36%), *Harttia* sp. (12.2%), *Poecilia reticulata* (14.34%) and *Trichomycterus* aff. *Alternatus*

(18.49%), flagging possible new MOTUs (i.e. hidden diversity) or problems related with taxonomic morphological identification.

Incongruences between morphological and barcode identifications (BIN, ABGD, bPTP) (i.e. one BIN/ABGD/bPTP cluster containing more than one morphological species, morphological species represented by more than one BIN/ABGD/bPTP cluster, and/or >2% of intraspecific genetic distance and <1% of interspecific divergence) were observed within species of the genus *Astyanax, Characidium*, *Crenicichla, Deuterodon*, *Gymnotus*, *Harttia*, *Hoplias*, *Hyphessobrycon*, *Hypostomus*, *Knodus*, *Neoplecostomus*, *Oligosarcus*, *Pareiorhaphis*, *Poecilia*, *Prochilodus*, *Rhamdia*, and *Trichomycterus* (Table S1 in the Supplementary material).

The NJ tree encompassing all species showed the occurrence of monophyletic clades and absence of shared haplotypes for 44 of the 69 analyzed species. The interspecific genetic distance showed that 63.2% of the analyzed species had a K2P divergence higher than 2% to their closest neighbor, with the exception of: *Astyanax* spp., *Deuterodon pedri*, *Hyphessobrycon eques*, *Characidium* sp. and *Characidium* gr. *timbuiense*, *Gymnotus* spp., *Oligosarcus argenteus* and *O. acutirostris*, *Poecilia reticulata* and *Poecilia vivipara,* and *Trichomycterus* aff. *Alternatus* and *T. longibarbatus* (Figure S1).

2.4.2 Identification of Molecular Operational Taxonomic Units (MOTUs)

The BIN analysis identified 81 clusters, including 48 taxonomically concordant, 17 discordant, and 16 singletons. The ABGD analysis detected 54-133 MOTUs when varying the prior maximal distance from P = 0.001 to P = 0.1000 (applying both the K2P and JC69

nucleotide evolution methods). The partition that recovered 81 groups (intraspecific distance P = 0.0077) was chosen due to its consistency with our BIN analysis. The bPTP analyses (Bayesian and ML approaches) resulted in the same number of clusters obtained by BIN, except for *Harttia* sp. (three BIN and ABGD clusters and one bPTP) and *Prochilodus costatus* (two BIN, and one ABGD and bPTP clusters). ABGD species delineation was in agreement with all the BIN clusters with the following exceptions, which contain more than one BIN for each morpho-species: *Astyanax scabripinnis* (BIN: AAC5910 and ACO5464, ABGD: 36 and 81), *Knodus moenkhausii* (BIN: AAM1485, ABGD: 46 and 49), *Prochilodus costatus* (BIN: ADC2568 and ADC2571, ABGD: 10), *Trichomycterus* sp./*Trichomycterus* aff. *Alternatus*/*Trichomycterus* aff. *Auroguttatus*/*Trichomycterus longibarbatus* (BIN: ACJ1164 and ACJ1161, ABGD: 64), *Trichomycterus* sp./*Trichomycterus* cf. *brasiliensis* (BIN: ACK5393 and ACT6325, ABGD: 65) (Table S2).

2.4.3 Identification of hidden biodiversity

Sequences from fifteen undescribed species or identified only at genus or family level were compared to other species available in BOLD database through NND and BIN analyses (Table 2). Within undescribed or unknown species, we recovered 9 new MOTUs from the following genera: *Astyanax*, *Characidium, Gymnotus, Harttia, Hisonotus, Neoplecostomus, Pareiorhaphis, Phalloceros* and *Trichomycterus*. The other six species were not considered new MOTUs (*Brycon* sp., *Hasemania* sp., *Hypostomus* sp., *Imparfinis* sp., Neoplecostominae

and *Pimelodella* sp.) since they were included in BINs composed by another nominal species and showed interspecific divergence <2% with the nearest neighbor.

Among species with deep intraspecific divergence (>2%) we recovered additionally at least 3 putative cryptic species (i.e. species containing low levels of morphological/phenotypic in contrast to their high genetic differentiation – Struck et al., 2018)due to the congruence among BIN, ABGD, bPTP and genetic distance methods for *Crenicichla lacustris*, *Hoplias malabaricus* and *Rhamdia* cf. *quelen* (Table 3). *Astyanax fasciatus* and *A. scabripinnis* despite showing a congruence of BIN and ABGD analyses were included in clusters comprising another species of the genus. *Knodus moenkhausii* had a maximum intraspecific divergence of 3.07% and two distinct ABGD numbers, however, only one clade and one BIN was recovered for this species.

TABLE 2 | List of undescribed species including the nearest neighbor, BIN and genetic similarity (%).

Species	NND (nearest neighbor species)	BIN	BIN Classification	Maximum similarity (%)
Astyanax sp.	*Astyanax fasciatus*	AAC5910	Discordant	99.2
	Deterodon pedri	ACJ9650	Discordant	99
	Astyanax intermedius	ACT0040	Singleton	93.7
	Astyanax fasciatus, A. bockmanni	AAY4812	Discordant	99.2
Brycon sp.	*Brycon ferox*	ACH8616	Concordant	100
Characidium sp.	*Characidium* sp.	ACS9348	Concordant	100
	Characidium cf. *timbuiense*	ACJ1226	Discordant	100
	Characidium cf. *timbuiense*	ACI3743	Discordant	100
Gymnotus sp.	*Gymnotus carapo*	AAB6216	Discordant	100
	Gymnotus sylvius	AAB6212	Concordant	100
	Gymnotus sp.	ACT0768	Concordant	100
Harttia sp.	*Harttia sp.*	ACJ1000	Singleton	100
	Harttia sp.	ACI6845	Concordant	100
	Harttia sp.	ACO6155	Singleton	100
Hasemania sp.	*Hasemania hanseni*	AAO6055	Concordant	100
Hisonotus sp.	*Hisotonus* sp.	ACW1732	Concordant	100
Hypostomus sp.	*Hypostomus auroguttatus*	AAB9690	Discordant	100
	Hypostomus heraldoi			98.52
	Hypostomus luetkeni			99.51
	Hypostomus strigaticeps			99.01
Imparfinis sp.	*Imparfinis minutus*	AAC2103	Concordant	99.28
	Imparfinis mirini			98.98
Neoplecostomus sp.	*Neoplecostomus* sp.	AAX6581	Concordant	100
	Neoplecostomus sp.	ACT2675	Concordant	100
Neoplecostominae	*Pareiohaphis* cf. *bahianus*	ACC0721	Concordant	98.3
Pareiorhaphis sp.	*Pareiohaphis scutula*	AAX0824	Discordant	99.8
	Pareiorhaphis sp.	ACI5663	Concordant	100
Phalloceros sp.	*Phalloceros* sp.	AAB7265	Concordant	100
Pimelodella sp.	*Pimelodella lateristriga*	AAC5327	Concordant	99.85
Trichomycterus sp.	*Trichomycterus* aff. *Immaculatus/ T.* cf. *pradensis*	ACI3868	Discordant	99.26
	Trichomycterus aff. *Auroguttatus*	ACJ1164	Discordant	100
	Trichomycterus sp.	ACJ9705	Singleton	98
	Trichomycterus cf. *brasiliensis*	ACK5393	Singleton	98.57
	Trichomycterus cf. *brasiliensis*	ACT6325	Discordant	99.8

TABLE 3 | List of described species with high intraspecific divergence (>2%), showing the maximum and mean intraspecific genetic distance, clades and number of BIN, ABGD and bPTP clusters.

Species	Maximum genetic distance (%)	Mean genetic distance (%)	Clades	BIN	ABGD	bPTP
Astyanax fasciatus	20.69	10.09	3	3	3	3
Astyanax lacustris	3.35	1.67	2	2	2	2
Astyanax scabripinnis	21.82	9.12	2	2	3	2
Astyanax taeniatus	3.96	1.48	2	2	2	2
Characidium sp./ *Characidium* cf. *timbuiense**	10.17/9.9	5.51/5.98	4	4	4	4
*Crenicichla lacustris**	21.36	10.76	2	2	2	2
*Hoplias malabaricus**	6.7	3.27	2	2	2	2
Knodus moenkhausii	3.07	1.21	1	1	2	1
Poecilia reticulata	14.34	9.48	2	2	2	2
Prochilodus costatus	2.6	1.32	2	1	1	2
Rhamdia cf. *quelen* *	3.48	1.25	2	2	2	2
Trichomycterus aff. *Alternatus*	18.49	10.8	2	2	2	2
Trichomycterus aff. *Immaculatus*	5.84	2.23	2	2	2	2

* Occurrence of cryptic species

2.5 Discussion

2.5.1 DNA Barcoding effectiveness

We analyzed 306 fish specimens obtained before the dam burst in 2015 and provided genetic data for the ichthyofauna of the DRB, highlighting the occurrence of cryptic and previously unrecognized biodiversity. Therefore, we significantly extend the knowledge on this river system, whose previous surveys mostly focused on the middle course of the river and in lakes located inside the Doce State Park and its surroundings (Sunaga and Verani, 1987; Vieira, 1994; Vono and Barbosa, 2001; Latini and Petrere, 2004). This baseline offers a more robust platform for any future attempt to restore biodiversity and ecosystem functions to a level comparable to pre-disaster conditions.

Using DNA barcoding, we observed an intraspecific genetic distance considerably higher than previously reported for freshwater fish species from other Brazilian basins. On the other hand, intrageneric divergences were found to be similar to previous studies (Carvalho et al., 2011; Pereira et al., 2011; Pugedo et al., 2016). These results suggest a higher occurrence of hidden biodiversity in DRB when compared to other studied Brazilian basins (Table 4).

TABLE 4 | Comparison among DNA barcoding studies conducted in Brazilian basins, including the number of sequences and species analyzed, and intraspecific and intrageneric distances (Minimum and maximum. The mean is inside the parentheses).

Reference	Basin	Number of sequences	Number of species	Intraspecific distance (%)	Intrageneric distance (%)
Pugedo et al., 2016	Jequitinhonha	260	52	0-11.43 (0.44)	1.09-21.55 (12.16)
Nascimento et al., 2016	Itapecuru	440	64	0-8.9 (0.80)	2.65-7.70 (5.13)
Benzaquem et al. 2015*	Amazon	110	14	0-9.8 (2.8)	2.2-22.5 (19.0)
Gomes et al., 2015	Mucuri	141	37	0-3.24 (0.74)	4.29-18.44 (9.5)
Pereira et al., 2013	Upper Parana	1244	254	0-8.5 (1.3)	0-24.9 (6.8)
Carvalho et al., 2011	São Francisco	431	101	0-10.54 (0.5)	0-22.88 (10.61)
Pereira et al., 2011	Paraíba do Sul	295	58	0-3.48 (0.13)	0.93-22.89 (10.36)
Present study	Doce	306	68	0-21.82 (2.59)	0-24.2 (11.4)

*Only *Nannostomus* spp.

2.5.2 Hidden Biodiversity

DNA barcoding has already been used to reveal hidden biodiversity such as cryptic species and new candidate fish species in the São Francisco (Carvalho et al., 2011), Mucuri (one species – Gomes et al., 2015) and Jequitinhonha (15 species – Pugedo et al., 2016) River catchments. In DRB, from 69 morphologically identified species, the barcode analyses recovered 12 putative cryptic species within *Astyanax* sp., *Characidium* sp., *Characidium* gr. *timbuiense*, *Crenicichla lacustris*, *Gymnotus* sp., *Harttia* sp. (2 putative cryptic species), *Hoplias malabaricus*, *Neoplecostomus* sp., *Rhamdia* cf. *quelen*, *Trichomycterus* sp. (2 putative cryptic species). The high intraspecific genetic distance estimation found for the DRB fish was related to the occurrence of cases of well-known species complexes – e.g. *Astyanax* spp. (maximum intraspecific distance reaching 21.82% in *A. scabripinnis*), *Gymnotus* sp. (6.32%),

Hoplias malabaricus (6.7%), *Rhamdia* cf. *quelen* (3.48%) and also due to the deep intraspecific barcode divergence found to putative overlooked cryptic MOTUs -e.g., *Crenicichla lacustris* (21.36%).

DNA barcoding allows for the identification of cryptic variation among morphologically similar species, indicating the occurrence of more than one species and reinforcing the need of an integrative approach combining molecular and morphological characters (Nascimento et al., 2016). By combining distinct species delimitation methods, we were able to identify new MOTUs from nine undescribed species *(Astyanax* sp*., Characidium* sp*., Gymnotus* sp*., Harttia* sp*., Hisonotus* sp*., Neoplecostomus* sp*., Pareiorhaphis* sp*., Phalloceros* sp. and *Trichomycterus* sp.). Other species showed a high similarity with already described species from another river basins (e.g. specimens of *Brycon* sp. were assigned as *B. ferox* from Mucuri River basin) and were not considered as possible new MOTUs (Table 2) as shown by the BIN analysis.

Among the undescribed species, we were able to highlight new MOTUs within five morpho-species due to their high intraspecific genetic divergence and based on BIN, ABGD and NND analyses. For instance, *Harttia* sp. showed mean divergence of 4.67% and 3 clades which were congruent within the BIN and ABGD clustering methods, suggesting the occurrence of 3 new MOTUs in this genus. Specimens of *Hisonotus* sp. were included in the same BIN/ABGD/bPTP cluster and had an exclusive BIN containing only specimens from DRB suggesting a new MOTU exclusive to this catchment. *Neoplecostomus doceensis* is the only loricariid from this genus described for DRB, however, we found 2 possible cryptic MOTUs

within this taxon, as the DNA barcodes from *Neoplecostomus* sp. did not cluster with barcodes available for this species and had two additional distinct BIN and ABGD clusters. Furthermore, exclusive BIN/ABGD clusters were recovered for *Pareiorhaphis* sp. and *Phallocer*os sp. suggesting at least one new MOTU for each genus endemic to the DRB.

Species with high intraspecific divergence were recovered within *Astyanax* spp. (*A. fasciatus*, *A. lacustris*, *A. scabripinnis,* and *A. taeniatus*). Despite showing a deep intraspecific divergence, and congruence of BIN/ABGD clusters, these species were not considered as comprising new MOTUs due to its high genetic similarity with another nominal species (e.g. *Astyanax parahybae*, *A.vermilion*, *Hyphesobrycon* spp., *Deuterodon* sp.) observed within the BIN and NND analysis, and also, because this highly diverse group is a well known complex of species in need of more systematic studies (Garutti, 1995; Froese and Pauly, 2010; Eschmeyer, 2015).

High intraspecific divergence was also found for *Trichomycterus* aff. *Alternatus* and *T.* aff*. Immaculatus*. These species, despite showing a high intraspecific distance (18.49% and 5.84%, respectively), were included in BINs comprised by another nominal species (e.g. *Trichomycterus longibarbatus*) indicating it may be a case of morphological misidentification and not the occurrence of new MOTUs. This genus has an extensive geographical range and its morphological identification is complex due to the lack of consistent synapomorphies (Barbosa and Costa, 2003). Therefore, further studies combining an integrative approach focusing in these species are required in order to investigate the occurrence of putative cryptic species.

Prochilodus costatus showed a high intraspecific divergence (2.6%) and occurrence of two clusters (NJ and BIN analyses). However, this non-native species was not considered as a putative cryptic species since it was included in BINs comprising another non-native species (e.g. *Prochilodus argenteus*, *P. hartii*). As suggested in previous studies, the incongruence between morphological and molecular identification of *Prochilodus costatus* may indicate the occurrence of *Prochilodus* hybrids and not due to new MOTUs (Gomes et al., 2015; Sales et al., 2018).

Poecilia reticulata is a species introduced worldwide, occuring in more than 69 countries outside of its native range (Deacon et al., 2011). A high intraspecific divergence (14.34%) was found for this species in the Doce River Basin. However, two specimens of *Poecilia reticulata* were assigned to a BIN comprising specimens of *P*. *vivipara* (BIN AAC0279) and the high intraspecific divergence was due to the incongruence between morphological and molecular identification and not due to the occurrence of new MOTUs. Hybridization process between congeneric species of *Poecilia (Poecilia velifera* or *P. petenensis* and *P. mexicana* or *P. orri)* and between different populations of *P. reticulata* have already been reported (Kittell et al., 2005; Lampert and Schartl, 2008; Sievers et al., 2012) and the incongruence detected in this study might be a case of hybridization between *P. reticulata* and *P. vivipara* or misidentification during the deposit in the museum collection and not due to the occurrence of cryptic species.

Hidden biodiversity was found within the genera *Characidium*, *Crenicichla*, *Gymnotus*, *Hoplias* and *Rhamdia* due to high intraspecific genetic divergence and congruence among

clustering methods BIN, ABGD, bPTP (Table 3). For instance, within the genera *Characidium* spp. we detected a mean intraspecific divergence of 5.82% and the occurrence of four clades, of which: two mixed clades comprising specimens identified as *Characidium* gr. *timbuiense* (n=3 and n=4) and *Characidium* sp. (n=1 and n=1), one clade exclusive to *Characidium* gr. *timbuiense* (n=1) and one clade exclusive to *Characidium* sp. (n=4). *Crenicichla lacustris* showed intraspecific divergence of 10.76% and presence of two different clades and BIN/ABGD/bPTP clusters (one for samples collected in Manhuaçu River and one for samples collected below the Baguari Dam). The electric knifefishes *Gymnotus* spp. had an intraspecific divergence above 2% and occurrence of 3 different clades corroborated by 3 BIN, ABGD, and bPTP clusters. All *Gymnotus* specimens were initially morphologically identified as *Gymnotus* sp. and *Gymnotus* cf. *carapo*. However, similarly to the findings obtained for this genus in Mucuri River Basin, these clusters may represent 2 different known species (*Gymnotus carapo* and the overlooked species *Gymnotus sylvius*) and a new MOTU yet to be analyzed and properly described (*Gymnotus* sp.). Two congruent BIN, ABGD, and bPTP clusters were identified for both *Hoplias malabaricus* and *Rhamdia* cf. *quelen* (mean intraspecific divergence of 3.27% and 1.25%, respectively) suggesting the occurrence of cryptic species for each of these taxa. The divergence found in *H. malabaricus* may be due to allopatric speciation resulting from geographical barriers enhanced by its sedentary habitat, since one cluster comprised exclusively specimens from Jose Pedro River and the other was exclusive for specimens from Corrente Grande River. High genetic diversity was already reported for this species in other studied systems (Paraná and Tibagi Rivers) suggesting distinct

evolutionary lineages, population structuring or occurrence of cryptic species (Dergam et al., 1998; Blanco et al., 2011; Oliveira et al., 2015).

Species of the genera *Rhamdia*, *Characidium*, *Pareiorhaphis*, *Gymnotus* were also flagged as cryptic and/or candidate species in other Brazilian basins (Carvalho et al., 2011; Gomes et al., 2015; Pugedo et al., 2016). Furthermore, genetic divergence was associated with the geographic location of some species (e.g. *Crenicichla lacustris*, *Hoplias malabaricus*) suggesting the occurrence of allopatric divergence between these populations.

The increase of available barcodes in BOLD database, including adjacent basins, may contribute to expose endemic cryptic species and reduce the risk of synonymies (Gomes et al., 2015). However, Pugedo et al. (2016) highlighted the concern of using solely DNA barcodes in defining species (e.g. using NND, BIN, ABGD and bPTP analyses) due to the fact that Neotropical DNA barcode libraries are not yet complete. Furthermore, specimens included in BINs composed by different nominal species should be re-evaluated by a taxonomist to verify the data and check for potential misidentifications (Díaz et al., 2016).

Thus, a thorough analysis should be done for each flagged species to verify the correspondence of new MOTUs with putative new candidate species based on accurate morphological taxonomy analysis and to evaluate the divergence causes and the correlation of speciation process to natural or anthropogenic causes (e.g. presence of dams).

2.5.3 Importance of DNA Barcoding library for the Doce River ichthyofauna

This newly developed DNA barcode reference library for the DRB fish detected the occurrence of new MOTUs and suggested the existence of hidden biodiversity. This baseline information will provide a platform for several applications and management efforts such as ichthyoplankton identification for the detection of fish recruitment areas, unambiguous choice of species to be used in restocking programs, and environmental DNA research. This data may contribute as a baseline for restoration programs in this catchment, by pointing out new MOTUs and suggesting the occurrence of overlooked and cryptic species among the DRB ichthyofauna, highlighting the complexity of Neotropical biodiversity.

The evidence presented here calls for a more robust, DNA-assisted cataloguing of biodiversity-rich ecosystems, in order to enable effective monitoring and informed actions to preserve and restore delicate habitats such as the DRB. Further studies should verify the extent to which fish biodiversity has been affected by the Doce dam collapse disaster, and what hotspots of diversity within the catchment can be identified as potential sources of replenishment. At the same time, the approaches used here, and additional high through-put methodologies (e.g. metabarcoding of water and sediment samples) should be increasingly employed to monitor biodiversity at a pace that can cater for the management needs of these increasingly impacted biodiverse habitats.

References

Agostinho, A. A., Thomaz, S. M., and Gomes, L. C. (2005). Conservação da biodiversidade em águas continentais do Brasil. Megadiversidade 1, 70-78.

Albert, J. S., Fernandes-Matioli, F. M. C., and Almeida-Toledo, L. F. (1999). New species of *Gymnotus* (Gymnotiformes, Teleostei) from south eastern Brazil: toward the deconstruction of *Gymnotus carapo*. *Copeia* 1999*,* 410-42. Doi: 10.2307/1447486

Aljanabi, S. M., and Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res.* 25, 4692-4693. Doi: 10.1093/nar/25.22.4692

ANA. (2016). Encarte especial sobre a bacia do rio Doce: rompimento da barragem em Mariana. Technical report. Brasília, Distrito Federal: Agência Nacional de Águas/ Ministério do Meio Ambiente.

ANEEL. (2010). Sistema de informações georreferenciadas do setor elétrico – SIGEL. Available at: http://sigel.aneel.gov.br [accessed March 2017].

Barbosa, M. A, and Costa, W. J. E. M. (2003). *Trichomycterus potschi* (Siluriformes: Loricarioidei): a new trichomycterid catfish from coastal streams of southeastern Brazil. *Ichthyol. Explor. Fres*. 14, 281-287 .

Becker, R. A., Sales, N. G., Santos, G. M., and Carvalho, D. C. (2015). DNA barcoding and morphological of neotropical ichthyoplankton from the Upper Paraná and São Francisco. *J. Fish Biol.* 87, 159-168. Doi: 10.1111/jfb.12707

Bensasson, D., Zhang, D., Hartl, D. L., and Hewitt, G. M. (2001). Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends Ecol. Evol.* 1, 314-321. Doi: 10.1016/S0169-5347(01)02151-6

Benzaquem, D. C., Oliveira, C., Batista, J. S., Zuanon, J., and Porto, J. I. R. (2015). DNA barcoding in pencilfishes (Lebiasinidae: *Nannostomus*) reveals cryptic diversity across the Brazilian Amazon. *PloS ONE* 10, e0112217, 1-14. Doi: 10.1371/journal.pone.0112217

Blanco, D. R., Lui, R. L., Vicari, M. R., Bertollo, L. A., and Moreira-Filho, O. (2011). Comparative cytogenetics of giant trahiras *Hoplias aimara* and *H. intermedius* (Characiformes, Erythrinidae): chromosomal characteristics of minor and major ribosomal DNA and cross-species repetitive centromeric sequences mapping differ among morphologically identical karyotypes. *Genome Res.* 132, 71-78. Doi: 10.1159/000320923

Carvalho, D. C., Oliveira, D. A. A., Santos, J. E., Teske, P., Beheregaray, L. B., Schneider, H., and Sampaio, I.. (2009). Genetic characterization of native and introduced populations of the neotropical cichlid genus *Cichla* in Brazil*. Genet. Mol. Biol.* 32, 601-607. Doi: 10.1590/S1415-47572009005000060

Carvalho, D. C., Oliveira, D. A. A., Pompeu, P. S., Leal, C. G., Oliveira, C., and Hanner, R. (2011). Deep barcode divergence in Brazilian freshwater fishes: the case of the São Francisco River Basin. *Mitochondrial DNA* 22, 80-86. Doi: 10.3109/19401736.2011.588214

Castro, R. M., and Vari, R. P. (2004). Detritivores of the South American fish family Prochilodontidae (Teleostei: Ostariophysi:Characiformes): a phylogenetic and revisionary study. Washington: Smithsonian Contributions to Zoology, 189.

Corin, S. E., Lester, P. J., Abbott, K. L., and Ritchie, P. A. (2007). Inferring historical introduction pathways with mitochondrial DNA: the case of introduced Argentine ants (*Linepithema humile*) into New Zealand. *Divers. Distrib.* 13, 510-518. Doi: 10.1111/j.1472-4642.2007.00355.x

Costa-Silva, G. J., Rodriguez, M. S., Roxo, F. F., Foresti, F., and Oliveira, C. (2015). Using different methods to access the difficult task of delimiting species in a complex Neotropical hyperdiverse group. *PloS ONE* 10, e0135075. Doi:10.1371/journal. Pone.0135075

da Silva, M. A., Silva, M. L. N., Curi, N., Avanzi, J. C., and Leite, F. P. (2011). Sistemas de manejo em plantios florestais de eucalipto e perdas de solo e água na região do vale do Rio Doce, MG. *Ciencia Florestal* 21, 765–776. Doi:10.5902/198050984520

Deacon, A. E., Ramnarine, I. W., and Magurran, A. E. (2011). How reproductive ecology contributes to the spread of a globally invasive fish. PloS ONE 6, 9, e24416. Doi: 10.1371/journal.pone.0024416

Dergam, J. A., Suzuki, H. I., Shibatta, O. A., Duboc, L. F., Júlio-Jr, H. F., Giuliano-Caetano, L., and Black IV, W. C. (1998). Molecular biogeography of the Neotropical fish *Hoplias malabaricus* (Erythrinidae: Characiformes) in the Iguaçu, Tibagi, and Paraná Rivers. *Genet. Mol. Biol.* 21*,* 4. Doi:10.1590/S1415-47571998000400015

Díaz, J., Villanova, G. V., Brancolini, F., del Pazo, F., Posner, V. M., Grimberg, A., et al. (2016). First DNA barcode reference library for the identification of South American freshwater fish from the lower Paraná river. *PloS ONE* 11, e0157419. Doi: 10.1371/journal.pone.0157419

Ecoplan-Lume. (2010). Plano integrado de recursos hídricos da bacia hidrográfica do rio Doce. Belo Horizonte: IGAM.

Ely, C. V., Bordignon, S. A. L., Trevisan, R., and Boldrini, I. I. (2017). Implications of poor taxonomy in conservation. *J. Nat. Conserv.* 36, 10-13. Doi: 10.1016/j.jnc.2017.01.003

Eschmeyer, W. N. (2015). *Catalog of fishes: World Wide Web electronic publication.* Available at: http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp [accessed September 2017].

Fernandes, G. W., Goulart, F. F., Ranieri, B. D., Coelho, M. S., Dales, K., Boesche, N., et al. (2016). Deep into the mud: ecological and socio-economic impacts of the dam breach in Mariana, Brazil. *Nat. Conserv.* 14, 35-45. Doi: 10.1016/j.ncon.2016.10.003

Frantine-Silva, W., Sofia, S. H., Orsi, M. L., and Almeida, F. S. (2015). DNA barcoding of freshwater ichthyoplankton in the Neotropics as a tool for ecological monitoring. *Mol. Ecol. Resour.* 1, 1226-1237. Doi: 10.1111/1755-0998.12385

Froese, R., and Paulay, D. (2010). *FishBase: World Wide Web electronic publication.* Available at: http://fishbase.org [accessed September 2017].

Garutti, V. (1995). Revisão taxonômica dos *Astyanax* (Pisces, Characidae), com mancha umeral ovalada e mancha no pedúnculo caudal, estendendo-se a extremidade dos raios caudais medianos, das bacias do Paraná, São Francisco e Amazônica. [thesis]. [São Paulo (SP)]: Universidade Estadual Paulista.

GFT. (2015). Avaliação dos efeitos e desdobramentos do rompimento da Barragem de Fundão em Mariana-MG. Available at: http://www.agenciaminas.mg.gov.br/ckeditor_assets/attachments/770/relatorio_final_ft_03_02_2016_15h5min.pdf [accessed September 2017].

Gomes, L. C., Pessali, T. C., Sales, N. G., Pompeu, P. S., and Carvalho, D. C. (2015). Integrative taxonomy detects cryptic and overlooked fish species in a Neotropical river basin. *Genetica* 143, 581-588. Doi: 10.1007/s10709-015-9856-z

Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. G., and Hallwachs, W. (2004a). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astrapes fulgerator*. *Proc. Natl. Acad. Sci. USA* 101, 14812-14817. Doi: 10.1073/pnas.0406166101

Hebert, P. D. N., Stoeckle, M. Y., Zemlak, T. S., and Francis, C. M. (2004b). Identification of birds through DNA barcodes**.** *PloS Biol*. 2, e312. Doi: 10.1371/journal.pbio.0020312

Hubert, N., Hanner, R., Holm, E., Mandrak, N. E., Taylor, E., Burridge, M., et al. (2008). Identifying Canadian freshwater fishes through DNA barcodes*. PloS ONE* 3, e2490. Doi: 10.1371/journal.pone.0002490

Hurst, G. D. D., and Jiggins, F. M. (2005). Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proc. R. Soc. B* 272, 1525-1534. Doi: 10.1098/rspb.2005.3056

Ivanova, V., Zemlak, T. S., Hanner, R. H., and Hebert, P. D. N. (2007). Universal primer cocktails for fish DNA barcoding. *Mol. Ecol. Resour.* 7, 544-548. Doi: 10.1111/j.1471-8286.2007.01748.x

Jukes, T. H., and Cantor, C. R. (1969). Evolution of protein molecules. In: Munro HN Mammalian protein metabolism. New York: Academic Press.

Kimura, M. (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16, 111-120.

Kittell, M. M., Harvey, M. N., Contreras-Balderas, S., Ptacek, M. B. (2005). Wild-caught hybrids between sailfin and shortfin mollies (Poeciliidae, *Poecilia*): morphological and molecular verification. *Hidrobiológica* 15, 2, 131-137.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol*. 33, 1870-1874. Doi:10.1093/molbev/msw054

Lampert, K. ., and Schartl, M. (2008). The origin and evolution of a unisexual hybrid: *Poecilia formosa*. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363,1505, 2901–2909. Doi: 10.1098/rstb.2008.0040

Latini, A. O., and Petrere-Jr., M. (2004). Reduction of a native fish fauna by alien species: an example from Brazilian freshwater tropical lakes. *Fish. Manag. Ecol.* 11, 71-79. Doi: 10.1046/j.1365-2400.2003.00372.x

Metcalf, J. L., Pritchard, V. L., Silvestri, S. M., Jenkins, J. B., Wood, J. S., Cowley, D. E. et al. (2007). Across the great divide: genetic forensics reveals misidentification of endangered cutthroat trout populations. *Mol. Ecol.* 16, 4445-4454. Doi: 10.1111/j.1365-294X.2007.03472.x

Meyer, C. P., and Paulay, G. (2005). DNA barcoding: error rates based on comprehensive sampling. *PloS Biol.* 3, 229-2238. Doi: 10.1371/journal.pbio.0030422.t002

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Fonseca, G. A. B., and Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature* 403, 853–858. Doi: 10.1038/35002501

Nascimento, M. H. S., Almeida, M. S., Veira, M. N. S., Limeira-Filho, D., Lima, R. C, Barros, M. C., and Fraga, E. C. (2016). DNA barcoding reveals high levels of genetic diversity in the fishes of the Itapecuru Basin in Maranhão, Brazil. *Genet. Mol. Research* 15, 3, 1-11. Doi: 10.4238/gmr.15038476

Neves, A. C. O., Nunes, F.P., Carvalho, F.A., and Fernandes, G. W. (2016**).** Neglect of ecosystems services by mining, and the worst environmental disaster in Brazil**.** *Nat. Conserv*. 14, 24-27. Doi: 10.1016/j.ncon.2016.03.002

Olds, A. D., Pitt, K. A., Maxwell, P. S., and Connolly, R. M. (2012). Synergistic effects of reserves and connectivity on ecological resilience. *J. Appl. Ecol*. 49, 1195-1203. Doi: 10.1111/jpe.12002

Oliveira, E. A., Bertollo, L. A. C., Yano, C. F., Liehr, T., and Cioffi, M. B. (2015). Comparative cytogenetics in the genus *Hoplias* (Characiformes, Erythrinidae) highlights contrasting karyotype evolution among congeneric species. *Mol. Cytogenet. 8*, 56, 1-10. Doi: 10.1186/s13039-015-0161-4

Padial, J. M., Miralles, A., De la Riva, I., and Vences, M. (2010). The integrative future of taxonomy. *Front. Zool.* 7, 16. Doi: 10.1186/1742-9994-7-16

Pereira, L. H., Pazian, M. F., Hanner, R., Foresti, F., and Oliveira, C. (2011). DNA barcoding reveals hidden diversity in the Neotropical freshwater fish *Piabina argentea* (Characiformes: Characidae) from the Upper Parana Basin of Brazil. *Mitochondrial DNA* 22, 87-96. Doi: 10.3109/19401736.2011.588213

Pereira, L. H. G., Hanner, R., Foresti, F., and Oliveira, C. (2013). Can DNA Barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? *BMC Genet.* 14, 14-20. Doi: 10.1186/1471-2156-14-20

Pfenninger, M., Nowak, C., Kley, C., Steinke, D., and Streit, B. (2007). Utility of DNA taxonomy and barcoding for the inference of larval community structure in morphologically cryptic *Chironomus* (Diptera) species. *Mol. Ecol.* 16, 1957-1968. Doi:10.1111/j.1365-294X.2006.03136.x

Pugedo, M. L., Andrade-Neto, F. R., Pessali, T. C., Birindelli, J. L. O., and Carvalho, D. C. (2016). Integrative taxonomy supports new candidate fish species in a poorly studied neotropical region: the Jequitinhonha River Basin. *Genetica* 144, 341-349. Doi: 10.1007/s10709-016-9903-4

Puillandre, N., Lambert, A., Brouillet, S., and Achaz, G. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Mol. Ecol.* 21, 1864-1877. Doi: 10.1111/j.1365-294X.2011.05239.x

Ramirez, J. L., Birindelli, J. L., Carvalho ,D. C., Affonso, P. R. A. M., Venere, P. C., Ortega, H., et al. (2017). Revealing hidden diversity of the underestimated Neotropical ichthyofauna: DNA Barcoding in the recently described genus *Megaleporinus* (Characiformes: Anostomidae). *Front. Genet*. 8,149. Doi: 10.3389/fgene.2017.00149

Ratnasingham, S., and Hebert, P. D. N. (2007). BOLD: the barcode of life data system (www.barcodingoflife.org). *Mol. Ecol. Notes* 3, 355-364. Doi: 10.1111/j.1471-8286.2007.01678.x

Ratnasingham, S., and Hebert, P. D. N. (2013). A DNA-based registry for all animal species: the barcode index number (BIN) system. *PloS ONE* 8, e66213. Doi: 10.1371/journal.pone.0066213

Reis, R. E., Albert, J. S., Di Dario, F., Mincarone, M. M., Petry, P., and Rocha, L. A. (2016). Fish biodiversity and conservation in South America. *J. Fish Biol.* 89, 12-47. Doi: 10.1111/jfb.13016

Ronquist, F., and Huelsenbeck, J. P. (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572-1574. Doi: 10.1093/bioinformatics/btg180

Rosa, R. S., and Lima, F. C. T. (2005). Livro da Fauna Brasileira Ameaçada de Extinção. Belo Horizonte: Biodiversitas.

Rossini, B. C., Oliveira, C. A. M., Melo, F. A. G. d, Bertaco V. A., Astarloa, J. M. D., Rosso, J. J., et al. (2016). Highlighting *Astyanax* species diversity through DNA Barcoding. *PloS ONE* 11, 12. Doi:10.1371/journal.pone.0167203

Rosso, J., Mabragaña, E., González-Castro, M., and Díaz de Astarloa, J. (2012). DNA barcoding Neotropical fishes: recent advances from the Pampa Plain, Argentina. *Mol. Ecol. Resour. 12,* 999-1011. Doi: 10.1111/1755-0998.12010

Saitou, N., and Nei, M. (1987). The Neighbor-Joining method—a new method for reconstructing phylogenetic trees**.** *Mol. Biol. Evol.* 4, 406–425.

Sales, N. G., Pessali, T. C., Andrade-Neto, F. R., and Carvalho, D. C. (2018). Introgression from non-native species unveils a hidden threat to the migratory Neotropical fish *Prochilodus hartii*. *Biol. Invasions* 20, 555-566*.* Doi: 10.1007/s10530-017-1556-4

Shum, P., Moore, L., Pampoulie, C., Di Muri, C., Vandamme, S., and Mariani S. (2017). Harnessing mtDNA variation to resolve ambiguity in 'Redfish' sold in Europe. *PeerJ* in press.

Sievers, C., Willing, E. M., Hoffmann, M., Dreyer, C., Ramnarine, I., and Magurran, A. (2012). Reasons for the Invasive Success of a Guppy (*Poecilia reticulata*) Population in Trinidad. PloS ONE 7, 5, e38404. Doi: 10.1371/journal.pone.0038404

Struck, T. H., Feder, J. L., Bendiksby, M., Birkeland, S., Cerca, J., Gusarov, V. I., et al. (2018). Finding Evolutionary Processes Hidden in Cryptic Species. Trends Ecol. Evol. 33, 153–163. Doi:https://doi.org/10.1016/j.tree.2017.11.007.

Sunaga, T., and Verani, J. R. (1987). "Second report of comparative study on fish community of the Rio Doce Valley lakes" In Limnological studies in Rio Doce Valley lakes and Pantanal wetland, Brazil, eds. Saijo, Y. and J. G. C. P. Tundisi (Nagoya University), 129-135.

Taylor, R. W. (1983). "Descriptive taxonomy: past, present, and future" In Australian Systematic entomology: a bicentenary perspective, eds. Highley E. and Taylor R. W. (Melbourne: CSIRO).

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997). The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic. Acids Res.* 24, 4876-4882.

Valdez-Moreno, M., Ivanova, N. V., Elias-Gutierrez, M., Contreras-Balderas, S., and Hebert, P. D. N. (2009). Probing diversity in freshwater fishes from Mexico and Guatemala with DNA barcodes*. J. Fish. Biol.* 74, 377-402. Doi: 10.1111/j.1095-8649.2008.02077.x

Vari, R. (1992). Systematics of the neotropical characiform genus *Cyphocharax* Fowler (Pisces:Ostariophysi). *Smithon. Contrib. Zool.* 529, 1-137.

Vieira, F. (1994) Estrutura de comunidade e aspectos da alimentação e reprodução dos peixes em dois lagos do médio rio Doce, MG. [dissertation]. [Belo Horizonte (MG)]. Universidade Federal de Minas Gerais

Vieira, F., Alves, C. B. M., and Santos, G. B. (2000). Rediscovery and first record of *Henochilus wheatlandii* (Teleostei: Characiformes) a rare Neotropical fish, in rio Doce basin of southeastern Brazil. *Ichthyol. Explor. Fres.* 11, 201-206.

Vieira, F., and Alves, C. B. M. (2001). Threatened fishes of the world: *Henochilus wheatlandii* Garman 1980 (Characidae). *Env. Biol. Fish.* 62, 414.

Vieira, F. (2009). Distribuição, impactos ambientais e conservação da fauna de peixes da bacia do rio Doce. *MG BIOTA* 2, 5-22.

Vinas, J. and Tudela, S. (2009). A Validated Methodology for Genetic Identification of Tuna Species (Genus *Thunnus*), *PloS ONE* 4, e7606. Doi:10.1371/journal.pone.0007606

Vono, V., and Barbosa, F, A, R. (2001). Habitats and littoral zone fish community structure of two natural lakes in southeast Brazil. *Env. Biol. Fish.* 61, 371-379.

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., and Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philos. T. R. Soc. B* 360, 1847-1857. Doi: 10.1098/rstb.2005.1716

Zanata, A. M., and Camelier, P. (2009). Astyanax vermilion and Astyanax burgerai: new characid fishes (Ostariophysi: Characiformes) from Northeastern Bahia, Brazil. *Neotrop. Ichthyol.* 7, 175-184. Doi: 10.1590/S1679-62252009000200007**.**

Zhang, J., Kapli, P., Pavlidis, P., and Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29, 2869–2876. Doi: 10.1093/bioinformatics/btt499

Supplementary Material

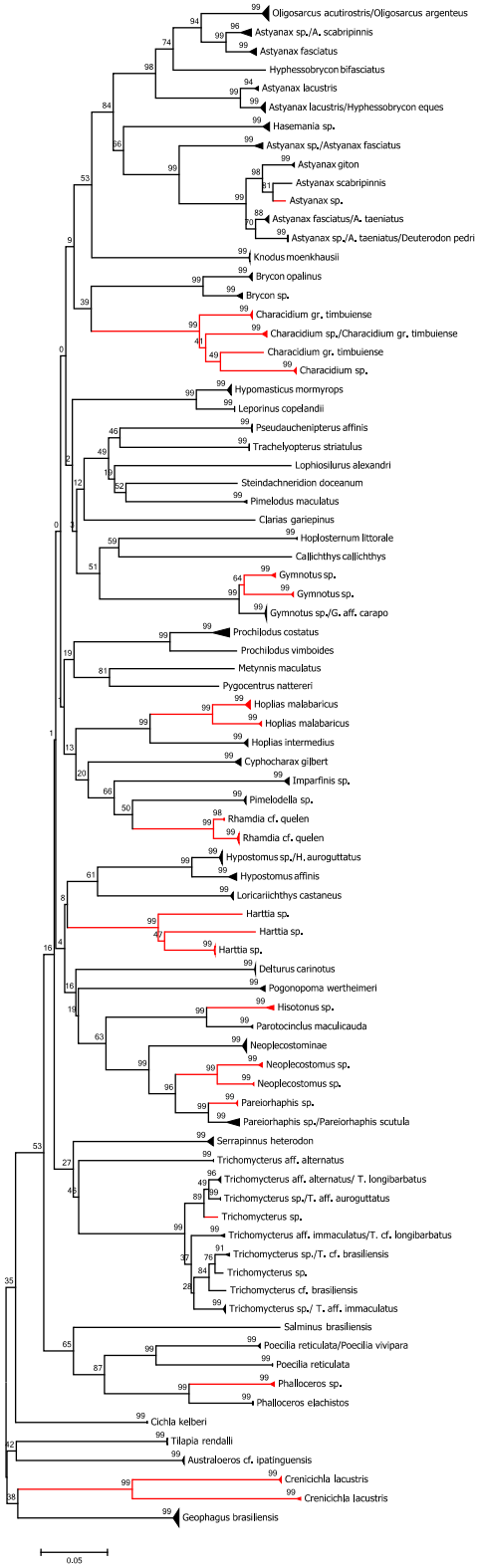


Figure 1: NJ tree based on K2P distance, encompassing all analyzed species to check the occurrence of putative new MOTUs (red branches).

Table S1: Sixty nine fish species barcoded from Doce River Basin and identified as a unique Barcode Index Number (BIN). LGC and RD– Laboratório de Genética da Conservação do Programa de Pós-graduação em Biologia de Vertebrados/PUC Minas, MBML – Museu de Biologia Professor Mello Leitão, MCNIP – Coleção de Ictiologia do Museu de Ciências Naturais da *PUC* Minas, ZUEC – Museu de Zoologia da Universidade Estadual de Campinas "Prof. Adão José Cardoso"

Process ID	Sample ID	Museum ID	Identification	Lat	Lon	BIN
RDOCE175-14	RD116	RD116	*Astyanax fasciatus*	-19.996	-41.739	BOLD:ABU7523
RDOCE229-14	LGC3525	BG-06VI-37	*Astyanax fasciatus*	-18.996	-42.225	BOLD:ACJ1542
RDOCE231-14	LGC3545	BG-12XII-06	*Astyanax fasciatus*	-18.996	-42.225	BOLD:ACJ1542
RDOCE234-14	LGC3688	LGC3688	*Astyanax fasciatus*	-19.004	-43.375	BOLD:ACJ9650
RDOCE284-14	LGC3567	LGC3567	*Astyanax fasciatus*	-18.996	-42.225	BOLD:ACJ1542
RDOCE285-14	LGC3568	LGC3568	*Astyanax fasciatus*	-18.996	-42.225	BOLD:ACJ1542
RDOCE286-14	LGC3569	LGC3569	*Astyanax fasciatus*	-18.996	-42.225	BOLD:ACJ1542
RDOCE212-14	LGC4145	MBML6827	*Astyanax giton*	-19.837	-40.555	BOLD:ACL8007
RDOCE214-14	LGC4153	MBML6842	*Astyanax giton*	-19.889	-40.576	BOLD:ACL8007
RDOCE321-15	LGC4147	MBML6831	*Astyanax giton*	-19.884	-40.575	BOLD:ACL8007
RDOCE091-13	LGC1819	MCNI-PUCMG-0476	*Astyanax lacustris*	-20.079	-41.733	BOLD:ABZ1711
RDOCE018-13	LGC153	MCNI-PUCMG-0476	*Astyanax lacustris*	-20.079	-41.733	BOLD:ABZ1711
RDOCE178-14	RD134	RD134	*Astyanax lacustris*	-19.996	-41.739	BOLD:ABZ1711
RDOCE179-14	RD136	RD136	*Astyanax lacustris*	-18.937	-42.045	BOLD:ABY8634
RDOCE180-14	RD137	RD137	*Astyanax lacustris*	-18.993	-42.225	BOLD:ABY8634
RDOCE152-13	RD139	RD139	*Astyanax lacustris*	-19.062	-42.162	BOLD:ABY8634
RDOCE221-14	RD132	RD132	*Astyanax lacustris*	-19.973	-41.725	BOLD:ABZ1711
RDOCE222-14	RD133	RD133	*Astyanax lacustris*	-19.973	-41.725	BOLD:ABZ1711
RDOCE250-14	LGC4598	MCNIP-1607	*Astyanax lacustris*	-19.985	-41.722	BOLD:ABZ1711
RDOCE176-14	RD121	RD121	*Astyanax scabripinnis*	-20.04	-41.93	BOLD:AAC5910
RDOCE218-14	RD120	RD120	*Astyanax scabripinnis*	-20.024	-43.460	BOLD:AAC5910
RDOCE219-14	RD123	RD123	*Astyanax scabripinnis*	-20.110	-43.400	BOLD:AAC5910
RDOCE220-14	RD124	RD124	*Astyanax scabripinnis*	-20.110	-43.400	BOLD:AAC5910
RDOCE237-14	LGC3727	83	*Astyanax scabripinnis*	-18.913	-43.439	BOLD:ACO5464
RDOCE142-13	RD68	RD68	*Astyanax* sp.	-20.083	-43.420	BOLD:ACJ9650
RDOCE143-13	RD69	RD69	*Astyanax* sp.	-20.083	-43.420	BOLD:AAY4812
RDOCE144-13	RD70	RD70	*Astyanax* sp.	-20.083	-43.420	BOLD:AAC5910
RDOCE272-14	RD159	RD159	*Astyanax* sp.	-19.014	-43.377	BOLD:ACT0040
RDOCE287-14	LGC3695	LGC3695	*Astyanax* sp.	-19.011	-43.372	BOLD:ACJ9650
RDOCE297-15	LGC3724	LGC3724	*Astyanax* sp.	-19.011	-43.372	BOLD:ACJ9650
RDOCE183-14	RD148	RD148	*Astyanax taeniatus*	-18.967	-42.318	BOLD:ABU7523
RDOCE184-14	RD149	RD149	*Astyanax taeniatus*	-18.967	-42.318	BOLD:ABU7523
RDOCE153-13	RD140	RD140	*Astyanax taeniatus*	-20.120	-43.400	BOLD:ABU7523

Process ID	Sample ID	Museum ID	Identification	Lat	Lon	BIN
RDOCE271-14	RD141	RD141	*Astyanax taeniatus*	-20.121	-43.401	BOLD:ABU7523
RDOCE203-14	LGC1533	LGC1533	*Australoheros* cf. *ipatinguensis*	-19.275	-42.425	BOLD:ACR9799
RDOCE204-14	LGC1534	LGC1534	*Australoheros* cf. *ipatinguensis*	-19.275	-42.425	BOLD:ACR9799
RDOCE205-14	LGC1535	LGC1535	*Australoheros* cf. *ipatinguensis*	-19.275	-42.425	BOLD:ACR9799
RDOCE206-14	LGC1536	LGC1536	*Australoheros* cf. *ipatinguensis*	-19.275	-42.425	BOLD:ACR9799
RDOCE207-14	LGC1537	LGC1537	*Australoheros* cf. *ipatinguensis*	-19.275	-42.425	BOLD:ACR9799
RDOCE189-14	LGC3678	LGC3678	*Brycon opalinus*	-19.289	-43.192	BOLD:ACL7114
RDOCE199-14	LGC3747	LGC3747	*Brycon opalinus*	-18.763	-43.459	BOLD:ACL7114
RDOCE232-14	LGC3677	1	*Brycon opalinus*	-19.289	-43.192	BOLD:ACL7114
RDOCE233-14	LGC3679	3	*Brycon opalinus*	-19.289	-43.192	BOLD:ACL7114
RDOCE238-14	LGC3745	6	*Brycon opalinus*	-18.969	-43.438	BOLD:ACL7114
RDOCE245-14	LGC4508	MCNIP-1602	*Brycon* sp.	-18.994	-42.226	BOLD:ACH8616
RDOCE252-14	LGC4635	MCNIP-1601	*Brycon* sp.	-19.156	-42.231	BOLD:ACH8616
RDOCE253-14	LGC4636	MCNIP-1601	*Brycon* sp.	-19.156	-42.231	BOLD:ACH8616
RDOCE254-14	LGC4637	MCNIP-1601	*Brycon* sp.	-19.156	-42.231	BOLD:ACH8616
RDOCE311-15	LGC5770	LGC5770	*Callichthys callichthys*	-18.91	-43.442	BOLD:AAB5066
RDOCE123-13	RD46	RD46	*Characidium* cf. *timbuiense*	-19.022	-42.122	BOLD:ACJ1226
RDOCE124-13	RD48	RD48	*Characidium* cf. *timbuiense*	-19.022	-42.122	BOLD:ACI3743
RDOCE125-13	RD51	RD51	*Characidium* cf. *timbuiense*	-19.022	-42.122	BOLD:ACI3743
RDOCE158-14	RD47	RD47	*Characidium* cf. *timbuiense*	-19.996	-41.739	BOLD:ACJ1226
RDOCE159-14	RD49	RD49	*Characidium* cf. *timbuiense*	-18.964	-42.318	BOLD:ACI3743
RDOCE140-13	RD45	RD45	*Characidium* cf. *timbuiense*	-19.022	-42.122	BOLD:ACJ9733
RDOCE141-13	RD50	RD50	*Characidium* cf. *timbuiense*	-19.002	-42.127	BOLD:ACI3743
RDOCE190-14	LGC3683	LGC3683	*Characidium* cf. *timbuiense*	-18.974	-43.372	BOLD:ACJ1226
RDOCE198-14	RD161	RD161	*Characidium* sp.	-18.754	-43.447	BOLD:ACS9348
RDOCE239-14	LGC4125	MBML4422	*Characidium* sp.	-19.788	-40.663	BOLD:ACI3743
RDOCE326-15	LGC5735	LGC5735	*Characidium* sp.	-18.974	-43.371	BOLD:ACJ1226
RDOCE327-15	LGC5719	LGC5719	*Characidium* sp.	-18.813	-43.413	BOLD:ACS9348
RDOCE328-15	LGC5752	LGC5752	*Characidium* sp.	-18.974	-43.371	BOLD:ACS9348
RDOCE230-14	LGC3540	BG-11XI-05	*Cichla kelberi*	-18.996	-42.225	BOLD:AAO9230
RDOCE258-14	LGC4697	3440	*Cichla kelberi*	-19.156	-42.231	BOLD:AAO9230
RDOCE094-13	LGC2674	LGC2674	*Clarias gariepinus*	-19.022	-42.122	BOLD:AAB2256
RDOCE053-13	LGC3555	ZUEC 8147	*Crenicichla lacustris*	-19.022	-42.122	BOLD:AAD6380
RDOCE054-13	LGC3556	ZUEC 8147	*Crenicichla lacustris*	-19.022	-42.122	BOLD:AAD6380
RDOCE248-14	LGC4579	17	*Crenicichla lacustris*	-19.979	-41.714	BOLD:ACO6050
RDOCE249-14	LGC4582	26	*Crenicichla lacustris*	-19.979	-41.714	BOLD:ACO6050
RDOCE294-14	LGC4978	ZUEC 8199	*Crenicichla lacustris*	-19.979	-41.714	BOLD:ACO6050
RDOCE295-14	LGC4979	ZUEC 8199	*Crenicichla lacustris*	-19.979	-41.714	BOLD:ACO6050
RDOCE090-13	LGC1810	MCNI-PUCMG-0458	*Cyphocharax gilbert*	-19.988	-41.72	BOLD:ACK1539

Process ID	Sample ID	Museum ID	Identification	Lat	Lon	BIN
RDOCE100-13	LGC3548	ZUEC 8153	*Cyphocharax gilbert*	-18.945	-42.363	BOLD:ACK1539
RDOCE101-13	LGC3550	ZUEC 8153	*Cyphocharax gilbert*	-18.945	-42.363	BOLD:ACK1539
RDOCE016-13	LGC143	MCNI-PUCMG-0458	*Cyphocharax gilbert*	-19.988	-41.72	BOLD:ACK1539
RDOCE050-13	LGC3549	ZUEC 8153	*Cyphocharax gilbert*	-18.945	-42.363	BOLD:ACK1539
RDOCE013-13	LGC129	MCNI-PUCMG-0444	*Delturus carinotus*	-20.016	-41.735	BOLD:ACC0184
RDOCE014-13	LGC130	MCNI-PUCMG-0446	*Delturus carinotus*	-20.048	-41.747	BOLD:ACC0184
RDOCE024-13	LGC164	LGC164	*Delturus carinotus*	-19.986	-41.716	BOLD:ACC0184
RDOCE065-13	LGC163	LGC163	*Delturus carinotus*	-19.986	-41.716	BOLD:ACC0184
RDOCE130-13	LGC165	LGC165	*Delturus carinotus*	-19.985	-41.716	BOLD:ACC0184
RDOCE224-14	RD58	RD58	*Deuterodon pedri*	-19.001	-42.231	BOLD:AAY4812
RDOCE273-14	RD56	RD56	*Deuterodon pedri*	-19.002	-42.231	BOLD:AAY4812
RDOCE274-14	RD57	RD57	*Deuterodon pedri*	-19.002	-42.231	BOLD:AAY4812
RDOCE086-13	LGC173	LGC173	*Geophagus brasiliensis*	-19.988	-41.72	BOLD:AAA8514
RDOCE088-13	LGC179	LGC179	*Geophagus brasiliensis*	-19.988	-41.72	BOLD:AAA8514
RDOCE110-13	RD29	RD29	*Geophagus brasiliensis*	-19.022	-42.122	BOLD:AAA8514
RDOCE111-13	RD30	RD30	*Geophagus brasiliensis*	-19.022	-42.122	BOLD:AAA8514
RDOCE112-13	RD31	RD31	*Geophagus brasiliensis*	-19.022	-42.122	BOLD:AAA8514
RDOCE113-13	RD33	RD33	*Geophagus brasiliensis*	-19.022	-42.122	BOLD:AAA8514
RDOCE114-13	RD34	RD34	*Geophagus brasiliensis*	-19.022	-42.122	BOLD:AAA8514
RDOCE021-13	LGC158	LGC158	*Geophagus brasiliensis*	-19.986	-41.716	BOLD:AAA8514
RDOCE032-13	LGC1817	MCNI-PUCMG-0476	*Geophagus brasiliensis*	-20.046	-41.735	BOLD:AAA8514
RDOCE062-13	LGC149	MCNI-PUCMG-0463	*Geophagus brasiliensis*	-20.046	-41.735	BOLD:AAA8514
RDOCE063-13	LGC159	LGC159	*Geophagus brasiliensis*	-19.986	-41.716	BOLD:AAA8514
RDOCE154-14	LGC150	MCNI-PUCMG-0463	*Geophagus brasiliensis*	-20.046	-41.735	BOLD:AAA8514
RDOCE264-14	LGC4961	ZUEC 8208	*Gymnotus* aff. *Carapo*	-19.985	-41.722	BOLD:AAB6216
RDOCE279-14	LGC4618	LGC4618	*Gymnotus* aff. *Carapo*	-20.046	-41.735	BOLD:AAB6216
RDOCE290-14	LGC4639	MCNIP-1604	*Gymnotus* aff. *Carapo*	-19.156	-42.231	BOLD:AAB6216
RDOCE292-14	LGC4704	MCNIP-1608	*Gymnotus* aff. *Carapo*	-19.156	-42.231	BOLD:AAB6216
RDOCE310-15	LGC4914	ZUEC 8208	*Gymnotus* aff. *Carapo*	-20.046	-41.735	BOLD:AAB6216
RDOCE322-15	LGC4955	ZUEC 8208	*Gymnotus* aff. *Carapo*	-19.985	-41.722	BOLD:AAB6216
RDOCE075-13	RD08	RD08	*Gymnotus* sp.	-20.110	-43.400	BOLD:AAB6212
RDOCE076-13	RD09	RD09	*Gymnotus* sp.	-19.996	-41.739	BOLD:AAB6216
RDOCE077-13	RD11	RD11	*Gymnotus* sp.	-19.996	-41.739	BOLD:AAB6216
RDOCE136-13	RD10	RD10	*Gymnotus* sp.	-19.996	-41.739	BOLD:AAB6216
RDOCE301-15	RD181	RD181	*Gymnotus* sp.	-19.996	-41.739	BOLD:ACT0768
RDOCE302-15	RD92	RD92	*Gymnotus* sp.	-20.109	-43.399	BOLD:AAB6212
RDOCE306-15	RD12	RD12	*Gymnotus* sp.	-20	-42	BOLD:ACT0768
RDOCE308-15	RD180	RD180	*Gymnotus* sp.	-18.99	-42.215	BOLD:ACT0768
RDOCE309-15	RD93	RD93	*Gymnotus* sp.	-20.109	-43.399	BOLD:AAB6212

Process ID	Sample ID	Museum ID	Identification	Lat	Lon	BIN
RDOCE115-13	RD35	RD35	*Harttia* sp.	-19.022	-42.122	BOLD:ACJ1000
RDOCE045-13	LGC3594	MCNIP-1637	*Harttia* sp.	-19.023	-42.125	BOLD:ACI6845
RDOCE235-14	LGC3700	54	*Harttia* sp.	-19.011	-43.372	BOLD:ACO6155
RDOCE247-14	LGC4531	MBML-PEIXES 7783	*Harttia* sp.	-19.979	-41.714	BOLD:ACI6845
RDOCE251-14	LGC4634	MCNIP-1606	*Harttia* sp.	-20.046	-41.735	BOLD:ACI6845
RDOCE265-14	LGC4976	ZUEC 8218	*Harttia* sp.	-19.979	-41.714	BOLD:ACI6845
RDOCE266-14	LGC4977	ZUEC 8218	*Harttia* sp.	-19.979	-41.714	BOLD:ACI6845
RDOCE267-14	LGC4985	ZUEC 8218	*Harttia* sp.	-19.979	-41.714	BOLD:ACI6845
RDOCE071-13	RD03	RD03	*Hasemania* sp.	-20.084	-43.415	BOLD:AAO6055
RDOCE072-13	RD04	RD04	*Hasemania* sp.	-20.084	-43.415	BOLD:AAO6055
RDOCE073-13	RD05	RD05	*Hasemania* sp.	-20.084	-43.415	BOLD:AAO6055
RDOCE074-13	RD06	RD06	*Hasemania* sp.	-20.084	-43.415	BOLD:AAO6055
RDOCE236-14	LGC3703	58	*Hasemania* sp.	-18.933	-43.447	BOLD:AAO6055
RDOCE166-14	RD85	RD85	*Hisonotus* sp.	-18.967	-42.318	BOLD:ACW1732
RDOCE134-13	RD01	RD01	*Hisonotus* sp.	-19.062	-42.162	BOLD:ACW1732
RDOCE323-15	RD86	RD86	*Hisonotus* sp.	-18.952	-42.36	BOLD:ACW1732
RDOCE102-13	LGC3552	ZUEC 8146	*Hoplias intermedius*	-19.022	-42.122	BOLD:AAB1734
RDOCE051-13	LGC3551	ZUEC 8146	*Hoplias intermedius*	-19.022	-42.122	BOLD:AAB1734
RDOCE052-13	LGC3553	ZUEC 8146	*Hoplias intermedius*	-19.022	-42.122	BOLD:AAB1734
RDOCE069-13	LGC3589	MCNIP-1638	*Hoplias intermedius*	-20.084	-43.415	BOLD:AAB1734
RDOCE097-13	LGC3532	ZUEC 8150	*Hoplias malabaricus*	-18.986	-42.216	BOLD:ACI3811
RDOCE098-13	LGC3533	ZUEC 8150	*Hoplias malabaricus*	-18.986	-42.216	BOLD:ACI3811
RDOCE099-13	LGC3541	MCNIP-1639	*Hoplias malabaricus*	-18.972	-42.286	BOLD:ACI3811
RDOCE022-13	LGC160	LGC160	*Hoplias malabaricus*	-19.986	-41.716	BOLD:AAY4779
RDOCE023-13	LGC162	LGC162	*Hoplias malabaricus*	-19.986	-41.716	BOLD:AAY4779
RDOCE031-13	LGC1814	MCNI-PUCMG-0461	*Hoplias malabaricus*	-20.046	-41.735	BOLD:AAY4779
RDOCE061-13	LGC147	MCNI-PUCMG-0461	*Hoplias malabaricus*	-20.046	-41.735	BOLD:AAY4779
RDOCE064-13	LGC161	LGC161	*Hoplias malabaricus*	-19.986	-41.716	BOLD:AAY4779
RDOCE068-13	LGC181	LGC181	*Hoplias malabaricus*	-19.988	-41.720	BOLD:AAY4779
RDOCE092-13	LGC1841	LGC1841	*Hoplosternum littorale*	-19.988	-41.72	BOLD:AAB5068
RDOCE037-13	LGC1845	LGC1845	*Hoplosternum littorale*	-19.988	-41.72	BOLD:AAB5068
RDOCE319-15	LGC4151	MBML6839	*Hyphessobrycon bifasciatus*	-19.888	-40.575	BOLD:ACT0106
RDOCE241-14	LGC4139	MBML6816	*Hyphessobrycon eques*	-19.47	-40.184	BOLD:ABZ1711
RDOCE280-14	LGC4963	ZUEC 8198	*Hypomasticus mormyrops*	-19.979	-41.714	BOLD:ACH5050
RDOCE281-14	LGC4964	ZUEC 8198	*Hypomasticus mormyrops*	-19.979	-41.714	BOLD:ACH5050
RDOCE282-14	LGC4965	ZUEC 8198	*Hypomasticus mormyrops*	-19.979	-41.714	BOLD:ACH5050
RDOCE283-14	LGC4966	ZUEC 8198	*Hypomasticus mormyrops*	-19.979	-41.714	BOLD:ACH5050
RDOCE293-14	LGC4967	ZUEC 8198	*Hypomasticus mormyrops*	-19.979	-41.714	BOLD:ACH5050
RDOCE296-15	LGC3715	LGC3715	*Hypomasticus mormyrops*	-19	-43	BOLD:ACH5050

Process ID	Sample ID	Museum ID	Identification	Lat	Lon	BIN
RDOCE303-15	LGC3713	LGC3713	*Hypomasticus mormyrops*	-19	-43	BOLD:ACH5050
RDOCE004-13	LGC08	MCNI-PUCMG-0197	*Hypostomus affinis*	-18.972	-42.286	BOLD:AAW9386
RDOCE006-13	LGC10	MCNI-PUCMG-0200	*Hypostomus affinis*	-19.114	-42.176	BOLD:AAW9386
RDOCE007-13	LGC11	MCNI-PUCMG-0200	*Hypostomus affinis*	-19.114	-42.176	BOLD:AAW9386
RDOCE008-13	LGC12	MCNI-PUCMG-0200	*Hypostomus affinis*	-19.114	-42.176	BOLD:AAW9386
RDOCE015-13	LGC138	MCNI-PUCMG-0451	*Hypostomus affinis*	-20.079	-41.733	BOLD:AAW9386
RDOCE002-13	LGC03	MCNI-PUCMG-0193	*Hypostomus auroguttatus*	-19.018	-42.121	BOLD:AAB9690
RDOCE003-13	LGC07	MCNI-PUCMG-0193	*Hypostomus auroguttatus*	-19.018	-42.121	BOLD:AAB9690
RDOCE009-13	LGC16	LGC16	*Hypostomus auroguttatus*	-19.018	-42.121	BOLD:AAB9690
RDOCE010-13	LGC17	LGC17	*Hypostomus auroguttatus*	-19.018	-42.121	BOLD:AAB9690
RDOCE011-13	LGC19	LGC19	*Hypostomus auroguttatus*	-19.018	-42.121	BOLD:AAB9690
RDOCE028-13	LGC1672	MCNI-PUCMG-0193	*Hypostomus auroguttatus*	-19.018	-42.121	BOLD:AAB9690
RDOCE194-14	LGC3712	LGC3712	*Hypostomus* sp.	-18.917	-43.462	BOLD:AAB9690
RDOCE313-15	LGC5786	LGC5786	*Hypostomus* sp.	-18.916	-43.461	BOLD:AAB9690
RDOCE129-13	RD55	RD55	*Imparfinis* sp.	-19.022	-42.122	BOLD:AAC2103
RDOCE080-13	RD16	RD16	*Imparfinis* sp.	-18.967	-42.318	BOLD:AAC2103
RDOCE156-14	RD17	RD17	*Imparfinis* sp.	-18.967	-42.318	BOLD:AAC2103
RDOCE227-14	RD76	RD76	*Imparfinis* sp.	-19.202	-42.361	BOLD:AAC2103
RDOCE276-14	RD77	RD77	*Imparfinis* sp.	-20.111	-43.4	BOLD:AAC2103
RDOCE197-14	RD160	RD160	*Knodus moenkhausii*	-18.933	-43.447	BOLD:AAM1485
RDOCE215-14	RD106	RD106	*Knodus moenkhausii*	-20.045	-43.444	BOLD:AAM1485
RDOCE216-14	RD107	RD107	*Knodus moenkhausii*	-20.045	-43.444	BOLD:AAM1485
RDOCE217-14	RD111	RD111	*Knodus moenkhausii*	-19.001	-42.231	BOLD:AAM1485
RDOCE268-14	RD104	RD104	*Knodus moenkhausii*	-19.002	-42.231	BOLD:AAM1485
RDOCE269-14	RD110	RD110	*Knodus moenkhausii*	-19.002	-42.231	BOLD:AAM1485
RDOCE095-13	LGC3033	LGC3033	*Leporinus copelandii*	-19.022	-42.122	BOLD:ACI6721
RDOCE030-13	LGC1811	MCNI-PUCMG-0459	*Leporinus copelandii*	-19.978	-41.714	BOLD:ACI6721
RDOCE034-13	LGC1823	LGC1823	*Leporinus copelandii*	-19.978	-41.714	BOLD:ACI6721
RDOCE155-14	LGC3544	BG-12XII-03	*Leporinus copelandii*	-18.945	-42.362	BOLD:ACI6721
RDOCE093-13	LGC2668	LGC2668	*Lophiosilurus alexandri*	-19.022	-42.122	BOLD:AAE4855
RDOCE128-13	RD54	RD54	*Loricariichthys castaneus*	-19.022	-42.122	BOLD:ACI6497
RDOCE038-13	LGC2671	LGC2671	*Loricariichthys castaneus*	-19.032	-42.126	BOLD:ACI6497
RDOCE040-13	LGC3036	LGC3036	*Loricariichthys castaneus*	-19.022	-42.122	BOLD:ACI6497
RDOCE208-14	LGC2672	LGC2672	*Loricariichthys castaneus*	-19.032	-42.126	BOLD:ACI6497
RDOCE209-14	LGC3034	LGC3034	*Loricariichthys castaneus*	-19.022	-42.122	BOLD:ACI6497
RDOCE133-13	LGC3537	LGC3537	*Metynnis maculatus*	-18.985	-42.216	BOLD:AAE7443
RDOCE168-14	RD96	RD96	Neoplecostominae	-19.002	-42.127	BOLD:ACC0721
RDOCE169-14	RD97	RD97	Neoplecostominae	-19.002	-42.127	BOLD:ACC0721
RDOCE170-14	RD98	RD98	Neoplecostominae	-19.002	-42.127	BOLD:ACC0721

Process ID	Sample ID	Museum ID	Identification	Lat	Lon	BIN
RDOCE173-14	RD103	RD103	Neoplecostominae	-19.002	-42.127	BOLD:ACC0721
RDOCE148-13	RD95	RD95	Neoplecostominae	-19.022	-42.125	BOLD:ACC0721
RDOCE149-13	RD100	RD100	Neoplecostominae	-19.022	-42.125	BOLD:ACC0721
RDOCE185-14	LGC1822	LGC1822	Neoplecostominae	-20.079	-41.733	BOLD:ACC0721
RDOCE195-14	LGC3717	LGC3717	Neoplecostominae	-18.917	-43.462	BOLD:ACC0721
RDOCE172-14	RD101	RD101	*Neoplecostomus* sp.	-20.047	-41.685	BOLD:AAX6581
RDOCE150-13	RD102	RD102	*Neoplecostomus* sp.	-20.047	-41.685	BOLD:AAX6581
RDOCE192-14	LGC3699	LGC3699	*Neoplecostomus* sp.	-19.011	-43.372	BOLD:ACT2675
RDOCE278-14	LGC3684	LGC3684	*Neoplecostomus* sp.	-18.974	-43.372	BOLD:ACT2675
RDOCE325-15	LGC5733	LGC5733	*Neoplecostomus* sp.	-18.974	-43.371	BOLD:ACT2675
RDOCE103-13	LGC3554	ZUEC 8149	*Oligosarcus acutirostris*	-19.022	-42.122	BOLD:AAI3590
RDOCE033-13	LGC1818	MCNI-PUCMG-01476	*Oligosarcus acutirostris*	-19.988	-41.72	BOLD:AAI3590
RDOCE048-13	LGC3546	LGC3546	*Oligosarcus acutirostris*	-18.945	-42.363	BOLD:AAI3590
RDOCE049-13	LGC3547	LGC3547	*Oligosarcus acutirostris*	-18.945	-42.363	BOLD:AAI3590
RDOCE055-13	LGC3557	ZUEC 8149	*Oligosarcus acutirostris*	-19.022	-42.122	BOLD:AAI3590
RDOCE316-15	LGC4518	MCNIP-1605	*Oligosarcus argenteus*	-19.978	-41.714	BOLD:AAI3590
RDOCE317-15	LGC4519	LGC4519	*Oligosarcus argenteus*	-19.978	-41.714	BOLD:AAI3590
RDOCE318-15	LGC4516	LGC4516	*Oligosarcus argenteus*	-19.978	-41.714	BOLD:AAI3590
RDOCE320-15	LGC3682	LGC3682	*Oligosarcus argenteus*	-18.974	-43.371	BOLD:AAI3590
RDOCE151-13	RD138	RD138	*Pareiorhaphis scutula*	-20.045	-43.444	BOLD:AAX0824
RDOCE188-14	RD127	RD127	*Pareiorhaphis scutula*	-19.275	-42.425	BOLD:AAX0824
RDOCE270-14	RD128	RD128	*Pareiorhaphis scutula*	-20.025	-43.46	BOLD:AAX0824
RDOCE177-14	RD129	RD129	*Pareiorhaphis* sp.	-19.01	-43.37	BOLD:AAX0824
RDOCE131-13	LGC1850	LGC1850	*Pareiorhaphis* sp.	-20.048	-41.746	BOLD:ACI5663
RDOCE210-14	LGC3710	LGC3710	*Pareiorhaphis* sp.	-18.917	-43.462	BOLD:AAX0824
RDOCE300-15	RD177	RD177	*Pareiorhaphis* sp.	-20	-42	BOLD:ACI5663
RDOCE307-15	RD178	RD178	*Pareiorhaphis* sp.	-20	-42	BOLD:ACI5663
RDOCE240-14	LGC4133	MBML4646	*Parotocinclus maculicauda*	-19.888	-40.574	BOLD:ACO5053
RDOCE242-14	LGC4148	MBML6834	*Parotocinclus maculicauda*	-19.885	-40.575	BOLD:ACO5053
RDOCE211-14	LGC4144	MBML6826	*Phalloceros elachistos*	-19.837	-40.555	BOLD:ACO4001
RDOCE243-14	LGC4149	MBML6835	*Phalloceros elachistos*	-19.885	-40.575	BOLD:ACO4001
RDOCE289-14	LGC4150	MBML6837	*Phalloceros elachistos*	-19.889	-40.576	BOLD:ACO4001
RDOCE193-14	LGC3704	LGC3704	*Phalloceros* sp.	-18.933	-43.447	BOLD:AAB7265
RDOCE312-15	LGC5776	LGC5776	*Phalloceros* sp.	-18.933	-43.446	BOLD:AAB7265
RDOCE314-15	LGC5788	LGC5788	*Phalloceros* sp.	-18.924	-43.465	BOLD:AAB7265
RDOCE126-13	RD52	RD52	*Pimelodella* sp.	-19.022	-42.122	BOLD:AAC5327
RDOCE127-13	RD53	RD53	*Pimelodella* sp.	-19.022	-42.122	BOLD:AAC5327
RDOCE039-13	LGC3016	LGC3016	*Pimelodella* sp.	-19.023	-42.125	BOLD:AAC5327
RDOCE132-13	LGC3015	LGC3015	*Pimelodella* sp.	-19.022	-42.125	BOLD:AAC5327

Process ID	Sample ID	Museum ID	Identification	Lat	Lon	BIN
RDOCE259-14	LGC4710	MCNIP-1609	*Pimelodus maculatus*	-19.081	-42.159	BOLD:AAB6504
RDOCE263-14	LGC4728	MCNIP-1610	*Pimelodus maculatus*	-18.951	-42.361	BOLD:AAB6504
RDOCE181-14	RD145	RD145	*Poecilia reticulata*	-19.77	-40.63	BOLD:AAC0279
RDOCE182-14	RD146	RD146	*Poecilia reticulata*	-19.77	-40.63	BOLD:AAC0279
RDOCE138-13	RD18	RD18	*Poecilia reticulata*	-20.042	-41.698	BOLD:ACE9037
RDOCE139-13	RD19	RD19	*Poecilia reticulata*	-20.04	-41.698	BOLD:ACE9037
RDOCE288-14	LGC4140	MBML6817	*Poecilia vivipara*	-19.47	-40.184	BOLD:AAC0279
RDOCE047-13	LGC3596	LGC3596	*Pogonopoma wertheimeri*	-19.023	-42.125	BOLD:ACI3792
RDOCE299-15	LGC4715	LGC4715	*Pogonopoma wertheimeri*	-19	-42	BOLD:ACI3792
RDOCE304-15	LGC4701	LGC4701	*Pogonopoma wertheimeri*	-19	-42	BOLD:ACI3792
RDOCE305-15	LGC4726	LGC4726	*Pogonopoma wertheimeri*	-19	-42	BOLD:ACI3792
RDOCE096-13	LGC3528	LGC3528	*Prochilodus costatus*	-19.032	-42.126	BOLD:ADK5931
RDOCE056-13	LGC3559	LGC3559	*Prochilodus costatus*	-19.081	-42.158	BOLD:ADK5931
RDOCE057-13	LGC3560	LGC3560	*Prochilodus costatus*	-19.081	-42.158	BOLD:ADK5931
RDOCE058-13	LGC3561	LGC3561	*Prochilodus costatus*	-19.081	-42.158	BOLD:ADK5929
RDOCE059-13	LGC3562	LGC3562	*Prochilodus costatus*	-19.081	-42.158	BOLD:ADK5931
RDOCE060-13	LGC3563	LGC3563	*Prochilodus costatus*	-19.081	-42.158	BOLD:ADK5929
RDOCE246-14	LGC4515	MCNIP-1603	*Prochilodus vimboides*	-18.951	-42.361	BOLD:ACN4578
RDOCE046-13	LGC3595	LGC3595	*Pseudauchenipterus affinis*	-19.023	-42.125	BOLD:AAH8177
RDOCE255-14	LGC4651	MCNIP-1598	*Pseudauchenipterus affinis*	-19.239	-42.306	BOLD:AAH8177
RDOCE256-14	LGC4652	MCNIP-1598	*Pseudauchenipterus affinis*	-19.239	-42.306	BOLD:AAH8177
RDOCE260-14	LGC4712	3503	*Pseudauchenipterus affinis*	-19.081	-42.159	BOLD:AAH8177
RDOCE261-14	LGC4713	3504	*Pseudauchenipterus affinis*	-19.081	-42.159	BOLD:AAH8177
RDOCE262-14	LGC4719	3525	*Pseudauchenipterus affinis*	-18.994	-42.226	BOLD:AAH8177
RDOCE291-14	LGC4656	MCNIP-1599	*Pygocentrus nattereri*	-19.114	-42.177	BOLD:ABZ7351
RDOCE105-13	LGC3572	LGC3572	*Rhamdia* cf. *quelen*	-20.106	-43.403	BOLD:AAA6322
RDOCE107-13	LGC3577	LGC3577	*Rhamdia* cf. *quelen*	-20.025	-43.46	BOLD:AAA6322
RDOCE036-13	LGC1844	LGC1844	*Rhamdia* cf. *quelen*	-19.988	-41.72	BOLD:AAA6323
RDOCE041-13	LGC3578	LGC3578	*Rhamdia* cf. *quelen*	-20.025	-43.46	BOLD:AAA6322
RDOCE042-13	LGC3579	LGC3579	*Rhamdia* cf. *quelen*	-20.025	-43.46	BOLD:AAA6322
RDOCE078-13	RD13	RD13	*Rhamdia* cf. *quelen*	-19.996	-41.739	BOLD:AAA6323
RDOCE079-13	RD15	RD15	*Rhamdia* cf. *quelen*	-20.110	-43.400	BOLD:AAA6322
RDOCE167-14	RD91	RD91	*Rhamdia* cf. *quelen*	-18.967	-42.318	BOLD:AAA6322
RDOCE137-13	RD14	RD14	*Rhamdia* cf. *quelen*	-20.110	-43.400	BOLD:AAA6322
RDOCE104-13	LGC3564	ZUEC 8145	*Salminus brasiliensis*	-19.235	-42.313	BOLD:AAD2790
RDOCE160-14	RD60	RD60	*Serrapinnus heterodon*	-18.967	-42.318	BOLD:AAE1686
RDOCE161-14	RD61	RD61	*Serrapinnus heterodon*	-18.967	-42.318	BOLD:AAE1686
RDOCE162-14	RD62	RD62	*Serrapinnus heterodon*	-18.967	-42.318	BOLD:AAE1686
RDOCE186-14	RD59	RD59	*Serrapinnus heterodon*	-19.275	-42.425	BOLD:AAE1686

Process ID	Sample ID	Museum ID	Identification	Lat	Lon	BIN
RDOCE226-14	RD65	RD65	*Serrapinnus heterodon*	-18.962	-42.273	BOLD:AAE1686
RDOCE108-13	RD27	RD27	Tilapia rendalli	-19.022	-42.122	BOLD:ABZ6465
RDOCE109-13	RD28	RD28	*Tilapia rendalli*	-19.022	-42.122	BOLD:ABZ6465
RDOCE081-13	RD25	RD25	*Tilapia rendalli*	-20.046	-41.735	BOLD:ABZ6465
RDOCE082-13	RD26	RD26	*Tilapia rendalli*	-20.046	-41.735	BOLD:ABZ6465
RDOCE298-15	LGC4612	LGC4612	*Tilapia rendalli*	-20	-42	BOLD:ABZ6465
RDOCE085-13	LGC170	LGC170	*Trachelyopterus striatulus*	-19.986	-41.716	BOLD:ACI3769
RDOCE035-13	LGC1833	LGC1833	*Trachelyopterus striatulus*	-19.986	-41.716	BOLD:ACI3769
RDOCE043-13	LGC3588	LGC3588	*Trachelyopterus striatulus*	-19.022	-42.122	BOLD:ACI3769
RDOCE044-13	LGC3593	LGC3593	*Trachelyopterus striatulus*	-19.032	-42.126	BOLD:ACI3769
RDOCE066-13	LGC169	LGC169	*Trachelyopterus striatulus*	-19.986	41.7164	BOLD:ACI3769
RDOCE116-13	RD36	RD36	*Trichomycterus* aff. *Immaculatus*	-19.022	-42.122	BOLD:ACJ1022
RDOCE117-13	RD39	RD39	*Trichomycterus* aff. *Immaculatus*	-19.022	-42.122	BOLD:ACI3868
RDOCE118-13	RD40	RD40	*Trichomycterus* aff. *Immaculatus*	-19.022	-42.122	BOLD:ACI3868
RDOCE157-14	RD37	RD37	*Trichomycterus* aff. *Immaculatus*	-19.062	-42.162	BOLD:ACI3868
RDOCE147-13	RD94	RD94	*Trichomycterus* aff. *Immaculatus*	-20.051	-43.397	BOLD:ACI3868
RDOCE120-13	RD42	RD42	*Trichomycterus* aff.*alternatus*	-19.022	-42.122	BOLD:ACJ1151
RDOCE121-13	RD43	RD43	*Trichomycterus* aff.*alternatus*	-19.022	-42.122	BOLD:ACJ1151
RDOCE122-13	RD44	RD44	*Trichomycterus* aff.*alternatus*	-19.022	-42.122	BOLD:ACJ1151
RDOCE187-14	RD88	RD88	*Trichomycterus* aff.*alternatus*	-18.953	-42.361	BOLD:ACL7294
RDOCE324-15	RD87	RD87	*Trichomycterus* aff.*alternatus*	-18.952	-42.36	BOLD:ACL7294
RDOCE196-14	RD150	RD150	*Trichomycterus* cf. *brasiliensis*	-19.011	-43.372	BOLD:ACT6325
RDOCE201-14	LGC4143	MBML6825	*Trichomycterus longibarbatus*	-19.837	-40.555	BOLD:ACJ1022
RDOCE202-14	LGC4152	MBML6841	*Trichomycterus longibarbatus*	-19.889	-40.576	BOLD:ACJ1151
RDOCE163-14	RD71	RD71	*Trichomycterus* sp.	-18.967	-42.318	BOLD:ACI3868
RDOCE164-14	RD72	RD72	*Trichomycterus* sp.	-18.967	-42.318	BOLD:ACJ1164
RDOCE165-14	RD73	RD73	*Trichomycterus* sp.	-18.967	-42.318	BOLD:ACK5393
RDOCE146-13	RD79	RD79	*Trichomycterus* sp.	-20.109	-43.399	BOLD:ACJ9705
RDOCE191-14	LGC3686	LGC3686	*Trichomycterus* sp.	-18.974	-43.372	BOLD:ACT6325

Species	Comparisons	Minimum distance (%)	Mean distance (%)	Maximum distance (%)	Number of MOTUs			Number of specimens	BIN_ID	ABGD_ID	BIN	Status
					BIN	ABGD	bPTP					
Astyanax fasciatus	21	0	10.09	20.69	3	3	3	1	ABU7523	35	Discordant	N
								5	ACJ1542	76	Concordant	
								1	ACJ9650	42	Discordant	
Astyanax giton	3	0	0.52	0.84	1	1	1	3	ACL8007	73	Concordant	N
Astyanax lacustris	36	0	1.67	3.35	2	2	2	3	ABY8634	59	Concordant	N
								6	ABZ1711	5	Discordant	
Astyanax scabripinnis	10	0	9.12	21.82	2	3	2	1	AAC5910	36	Discordant	N
										81	Discordant	
								1	ACO5464	74	Singleton	
Astyanax sp.	15	0.79	8.77	20.5	4	4	4	1	AAC5910	81	Discordant	N
								3	ACJ9650	42	Discordant	
								1	ACT0040	75	Singleton	
								1	AAY4812	80	Discordant	
Astyanax taeniatus	10	0	1.48	3.96	2	2	2	1	AAY4812	80	Discordant	N
								4	ABU7523	35	Discordant	
Australoeros cf. *ipatinguensis*	10	0	0.1	0.17	1	1	1	5	ACR9799	48	Concordant	N.E
Brycon opalinus	10	0	0.22	0.47	1	1	1	5	ACL7114	44	Concordant	N.E.T
Brycon sp.	6	0.18	0.7	1.43	1	1	1	4	ACH8616	78	Concordant	N.E.T
Callichthys callichthys	-	-	-	-	1	1	1	1	AAB5066	56	Singleton	N
Characidium cf. timbuiense	28	0	5.98	9.9	3	3	3	4	ACI3743	20	Discordant	N
								3	ACJ1226	19	Discordant	
								1	ACJ9733	41	Singleton	
Characidium sp.	10	0	5.51	10.17	3	3	3	4	ACS9348	47	Concordant	NR
								1	ACI3743	20	Discordant	
								1	ACJ1226	19	Discordant	
Cichla kelberi	-	-	-	-	1	1	1	1	AAO9230	50	Concordant	I
Clarias gariepinus	-	-	-	-	1	1	1	1	AAB2256	8	Singleton	I
Crenicichla lacustris	15	0	10.76	21.36	2	2	2	2	AAD6380	28	Concordant	N
								4	ACO6050	53	Concordant	
Cyphocharax gilbert	10	0	0.56	1.14	1	1	1	5	ACK1539	4	Concordant	N
Delturus carinotus	15	0	0.15	0.33	1	1	1	6	ACC0184	3	Concordant	N.E
Duterodon pedri	3	0	0.19	0.29	1	1	1	3	AAY4812	80	Discordant	N.E

Table S2: Species from Doce river basin including intraspecific distance and additional analyses. E=endemic, I=Invasive, N=native, NR=New Record, T=Threatened.

Species	**Comparisons**	**Minimum distance (%)**	**Mean distance (%)**	**Maximum distance (%)**	**Number of MOTUs**			**Number of specimens**	**BIN_ID**	**ABGD_ID**	**BIN**	**Status**
					BIN	ABGD	bPTP					
Geophagus brasiliensis	66	0	0.39	1.22	1	1	1	12	AAA8514	2	Concordant	N
Gymnotus aff. *Carapo*	15	0	0.14	0.46	1	1	1	6	AAB6216	69	Discordant	N
Gymnotus sp.	36	0	3.96	6.32	3	3	3	3	AAB6216	69	Discordant	NR
								3	AAB6212	30	Concordant	
								3	ACT0768	70	Concordant	
Harttia sp.	28	0	4.67	12.2	3	3	1	1	ACJ1000	17	Singleton	N
								6	ACI6845	25	Concordant	
								1	ACO6155	51	Singleton	
Hasemania sp.	15	0	0.28	0.75	1	1	1	6	AAO6055	29	Concordant	N
Hisonotus sp.	3	0	0.55	0.83	1	1	1	3	ACW1732	32	Concordant	N
Hoplias intermedius	6	0	0.28	0.64	1	1	1	5	AAB1734	12	Concordant	N
Hoplias malabaricus	36	0	3.27	6.7	2	2	2	6	AAY4779	61	Concordant	N
								3	ACI3811	11	Concordant	
Hoplosternum littorale	-	0.3	0.3	0.3	1	1	1	2	AAB5068	6	Concordant	I
Hyphessobrycon bifasciatus	-	-	-	-	1	1	1	1	ACT0106	57	Singleton	NR
Hyphessobrycon eques	-	-	-	-	1	1	1	1	ABZ1711	5	Discordant	I
Hypomasticus mormyrops	21	0	0.13	0.49	1	1	1	7	ACH5050	60	Concordant	N
Hypostomus affinis	10	0	0.51	1.23	1	1	1	5	AAW9386	68	Concordant	N
Hypostomus auroguttatus	21	0	0.14	0.7	1	1	1	7	AAB9690	24	Discordant	N
Hypostomus sp.	-	-	-	-	1	1	1	2	AAB9690	24	Discordant	NR
Imparfinis sp.	10	0	0.33	0.76	1	1	1	5	AAC2103	23	Concordant	N
Knodus moenkhausii	15	0	1.21	3.07	1	2	1	1	AAM1485	46	Concordant	N
								5	AAM1485	79	Concordant	
Leporinus copelandii	6	0	0.82	2.21	1	1	1	4	ACI6721	9	Concordant	N
Lophiosilurus alexandri	-	0	-	-	1	1	1	1	AAE4855	7	Singleton	I
Loricariichthys castaneus	10	0.2	0.85	2.16	1	1	1	5	ACI6497	22	Concordant	N
Metynnis maculatus	-	-	-	-	1	1	1	1	AAE7443	39	Singleton	I
Neoplecostominae	36	0	0.3	0.8	1	1	1	10	ACC0721	33	Concordant	NR
Neoplecostomus sp.	10	0.16	3.46	6.12	2	2	2	2	AAX6581	34	Concordant	N
								3	ACT2675	72	Concordant	
Oligosarcus acutirostris	10	0	0.34	0.76	1	1	1	8	AAI3590	13	Discordant	NR
Oligosarcus argenteus	10	0.17	0.7	1.24	1	1	1	5	AAI3590	13	Discordant	N
Pareiorhaphis scutula	3	0.16	0.46	0.75	1	1	1	3	AAX0824	37	Discordant	NR.E.T
Pareiorhaphis sp.	10	0	2.74	4.44	2	3	2	2	AAX0824	37	Discordant	N.E
								3	ACI5663	77	Concordant	

Species	**Comparisons**	**Minimum distance (%)**	**Mean distance (%)**	**Maximum distance (%)**	**Number of MOTUs**			**Number of specimens**	**BIN_ID**	**ABGD_ID**	**BIN**	**Status**
					BIN	**ABGD**	**bPTP**					
Parotocinclus maculicauda	-	-	-	-	1	1	1	2	ACO5053	71	Concordant	NR
Phalloceros elachistos	3	0	0.14	0.43	1	1	1	3	ACO4001	49	Concordant	NE
Phalloceros sp.	3	0.16	0.42	0.72	1	1	1	3	AAB7265	45	Concordant	NR
Pimelodella sp.	10	0	0.25	0.48	1	1	1	5	AAC5327	21	Concordant	N
Pimelodus maculatus	-	-	-	-	1	1	1	1	AAB6504	54	Concordant	I
Poecilia reticulata	6	0	9.48	14.34	2	2	2	2	AAC0279	38	Discordant	I
								2	ACE9037	40	Concordant	
Poecilia vivipara	-	-	-	-	1	1	1	1	AAC0279	38	Discordant	N
Pogonopoma wertheimeri	6	0.33	0.48	0.78	1	1	1	4	ACI3792	27	Concordant	I
Prochilodus costatus	15	0	1.32	2.6	2	1	1	4	ADC2568	10	Concordant	I
								2	ADC2571	10	Concordant	
Prochilodus vimboides	-	-	-	-	1	1	1	1	ACN4578	52	Singleton	N.T
Pseudauchenipterus affinis	15	0	0.52	1.16	1	1	1	6	AAH8177	26	Concordant	N
Pygocentrus nattereri	-	-	-	-	1	1	1	1	ABZ7351	55	Singleton	I
Rhamdia quelen	36	0	1.25	3.48	2	2	2	7	AAA6322	15	Concordant	N
								2	AAA6323	62	Concordant	
Salminus brasiliensis	-	-	-	-	1	1	1	1	AAD2790	14	Singleton	I
Serrapinus heterodon	15	0.32	0.76	1.36	1	1	1	6	AAE1686	31	Concordant	N
Steindachneridion doceanum	-	-	-	-	1	1	1	1	ACT0106	58	Singleton	N. E. T
Tilapia rendalli	10	0	0	0	1	1	1	5	ABZ6465	16	Concordant	I
Trachelyopterus striatulus	10	0	0.26	1.23	1	1	1	5	ACI3769	1	Concordant	N
Trichomycterus aff. Alternatus	10	0	10.8	18.49	2	2	2	3	ACJ1161	64	Discordant	N
								2	ACL7294	43	Concordant	
Trichomycterus aff. *Auroguttatus*	-	-	-	-	1	1	1	1	ACJ1164	64	Discordant	NR
Trichomycterus brasiliensis	-	-	-	-	1	1	1	2	ACT6325	65	Discordant	N
Trichomycterus immaculatus	-	0.15	2.23	5.84	2	2	2	4	ACI3868	63	Discordant	N
								1	ACJ1022	18	Discordant	
Trichomycterus longibarbatus	-	-	-	-	2	2	2	1	ACJ1022	18	Discordant	NR
								1	ACJ1161	64	Discordant	
Trichomycterus sp.	10	1.54	4.17	5.8	5	5	5	1	ACI3868	63	Discordant	N
								1	ACJ1164	64	Discordant	
								1	ACT6325	67	Discordant	
								1	ACJ9705	66	Singleton	
								1	ACK5393	65	Singleton	

Chapter III

INFLUENCE OF PRESERVATION METHODS, SAMPLING MEDIUM AND SAMPLING TIME ON eDNA RECOVERY IN NEOTROPICAL LOTIC ENVIRONMENTS

Chapter III

INFLUENCE OF PRESERVATION METHODS, SAMPLING MEDIUM AND SAMPLING TIME ON eDNA RECOVERY IN NEOTROPICAL LOTIC ENVIRONMENTS

Naiara Guimaraes Sales[1], Owen Simon Wangensteen[2], Daniel Cardoso Carvalho[3], Stefano Mariani[1]

Study design: NGS and SM. Field work and sample collection: NGS. Laboratory experiment: NGS and OSW. Data analyses: NGS, OSW, SM. Manuscript writing: NGS, DCC, OSW, and SM.

Submitted to Environmental DNA Journal (December/2018).

Running title: eDNA recovery in Neotropical rivers.

3.1 Abstract

Environmental DNA (eDNA) metabarcoding has rapidly emerged as a promising biodiversity assessment technique, proving to be a sensitive and cost-effective method. Despite the increasing popularity of using eDNA to survey species, several questions regarding its limitations remain to be addressed. We investigated the effect of sampling medium, time, and preservation methods, on fish detection performance based on eDNA metabarcoding of neotropical freshwater samples. Water and sediment samples were collected from 11 sites along the Jequitinhonha River, Southeastern Brazil. Sediment samples were stored in ethanol, while the same amounts of water per sample (3L) were stored in a cool box with ice, as well as by adding the cationic surfactant Benzalkonium chloride (BAC). Sediment and water samples yielded a similar amount of fish MOTUs (237 vs 239 in the first sampling event, and 153 vs 142 in the second sampling event). Water stored in ice provided better results than those preserved in BAC (239 and 142 vs 194 and 71 MOTUs). While documenting the effectiveness of eDNA surveys as practical tools for fish biodiversity monitoring in poorly accessible areas, we showed that keeping water samples cooled results in greater eDNA recovery and taxon detection than by adding cationic surfactants as sample preservatives. Furthermore, by comparing two sets of samples collected from the same locations at two distinct sampling events, we highlight the importance of conducting multiple sampling events when attempting to recover a realistic picture of fish assemblages in lotic systems.

Key-words: environmental DNA, freshwater, metabarcoding, ichthyofauna, neotropical.

3.2 Introduction

Environmental DNA metabarcoding has been hailed as a promising tool for biodiversity assessment and monitoring worldwide, in both marine and freshwater ecosystems (Bohmann et al., 2014; Boussarie et al., 2018; Deiner et al., 2017; Hänfling et al., 2016; Pont et al., 2018; Thomsen & Willerslev, 2015). This method relies on obtaining the DNA shed by organisms in the surrounding environment (e.g. water, soil), amplifying it with primers targeting the taxonomic spectrum of interest, and sequencing it to reconstruct community composition (Bohmann et al., 2014; Handley et al., 2018; Valdez-Moreno et al., 2018; Valentini et al., 2016).

Despite the increased number of publications in the past decade, the application of eDNA techniques is still not considered straightforward (Taberlet, Bonin, Zinger, & Coissac, 2018). Molecular and bioinformatics protocols continue to be revised and optimized, while uncertainties remain as to how to streamline and rationalize sampling and sample preservation (Dickie et al., 2018). The usefulness of eDNA approaches depend on their ability to provide effective and accurate detection of species, thus requiring a better understanding of the factors influencing detection rates (Lodge, 2012). Detectability of eDNA in environmental samples is limited mainly by three processes: i) eDNA production (i.e. rate of DNA shedding), ii) degradation, iii) removal and transport (Barnes and Turner, 2016; Strickler, Fremier & Goldberg, 2015). Several factors can affect eDNA production, such as the type of organism/species (with some species showing a higher eDNA release rate than others – Maruyama, Nakamura, Yamanaka, Kondoh, & Minamoto, 2014; Sassoubre, Yamahara,

Gardner, Block, & Boehm, 2016), biomass, density and life stage of specimens (Maruyama et al., 2014; Takahara, 2012), season (Buxton, Groombridge, Zakaria, & Griffiths, 2017), and water oxygen and temperature which can cause behavioral and physiological changes (e.g. stress) and affect metabolic rates, hence influencing eDNA production (Jo et al., 2010; Maruyama et al., 2014; Pilliod, Goldberg, Arkle, & Waits, 2014). After eDNA is released in the water it starts to be removed through transport and/or degradation. eDNA can settle and bind to sediment, and/or be transported by long distances depending on the type of environment (e.g. lotic, lentic), and thus, degrade and become diluted during the transport downstream (Strickler et al., 2015).

The DNA released in the environment can be degraded at a fast pace, hampering the identification of rare species and providing false negatives (Barnes et al., 2014; Dejean et al., 2011; Pilliod et al., 2014; Strickler et al., 2015), which leads to the need for improved preservation systems that can maximize eDNA recovery (Fonseca, 2018; Hansen, Bekkevold, Clausen, & Nielsen, 2018). The persistence of DNA in environmental samples can be influenced by many factors (e.g. temperature, microbial activity, pH, salinity, solar radiation), and detectability of eDNA in water has been shown to be associated with cold temperatures, alkaline conditions, and low UV-B levels (Strickler et al., 2015; Tsuji, Ushio, Sakurai, Minamoto, & Yamanaka, 2017). Even though several studies suggest a negligible role of temperature, UV levels or seasonality on DNA degradation (Andruszkiewicz, Sassoubre, & Boehm, 2017; Collins et al., 2018; Robson et al., 2016; Seymour et al., 2018).

The most recommended approach to reduce degradation is to extract the DNA as quickly as possible after sampling. However, due to the constraints of field work conducted in remote sites located far from laboratory facilities (e.g. difficulties for on-site filtration due to lack of equipment, and risk of contamination), the filtering process and subsequent DNA extraction might not be possible or advisable, and a preservation method for the medium sampled must be employed in order to block biological activities and minimize DNA degradation.

Different approaches have been tested to preserve water samples before the filtering process, showing distinct benefits and drawbacks. Storing the samples at low temperatures, including freezing the samples or cooling using a cool box, are widely employed; however, these approaches entail equipment requirement increase; whereas the efficiency of cooling the samples has also been questioned (Eichmiller, Best, & Sorensen, 2016; Pilliod et al., 2014). Inclusion of buffers, such as EtOH–NaAc (ethanol-sodium acetate) solution, have been reported to show an eDNA persistence rate similar to samples stored in ice (Ladell, Walleser, McCalla, Erickson, & Amberg, 2018), however, when sampling larger volumes of water the increased final volume obtained (i.e. addition of over 2x of solution) might be considered as a problem during long sampling campaigns. Recently, Yamanaka et al. (2017) tested the addition of cationic surfactants as preservatives to suppress DNA degradation at ambient temperatures and demonstrated the efficiency of Benzalkonium chloride (0.01%) in retaining eDNA concentration even after 10-day incubation at 21°C. Still, despite being considered as an effective eDNA preservative, this preservation method was restricted to a species specific

eDNA recovery test and the effectiveness of the cationic surfactant in preserving eDNA samples for metabarcoding analysis has not yet been evaluated.

The application of eDNA as a biodiversity assessment tool requires the development, field validation and optimization of protocols in order to minimize bias and tailor procedures to the variety of environments and habitats investigated (Taberlet et al., 2018). Furthermore, the occurrence of a time lag between species presence and sampling event can contribute to DNA degradation, leading to an erroneous inference of species absence (i.e. short time frame detection due to high degradation rates may hamper the eDNA efficiency in detecting species where they are present). Sediment samples have shown to contribute to tackling this issue once DNA attached to sediments can be detected longer than in the water column. In addition, sediment samples can provide a higher concentration and longer persistence of genetic material for studying past and current species presence, also contributing to understand issues associated with eDNA transport and removal (Turner, Uy, & Everhart, 2015).

Neotropical freshwaters harbor high, and often understudied biodiversity (Sales, Mariani, Salvador, Pessali, & Carvalho, 2018), and eDNA could assist biodiversity assessment and monitoring programs, with the ultimate aim to contribute to conservation and management strategies. Higher temperatures,solar radiation, and associated turbidity in tropical waters might contribute to make rivers in the tropics a challenge for eDNA studies due to hypothesized higher degradation rates (Barnes et al., 2014; Matheson, Gurney, Esau, & Lehto, 2014; Pilliod et al., 2014). A rapid removal of eDNA (through transport and

degradation) might hamper the detection of species and lead to false negatives (Hansen et al., 2018), compromising the use of this method for biodiversity assessment and monitoring. In this context, testing effectiveness of sampling methods is particularly important in remote and tropical locations (Ladell et al., 2018). Furthermore, the knowledge regarding the use of eDNA in tropical rivers remains scarce and despite being considered as a promising tool for fish biodiversity assessment in this region, this approach still requires the optimization of field and laboratory protocols (Cilleros et al., 2018). Here, we hypothesized that: I) preservation method effects eDNA recovery and MOTU detection; II) sample medium detects different communities due to presumed preservation time; III) time of sampling detects different communities To advance our knowledge in how to collect, preserve and obtain eDNA samples in Neotropical catchments we obtained water and sediment samples from 11 sites located along the main stem of River Jequitinhonha (South-Eastern Brazil), and: a) compared two preservation methods for water samples (cooling the samples using ice and adding the cationic surfactant Benzalkonium chloride – BAC); b) compared MOTU recovery from water vs sediment samples, and c) examined the influence of short-term temporal sample replication by sampling the same locations between a three-week interval.

3.3 Material and Methods

Study Site

The Jequitinhonha River Basin, located in Southeast Brazil, flows through two biodiversity hotspots (Atlantic Forest and Cerrado) encompassing an area of 70,315 km^2 and running over 1082 km. This region is characterized by tropical climate and environmental heterogeneity, including semi-arid regions with high temperatures (annual mean of 24.9°C) and dry period extending over six months per year (Climate-Data, 2018, Bilibio, Hensel, & Selbach, 2011). This catchment, located in one of the poorest and least studied regions of Brazil, is part of an ecoregion (Coastal Drainages of Eastern Brazil) that harbors considerable fish biodiversity and one of the highest numbers of endemic and threatened fish species in Brazil (Machado, Drummond, & Paglia, 2008, Pugedo, Andrade-Neto, Pessali, Birindelli, & Carvalho, 2016, Rosa & Lima, 2008).

eDNA sampling and processing

Sediment and water samples were obtained from 11 sample sites, in the Jequitinhonha River Basin, during two replicated sampling events conducted with a three-week interval (Figure 1, Table S1 Supplementary Material). In each sampling event, 6 liters of water were collected from each sample site (i.e. 3 samples each of 1 liter each, per treatment as described as following) and before the filtering process the water was preserved using two different methods to compare their efficiency. One set of samples (N=3) was stored at low

temperatures (using a cooling box with ice), while in the other samples (N=3) the cationic surfactant benzalkonium chloride (BAC) was added at a final concentration of 0.01% (Yamanaka et al. 2017). Water samples were filtered approximately 8 hours after collection, using Microfil V, 100mL, mixed cellulose esters (MCE) filters (diameter: 47 mm, pore size: 0.45 µm, Merck Millipore) (Bakker et al. 2017) in combination with an automatic vacuum pump. Filters were stored in microcentrifuge tubes containing silica beads (Bakker et al. 2017). Sediment samples (2 samples/locality) were obtained in the shores, from the superficial layer (approximately 5cm), and were stored in 50mL centrifuge tubes and preserved in 100% ethanol.

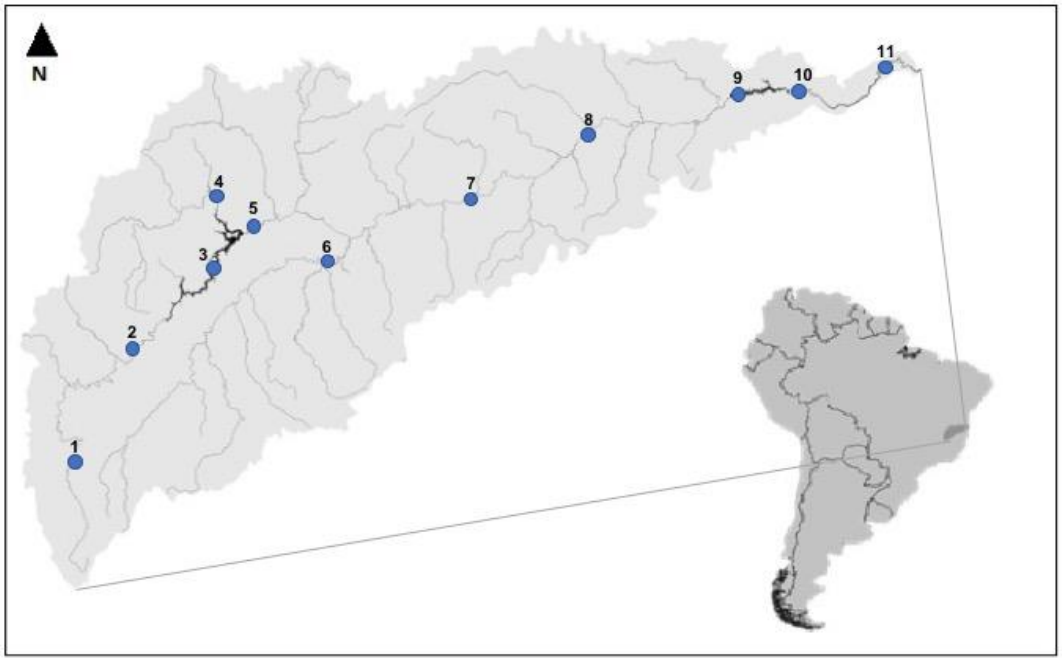


FIGURE 1 | Map of Jequitinhonha river basin sampling locations.

DNA extraction from the filters was conducted using the Dneasy PowerWater Kit (Qiagen) and DNA from the sediments was extracted from 10g of sediment using Dneasy PowerMax Soil Kit (Qiagen), following the manufacturer's protocol. Purified extracts were checked for DNA concentration in a Qubit fluorometer (Invitrogen).

A contamination control procedure was applied in both field and laboratory works to avoid the occurrence of contamination. All samples were stored in disposable sterile collection bottles, disposable gloves were worn at all times, sampling and laboratory equipment and surfaces were treated with 50% bleach solution for 10 minutes, followed by rinsing in water after each use. Filtration blanks were run between every sample site, immediately before the next filtration in order to test for potential contamination during the filtration stage.

Amplification, library preparation and sequencing

The amplification of eDNA metabarcoding markers was conducted using a previously published fish-specific 12S primer set (Miya et al., 2015). Amplicons of ~172bp from a variable region of the mitochondrial 12S rRNA gene were obtained with the primers (MiFish-U-F, 5'-GCCGGTAAAACTCGTGCCAGC-3'; MiFish-U-R, 5'-ACATTATCATAGTGGGGTATCTAATCCCAGTTTG -3').

A total of 183 samples including collection blanks (N=3) and laboratory negative controls -DNA extraction blanks (N=2) and PCR blanks (N=2) were sequenced in a single

multiplexed Illumina MiSeq run using 2 sets of 96 primers with seven-base sample-specific oligo-tags and a variable number (2-4) of leading Ns (fully degenerate positions) to increase variability in amplicon sequences. DNA extractions were not normalised prior PCR reactions and PCR amplification was conducted using a single-step protocol and to minimize stochasticity in individual reactions, PCRs were replicated three times for each sample and the products subsequently pooled into single samples. The PCR reaction consisted of a total volume of 20 µL including 10 µl AmpliTaq Gold™ 360 Master Mix (Applied Biosystems); 0.16 µl of bovine serum albumin; 1 µl of each of the two primers (5 µM); 5.84 µl of ultra-pure water and 2 µl of eDNA template. The PCR profile included an initial denaturing step of 95°C for 10 min, 40 cycles of 95°C for 30s, 60°C for 45s, and 72°C for 30s and a final extension step of 72°C for 5 min. Amplifications were checked through electrophoresis in a 1.5% agarose gel stained with GelRed (Cambridge Bioscience). PCR products were pooled in two different sets and purified using MinElute columns (Qiagen), and Illumina libraries were built from each set, using a NextFlex PCR-free library preparation kit (Bioo Scientific) with unique 6-bp library tags. A left-sided size selection was performed using 1.1x Agencourt AMPure XP (Beckman Coulter). Libraries were then quantified by qPCR using a NEBNext qPCR quantification kit (New England Biolabs) and pooled in equimolar concentrations along with 1% PhiX (v3, Illumina). The libraries were run at a final molarity of 10pM on an Illumina MiSeq platform in a single MiSeq flow cell using the 2x 150bp v2 chemistry.

Bioinformatics analyses

Bioinformatic analyses were based on the OBITools 1.2.2 metabarcoding package (Boyer et al. 2016). FastQC was used to assess the quality of the reads, paired-end reads were aligned using illuminapairedend, and dataset demultiplexing and primer removal were then conducted using ngsfilter command. A bespoke filter using obigrep was used to select fragments of 140-190bp and remove short fragments originated from library preparation artefacts (primer-dimer, non-specific amplifications) and reads containing ambiguous bases. Clustering of strictly identical sequences was performed using obiuniq and a chimera removal step was applied in vsearch 2.7.1 (Rognes, Flouri, Nichols, Quince, & Mahé, 2016) through the uchime-denovo algorithm (Edgar, Haas, Clemente, Quince, & Knight, 2011). Molecular Operational Taxonomic Unit (MOTU) delimitation was performed using SWARM 2.0 algorithm (Mahé, Rognes, Quince, de Vargas, & Duthorn, 2015) with a distance value of d=3 (Siegenthaler et al., 2018) and ecotag (Boyer et al. 2016) was used for the subsequent taxonomic assignment, with a custom reference database including all known vertebrate sequences for the sequenced 12S fragment (Siegenthaler et al., 2018). Ambiguous taxonomic assignments (more than one species assigned per MOTU) after ecotag were checked using BLAST against the Genbank nucleotide database.

A conservative approach was applied to our analyses to avoid false positives and exclude MOTUs/reads putatively belonging to sequencing errors or contamination. Reads detected in the negative controls were removed from all samples, and MOTUs containing less

than 5 reads were excluded from subsequent analyses.

Statistical analyses

Samples were grouped according to the treatments analyzed (Table 1) and all statistical analyses were performed in R v3.5.1 (https://www.R-project.org/). Due to differences in the sequencing depth for each sample, relative read abundances were used for all statistical analyses (i.e. for each sample the MOTU counts were divided by the total amount of reads obtained for that sample). The vegan package was used to perform the nonparametric method permutational multivariate analysis of variance (PERMANOVA - Anderson, 2017), through the 'adonis' function (Bray-Curtis dissimilarities, 1000 permutations). Separate testes were used and pairwisecomparisons were performed on relative abundances calculated for MOTUs in each sample site, per preservation method (BAC vs ICE), sampling time (1st round vs 2nd round), and per sampling medium (water vs sediment), to verify the influence of these factors over eDNA recovery. A significance threshold of $p < 0.05$ was applied at all analyses.

TABLE 1 | Treatments analyzed according to sampling medium, preservation method used and sampling event.

CODE	Sampling Medium	Preservation method	Sampling event	N	Sampling Period
SED1	Sediment	Ethanol	1	22 (2x11)	22/01-01/02/2017
SED2	Sediment	Ethanol	2	22(2x11)	19/02-01/03/2017
BAC1	Water	Benzalkonium chloride	1	33 (3x11)	22/01-01/02/2017
BAC2	Water	Benzalkonium chloride	2	33 (3x11)	19/02-01/03/2017
ICE1	Water	ICE	1	33 (3x11)	22/01-01/02/2017
ICE2	Water	ICE	2	33 (3x11)	19/02-01/03/2017

Non-metric multidimensional scaling plots were obtained using Bray-Curtis dissimilarity, through PAST3 software (Hammer, Harper, & Ryan, 2001). Ggplot2 and esquisse packages were used to build ggplot charts in R, and due to an incomplete reference database and a relatively low taxonomic resolution of the 12S fragment we used the taxonomic assignment down to family level to compare those methods regarding their performance in detecting teleost fish communities. Venn diagrams were obtained with BioVenn (Hulsen, Vlieg, & Alkema, 2008).

3.4 Results

Library quality and raw data

A total of 16,104,492 raw reads were obtained in one Illumina MiSeq run (Library 1: 6,399,823 reads, Library 2: 9,704,669 reads), including 44 sediment samples and 132 water samples. 10,064,034 reads were kept after initial quality filtering and removal of chimaeras. After applying a subsequent conservative filtering step (retaining only reads taxonomically assigned to Actinopterygii, and removal of MOTUs containing less than 5 reads) the number of reads per sample ranged from 0 (sample 10 – sediment; second sampling event) to 127,250. The final dataset comprised 311 MOTUs distributed differently in each treatment analyzed (Figure 2).

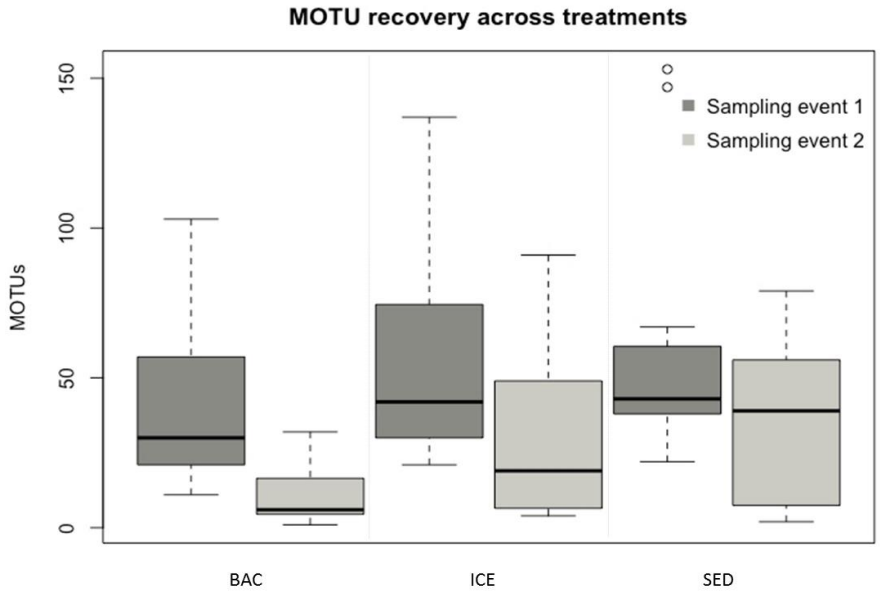


FIGURE 2 | Number of MOTUs recovered per sampling medium and preservation method (sediment vs water – BAC and ICE) and sampling event.

Taxonomic assignment

All MOTUs from the sediment samples could be taxonomically assigned at order level (see Appendix A1) whereas at family level the assignment rate was 96.4% (SED1) and 95.68% (SED2). Regarding the water samples, at order and family levels the assignment rates were, respectively, 98.97% and 95.88% for BAC1, 97.47% and 93.68% for BAC2, 100% and 96.83% for ICE1, and 98.72% and 94.17% for ICE2.

Influence of preservation method, sampling medium, and sampling time

All results of the PERMANOVA analyses (Bray-Curtis, $p<0.005$), including effect size (R^2) and significance (p-value) are summarized in Table S2, Supplementary Material. A significant difference ($p<0.05$) in MOTU composition among all the treatments was found and to verify the influence of preservation methods, sampling medium, and sampling time we performed pairwise comparisons for all combinations of treatments.

The influence of preservation method on MOTU diversity recovery was small (around 2% variance explained) but significant between samples collected during the first sampling event (BAC1 vs ICE1, $p=0.016$). However, no significant effect was detected for the preservation methods in the second sampling event (BAC2 vs ICE2, $p=0.06$, Table S2).

Overall and also in all pairwise comparisons, a significant difference between sediment and water samples was detected. Non-metric multidimensional scaling (nMDS) (Figure 3) showed a much greater variability among the water samples when compared to

the sediment samples, and a greater separation of water samples was apparent for the first sampling event (Figure 3A). During the second sampling, a higher similarity between sediment and water samples preserved cooled was found (Figure 3B), and the highest effect size (R^2=0.08) was found between SED2 and BAC2 (sediment and water samples preserved in BAC, collected during the second sampling event).

When testing for the effect of sampling event, the community composition differed from the two events for all treatments analyzed, showing a highest effect size for the sediment samples (R^2=0.07) and a lower effect size for the water samples preserved in BAC (R^2=0.04). A smaller effect was found for preservation method than sampling medium and time. Despite showing significant differences, overall, the R^2 effect sizes never accounted for any more than 8% of the variance, with a mean around 6%.

The Venn diagram overlaps showed a high similarity between the treatments in the first sampling event with 56.8% of the MOTUs detected in all of them (Figure 4). However, for the second sampling event a higher dissimilarity was detected when comparing the methods applied with only 27.55% of the MOTUs recovered being detected in all three methods (sediment, BAC, ICE).

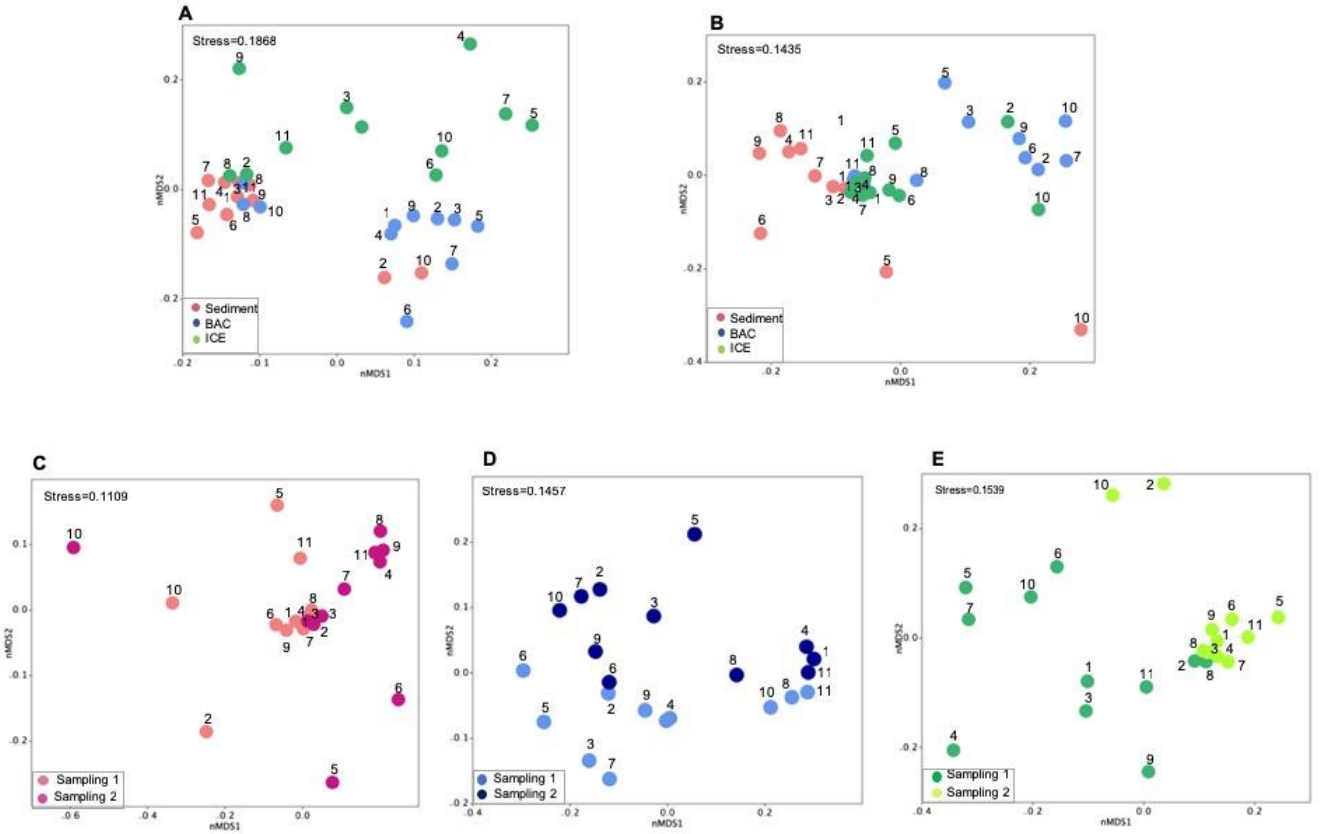
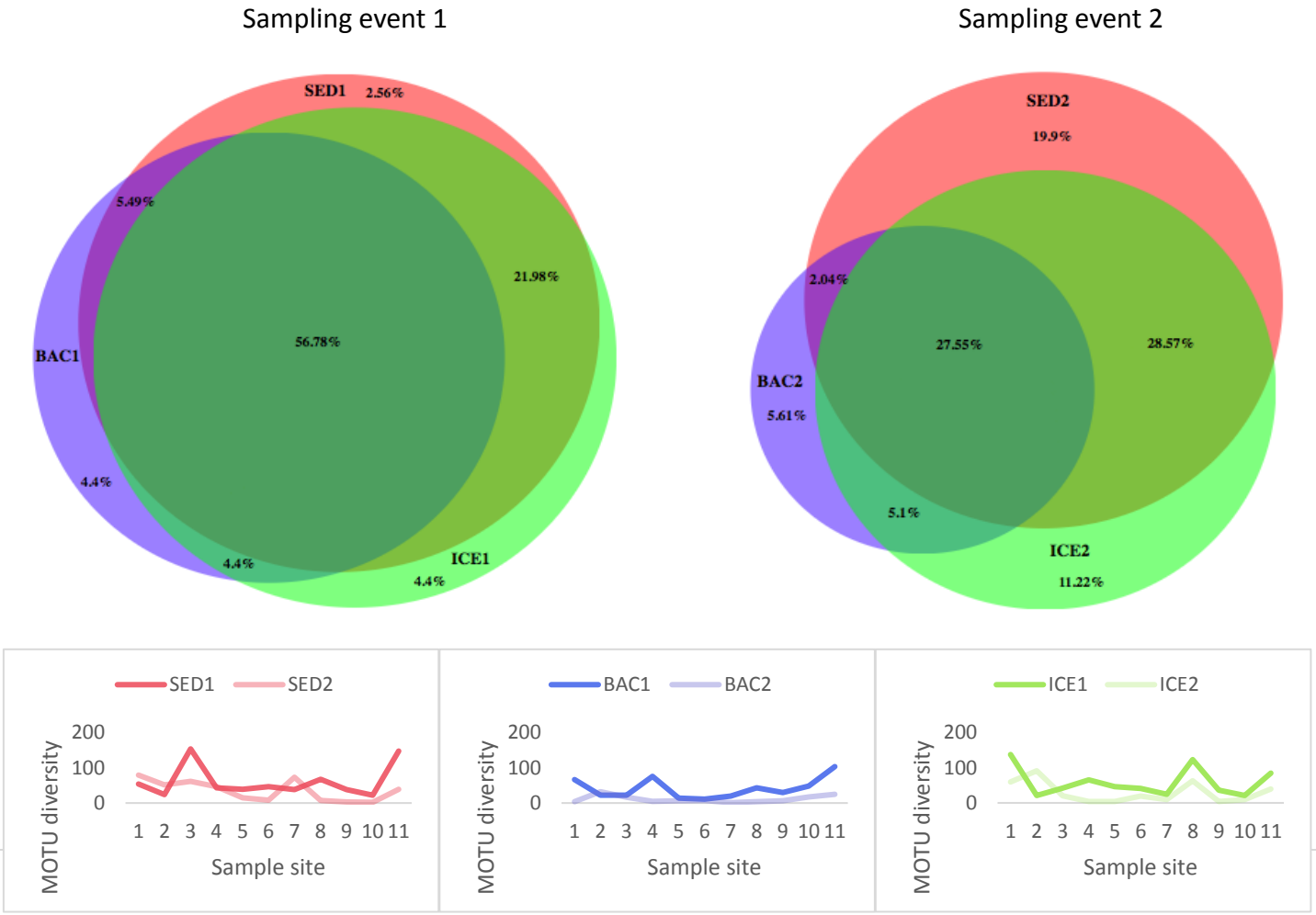


FIGURE 3 | Non-metric multidimensional scaling (nMDS) plots showing Bray-curtis dissimilarities of sample sites per sampling event. Analyses based on A) Sampling event 1; B) Sampling event 2; C) Sediment samples; D) Water samples preserved using BAC; and E) Water samples preserved using ICE.

FIGURE 4 | Comparison of MOTU recovery between sampling events.



Community composition across treatments

In total, we detected 7 orders (Characiformes, Cichliformes, Clupeiformes, Cypriniformes, Cyprinodontiformes, Gymnotiformes, and Siluriformes) and 20 families. Order and family richness obtained were compared using ggplot charts (Figure 5) and showed a slight difference across all treatments. As for preservation methods, the relative read abundance (%) was similar between water samples preserved in BAC and ICE for the first sampling, however, eDNA from two families of Siluriformes (Callichthyidae and Auchenipteridae) was not recovered from samples preserved using the cationic surfactant.

During the second sampling, the relative read abundance slightly differed between these two methods with a highest amount of reads from Trichomycteridae (Order Siluriformes) and also absence of reads from Pimelodidae (Order Siluriformes) in samples with added BAC. Thus, samples stored in ICE outperformed samples preserved with BAC in both MOTUs recovery and order/family richness.

Regarding the sampling medium, sediment samples provided similar results to water samples, except in the order Siluriformes, where it outperformed water samples preserved with BAC by detecting the family Auchenipteridae, and was surpassed by water samples preserved in ICE in detecting the family Callichthyidae, during the first sampling event. Whereas during the second sampling, the sediment samples did not recover MOTUs from two orders (Gymnotiformes and Cypriniformes) but detected one order (Clupeiformes) not identified in the water samples.

In contrast with results obtained for MOTUs recovery, despite showing a lower amount of MOTUs when compared to samples obtained in the first sampling event, samples obtained in the second event allowed the detection of additional orders and families. For the sediment samples, two orders were not detected (Cypriniformes and Gymnotiformes) but one order (Clupeiformes) and one additional family of Siluriformes (Callichthyidae) were only detected in sediments collected at the second sampling time. Regarding the samples preserved in BAC, two families of the order Siluriformes were not detected during the second sampling (Claridae and Pimelodidae) and two additional families of the same order were included (Callichthyidae and Auchenipteridae), while samples stored in ICE detected one fewer family (Callichthyidae) when compared to the first sampling.

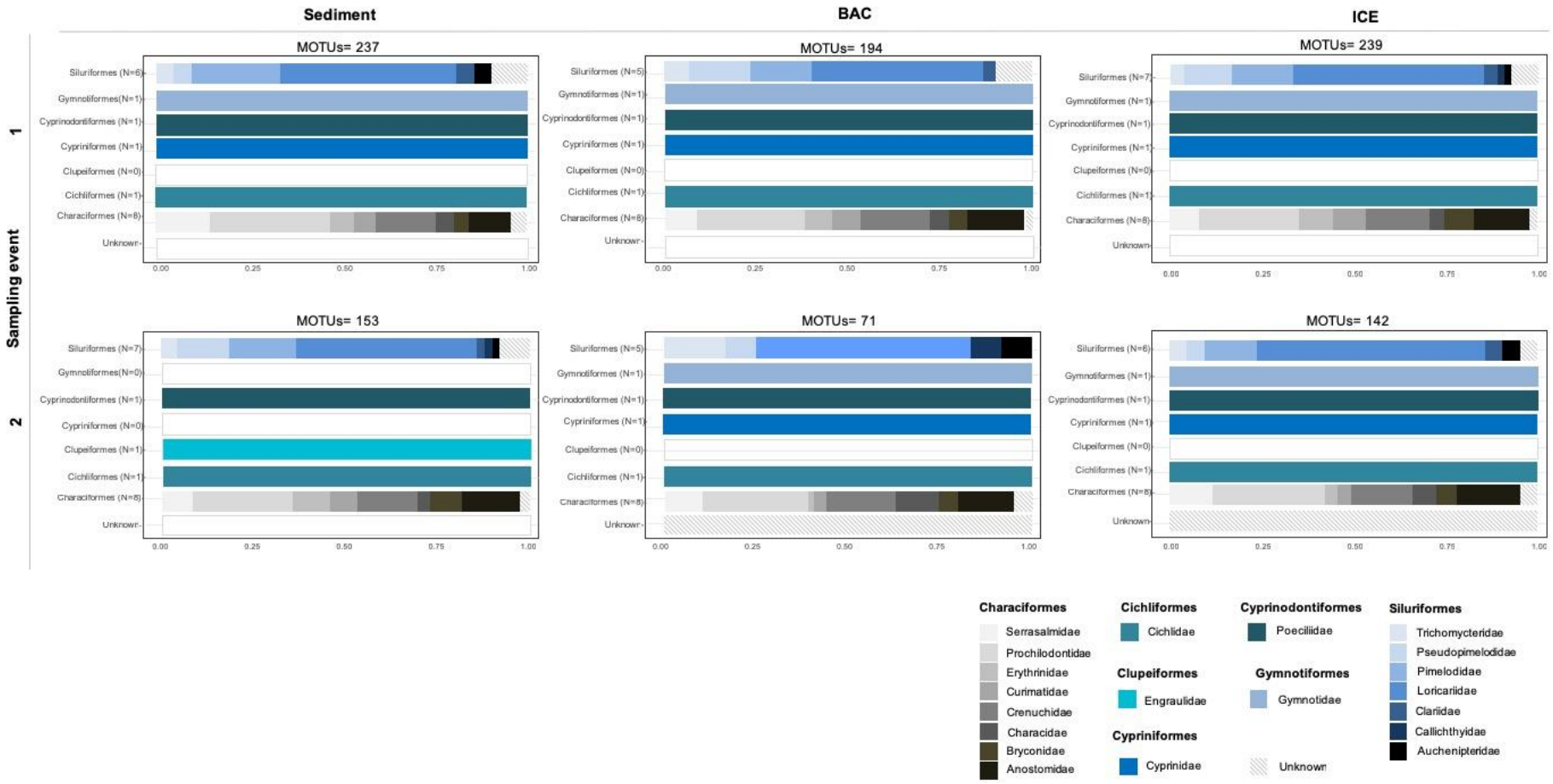


FIGURE 5 | Relative read abundance per order and family.

3.5 Discussion

Despite the exponential increase of eDNA publications, most of the studies have been conducted in temperate regions and in fairly well accessible areas. To date, few studies have tested the use of eDNA metabarcoding in remote tropical sites. Here, we tested two preservation methods for neotropical water samples (cooling the samples vs adding a cationic surfactant as preservative) and also, we tested the influence of sampling medium (water vs sediment) and time on eDNA recovery to evaluate the most suitable method and provide a framework for downstream studies in tropical catchments.

Overall, comparisons between preservation methods showed a smaller effect on eDNA recovery than sampling medium and time (Table S2). Sediment and water samples kept in cooling boxes outperformed water samples preserved with the cationic surfactant solution (237 and 239 against 194 MOTUs, respectively), while the highest amount of MOTUs was detected during the first sampling event for all treatments. Most of the variance found resides within the treatments analyzed, this variance may be due to: i) the distribution of eDNA might be heterogeneous in rivers showing different spatial structures ; ii) eDNA transport distances may vary between species (Deiner & Altermatt, 2014); iii) natural differences found in community composition across samples sites, as the structure of freshwater fish communities are influenced by complex interactions and by heterogeneity of freshwaters along the river gradient (e.g. geomorphic and hydrologic conditions, microbiota, temperature, pH, acidity, and chemical composition - Spurgeon, Pegg, Parasiewicz, & Rogers, 2018). Also, as shown by Macher and Leese (2018) community composition can change even when sampling the same

location in a time frame shorter than one minute and our findings also agree with earlier authors in that patterns of persistence of eDNA in rivers can be irregular.

Despite showing a significant difference, a small effect size was found for comparisons between preservation methods. The effect of preservation method might be related to the physical state of DNA molecules in the sample, free DNA can bind to humic substances and thus, be protected from enzymatic degradation and show a decreased rate on eDNA removal (Crecchio & Stotzky, 1998). Environmental DNA persistence can also be affected by the trophic state of the aquatic environment, showing a higher detectability in dystrophic and eutrophic waters than in oligotrophic systems (Eichmiller et al., 2016). The Jequitinhonha River is characterized by acid waters and contains mostly dystrophic and eutrophic soils (Intertechne, 2010). A faster degradation of eDNA throught chemical process is known to occur in acidic environments, and perhaps, in this case, low temperatures could better preserve the eDNA molecules on water samples and might be more important to eDNA preservation than adding the cationic surfactant which are mostly used to decrease/stop degradation caused by microbial activities. However, degradation rates at complex tropical environments, such as the Jequitinhonha River, have not been evaluated and the trends for eDNA persistence remain unknown in this realm. A similar result was found by Laddel et al. (2018), who compared lowering the temperature of samples to adding EtOH–NaAc, where cooling of the samples outperformed the use of a buffer solution. It should also be noted that some of the discrepancies between ICE and BAC detections may simply be due to the reduction of stochasticity afforded by the additional PCRs conducted on each water sample (18 replicates in total) (Leray & Knowlton, 2017).

Thus, despite increasing the equipment need, cooling may be considered as the first option to decrease DNA degradation in water samples during field collection. Unless no other option is available, cationic surfactant solutions might not be worthwhile for field sampling in remote areas due to the difficulties in accessing these specific laboratory reagents and the significant safety hazard posed by these chemicals (Ladell et al., 2018). However, if neither filtering nor cooling is feasible for a few hours after sampling, the use of some form of preserving buffer should remain a requirement.

Community composition is expected to differ between sampling media, as previous eDNA studies have found sediment to show a higher DNA concentration and a longer detectability than surface water (Turner et al., 2015). Since DNA can persist longer when incorporated into the sediment, temporal inference may be challenging (Turner et al., 2015); on the other hand, a higher degradation rate and lower detection lag time in aqueous eDNA samples provide a contemporary snapshot of the biodiversity being assessed (Hansen et al., 2018). Here, we have found a significant difference ($p<0.05$) and a higher size effect (R^2=0.06-0.08) on MOTU recovery between sediment and water samples (Table 2). Sediment samples outperformed water samples preserved with BAC by detecting the family Auchenipteridae (Order Siluriformes), and was surpassed by water samples preserved in ICE in detecting the family Callichthyidae, during the first sampling event. In the second sampling event, sediment samples failed to detect the family Callichthyidae and the orders Gymnotiformes and Cypriniformes, however, the order Clupeiformes was only found in the sediments. MOTUs obtained (19.9%) for the second sampling event were exclusive to sediment. MOTUs detected only in water samples might indicate the contemporary presence of those while their absence

in sediments samples may be due to a short time frame for those to settle and bind to the substrate. MOTUs belonging to the order Clupeiformes were detected only in sample site 11, located at the river mouth and refer to marine species that occasionally venture into the river to feed (Andrade-Neto, 2010). Although these species might not have been there at the time of sampling, they might have shed DNA during their incursions and the eDNA bound to sediment can have persisted longer than the eDNA in the surface water, contributing to its later detection. Thus, combining sediment and water samples may contribute to obtain a snapshot of the fish community that can distinguish between resident and transient species.

TABLE 2 | PERMANOVA results (R^2-effect sizes and significance level) showing the effect of sampling medium on MOTU diversity recovery.

Sampling medium	R^2	Effect	Significance (*p*-value)
SED *vs* WAT	0.03626	*	0.00099
SED1 *vs* WAT1	0.07234	*	0.00999
SED2 *vs* WAT2	0.08183	**	0.00299

Sampling time influenced MOTU recovery and community composition in all treatments analyzed, showing a highest effect size in sediment samples and a lowest effect size in water samples preserved in BAC. An correlation between the number of MOTUs and effect size was found, as the higher amount of MOTUs obtained, the higher was also the effect size of sampling event. Despite showing a lower amount of MOTUs detected, samples obtained in the second event allowed the detection of additional orders and families. During the second sampling event 19.9% of the MOTUs were only detected in sediment samples when contrasted to 2.56% in the first sampling. Sediments can act as eDNA molecules reservoirs, since eDNA can settle and bind to the substrate and when incorporated its persistence can be much longer (Eichmiller et al., 2014; Turner et al., 2014).

Environmental DNA concentration can change seasonally, as well as changes in community composition over time should be expected due to natural (e.g. environmental changes, such as variation in water temperature and flow) or anthropogenic factors (e.g. pollution, introduction of physical barriers) and this variation has already been documented through metabarcoding in estuaries (Stoeckle, Soboleva, & Charlop-Powers, 2017), lakes (Bista et al., 2017) and rivers, even over a small temporal scale (Macher & Leese, 2018). The Jequitinhonha Valley is a dry region that is under the risk of desertification and by the beginning of 2017, when the first sampling event was undertaken, the area was facing the worst drought in the past 80 years. However, the sampling was conducted during the rainy season and the average accumulated rainfall increased from 2.1-50 mm (first sampling time) to 100-250 mm (second sampling event) per month (CPTEC/INPE, 2018). The increase in the precipitation level in this region, with heavy rainfall causing floods in several sites and this seasonal change might have impacted the MOTU recovery during the second sampling, as the increase in water level can dilute the eDNA, change the water temperature and flow, and also cause fluctuations in community composition. Increased water volume after the rainfall contributes to a higher velocity and affects eDNA concentrations in water columns, as eDNA containing particles are transported and dispersed towards downstream river (Shogren et al., 2017). Furthermore, an increase in water flow caused by rainfall might lead to eDNA particle resuspension, which could explain a higher similarity detected by the nMDS between sampling medium in the second sampling event.

Understanding the effect of abiotic and biotic factors on eDNA recovery in tropical lotic environments is crucial to improve the interpretation of results and assure the

effectiveness of eDNA as a biodiversity assessment tool. Here, we showed the first results on effect of sampling medium, time, and preservation methods in lotic environments and our findings suggest that the interaction between preservation method and MOTU recovery might be less significant than the influence of sampling medium and sampling event. Cooling the water samples before the filtering might be a better option in field work conducted in remote areas due to logistical issues and to an increased eDNA recovery when compared to addition of cationic surfactants as sample preservatives.

We also highlight the importance of a better interpretation of eDNA results when comparing sediment and water samples due to distinct temporal intervals covered, and comparing two sets of samples obtained in a short time interval we demonstrate the importance of applying multiple sampling collections when planning a realistic screening of fish biodiversity in lotic environments. The recovery of a high amount of MOTUs allowed the detection of a high degree of fish biodiversity, including changes in community composition, demonstrating the effectiveness of eDNA as a biodiversity assessment tool in neotropical lotic rivers. However, this study was method-focused and detailed ecological analysis of the recovered biodiversity is the natural next step. This will require an improved reference database, as the data obtained here (i.e. potentially hundreds of fish species) suggests that the biodiversity of this catchment is grossly underestimated (Andrade-Neto, 2010).

DATA ACCESSIBILITY

Data will be made public on the DRYAD repository upon acceptance. A list of all samples analysed (including additional information) is provided in Table S3, Supplementary Material.

References

Anderson, M. J. (2017). Permutational Multivariate Analysis of Variance (PERMANOVA). *Wiley Statistic Reference Online*, 1–15. doi:10.1002/9781118445112.stat07841.

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

Andruszkiewicz, E. A., Sassoubre, L. M., and Boehm, A. B. (2017). Persistence of marine fish environmental DNA and the influence of sunlight. *PLoS ONE, 12,* e0185043. doi: 10.1371/journal.pone.0185043

Bakker, J., Wangensteen, O. S., Chapman, D. D., Boussarie, G., Buddo, D., Guttridge, T. L., et al. (2017). Environmental DNA reveals tropical shark diversity in contrasting levels of anthropogenic impact. *Scientific Reports 4(7)*, 16886. doi: 10.1038/s41598-017-17150-2

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

Barnes, M. A., and Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, *17(*1), 1-17. doi: 10.1007/s10592-015-0775-4

Bilibio, C., Hensel, O., and Selbach, J. (2011). Sustainable water management in the tropics and subtropics - and case studies in Brazil. V1. Ed: Fundação Universidade Federal do Pampa, UNIKASSEL, PGCult-UFMA. 358p.

Bista, I., Carvalho, G. R., Walsh, K., Seymour, M., Hajibabaei, M., Lallias, D., et al. (2017). Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. *Nature Communications,* 8:14087. doi: 10.1038/ncomms14087

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knap, M., et al. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology and Evolution*, 29(6), 358-367. doi: 10.1016/j.tree.2014.04.003

Boussarie, G., Bakker., J., Wangensteen, O. S., Mariani, S., Bonnin, L., Juhel, J., et al. (2018). Environmental DNA illuminates the dark diversity of sharks. *Science Advances, 4*, 1-8. doi: 10.1126/sciadv.aap9661

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

Buxton, A. S., Groombridge, J. J., Zakaria, . B., and Griffiths, R. A. (2017). Seasonal variation in environmental DNA in relation to population size and environmental factors. *Scientific Reports, 7*, 1-9. doi: 10.1038/srep46294

Cilleros, K., Valentini, A., Allard, L., Dejean, T., Etienne, R., Grenouillet, G., et al. (2018).

Unlocking biodiversity and conservation studies in high diversity environments using environmental DNA (eDNA): a test with Guianese freshwater fishes. *Molecular Ecology Resources*, 1–20. doi:10.1111/1755-0998.12900

Climate data. 2018. Available at: https://pt.climate-data.org/location/24907/

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

Collins, R. A., Wangensteen, O. S., G'orman, E. J., Mariani, S., Sims, D. S., and Genner, M. J. (2018). Persistence of environmental DNA in marine systems. *Communications Biology,* 1(185), 1-11. doi: 10.1038/s42003-018-0192-6

Crecchio, C., and Stotzky, G. (1998). Binding of DNA on humic acids: effect on transformation of Bacillus subtilis and resistance to DNase. *Soil Biology and Biochemistry,* 30(8-9), 1061-1067. doi: 10.1016/S0038-0717(97)00248-4

Deiner, K., Bik, H. M., Machler, E., Seymour, M., Lacoursiere-Roussel, A., Altermatt, F., et al. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology,* 26, 1-24. doi:10.1111/mec.14350

Deiner, K., Lopez, J., Bourne, S., Holman, L. E., Seymour, M., Grey, E. K., et al. (2018). Optimising the detection of marine taxonomic richness using environmental DNA metabarcoding: the effects of filter material, pore size and extraction method. *Metabarcoding and Metagenomics,* 2, 1-15. doi: 10.3897/mbmg.2.28963

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

Dejean, T., Valentini, A., Duparc, A., Pellier-Cuit, S., Pompanon, F., Taberlet, P., et al. (2011). Persistence of environmental dna in freshwater ecosystems. *PLoS ONE,* 6: e23398. doi: 10.1371/journal.pone.0023398

Dickie, I. A., Boyer, S., Buckley, H. L., Duncan, R. P., Gardner, P. P., Hogg, I. D., et al. (2018). Towards robust and repeatable sampling methods in eDNA-based studies. *Molecular Ecology Resources,* 18(5), 940-952. doi: 10.1111/1755-0998.12907

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics,* 27, 2194–2200. doi: 10.1093/bioinformatics/btr381

Eichmiller, J. J, Best, S. E., and Sorensen, P. W. (2016). Effects of temperature and trophic state on degradation of environmental DNA in lake water. *Environmental Science and Technology,* 50, 1859-1867. doi: 10.1021/acs.est.5b05672

Fonseca, V. G. (2018). Pitfalls in relative abundance estimation using eDNA metabarcoding. *Molecular Ecology Resources,* 18(5), 923-926. doi: https://doi.org/10.1111/1755-0998.12902

Hammer, Ø., Harper, D.A.T., and Ryan, P.D. 2001. PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica,* 4(1). 9pp.

Handley, L. L., Read, D. S., Winfield, I. J., Kimbell, H., Li, J., Hahn, C., Blackman, R., et al. (2018). Temporal and spatial variation in distribution of fish environmental DNA in England's largest lake. *bioRxiv.* doi: 10.1101/376400

Hänfling, B., Handley, L. L., Read, D. S., Hahn, C., Li, J., Nichols, P., et al. (2016). Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. *Molecular Ecology*, 25(13), 3101–3119. doi: 10.1111/mec.13660

Hansen, B. K., Bekkevold, D., Clausen L. W., and Nielsen, E. E. (2018). The sceptical optimist: challenges and perspectives for the application of environmental DNA in marine fisheries. *Fish and Fisheries,* 2018, *19*(5) 751-768. doi: 10.1111/faf.12286

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

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

Jo, T., Murakami, H., Yamamoto, S., Masuda, R., and Minamoto, T. (2019). Effect of water temperature and fish biomass on environmental DNA shedding, degradation, and size distribution. *Ecology and Evolution*, 1–12. doi:10.1002/ece3.4802.

Leray, M. and Knowlton, N. (2017). Random sampling causes the low reproducibility of rare eukaryotic OTUs in Illumina COI metabarcoding. *PeerJ,* 5*,* e3006. doi: 10.7717/peerj.3006.

Ladell, B. A., Walleser, L. R., McCalla, S. G., Erickson R. A., and Amberg, J. J. (2018). Ethanol and sodium acetate as a preservation method to delay degradation of environmental DNA. *Conservation Genetics Resources*, 2-7. doi: 10.1007/s12686-017-0955-2

Lodge, D. M. (2012). Conservation in a cup of water: estimating biodiversity and abundance from environmental DNA. *Molecular Ecology*, 21(11), 2555-2558. doi: 10.1111/j.1365-294X.2012.05600.x.

Machado, A. B. M., Drummond, G. M., and Paglia, A. P. (2008). Livro vermelho da fauna brasileira ameaçada de extinção. Belo Horizonte, Fundação Biodiversitas.

Macher, J., and Leese, F. (2017). Environmental DNA metabarcoding of rivers: Not all DNA is everywhere, and not all the time. *bioRxiv*. doi: 10.1101/164046.

Mahé, F., Rognes, T., Quince, C., de Vargas, C., and Dunthorn, M. (2015). Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ,* 3, e1420. doi: 10.7717/peerj.1420.

Maruyama, A., Nakamura, K., Yamanaka, H., Kondoh, M., and Minamoto, T. (2014). The release rate of environmental DNA from juvenile and adult fish. *PLoS ONE,* 9 (12), e114639. doi:10.1371/ journal.pone.0114639

Matheson, C. D., Gurney, C., Esau, N., and Lehto, R. (2014). Assessing PCR inhibition from humic substances. *The Open Enzyme Inhibition Journal,* 3(1), 38–45. doi: 10.2174/1874940201003010038

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J., Y., Sato, K., et al. (2015). MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. *Royal Society Open Science*, *2*(7). doi: 10.1098/rsos.150088

Pilliod, D. S., Goldberg, C. S., Arkle, R. S., and Waits, L. P. (2014). Factors influencing detection of eDNA from a stream-dwelling amphibian. *Molecular Ecology Resources*, 14(1), 109-116. doi. 10.1111/1755-0998.12159

Pugedo, M. L., de Andrade Neto, F. R., Pessali, T. C., Birindelli, J. L. O., and Carvalho, D. C. (2016). Integrative taxonomy supports new candidate fish species in a poorly studied neotropical region: the Jequitinhonha River Basin. *Genetica, 144*(3), 341-349. doi: 10.1007/s10709-016-9903-4

Pont, D., Rocle, M., Valentini, A., Civade R., Jean, P., Maire, A., et al. (2018). Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. *Scientific reports,* 8, 10361. doi: 10.1038/s41598-018-28424-8

Robson, H. L. A., Noble, T. H., Saunders, R. J., Robson, S. K. A., Burrows, D. W., and Jerry, D. R. (2016). Fine tuning for the tropics: application of eDNA technology for invasive fish detection in tropical freshwater ecosystems. *Molecular Ecology Resources,* 16, 922–932.

Rognes, T., Flouri, T., Nichols, B., Quince, C. and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ,* 4, e2584. doi:10.7717/peerj.2584

Rosa, R. S. and Lima, F. C. T. (2008). Os peixes ameaçados de extinção. In: Machado, A. B. M., Drummond, G. M., Paglia, A. P. (eds). Livro vermelho da fauna brasileira ameaçada de extinção. Ministério do Meio Ambiente/ Fundação Biodiversitas, Brasilía, pp9-285.

Rosa, R. S., and Lima, F. C. T. (2008). Peixes. In: Machado, A. B. M., Drummond, G. M., Paglia, A. P. (eds) Livro Vermelho da Fauna Brasileira Ameaçada de Extinção, 1st edn. Ministério do Meio Ambiente, Brasília, pp 9–275

Sales, N. G., Mariani, S., Salvador, G. N., Pessali, T. C., and Carvalho, D. C. (2018). Hidden diversity hampers conservation efforts in a highly impacted Neotropical river system. *Frontiers in Genetics,* 9*,* 27*1.* doi: 10.3389/fgene.2018.00271

Sassoubre, L. M., Yamahara, K. M., Gardner, L. D., Block, B. A., Boehm, A. B. (2016). Quantification of Environmental DNA (eDNA) Shedding and Decay Rates for Three Marine Fish. *Environmental Science and Technology,* 50, 10456–10464. doi:10.1021/acs.est.6b03114

Shogren, A. J., Tank, J. L., Andruszkiewicz, E., Olds, B., Mahon, A. R., Jerde, C. L., et al. (2017). Controls on eDNA movement in streams: transport, retention, and resuspension. *Scientific Reports, 7*, 1-11. doi:10.1038/s41598-017-05223-1

Siegenthaler, A., Wangensteen, O. S., Soto, A. Z., Benvenuto, C., Corrigan, L., and Mariani, S. (2018). Metabarcoding of shrimp stomach content: harnessing a natural sampler for fish biodiversity monitoring. *Molecular Ecology Resources,* 8, 19(1), 206-220. doi: *10.1111/1755-0998.12956*

Spurgeon, J. J., Pegg, M. A., Parasiewicz, P., and Rogers, J. (2018). Diversity of river fishes influenced by heterogeneity across hydrogeomorphic divisions. *River Research and applications,* 34(7), 797-806. doi: 10.1002/rra.3306

Stoeckle, M. Y., Soboleva, L. and Charlop-Powers, Z. (2017). Aquatic environmental DNA detects seasonal fish abundance and habitat preference in an urban estuary. *PLoS ONE, 12*, e0175186. doi: 10.1371/journal.pone.0175186

Strickler, K.M., Fremier, A.K., and Goldberg C.S. (2015). Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. *Biological Conservation,* 183*,* 85–92. doi:10.1016/j.biocon.2014.11.038

Taberlet, P., Bonin, A., Zinger, L., and Coissac, E. (2018). Environmental DNA for biodiversity research and monitoring. Oxford University Press. 253p.

Takahara, T.; Minamoto, T.; Yamanaka, H.; Doi, H.; and Kawabata, Z. (2012). Estimation of fish biomass using environmental DNA. *PLoS ONE,* 7(4), e35868. doi: 10.1371/journal.pone.0035868

Thomsen, F. P., and Willerslev, E. (2015). Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation,* 183*,* 4-18. doi: 10.1016/j.biocon.2014.11.019

Tsuji, S., Ushio, M., Sakurai, S., Minamoto, T., and Yamanaka, H. (2017). Water temperature-dependent degradation of environmental DNA and its relation to bacterial abundance. *PLoS ONE*, 12, 4. doi: 10.1371/journal.pone.0176608

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

Valdez-Moreno, M., Ivanova, N. V., Elias-Gutierrez, M., Pedersen, S. L., Bessonov, K., and Hebert, P. D. N. (2018). Using eDNA to biomonitor the fish community in a tropical oligotrophic lake. *bioRxiv*, 375089. doi:10.1101/375089.

Valentini, A. Taberlet, P., Miaud, C., Civade, R., Herder, J., Thomsen, P. F., et al. (2016). Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology,* 25(4), 929–942. doi:10.1111/mec.13428.

Yamanaka, H., Minamoto, T., Matsuura, J., Sakurai, S., Tsuji, S., Motozawa, H., et al. (2017). A simple method for preserving environmental DNA in water samples at ambient temperature by addition of cationic surfactant. *Limnology,* 18(2), 233–24. doi:10.1007/s10201-016-0508-5

Supplementary material

Table S1: Sample sites, including GPS coordinates.

ID	CODE	Sample site	GPS Coordinates	
1	MED	Medanha	18° 7'15.06"S	43°30'59.16"W
2	TB	Terra Branca	17°18'48.34''S	43°12'26.61''W
3	JGON	José Gonçalves (upstream the UHE Irapé dam)	16°44'25.89''S	42°34'16.34''W
4	ITAC	Itacambiruçu	16°36'24.00''S	42°49'46.00''W
5	CM	Coronel Murta (downstream the UHE Irapé dam)	16°44'26.85''S	42°34'11.78''W
6	ARA	Araçuaí	16°51'10.47"S	41°51'33.53"W
7	JEQ	Itaobim/Jequitinhonha	16°26'16.74"S	41°1'1.45"W
8	ALM	Almenara/Jacinto	16° 8'26.20"S	40°35'4.64"W
9	SD	Salto da Divisa	15°59'51.07"S	39°53'29.76"W
10	ITAP	Itapebi	15°56'57.69"S	39°31'27.08"W
11	BEL	Belmonte	15°51'0.02"S	38°52'13.66"W

Table S2: PERMANOVA results (R^2-effect sizes and significance level) showing the effect of preservation method, sampling medium, and sampling time on MOTU richness.

	R^2	Effect	Significance (p-value)
Preservation method			
BAC1 vs ICE1	0.02636	*	0.01698
BAC2 vs ICE2	0.0278	*	0.06493
Sampling medium			
SED vs WAT	0.03626	*	0.00099
SED1 vs WAT1	0.07234	*	0.00999
SED2 vs WAT2	0.08183	**	0.00299
SED1 vs BAC1	0.06006	*	0.00099
SED1 vs ICE1	0.0598	*	0.00099
SED2 vs BAC2	0.0841	**	0.00099
SED2 vs ICE2	0.07056	*	0.00099
Sampling time			
SED1 vs SED2	0.07762	*	0.00099
BAC1 vs BAC2	0.04192	*	0.00099
ICE1 vs ICE2	0.06436	*	0.00099

Table S3: List of samples including primers and tags used.

LIBRARY	SAMPLE	TAGS	PRIMER FORWARD	PRIMER REVERSE	CODE
BNA8	B1	TATCATT:TATCATT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B6	AAAGACC:AAAGACC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B9	GGTAGGG:GGTAGGG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B2	AGCCCTC:AGCCCTC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B7	CAAAGCG:CAAAGCG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B10	CCGCTAA:CCGCTAA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B13	GCTCAGA:GCTCAGA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B14	TTAGAAC:TTAGAAC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B15	CGGAAAC:CGGAAAC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B16	ATCCCGG:ATCCCGG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B19	AAAGGTA:AAAGGTA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B22	TTAAACT:TTAAACT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B24	GAGTCTA:GAGTCTA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B26	GGTGACG:GGTGACG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B28	AGCGTGC:AGCGTGC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B30	GTTTGAT:GTTTGAT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B31	TGTGGGT:TGTGGGT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B32	TATCTAC:TATCTAC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B33	TCTGTGC:TCTGTGC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B35	GAGTAGC:GAGTAGC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B37	CCCTGTG:CCCTGTG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B39	AACACCA:AACACCA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B44	TCAAATC:TCAAATC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B47	GTCATTC:GTCATTC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B50	AAACGGC:AAACGGC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B54	AAAGCAT:AAAGCAT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B57	CAGATCT:CAGATCT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B60	TCTAGGA:TCTAGGA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B63	AAATTCA:AAATTCA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	R66	GATAACT:GATAACT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B69	GATAGAC:GATAGAC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B72	ATCCGAC:ATCCGAC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	B75	GAGCTAT:GAGCTAT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE1
BNA8	R1	GGTACCC:GGTACCC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R2	TATGCCC:TATGCCC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R3	AAATCTC:AAATCTC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R4	GTTGAGC:GTTGAGC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R9	TATTGTC:TATTGTC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R13	AGCTAAA:AGCTAAA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R18	AACCTAG:AACCTAG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R19	TAGCGTG:TAGCGTG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R20	CTGCATA:CTGCATA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R21	ACCAATT:ACCAATT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R23	ATCATCG:ATCATCG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R25	ACCCAGC:ACCCAGC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R28	CGGGCGC:CGGGCGC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R30	ACCGCCC:ACCGCCC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R32	ACCTACG:ACCTACG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R34	GCGGGAG:GCGGGAG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1

LIBRARY	SAMPLE	TAGS	PRIMER FORWARD	PRIMER REVERSE	CODE
BNA8	R36	TGTTATG:TGTTATG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R38	AAAGTGG:AAAGTGG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R40	AGCCGGT:AGCCGGT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R42	CGGCTTG:CGGCTTG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R44	AGCGGCG:AGCGGCG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R46	ATGAAGA:ATGAAGA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R48	CCGTATT:CCGTATT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R50	AGCACAT:AGCACAT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R53	ACATTAT:ACATTAT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R55	CACTATA:CACTATA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R57	ACCATAA:ACCATAA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R60	AACAAAC:AACAAAC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R61	GAGGAAA:GAGGAAA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R62	CAGCAAG:CAGCAAG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R65	GGTTCTT:GGTTCTT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R66	TATCGCA:TATCGCA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	R68	TGTTCAC:TGTTCAC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC1
BNA8	NS2_1_1	CGGCAGT:CGGCAGT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_2	AAATGAG:AAATGAG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_3	TTAATAA:TTAATAA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_4	GATACGA:GATACGA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_5	TGACACC:TGACACC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_6	GATCCTC:GATCCTC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_7	CTCCTGA:CTCCTGA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_8	GAGGCCG:GAGGCCG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_9	TATGGAG:TATGGAG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_10	GTCCCTA:GTCCCTA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_11	TATTCGG:TATTCGG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_12	AACAGGG:AACAGGG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_13	AACCACT:AACCACT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_14	TCACAGT:TCACAGT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_15	GTTAGCA:GTTAGCA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_16	CACGTAT:CACGTAT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_17	AAATAGT:AAATAGT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_18	CACTGGT:CACTGGT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_19	ACCCGCA:ACCCGCA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_20	TCCGAGG:TCCGAGG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_21	TTACTCG:TTACTCG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NS2_1_22	AACGAGA:AACGAGA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED1
BNA8	NegContField2	AAACTTT:AAACTTT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	CONTROL
BNA8	NegContField3	CGGAGTT:CGGAGTT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	CONTROL
BNA8	BlankPCR3	CACTCCG:CACTCCG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	CONTROL
BNA8	BlankPCR4	TACCCAA:TACCCAA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	CONTROL
BNA9	G1	TATCATT:TATCATT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G2	AAAGACC:AAAGACC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G3	GGTAGGG:GGTAGGG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G4	AGCCCTC:AGCCCTC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G5	CAAAGCG:CAAAGCG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G6	CCGCTAA:CCGCTAA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G7	GCTCAGA:GCTCAGA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G8	TTAGAAC:TTAGAAC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2

LIBRARY	SAMPLE	TAGS	PRIMER FORWARD	PRIMER REVERSE	CODE
BNA9	G9	CGGAAAC:CGGAAAC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G10	ATCCCGG:ATCCCGG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G12	AAAGGTA:AAAGGTA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G14	TTAAACT:TTAAACT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G16	GAGTCTA:GAGTCTA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G17	GGTGACG:GGTGACG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G18	AGCGTGC:AGCGTGC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G19	GTTTGAT:GTTTGAT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G21	TGTGGGT:TGTGGGT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G22	TATCTAC:TATCTAC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G24	TCTGTGC:TCTGTGC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G25	GAGTAGC:GAGTAGC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G26	CCCTGTG:CCCTGTG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G28	AACACCA:AACACCA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G30	TCAAATC:TCAAATC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G32	GTCATTC:GTCATTC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G34	AAACGGC:AAACGGC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G35	AAAGCAT:AAAGCAT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G38	CAGATCT:CAGATCT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G39	TCTAGGA:TCTAGGA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G40	AAATTCA:AAATTCA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G42	GATAACT:GATAACT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G43	GATAGAC:GATAGAC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G44	ATCCGAC:ATCCGAC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	G45	GAGCTAT:GAGCTAT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	ICE2
BNA9	P1	GGTACCC:GGTACCC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P2	TATGCCC:TATGCCC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P3	AAATCTC:AAATCTC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P5	GTTGAGC:GTTGAGC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P6	TATTGTC:TATTGTC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P7	AGCTAAA:AGCTAAA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P8	AACCTAG:AACCTAG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P9	TAGCGTG:TAGCGTG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P10	CTGCATA:CTGCATA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P11	ACCAATT:ACCAATT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P13	ATCATCG:ATCATCG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P15	ACCCAGC:ACCCAGC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P17	CGGGCGC:CGGGCGC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P18	ACCGCCC:ACCGCCC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P19	ACCTACG:ACCTACG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P20	GCGGGAG:GCGGGAG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P22	TGTTATG:TGTTATG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P24	AAAGTGG:AAAGTGG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P26	AGCCGGT:AGCCGGT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P28	CGGCTTG:CGGCTTG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P30	AGCGGCG:AGCGGCG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P32	ATGAAGA:ATGAAGA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P34	CCGTATT:CCGTATT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P36	AGCACAT:AGCACAT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P39	ACATTAT:ACATTAT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P41	CACTATA:CACTATA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2

LIBRARY	SAMPLE	TAGS	PRIMER FORWARD	PRIMER REVERSE	CODE
BNA9	P44	ACCATAA:ACCATAA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P47	AACAAAC:AACAAAC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P49	GAGGAAA:GAGGAAA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P51	CAGCAAG:CAGCAAG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P53	GGTTCTT:GGTTCTT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P54	TATCGCA:TATCGCA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	P55	TGTTCAC:TGTTCAC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	BAC2
BNA9	NS2_2_1	CGGCAGT:CGGCAGT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_2	AAATGAG:AAATGAG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_3	TTAATAA:TTAATAA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_4	GATACGA:GATACGA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_5	TGACACC:TGACACC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_6	GATCCTC:GATCCTC	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_7	CTCCTGA:CTCCTGA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_8	GAGGCCG:GAGGCCG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_9	TATGGAG:TATGGAG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_10	GTCCCTA:GTCCCTA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_11	TATTCGG:TATTCGG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_12	AACAGGG:AACAGGG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_13	AACCACT:AACCACT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_14	TCACAGT:TCACAGT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_15	GTTAGCA:GTTAGCA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_16	CACGTAT:CACGTAT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_17	AAATAGT:AAATAGT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_18	CACTGGT:CACTGGT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_19	ACCCGCA:ACCCGCA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_20	TCCGAGG:TCCGAGG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_21	TTACTCG:TTACTCG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NS2_2_22	AACGAGA:AACGAGA	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	NegContField4	AAACTTT:AAACTTT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	BlankPCR5	CGGAGTT:CGGAGTT	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2
BNA9	BlankPCR6	CACTCCG:CACTCCG	GCCGGTAAAACTCGTGCCAGC	CATAGTGGGGTATCTAATCCCAGTTTG	SED2

Chapter IV

VARIATION IN SPACE AND TIME OF NEOTROPICAL RIVERINE ICHTHYOFAUNA INFERRED FROM eDNA ANALYSIS

Chapter IV

VARIATION IN SPACE AND TIME OF NEOTROPICAL RIVERINE ICHTHYOFAUNA INFERRED FROM eDNA ANALYSIS

Naiara Guimarães Sales[1], Owen Simon Wangensteen[2], Daniel Cardoso Carvalho[3], Ilaria Coscia[1], Allan D. McDevitt[1], Stefano Mariani[1]

Study design: NGS, S.M. Field Work and Sample collection: NGS. Laboratory experiment: NGS, OWS. Data analyses: NGS, SM. Paper writing: NGS, OSW, DCC, IC, ADM, SM.

Running title: Fish diversity assessment in Brazil via eDNA

4.1 Abstract

The Neotropical region harbours a vast amount biodiversity that remains insufficiently assessed, poorly managed, and threatened by unregulated human activities. Novel, rapid and cost-effective DNA-based approaches may represent a valuable avenue towards an improved understanding of the biological communities that underpin essential ecosystem services, especially in rural areas. Here, we evaluate eDNA metabarcoding as a biodiversity assessment and ecological analysis tool in Brazilian rivers. We obtained sediment and water samples from 11 locations along the Jequitinhonha River catchment (South-eastern Brazil), with each site sampled twice in two independent sampling events. The fish-specific ribosomal 12S mitochondrial marker (˜172bp) was amplified and the sequences obtained allowed the detection of 252 Molecular Operational Taxonomical Units (MOTU), of which at least 34 were assigned to the species level, including endemic (*Wertheimeria maculata)* and introduced (*Astronotus ocellatus, Moenkhausia costae*) species, as well as new records for this basin (*Salminus brasiliensis, Lophiosilurus alexandri*). Short-term spatio-temporal variation of fish assemblages demonstrated that communities can vary even within weeks. Species richness during the first campaign was nearly twice as high as the second sampling series, though peaks of diversity were primarily associated with 4 locations, while a much more homogeneous trend was observed during the second campaign. Although no correlation between β-diversity and longitudinal distance or presence of dams (barriers) was detected, low species richness at sites located near the dams may still be the result of anthropogenic impacts.. Environmental DNA can contribute to fish biodiversity assessment in Brazil, by detecting introduced species and provides data from localities often neglected by traditional sampling surveys such as those sampled here.

Keywords: eDNA, biodiversity assessment, fish, freshwater, Neotropical

4.2 Introduction

Despite covering less than 1% of the Earth's surface, freshwater habitats harbour over 40% of the global fish diversity (Nelson, 2006; Eschmeyer, 2005; Dudgeon et al., 2006). Fish from rivers, lakes, and wetlands are important for the provision of essential protein subsistence for human populations worldwide (FAO, 2012), and are increasingly affected by anthropogenic impacts (e.g. habitat modification, fragmentation, climate change- Vörösmarty et al., 2010). In order to provide ecosystem services, a "good ecological status" of these habitats is crucial. Due to the trade-off between socioeconomic development and aquatic habitats preservation, several regulations and restoration programmes have been implemented to protect water resources worldwide and an accurate assessment is required to monitor the impacts and guarantee habitat recovery (Pawlowski et al., 2018; Friberg et al., 2016; Palmer, 2010; Vorosmarty et al., 2010).

Freshwater populations are declining at alarming rates (83% decline since 1970 WWF, 2018) and their conservation and management are a priority for global biodiversity. Nevertheless, despite broad agreement on the requirements to understand and monitor biodiversity and ecological networks in freshwater habitats (Socolar et al., 2015), our comprehension of this realm is lagging behind, compared to marine or terrestrial environments (Jucker et al., 2018).

The Neotropical region comprises one of the greatest freshwater fish diversities in the world (approximately 30% of all described fish species), which is currently facing unprecedented levels of anthropogenic pressure. Conservation and management actions in

freshwater habitats are challenging in this region due to infrastructure problems and sampling constraints, shortage of taxonomic expertise, and the limited amount of agreed descriptors to fully characterise this megadiverse ichthyofauna, especially when dealing with early life stages (Reis et al., 2016). Fish biodiversity assessment in Neotropical countries (e.g. Brazil) relies on the use of traditional methods and these, when not conducted extensively and meticulously, might be selective and fail to detect species (unable to detect organisms in early life stage, cryptic species, rare and/or elusive species; Becker et al., 2015; Sales et al., 2018). Furthermore, as a result of the effort required to apply these methods in wide geographical regions, the sampling might be punctual and cover only a small portion of the studied area. Underestimation of fish biodiversity resulting from low sampling efficiency may provide biased metrics and hamper management and conservation plans (Trimble & van Aarde, 2012). One of the most effective approaches to circumvent the limitations of traditional surveys in mega-diverse systems is the use of DNA methods (e.g. DNA barcoding and metabarcoding; Gomes et al., 2015; Sales et al., 2018, Shimabukuru-Dias et al., 2016, Silva-Santos et al., 2018). However, rapid technological advances in this field still make it difficult to agree on a set of standardized methodologies to be tailored and implemented in a range of monitoring schemes (Hering et al., 2018).

Molecular approaches offer a universal key to identify, assess and quantify biodiversity, especially in biodiversity-rich and understudied ecosystems and regions (Schwartz et al., 2006). Recent advancements have now opened new opportunities for studying biodiversity by sequencing trace DNA present in the water – so-called "environmental DNA" (eDNA) – to identify species presence and, to some extent relative

abundance (Hajibabaei et al. 2011; Tillotson et al., 2018). DNA can be retrieved from water and sediment samples to reconstruct communities at various taxonomic levels, depending on the primers used (Deiner et al., 2017;Rees et al., 2014; Thomsen & Willerslev, 2015).

The vast majority of environmental DNA studies have, focused on temperate regions, in established and fairly well-accessible environments (Handley et al., 2018; Holman et al., 2018; Bracken et al., 2018; Parsons et al., 2018). Recently, Cilleros et al. (2018) demonstrated the efficiency of eDNA metabarcoding in providing spatially extensive data on freshwater fish biodiversity (in French Guyana) and a better discrimination of assemblage compositions when compared to traditional sampling. We recently showed (Sales et al., 2018) the influence of sampling medium, as well as sampling preservation and time, on the reconstruction of ichthyofaunal assemblages in a Brazilian catchment, inferred through eDNA. Here we delve deeper into the biodiversity of the Jequitinhonha river system and attempt to use the wealth of DNA detection data from both water and sediment, in order to assess fish biodiversity, spatially and temporally, and explore community structure along the course of the river.

We hypothesise that i) eDNA metabarcoding can be used as a biodiversity assessment tool in neotropical freshwater ecosystems allowing the detection of multiple fish species, ii) sampling medium (sediment and water samples) provide different community composition, iii) community composition can vary even within short time frames; iv) biodiversity estimates (alpha and beta-diversities) can be obtained in the absence of taxonomic assignments, vi) spatio temporal fluctuation of fish assemblages can be associated to anthropogenic impacts and natural seasonal changes. To address these questions we applied the following methodological framework: i) constructed a recent species list for the Jequitinhonha River

Basin compiling data already published reporting fish species occurrence in this river, ii) built a custom reference database for Jequitinhonha freshwater fish including new sequences for Neotropical species, iii) obtained eDNA metabarcoding data from 176 sediment and water samples collected during two sampling campaigns from 11 sample sites distributed along the entire river basin, iv) compared metabarcoding data with known ichthyofauna in this catchment, v) obtained ecological measures of diversity patterns (alpha, beta diversity), vi) evaluated the ecological communities dynamics by comparing two sets of samples obtained at the same localities at a three-week interval.

4.3 Material and Methods

Study Area

The study area was the Jequitinhonha River basin (Figure 1), Southeast Brazil (17°, 43° W), inserted in two biodiversity hotspots (Cerrado and Atlantic Forest) and characterised by tropical climate and environmental heterogeneity. The main river flows over 1,082 km, from its source in Serro (MG) at an elevation of 1200 m, to drain its waters in the Atlantic Ocean at the locality of Belmonte (BA). According to its hydrology, this catchment is subdivided in three main regions: I) Headwaters, influenced by a tropical continental climate and a rainfall index of 1600 mm. The topography is high and rugged, influencing the climate with lower temperatures and creating rainfall events of higher intensity but shorter duration. Thus, there is a decrease in the rainfall towards the river mouth. II) The following region represents the confluence between the Jequitinhonha river and the Araçuaí tributary and is characterised by a tropical marine climate with a rainfall index of 1.000 mm-1600 mm, year. III) The final stretch comprises the Araçuaí headwaters until the river mouth (Belmonte) and is characterised by a climate similar to region II. The main river stem is interrupted by two large dams built for hydroelectric power generation: the Irapé, the tallest dam in Brazil, built in 2006 (located at region I), and the Itapebi, established in 2002 (region III).

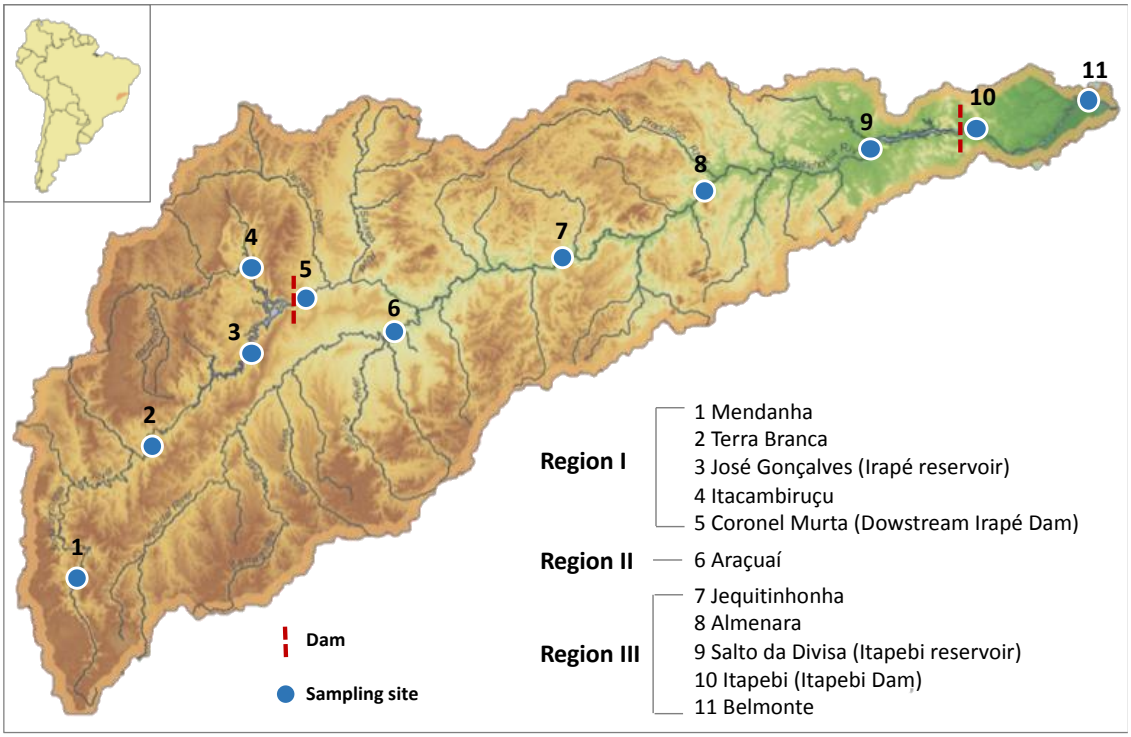


FIGURE 1 | The Jequitinhonha river basin, including sampling sites used in the study, dams and respective hydrological regions

The Jequitinhonha River Basin belongs to the east Atlantic basin complex, characterised by a high number of endemism (67% of its fish species are endemic) (e.g. *Wertheimeria* sp., *Delturus* sp.; Reis et al., 2016, Vono & Birindelli, 2007). The Jequitinhonha is known to harbour a substantial number of endangered (*Steindachnerion amblyurum, Rhamdia jequitinhonha*, *Nematocharax venustus*) and endemic species (*Rhamdia jequitinhonha*) (Rosa & Lima, 2008). Until 2010, the known ichthyofauna of this catchment included 63 described fish species (including 10 introduced species; Andrade-Neto, 2010). This river has long been seen as a low biodiversity ecosystem when compared to neighbouring basins, and its reduced species richness had been linked to historical geological and geographical features (Andrade-Neto, 2010). Yet, the geological history of the Jequitinhonha

is very similar to that of adjacent basins (e.g. Doce river, Mucuri river), which led to the consideration of more recent factors to explain low biodiversity in the catchment, namely: the lack of adequate surveys, and anthropogenic activities (e.g. mining, exploitation fisheries, and deforestation). The Jequitinhonha is particularly affected by the impact of dams on the main river course and tributaries, and recent studies highlighted the occurrence of introduced, undescribed and cryptic species (Pugedo et al., 2016, Sales et al., 2017), leading to the description of several new species (Jerep et al., 2016; Dutra et al., 2016; Nielsen, Pessali & Dutra, 2017; Zawadzki et al., 2016, Pereira et al., 2017). Thus, the lack of adequate sampling might still account for a great number of native species yet to be described for this catchment.

A compiled species list was built by retrieving all papers available at Google Scholar, published in international journals (using the terms “fish”, “Jequitinhonha”), published in local Brazilian journals (applying the terms “peixe”, “Jequitinhonha”, “ictiofauna”) and also, we included non-published data available only in environmental reports. Please see Table S2 in Supporting information for references and additional data.

Local reference database

To obtain a better taxonomic assignment, we retrieved all 12S rRNA mitochondrial gene fish sequences available from GenBank and improved the Brazilian fish 12S sequence reference data from species hitherto missing from the repository (Table S7, Supplementary Material). Tissue samples of 108 specimens belonging to 55 neotropical fish species were obtained from the Laboratório de Genética da Conservação tissue collection (LGC), at Pontifícia Universidade Católica de Minas Gerais (PUC Minas). DNA was extracted from fin

clips, using DNeasy Animal tissue DNA extraction kit (Qiagen). Fragments of the mitochondrial 12S gene were amplified using the MiFish primers (Miya et al., 2015) and the polymerase chain reaction (PCR) consisted in: 1.0µl of buffer $MgCl_2$, 0.3µl of dNTP (total of 10mM), 0.25µl of each MiFish primer (10µM) (Miya et al., 2015), 0.2µl of BIOTAQ DNA polymerase (5U/µl) (Bioline), 7.0µl of ultrapure water, and 1.0µl of DNA template (10 ng/µl). PCR conditions consisted of an initial step of 10 min at 95 °C followed by 35 cycles of 30s at 95 °C, 45s at 60°C, and 30s at 72°C and one final step of 5 min at 72 °C. PCR products were visualised on 1% agarose gels and successfully amplified samples were sequenced by Macrogen Laboratories (www.macrogen.com).

eDNA sampling and processing

Two sampling campaigns were conducted covering a three-week interval (First sampling period: 22/01 to 01/02/2017; Second sampling: 19/02 to 01/03/2017) and 11 sample sites including the main river and two tributaries (nine sites located in the main river course, one site in the Itacambiruçu river and one in Araçuaí river) (Figure 1). Six samples of one liter each of water and two sediment samples were collected in each sample site for each campaign, one set of samples (N=3) was stored at low temperatures (using a cooling box with ice), while in the other samples (N=3) the cationic surfactant benzalkonium chloride (BAC) was added at a final concentration of 0.01% (Yamanaka et al. 2017). In total: 132 water samples and 44 sediment samples were analysed. Water was filtered filtered approximately 8 hours after collection, using Microfil V, 100mL, mixed cellulose esters (MCE) filters (diameter: 47 mm, pore size: 0.45 mm, Merck Millipore - Bakker et al, 2018) using an

automatic vacuum pump and stored at −20°C in microcentrifuge tubes containing silica beads (Majaneva et al., 2018) and sediment samples were stored in 50mL centrifuge tubes and preserved in 100% ethanol.

DNA extraction from the filters was conducted using the DNeasy PowerWater Kit (Qiagen) and DNA from the sediments was extracted using DNeasy PowerMax Soil Kit (Qiagen), following the manufacturer's protocol. Purified extracts were checked for DNA concentration in a Qubit fluorometer (Invitrogen).

Field and laboratory works were conducted following a contamination control procedure, including the use of disposable sterile collection bottles, disposable gloves, and all equipment and surfaces were treated with 50% bleach solution for 10 minutes, followed by rinsing in distilled water after each use. Filtration blanks were run between every sample site, immediately before the next filtration to test for potential contamination during the filtration stage.

Amplification, library preparation and sequencing

Amplicons of 169-172bp from a variable region of the mitochondrial 12S rRNA gene were obtained with the MiFish primers (MiFish-U-F, 5'- GCCGGTAAAACTCGTGCCAGC-3'; MiFish-U-R, 5'- ACATTATCATAGTGGGGTATCTAATCCCAGTTTG -3', Miya et al., 2015).

Samples were sequenced in a single multiplexed Illumina MiSeq run, along with 54 additional samples belonging to a non-related project (not included in this study). For the present study, two libraries were sequenced, containing a total of 183 samples including

collection blanks (N=3) and laboratory negative controls (N=4), using two sets of 96 primers with seven-base sample-specific oligo-tags and a variable number (2-4) of leading Ns (fully degenerate positions) to increase variability in amplicon sequences. PCR amplification was conducted using a single-step protocol and to minimize bias in individual reactions, PCRs were replicated three times for each sample and the products subsequently pooled into single samples. The PCR reaction consisted of a total volume of 20 µL including 10 µl AmpliTaq Gold™ 360 Master Mix (Applied Biosystems); 0.16 µl of BSA; 1 µl of each of the two primers (5 µM); 5.84 µl of ultra-pure water and 2 µl of eDNA template. The PCR profile included an initial denaturing step of 95°C for 10 min, 40 cycles of 95°C for 30s, 60°C for 45s, and 72°C for 30s and a final extension step of 72°C for 5 min. Amplifications were checked through electrophoresis in a 1.5% agarose gel stained with GelRed (Cambridge Bioscience). PCR products were pooled in two different sets and purified using MinElute columns (Qiagen), and Illumina libraries were built from each set, using a NextFlex PCR-free library preparation kit (Bioo Scientific). Size selection was performed using 1.1x Agencourt AMPure XP (Beckman Coulter), libraries were then quantified by qPCR using a NEBNext qPCR quantification kit (New England Biolabs) and pooled in equimolar concentrations along with 1% PhiX (v3, Illumina). The libraries were run at a final molarity of 10pM on an Illumina MiSeq platform in a single MiSeq flow cell using the 2x 150bp v2 chemistry.

Bioinformatic analyses

The metabarcoding pipeline used for data analysis was based on the OBITools software suite (Boyer et al., 2016) following the protocol described in Sales et al. (2018). However, due to the high fish biodiversity expected in Neotropical realms, to optimize the clustering step we applied a clustering method based on variable cut-off (SWARM, Mahé et al., 2014) and evaluate the influence of three *d* (clustering threshold) values (*d*=1, *d*=2, and *d*=3) by comparing the MOTUs (Molecular Operational Taxonomic Unit) and species (identity >0.97) richness in the three datasets obtained.

The taxonomic assignment was performed using ecotag with a custom reference database built retrieving all 12S sequences available from GenBank and the local reference database built in this study. MOTUs of other origin than teleost were removed, each MOTU was assigned to species based on 97% sequence similarity to references (identity 0.97). The *cut off* was established based on published eDNA metabarcoding studies which applied this threshold (>97%) for detecting a great variety of fish species (Li et al., 2018; Nakagawa et al., 2018); MOTUs well represented but showing <97% similarity to references were presented and discussed as putatively belonging to species still absent in the database.

The risk of false positives due to contamination or tag jumping still challenges the application of eDNA metabarcoding analyses (Schnell et al, 2015), to take this issue into account we adopted a stringent approach to guarantee the removal of false positives and MOTUs putatively originated by sequencing error or contamination. Thus, for each MOTU the total number of reads detected in the negative controls (corresponding to this MOTU) were subtracted from all samples, then MOTUs containing less than 5 reads in total were not

included in the final dataset.

Statistical analyses

The use of eDNA metabarcoding data for inferring abundance estimates remains a conundrum since some studies reported that the number of reads might not correspond to the amount of species or their biomass, due to multiple factors (e.g. eDNA degradation, primer bias, distinct DNA shedding rates across species) (Jo et al. 2019; Shaw et al, 2016) while some reports demonstrated similar patterns in molecular and morphological abundances recovered (Evans et al., 2016; Hänfling et al., 2016). Thus, for the diversity analyses (species richness and β-diversity) we applied a recommended conservative approach and treated our results as incidence-based (Li et al., 2018). Molecular Operational Taxonomic Units (MOTUs) are often used as substitute for species, however, the correlation between these two is not straightforward. Biotic indices may be obtained based on both MOTUs or species richness, however each of these approaches have their own drawbacks. MOTUs richness is highly influenced by the occurrence of cryptic species and by the thresholds applied during the bioinformatic analyses (Pawlowski et al., 2018), still it may cause richness overestimation (e.g. inflation of different MOTUs belonging to the same species due to natural intraspecific variability, PCR amplification or sequencing errors) whilst richness based on species may be underestimate due the lack of a complete reference database or due to a low taxonomic resolution of the target fragment analysed.

To verify whether the biodiversity patterns varied significantly due to the species

assignment process, two datasets were analysed individually applying the same methods. The filtered dataset included only MOTUs that could be identified up to species level, whereas the non-filtered dataset included all MOTUs retrieved after the initial filtering steps. The dataset including only species information is considered as a sub dataset of the total MOTU diversity recovered, and thus is expected to provide a more conservative overview (Li et al., 2018).

Statistical analyses were performed in R v3.5.1 (https://www.R-project.org/). Sample replicates were pooled for each site (N=6 samples per site) prior the following statistical analyses. Alpha-diversity (species richness) was estimated as the total number of MOTUs (unfiltered dataset), or number of MOTUs assigned to species level (filtered dataset), at each sample site. β-diversity was obtained applying the Jaccard distance using the vegan package version 2.5-2 using the "vegdist" command (Oksanen et al. 2013). To visualize the relationships amongst sampling sites we obtained PCoA plots using the β-diversity matrix ("cmdscale" command) and the correlation between β-diversity and longitudinal distance and the β-diversity and presence of physical barriers (dams) was tested using the Mantel test ("mantel.rtest" command). The distance matrix was reconstructed using the distance between sample sites estimated using road route as the road follows the river course and thus, this distance would provide a better estimate when compared to linear river distances. The matrix used for testing the influence of physical barriers was constructed attributing distance values between sites according to the existence of barriers (e.g. 0 – no physical barrier between sites, 1- one barrier between sites and 2 – two barriers).

Due to a still incomplete reference database, most of the MOTUs recovered were not identified up to species level and thus a great portion of information regarding the

biodiversity is lost after taxonomic assignment of MOTUs. To verify the total diversity recovered and visualize the community data using a hierarchical structure of taxonomic classifications we used the R package Metacoder (Foster et al., 2017). This package, designed for metabarcoding data, provides "heat tree" plots using statistics associated with taxa (e.g. read abundances) and includes, a pairwise comparison between samples or groups analysed. Venn diagrams were obtained by comparing the orders and families included in the compiled species list, and orders and families detected in each of the eDNA datasets (filtered and non-filtered), using BioVenn (Hulsen, Vlieg, & Alkema, 2008).

4.4 Results

The compiled list of species, including different reports for the Jequitinhonha river basin, resulted in 111 species records (90 already described and 21 identified up to the genus level and/or not described yet) (Table S1 and Table S2, Supplementary Material).

We obtained 16.1 million raw reads (LIB1-6,399,823; LIB2-9.704.699) in one Illumina MiSeq run. To verify the influence of different thresholds in the MOTU diversity recovered we compared three thresholds (*d*=1, *d*=2, *d*=3) for the SWARM clustering. The dataset obtained applying the *d*=1 threshold was used for subsequent analysis due to the loss of species richness when using higher thresholds (e.g. *d*=3 recovered less than 15 species) (Table S3, Supplementary Material). After quality control, clustering and all initial filtering steps, 2056 and 967 MOTUs were kept for library 1 and library 2, with 154 and 59 MOTUs being assigned to species with >0.97 min-identity, respectively. A great difference on number of MOTUs retained for each dataset was obtained and for several species more than one MOTU was also recovered (Figure 2 and Table S4, Supplementary Material).

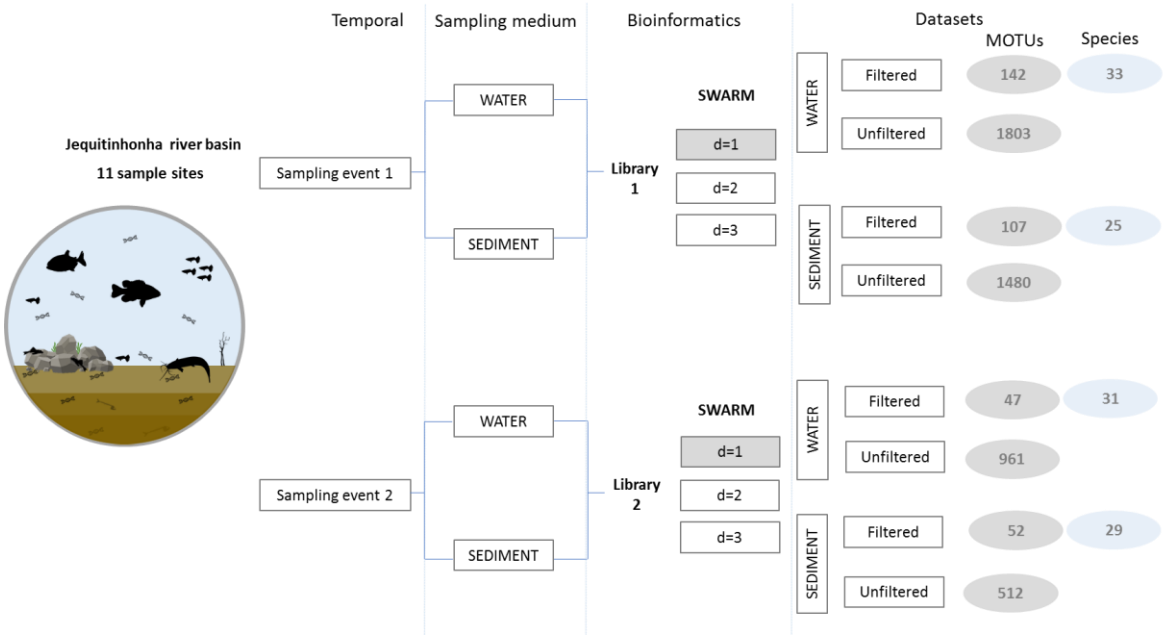


FIGURE 2 | Workflow illustrating the methods used in this paper and respective number of MOTUs retrieved in each dataset analysed, and the final number of species assigned with >0.97 identity.

Taxonomic assignment

Based on the combined data (including all filtered datasets - species >0.97 identity) taxonomical diversity included six orders, 20 families, 28 genera and at least 34 fish "species" (Figure 2, Table S1, Supplementary material). As expected, Characiformes (n=12) and Siluriformes (n=12) were the two orders represented by the largest number of species identified and all the remaining orders were comprised by less than 5 species. Due to the conservative criteria applied to analyse the data, the number of species detected is surely underestimated and many congeneric species might have been clustered together as one single species due to the low taxonomic resolution of the fragment analysed.

The species name herein used might not correspond exactly to the species occurring in the Jequitinhonha river basin (based on the compiled species list) as when the correct

species is not present in the reference database the taxonomic assignment is based on the nearest neighbour species (e.g. *Hypostomus gymnorhynchus* may refer to a closely related species of the same genus). Thus, these species are referenced only up to the genus level, whereas species not previously reported for this basin are marked with an asterisk. Both cases are further discussed below. A comparison between species identified by eDNA and closely related species reported for the Jequitinhonha river basin suggests that several species might have been clustered together (e.g. *Leporinus, Prochilodus, Trichomycterus*) reducing the number of species detected. For instance, *Prochilodus* spp. showed low genetic divergence for the 12S fragment (maximum genetic divergence of 1.8%) and could not thus be reliably subdivided into the likely multiple species present in the catchment (Sales et al, 2018) (Figure S1, Supplementary material).

Comparing the data obtained for both sampling times (Figure 3, Table S5 Supplementary Material), four species were detected only during the first sampling (*Australoheros facetus*, *Cyprinus carpio*, Hypostomus* sp.**, Trichomycterus* sp.), whilst *Coptodon zilli* and Hoplias intermedius* were detected only in the second sampling. Sediment samples failed to detect five species (*Australoheros facetus*, *Cyprinus carpio**, *Hypostomus gymnorhyncus**, *Poecilia reticulata*, *Trichomycterus* sp.), whilst water samples detected all species present in the sediments. Analyses of water and sediment samples demonstrated the occurrence of widely distributed and less abundant species. Several taxa (e.g. *Leporinus* sp., *Prochilodus* sp., *Rhamdia quelen*) were detected in both water and sediment samples in most of sampling sites, in at least one sampling campaign, and therefore seem to have a broad geographic distribution in the Jequitinhonha river basin. Due to marker resolution, these likely

correspond to more than three species as species might be clustered together due to a low taxonomic resolution (please see Figure S1, Supplementary Material) Yet, despite recovering less species than water samples, the sediments showed a higher proportion of eDNA for Prochilodontidae during the first campaign and for Anostomidae and Serrasalmidae in the second campaign (Figure 4).

Some species, including native and non-indigenous species, were restricted to a few locations (e.g. non-indigenous: the oscar - *Astronotus ocellatus*, chameleon cichlid- *Australoheros facetus*, tilapias *Coptodon* sp.* and *Oreochromis* sp., native: roncador - *Wertheimeria maculata*) or were detected in only one campaign (e.g. *Australoheros facetus*, *Coptodon* sp., carp - *Cyprinus carpio**, wolf fish - *Hoplias intermedius*, pleco - *Hypostomus gymnorhyncus**, pencil catfish - *Trichomycterus* sp.). Furthermore, a notable result obtained by eDNA included the detection of species rarely reported in traditional sampling studies (e.g. *Crenicichla* sp.) in all sites analysed and, suggested the occurrence of putative new records for this basin including possibly invasive species such as the dourado - *Salminus brasiliensis** and pacamã - *Lophiosilurus alexandri**.

Despite detecting over 30 species the data here surely represent an underestimation. The filtered dataset provides a more reliable data by assigning species with a minimum of 0.97 identity, however, this conservative approach might hamper the detection of several taxa. Fish diversity depicted by the heat trees based on the unfiltered data shows that a hidden diversity might be present, especially for the Order Characiformes, as many families appears to comprise several MOTUs (e.g. Anostomidae, Prochilodontidae Figure 4). In all comparisons between the datasets analysed, the unfiltered dataset surpassed the filtered

one by detecting orders and families known to occur in this catchment and not identified up to the species level (Figure 5).

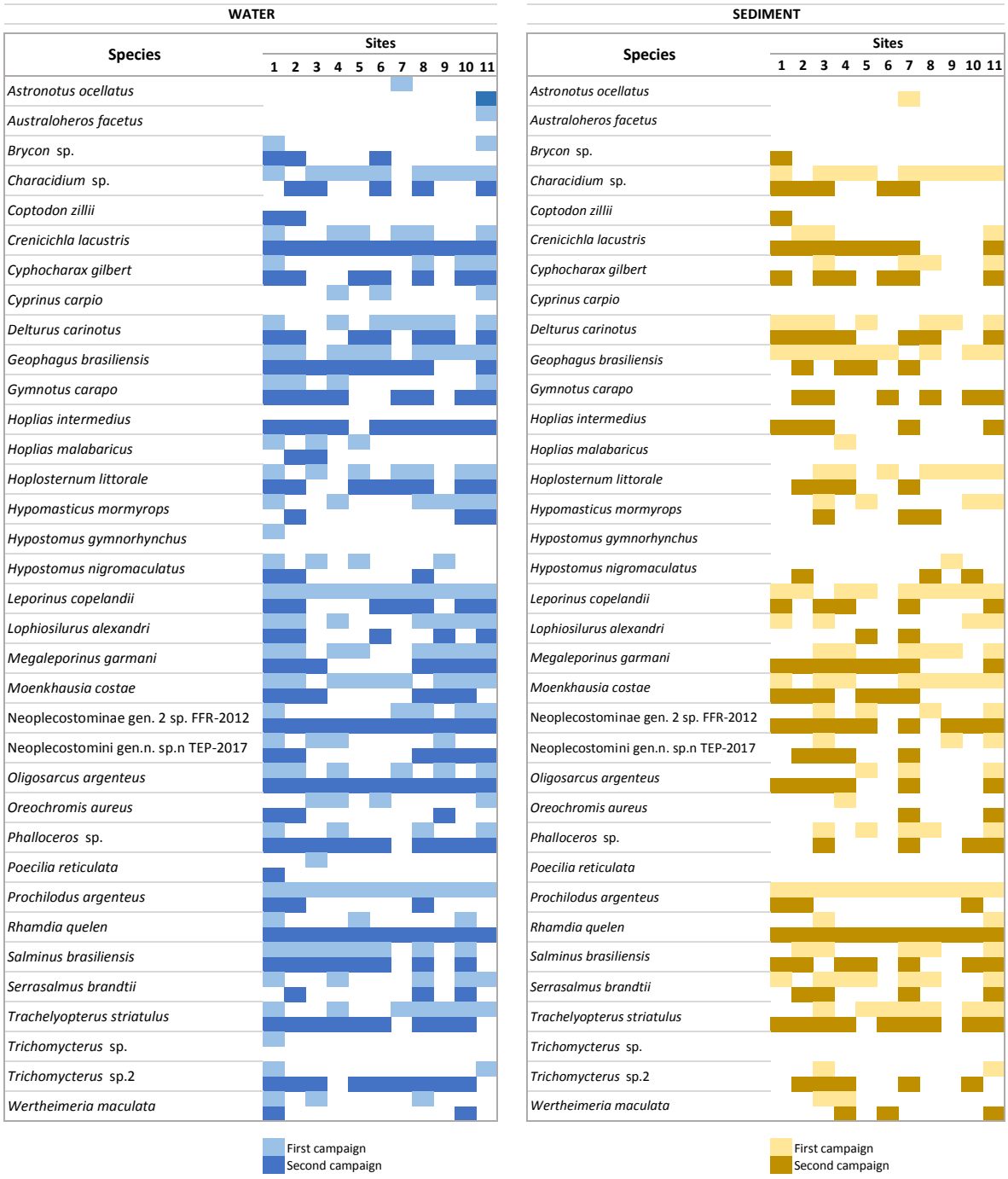


FIGURE 3 | Species distribution in the Jequitinhonha river basin, according to sampling media and campaign.

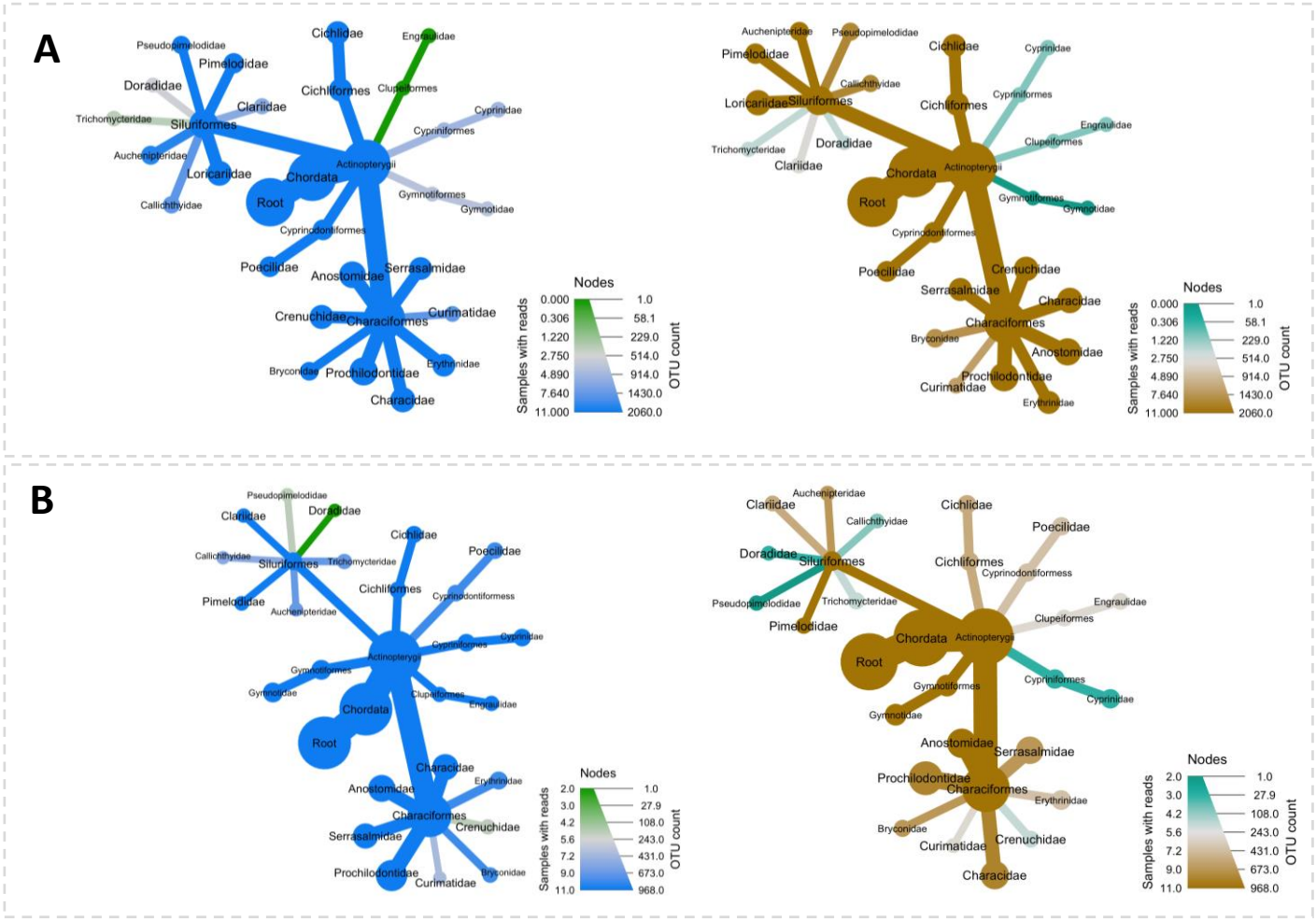


FIGURE 4 | Heat trees displaying the fish diversity recovered for Jequitinhonha river Basin using eDNA metabarcoding, during the first (A) and second (B) campaigns. Blue = Water samples; Brown = Sediment samples.

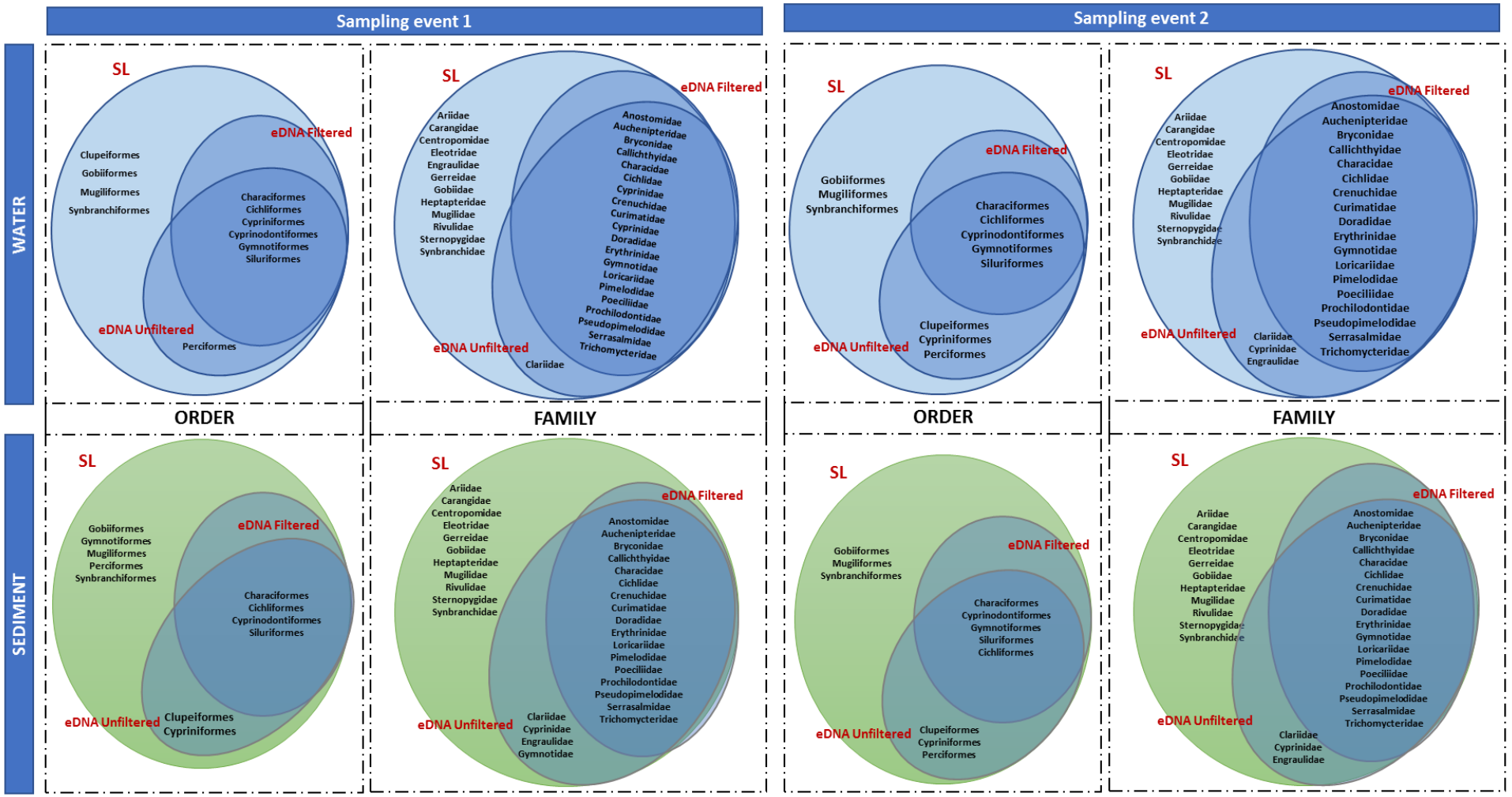


FIGURE 5 | Venn diagram of fish orders and families comparing the data included in the species list based on traditional sampling (SL) to eDNA detected in distinct sampling media (water *vs* sediment); sampling campaign; and datasets analysed (unfiltered *vs* filtered).

Species richness and Beta diversity

During the first campaign, according to the data obtained from water samples (Figure 6A), the most upstream (site 1) and downstream (site 11) sampling sites harboured the highest MOTU richness, followed by locations 4 and 8. The lowest number of MOTUs was recovered for location 7. Beta diversity was similar between sites 4 and 11, and 1 and 8, whereas site 7 showed the most distinct fish assemblage when compared to all locations. Environmental DNA recovered from water samples after a three-week interval demonstrated fluctuations in species richness across time in this catchment (Figure 6B), with a relative increase of stability in the species richness amongst all sample sites. Still, the most upstream and downstream locations (1, 2, 10, 11), alongside location 8, harboured the highest number of species.

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

Over time, the pattern of harbouring the highest species richness appeared relatively constant in sites 1 and 11 for both sampling media, except in the first campaign where few

species were detected in location 1 for sediment. Yet, the most downstream location kept an almost stable species richness in both sampling media for both sampling campaigns.

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

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)

For both sampling media, despite the variation in taxa richness showed by both datasets, the pattern of alpha diversity variation amongst sample sites obtained for filtered (species) and unfiltered (MOTUs) datasets were still quite congruent (Figure 7). However, for sediment samples collected in the first campaign, sites 3 and 11 had a greater MOTU diversity when compared to all nine remaining locations (Figure 7C). Despite also being the most species richness sites, the great amount of MOTUs obtained and not assigned with more than 0.97 identity indicates that a great diversity remains hidden in this sampling medium. Also, as demonstrated by the PCoA (Figure 7C) these sites had a more distinct fish assemblage when compared to the others.

Furthermore, a higher resolution was obtained for the unfiltered dataset as a greater similarity was obtained for the clusters in the PCoA based in the beta diversity. The only great variation was detected between sediment samples from the first campaign, where a relatively lower number of MOTUs was recovered for five sample sites (4, 5, 7, 8 and 9) in comparison to the other locations, when compared to the filtered dataset.

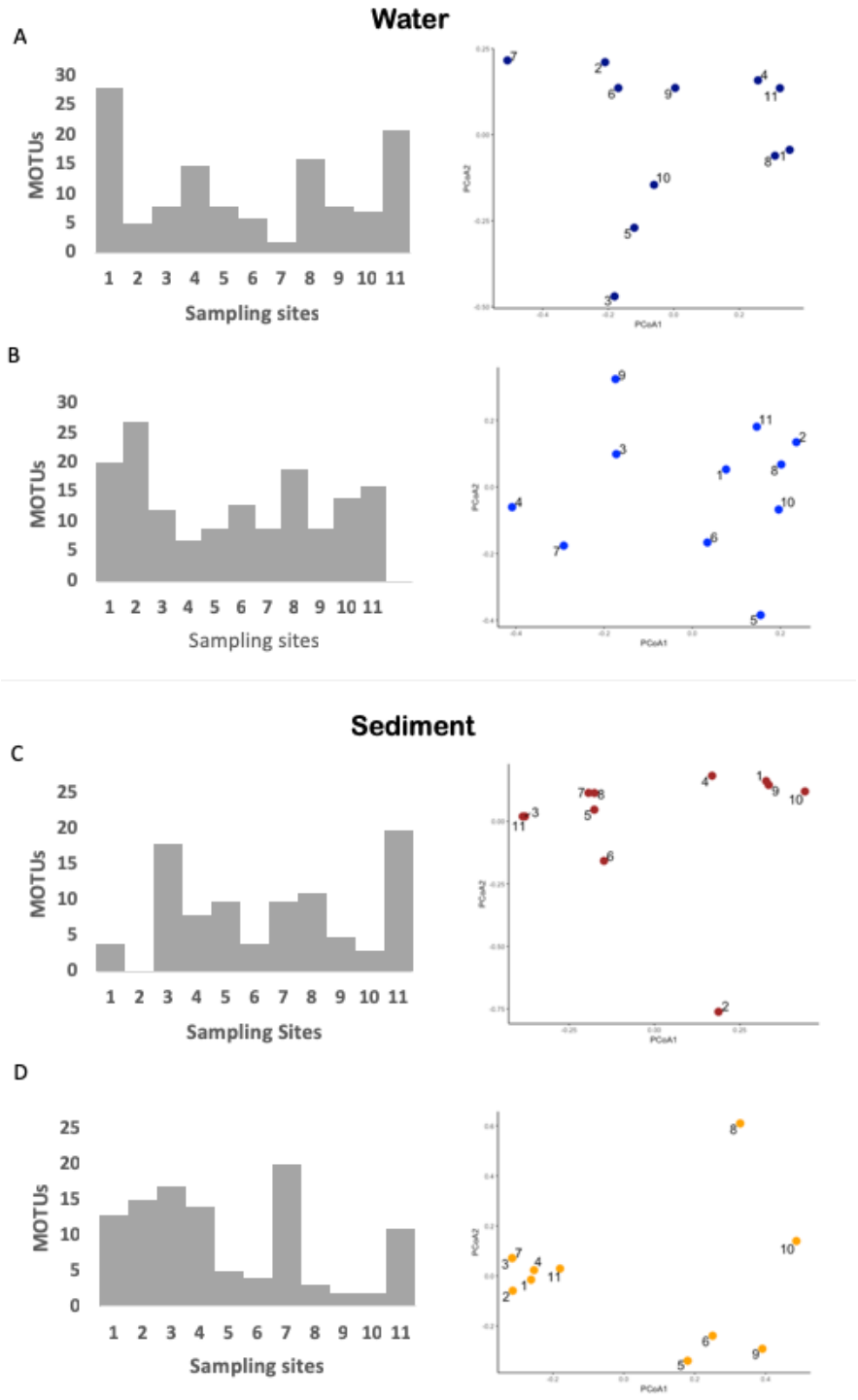


FIGURE 6 | Filtered dataset, showing the species richness distribution along the Jequitinhonha river basin and Principal Coordinates Analysis (PCoA) of β-diversity of sampling locations (Jaccard distance). A) Water samples obtained in the first campaign; B) Water samples obtained in the second campaign;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

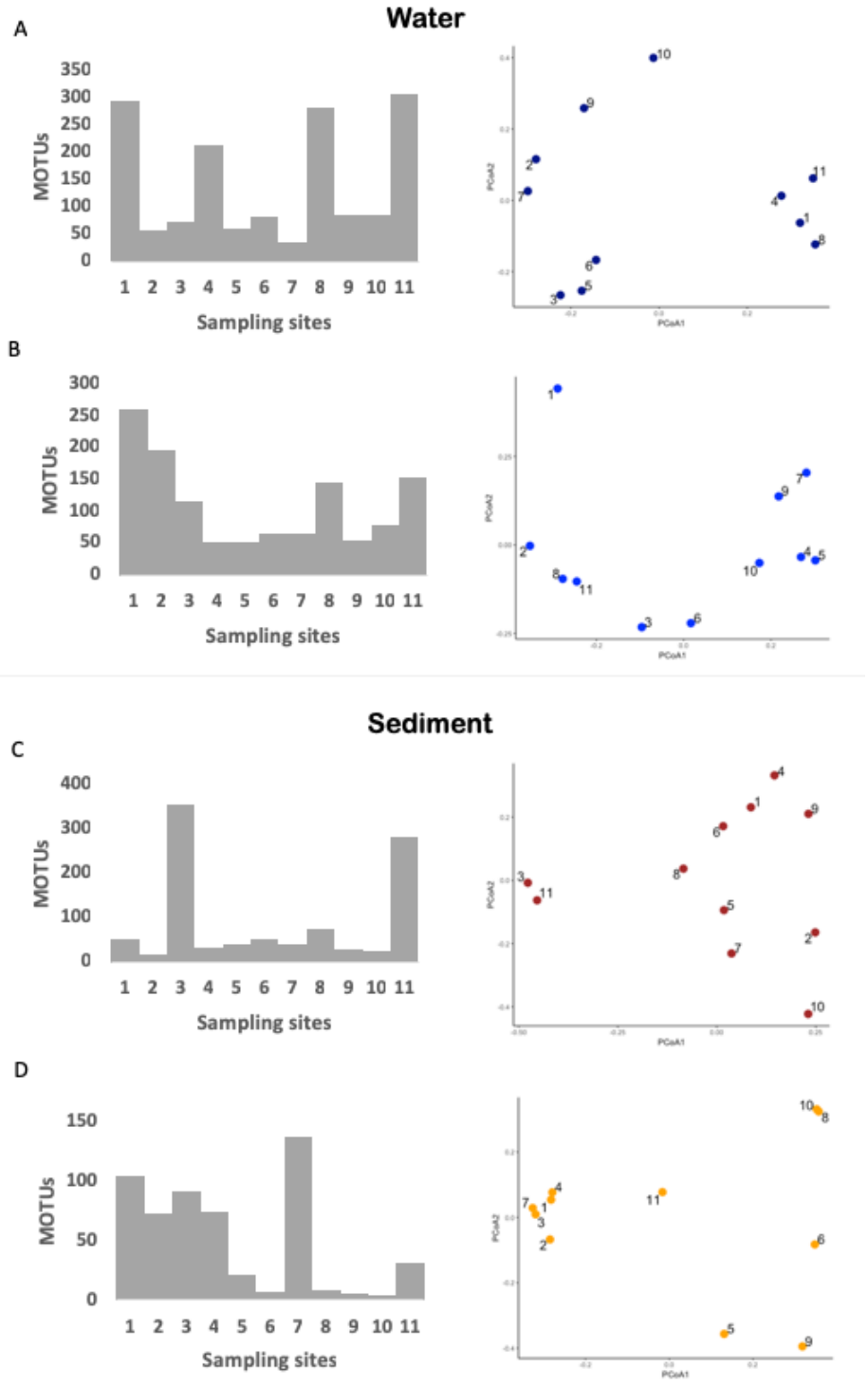


FIGURE 7 | Unfiltered dataset, showing the species richness distribution along the Jequitinhonha river basin and Principal Coordinates Analysis (PCoA) of β-diversity of sampling locations (Jaccard distance). A) Water samples obtained in the first campaign; B) Water samples obtained in the second campaign;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

4.5 Discussion

The identification of species distribution and the understanding of processes shaping spatial variation and community composition are crucial for management and conservation purposes, and rapid biodiversity survey methods are required to measure these fluctuations and support management schemes (Kelly et al., 2014). Here, we provided a list of fish species identified by eDNA and documented the variation in both species richness and beta diversity of different sampling media and time frames, for 11 sites located in the Jequitinhonha river basin, Brazil.

4.5.1 Taxonomic assignment

First, as expected, the compiled list of species reported for the Jequitinhonha river basin was higher than previously recorded in 2010, which corroborates previous estimates suggesting the occurrence of more than 80 species in this catchment (Andrade-Neto, 2010; Godinho et al., 1999). We conservatively contrasted the taxonomic assignment obtained by eDNA with a list of species reported for this basin; however, thes Jequitinhonha, like many other catchments in the Neotropical region, remains poorly studied and the knowledge regarding its fish diversity is still incomplete. Furthermore, an additional issue reported worldwide, is that even when monitoring programmes are conducted, most of the data obtained are often not published or made available and thus remain inaccessible to further scientific studies (Lindenmayer & Likens, 2009; Revenga et al., 2005). Taxonomic issues are often present in monitoring programs and the risk of misidentification exists regardless of the

method applied (i.e. traditional sampling, morphological identification, eDNA) (Radinger et al., 2018). Erroneous identifications might also be present in the reference databases, especially in biodiversity rich regions such as the Neotropics, where the amount of unknown and undescribed taxa and the occurrence of cryptic species are substantial issues.

As suggested by Li et al. (2018) the filtered dataset provided a more conservative overview compared to the unfiltered dataset and thus did not detect several families and orders known to be present in this catchment. Fish diversity depicted by the heat trees based on the unfiltered data shows that a hidden diversity might be present, especially for the Order Characiformes, as many families appears to comprise several MOTUs (e.g. Anostomidae, Prochilodontidae). This likely reflects the presence of multiple genera/species such as Anostomidae, which includes at least seven species known to occur in this basin.

4.5.2 Introduced and autoctonous species

Environmental DNA metabarcoding allows the detection of multiple species simultaneously, including species not expected to occur in the area (Deiner et al., 2017), which makes it a great tool for tracking biological invasions and providing an early warning of species introduction. Here, almost 30% of the taxa detected by eDNA are represented by non-indigenous species, including species not reported yet for this catchment. To our knowledge, records of *Salminus brasiliensis* and *Lophiosilurus alexandri* are absent in the literature. These are commercially important species, already introduced for fishery purposes in several Brazilian basins (Alves et al., 2007; Vitule et al., 2014), hence their occurrence in the

Jequitinhonha is not a surprise but raises the question: what are the ecosystem consequences of such unmanaged introductions? The only cyprinid documented in this basin was *Hypophthalmichthys molitrix* and here we registered the presence of *Cyprinus carpio*, another species that has been widely introduced to Brazilian waters (Alves et al., 2007). Furthermore, eDNA also allowed the detection of distinct species of tilapia (*Oreochromis* sp. and *Coptodon zilli*). The impacts of tilapia invasion are well known worldwide, and all species show high invasive potential, including in Neotropical countries. *Coptodon zillii* invasion is almost restricted to North America (Cassemiro et al., 2017), however, closely related species have been reported in the Jequitinhonha river Basin (*Oreochromis niloticus*, *Tilapia* sp.) and neighbouring drainages (*Tilapia rendalli* – Doce river basin).

Regarding the native species, eDNA allowed the detection of endemic species (*Wertheimeria maculata*) and other remarkable cases, such as *Crenicichla* sp. The cichlid genus *Crenicichla* is one of the most species rich among the Cichlids and know to widely occur in South America and still lacking an improved taxonomic resolution and conservation status evaluation (Kullander & de Lucena, 2006). In 2006, an expedition applied extensive sampling efforts to collect *Crenicichla* sp. in the Jequitinhonha, without any success and this species were only documented in 2009 by an environmental report (Kullander & Lucena, 2006; Intertechne, 2009), whilst by using eDNA metabarcoding this species was recovered from several locations, indicating a possible large geographical distribution.

As demonstrated in previous studies, identification of some species might be problematic when using eDNA metabarcoding based on the 12S fragment employed here, due to its lack of phylogenetic resolution and the incompleteness of the reference database

(Yu et al. 2012; Carew et al. 2013; Eiler et al. 2013). Most of the MOTUs belonging to *Prochilodus* sp. could not be assigned to species level due to the low taxonomic resolution. Despite the taxonomic assignment to *Prochilodus argenteus,* this represents an invasive species in the Jequitinhonha and due to a low genetic divergence from the endemic species *P. hartii* (Melo et al., 2018) these species are indistinguishable over the 172bp stretch of the 12S fragment (Figure S1, Supplementary Material); therefore, these species are certainly clustered together in this study.

Six anostomids are described for the Jequitinhonha, and here we identified one of these species (*Megaleporinus garmanii*) and two not previously reported (*Leporinus copelandii* and *Hypomasticus mormyrops*). The only previous record of *Leporinus copelandii* was deemed as an historical error (Andrade-Neto, 2010). Cilleros et al. (2018), despite using a different 12S fragment, also reported the limitations in the taxonomic assignment of species belonging to the genus *Leporinus*, therefore our data set is unable to clarify the nuances within this group.

4.5.3 Caveats to species detection

Despite providing important initial information regarding species occurrence and introduction, the data provided here should be used with caution. The effectiveness of this innovative method in recovering eDNA of rare and elusive species from the environment is widely recognized, however, some drawbacks might be discussed before drawing final conclusions. i) eDNA persistence and transport: DNA molecules have a relative fast

degradation and thus, the detection of species suggests their recent presence and a contemporary snapshot of the fish community is expected. However, eDNA persistence varies according to the sampling media, from days in the water column to many weeks in the sediment and as DNA can remain in the water column for more than few days it can also allow the detection of transient species not exactly present in the environment at the sampling time (Dejean et al., 2011, Thomsen et al., 2012). In addition to that, eDNA particles can also travel long distances and be detected far from their original source (Barnes & Turner, 2016; Deiner et al., 2014). However, despite having a high discharge rate (average of 409 m^3/s), the approximate distance between sites was 100km and thus, the influence of eDNA transport on species recovered might not be considered as a great concern in species detection; ii) Origin: besides allowing the detection of fish relatively far from their natural occurrence (Jane et al., 2015), eDNA recovered does not distinguish between dead and live animals and can be even originated from different sources (e.g. disposal of fish products, carcasses, or fishing baits) and this eDNA of exogenous origin should be considered as a potential source of contamination (Merkes et al., 2014). Thus, the presence of species not previously reported should be carefully analysed to verify the origin of their eDNA recovered; iii) Incorrect taxonomic assignment: although 108 new sequences were included in this study, this drainage exhibits high endemism and most of its species are still absent in the database. The incompleteness of genetic database and the lack of phylogenetic resolution of the 12S fragment analysed hampers full recovery of the diversity of neotropical fish, hence underestimating biodiversity. Therefore, efforts to screen longer DNA stretches and to

complete local reference libraries are required to improve the taxonomic assignment quality and take eDNA biomonitoring approach to the next level.

4.5.4 Species richness *vs* anthropogenic impacts and seasonal changes

Ecological communities vary in time and space, and the monitoring of these dynamics is essential for conservation purposes (Bálint et al., 2018). In the Jequitinhonha River basin, significant spatial and temporal fluctuations in fish assemblages were detected. The longitudinal distance and presence of barriers did not explain the variation; however, anthropogenic impacts might still have an influence of fish diversity distribution in this river basin. The sites showing the lowest species richness were represented by the reservoirs (3 – José Gonçalves/Irapé reservoir, 9 – Salto d Divisa/Itapebi reservoir) and the first sites downstream the dams (4 – Coronel Murta and 10 – Itapebi). The presence of dams impacts the environment due to modification of physical and ecological characteristics of the habitats (e.g. water flow, nutrient dynamics, water quality and temperature, increased predation pressure, habitat loss) and is well known as a fish diversity reduction factor (Pelicice & Agostinho, 2007; Pompeu et al., 2012).

The sites comprising most of the fish diversity in this basin were represented by locations characterized by different influences. The most upstream site is located in a less populated and impacted region (Table S7, Supplementary Material), near two areas of natural preservation (State Parks Biribiri and Rio Preto) whilst the other two (Almenara and Belmonte) are included in more populated and impacted cities. Due to the deforestation and

mining activities, the siltation represents one of the greatest impacts in the Jequitinhonha river and it increases towards the river mouth (IBGE, 1997). Almenara is a particularly impacted area and during the sampling this location showed a low water level and accumulation of sediments, which might have contributed to increase the eDNA concentration and accumulation increasing the species diversity.

Furthermore, the high alpha diversity values found for the site located at the river mouth deserves some consideration. This region has marine influence and its abiotic characteristics (e.g. increased salinity) would be expected to restrict the occurrence of some species. Still, most of the species were detected at this site during the first campaign. A hypothesis that could explain this result includes eDNA transport and accumulation. Species shed DNA constantly, which can be available in the water column or bound to sediment, with the latter showing a longer persistence than the eDNA in the surface water. DNA is known to be transported by long distances, in lotic systems the rainfall may lead to an increased water volume contributing to a higher velocity and thus, affect eDNA molecules transport and dispersal towards downstream river (Shogren et al., 2018).

In addition to that, an increase in water flow can also cause the eDNA particles resuspension, which associated with the resistance applied by the incursion of the marine waters into the river, can contribute to retain and resuspend the eDNA accumulated in this area, making it available in the water column. Another fact we need to take into account is that species richness recovered for each site might possibly reflect an overestimation, as eDNA transport from a different location upstream might be detected and thus does not mean that the species themselves are present there at the collection time. Still, eDNA

transport distances may vary between river systems due to abiotic factors (e.g. temperature, pH, UV light) or seasonal changes such as drought or intense rainfall periods (Deiner et al., 2016). As no study has been conducted in Brazilian lotic environments focusing on understanding eDNA transport and diffusion, it is therefore difficult to draw sound conclusions regarding this matter.

Seasonal changes driven by natural factors (e.g. water flow, rainfall) could also contribute to explain assemblage variation even over a short time frame (i.e. weeks) as mobile species such as fish can rapidly disperse and vary their distribution in response to changing abiotic conditions (Arrington & Winemiller, 2006; Fitzgerald et al., 2017). Water availability shows a great temporal variability in semi-arid and arid regions, with short but intense rainfall episodes followed by long dry periods (Leite et al., 2010). The Jequitinhonha river basin is inserted in a semi-arid region and in the first campaign it was facing a severe drought. Before the second sampling the increase in the average accumulated rainfall (2.1-50mm to 100-250 mm) (CPTEC/INPE, 2018), might have contributed to a higher stability amongst sample sites, regarding the contemporary species richness (inferred through water samples), when compared to the first sampling campaign. An increased volume and subsequently higher connectivity of aquatic habitats might stimulate the dispersal and result in reduced densities of organisms (Fitzgerald et al., 2017). Thus, this result might suggest that freshwater fish assemblages in tropical habitats may vary significantly between dry and wet seasons.

4.5.5 Sampling media

Studies conducted in marine and lotic freshwater environments reported a lower detection rate of fish eDNA in sediment compared to water samples (Holman et al., 2018; Koziol et al., 2018, McDevitt et al., 2018; Shaw et al., 2016). Our results were congruent with that, as sediment samples yielded less MOTUs in both filtered and unfiltered datasets and failed in detecting 5 species (*Australoheros facetus*, *Cyprinus carpio*, *Hypostomus gymnorhyncus**, *Poecilia reticulata*, *Trichomycterus* sp.). The influence of substrate choice in eDNA detection rates have been correlated to many factors, including organisms' biological characteristics (habitat preference, life history traits, etc). However, the species not detected by this sampling medium display a wide range of habitat preferences and thus, this might not be a major issue, at least in rivers with these features. Still, those species were detected in very few sites and might have a restricted distribution and occur at low abundances in this basin, which could explain the failed detection by sediments. Furthermore, as the number of samples analysed were different (two sediment samples *vs* 6L of water) the lower yield obtained from sediments compared to water samples might also be due to the lower number of samples collected.

More importantly, both sampling media should be analysed differently since they provide distinct temporal information. While eDNA in the water surface reflects recent presence of species, eDNA bound to the sediment may correspond to an accumulation of DNA molecules during a longer period time and provide information regarding species occurrence over a longer time frame (Deiner et al., 2017).

Regarding the datasets analysed, the filtered data is considered as a sub dataset of the total diversity recovered and showed a lower diversity at the order and family levels. However, the significant positive correlation between datasets demonstrated that beta-diversity is not influenced by the filtering criteria applied as much as sampling medium or sampling time.

Given the unprecedent rates of populations and species decline and the increasing anthropogenic impacts on freshwater communities, the importance of a rapid, robust and efficient monitoring program has never been more in need. Here we demonstrate the advantages of applying eDNA metabarcoding in spatio-temporal ecological studies, as suggested by Bista et al. (2017). Environmental DNA metabarcoding used as a complementary monitoring tool, can extend the data recovered by traditional methods and could greatly contribute to improve biomonitoring in Brazilian freshwaters by providing data for difficult to access localities and allowing the detection of elusive, rare or patchily-distributed species, as shown here. However, to avoid underestimating the biodiversity and reduce ambiguity in eDNA-based species detection, we stress the importance of coordinating morphological surveys with DNA assessments, and increasing the efforts towards building complete genetic databases, ideally composed of whole mitochondrial genomes.

References

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

Bakker, J., Wangensteen, O. S., Chapman, D. D., Boussarie, G., Buddo, D., Guttridge, T. L., et al. (2017). Environmental DNA reveals tropical shark diversity in contrasting levels of anthropogenic impact. *Scientific Reports*, 4(7), 16886. doi: 10.1038/s41598-017-17150-2.

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

Barnes, M. A., and Turner, C. R. (2016). The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, 17(1), 1–17. doi: 10.1007/s10592-015-0775-4

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

Bik, H. M., Porazinska, D. L., Creer, S., Caporaso, G., Knight, R. and Thomas, W. K. (2012). Sequencing our way towards understanding global eukaryotic biodiversity. *Trends in Ecology and Evolution*, 27 (4), 233-243. doi: 10.1016/j.tree.2011.11.010

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knap, M., et al. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology and Evolution*, 29(6), 358-367. doi: 10.1016/j.tree.2014.04.003

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

Carew, M. E., Miller, A. D., and Hoffmann, A. A., (2011). Phylogenetic signals and ecotoxicological responses: potential implications for aquatic biomonitoring. *Ecotoxicology*, 20(3), 595-606. doi: 10.1007/s10646-011-0615-3

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

Chown, S. L., Sinclair, B. J., and van Vuuren, B. J. (2008). DNA barcoding and the documentation of alien species establishment on sub-Antarctic Marion Island. *Polar Biology*, 31, 651–655. doi: 10.1007/s00300-007-0402-z

Cilleros, K., Valentini, A., Allard, L., Dejean, T., Etienne, R., Grenouillet, G., et al. (2018). Unlocking biodiversity and conservation studies in high diversity environments using

environmental DNA (eDNA): a test with Guianese freshwater fishes. *Molecular Ecology Resources*, 7, 1–20. doi: 10.1111/1755-0998.12900

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

Deiner, K. and Altermatt, F. (2014). Transport distance of invertebrate environmental DNA in a natural river. *PLoS ONE*, 9: e88786. doi: 10.1371/journal.pone.0088786

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

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

Deiner, K., Lopez, J., Bourne, S., Holman, L., Seymour, M., Grey, E. K., et al. (2018). Optimising the detection of marine taxonomic richness using environmental DNA metabarcoding: the effects of filter material, pore size and extraction method. *Metabarcoding and Metagenomics*, 2, e28963. doi: 10.3897/mbmg.2.28963

Dejean, T., Valentini, A., Duparc, A., Pellier-Cuit, S., Pompanon, F., Taberlet, P., Miaud, C. (2011). Persistence of Environmental DNA in Freshwater Ecosystems. *PLoS ONE*, 6: e23398. doi: 10.1371/journal.pone.0023398

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

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

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

Elbrecht, V., and Leese, F. (2017). Validation and Development of COI Metabarcoding Primers for Freshwater Macroinvertebrate Bioassessment. *Frontiers in Environmental Science*, 5, 1-11. doi: 10.3389/fenvs.2017.00011

Fricke, R., Eschmeyer, W. N. and van der Laa, R. (eds). (2018). Catalog of Fishes: Genera, Species, References. Available at: (http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp).

Evans, N. T., Olds, B. P., Renshaw, M. A., Turner, C. R., Li, Y., Jerde, C. L., et al. (2016). Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. *Molecular Ecology Resources*, 16, 29-41. doi: 10.1111/1755-0998.12433

Evans, N. T., and Lamberti, G. A. (2017). Freshwater fisheries assessment using environmental DNA: A primer on the method, its potential, and shortcomings as a conservation tool. *Fisheries Research*, 197, 60-67. doi: 10.1016/j.fishres.2017.09.013

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

Ficetola, G. F., Miaud, C., Pompanon, P. and Taberlet, P. (2008). Species detection using environmental DNA from water samples. *Biology Letters*, 4, 423–425.

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

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

Friberg, N., Angelopoulos, N. V., Buijse, A. D., Cowx, I. G., Kail, J., Moe, T. F., et al. (2016). Effective river restoration in the 21st century: from trial and error to novel evidence-based approaches. *Advances in Ecolological Research*, 55, 535–611. doi: 10.1016/bs.aecr.2016.08.010

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

Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., et al. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution*, 7(11), 1299–1307. doi: 10.1111/2041-210X.12595

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

Handley, L. L., Read, D. S., Winfield, I. J., Kimbell, H., Li, J., Hahn, C., et al. (2018). Temporal and spatial variation in distribution of fish environmental DNA in England's largest lake. *bioRxiv.* doi: 10.1101/376400

Hänfling, B., Handley, L. L., Read, D. S., Hahn, C., Li, J., Nichols, P., et al. (2016). Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. *Molecular Ecology*, 25(13), 3101–3119. doi: 10.1111/mec.13660

Hering, D., Borja, A., Jones, J. I., Pont, D., Boets, P., Bouchez, A., et al. (2018). Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive. *Water Research*, 138, 192–205. doi: 10.1016/j.watres.2018.03.003

Holman, L. E., de Bruyn, M., Creer, S., Carvalho, G., Robidart, J., and Rius, M. (2018). The detection of novel and resident marine non-indigenous species using environmental DNA metabarcoding of seawater and sediment. bioRxiv. doi: 10.1101/440768

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

Jane, S. F., Wilcox, T. M., McKelvey, K. S., Young, M. K., Schwartz, M. K., Lowe, W. H., et al. (2015). Distance, flow and PCR inhibition: eDNA dynamics in two headwater streams. *Molecular Ecology Resources*, 15(1), 216-227. doi: 10.1111/1755-0998.12285

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

Jo, T., Murakami, H., Yamamoto, S., Masuda, R., and Minamoto, T. (2019). Effect of water temperature and fish biomass on environmental DNA shedding, degradation, and size distribution. *Ecology and Evolution*, 1–12. doi:10.1002/ece3.4802.

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

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

Koziol, A., Stat, M., Simpson, T., Jarman, S., Di Battista, J. D., Harvey, E. S., et al. (2018). Environmental DNA metabarcoding studies are critically affected by substrate selection. *Molecular Ecology Resources*, 1–11. doi: 10.1111/1755-0998.12971

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

Leese, F., Bouchez, A., Abarenkov, K., Altermatt, F., Borja, A., Bruce, K., et al. (2018). Why we need sustainable networks bridging countries, disciplines, cultures and generations for aquatic biomonitoring 2.0: a perspective derived from the DNAqua-Net COST Action. *Advances in Ecological Research*, 58, 63-99. doi: 10.1016/bs.aecr.2018.01.001

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

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

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

Machado, V. N., Collins, R. A., Ota, R. P., Andrade, M. C., Farias, I. P., and Hrbek, T. (2018). One thousand DNA barcodes of piranhas and pacus reveal geographic structure and unrecognised diversity in the Amazon. *Scientific Reports*, 8(1), 8387. doi: :10.1038/s41598-018-26550-x

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

Majaneva, M., Diserud, O. H., Eagle, S. H. C., Boström, E., Hajibabaei, M. and Ekrem, T. (2018). Environmental DNA filtration techniques affect recovered biodiversity. *Scientific Reports*, 8(1), 1-11. doi: 10.1038/s41598-018-23052-8

McDevitt, A. D., Sales, N. G., Browett, S. S., Sparnenn, A. O., Mariani, S., Wangensteen, O. S., et al. (2018). Environmental DNA metabarcoding as an effective and rapid tool for fish monitoring in canals. *bioRxiv*. doi: 10.1101/498451.

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

Merkes, C. M., McCalla, S. G., Jensen, N. R., Gaikowski, M. P., Amberg, J. J. (2014). Persistence of DNA in carcasses, slime and avian feces may affect interpretation of environmental DNA data. *PLoS ONE,* 9:e113346. doi: 10.1371/journal.pone.0113346

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

Nakagawa, H., Yamamoto, S., Sato, Y., Sado, T., Minamoto, T., and Miya, M. (2018). Comparing local and regional scale estimations of the diversity of stream fish using eDNA metabarcoding and conventional observation methods. *Freshwater Biology*, 63, 569-580. doi: 10.1111/fwb.13094

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

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

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., Ohara, R., et al. (2013). Package "vegan." Community ecology package.

Olds, B. P., Jerde, C. L., Renshaw, M. A., Li, Y., Evans, N. T., Turner, C. R., et al. (2016). Estimating species richness using environmental DNA. *Ecology and Evolution*, 6, 4214-4226. doi: 10.1002/ece3.2186

Palmer, M. A. (2010). Beyond infrastructure. *Nature*, 467, 534-535.

Pawlowski, J., Kelly-Quinn, M., Altermatt, F., Apothéloz-Perret-Gentil, L., Beja, P., Boggero, A., et al. (2018). The future of biotic indices in the ecogenomic era: Integrating (e)DNA metabarcoding in biological assessment of aquatic ecosystems. *Science of the Total Environment*, 637, 1295–1310. doi: 10.1016/j.scitotenv.2018.05.002

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

Pereira, E. H. L., Pessali, T. C., Andrade, F., and Reis, R. E. (2017). Description of a new species of Pareiorhaphis (Loricariidae: Neoplecostominae) from the rio Jequitinhonha basin, Minas Gerais, eastern Brazil. *Neotropical Ichthyology*, 15(3), e170007. doi: 10.1590/1982-0224-20170007.

Piaggio, A. J., Engeman, R. M., Hopken, M. W., Humphrey, J. S., Keacher, K. L., Bruce, W. E. and Avery, M. L. (2013). Detecting an elusive invasive species: a diagnostic PCR to detect Burmese python in Florida waters and an assessment of persistence of environmental DNA. *Molecular Ecology Resources*, 14 (2), 374-380. doi: 10.1111/1755-0998.12180

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

Pugedo, M. L., Andrade-Neto, F. R., Pessali, T. C., Birindelli, J. L. O., and Carvalho, D. C. (2016). Integrative taxonomy supports new candidate fish species in a poorly studied neotropical region: the Jequitinhonha River Basin. *Genetica*, 144, 341-349. doi: 10.1007/s10709-016-9903-4

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

Rees, H. C., Maddison, B. C., Middleditch, D. J. Patmore, J. R. M., and Gough, K. C. (2014). The detection of aquatic animal species using environmental DNA – a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology*, 51(5), 1450-1459. doi: 10.1111/1365-2664.12306

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

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

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

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

Sales, N. G., Wangensteen, O. S., Carvalho, D. C., and Mariani, S. (2018). Influence of preservation method, sampling medium and time on eDNA recovery in a Neotropical river. *bioRxiv*. doi: 10.1101/489609

Sales, N.G., Mariani, S., Pessali, T. C., Salvador, G. N., and Carvalho, D. C. (2018). Hidden diversity hampers conservation efforts in a highly impacted Neotropical river system. *Frontiers in Genetics*, 9(271), 1-10. doi: 10.3389/fgene.2018.00271

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

Shimabukuro-Dias, C. K., Costa Silva, G. J. da, Ashikaga, F. Y., Foresti, F., and Oliveira, C. (2017). Molecular identification of the fish fauna from the pantanal flood plain area in Brazil. Mitochondrial DNA Part A, 28(4), 588–592. doi: 10.3109/24701394.2016.1149826

Simberloff, D., Martin, J. L., Genovesi, P., Maris, V., Wardle, D. A., Aronson, J., et al. (2013). Impacts of biological invasions: what's what and the way forward. *Trends in Ecology and Evolution*, 28(1), 58-66. doi: 10.1016/j.tree.2012.07.013

Silva-Santos, R., Ramirez, J. L., Galetti, P. M., and Freitas, P. D. (2018). Molecular evidences of a hidden complex scenario in *Leporinus cf. friderici*. *Frontiers in Genetics*, 9, 1-9. doi: 10.3389/fgene.2018.00047

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

Takahara, T., Minamoto, T., and Doi, H. (2013). Using Environmental DNA to estimate the distribution of an invasive fish species in ponds. *PLoS ONE,* 8(2): e56584. doi: 10.1371/journal.pone.0056584

Thomsen, P. F., Kielgast, J., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T., et al.(2012), Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular Ecology*, 21, 2565-2573. doi: 10.1111/j.1365-294X.2011.05418.x

Thomsen, F. P. and Willerslev, E. (2015). Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, 183, 4-18 . doi: 10.1016/j.biocon.2014.11.019

Tillotson, M. D., Kelly, R. P., Duda, J. J., Hoy, M., Kralj, J., and Quinn, T. P. (2018). Concentrations of environmental DNA (eDNA) reflect spawning salmon abundance at fine spatial and temporal scales. *Biological Conservation*, 220, 1–11. doi:10.1016/j.biocon.2018.01.030.

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

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

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

Vono, V. and Birindelli, J. L. (2007). Natural history of *Wertheimeria maculata*, a basal doradid catfish endemic to eastern Brazil (Siluriformes: Doradidae). *Ichthyological Exploration of Freshwaters*, 18(2), 183.

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

Wilcox, T. M., McKelvey, K. S., Young, M. K., Jane, S. F., Lowe, W. H., Whiteley, A. R. et al. (2013). Robust detection of rare species using Environmental DNA: The Importance of Primer Specificity. *PLoS ONE*, 8 (3), e59520. doi: 10.1371/journal.pone.0059520

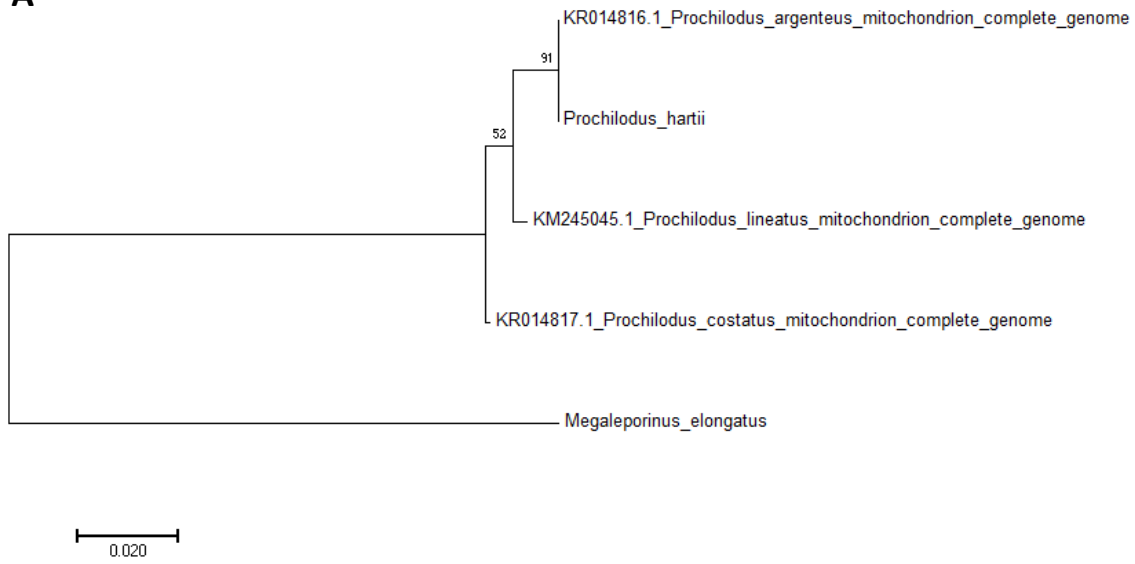
WWF. 2018. Living Planet Report – 2018: Aiming Higher. Grooten, M. and Almond, R. E. A. (Eds.) WWF, Gland, Switzerland.

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

Zawadzki, C. H., Carvalho, P. H., Birindelli, J. O., and Azevedo, F. M. (2016). *Hypostomus nigrolineatus*, a new dark-striped species from the rio Jequitinhonha and rio Pardo basins, Brazil (Siluriformes, Loricariidae). *Ichthyological Exploration of Freshwaters*, 27(3), 263-274

Supplementary material

A



B

Species 1	Species 2	Dist
KR014816.1_Prochilodus_argenteus_	KR014817.1_Prochilodus_costatus	0.018
KR014816.1_Prochilodus_argenteus_	KM245045.1_Prochilodus_lineatus	0.012
KR014817.1_Prochilodus_costatus	KM245045.1_Prochilodus_lineatus	0.006
KR014816.1_Prochilodus_argenteus	Megaleporinus_elongatus	0.216
KR014817.1_Prochilodus_costatus	Megaleporinus_elongatus	0.205
KM245045.1_Prochilodus_lineatus	Megaleporinus_elongatus	0.216
KR014816.1_Prochilodus_argenteus	Prochilodus_hartii	0.000
KR014817.1_Prochilodus_costatus	Prochilodus_hartii	0.018
KM245045.1_Prochilodus_lineatus	Prochilodus_hartii	0.012
Megaleporinus_elongatus	Prochilodus_hartii	0.216

Figure S1: Phylogenetic tree (A) and pairwise genetic distance (B) recovered for *Prochilodus* spp.

Table S1: Taxa detected by eDNA metabarcoding and correspondent nearest neighbor species reported for Jequitinhonha river basin.

Order	Family	Genus	eDNA taxon	Nearest neighbor species reported for JRB			
Characiformes	Anostomidae	Hypomasticus	*Hypomasticus mormyrops*	*Megaleporinus garmani)*			
		Leporinus	*Leporinus copelandii*	*Leporinus bahiensis*	*Leporinus* spp.	*Leporinus steindachneri*	*Leporinus taeniatus*
		Megaleporinus	*Megaleporinus garmani*	*Megaleporinus garmani*	*Megaleporinus elongatus*		
	Bryconidae	Brycon	*Brycon* sp.	*Brycon devillei*	*Brycon* sp.		
		Salminus	*Salminus brasiliensis**				
	Characidae	Moenkhausia	*Moenkhausia costae*	*Moenkhausia costae*	*Moenkhausia intermedia*		
		Oligosarcus	*Oligosarcus argenteus*	*Oligosarcus hepsetus*	*Oligosarcus macrolepis*		
	Crenuchidae	Characidium	*Characidium* sp.	*Characidium spp.*	*Characidium* cf. *fasciatum*		
	Curimatidae	Cyphocharax	*Cyphocharax gilbert*	*Cyphocharax jagunco*	*Cyphocharax lundi*	*Cyphocharax naegelii*	
	Erythrinidae	Hoplias	*Hoplias malabaricus*	*Hoplias brasiliensis*	*Hoplias malabaricus*		
	Prochilodontidae	Prochilodus	*Prochilodus argenteus*	*Prochilodus hartii*	*Prochilodus argenteus*	*Prochilodus costatus*	*Prochilodus lineatus*
	Serrasalmidae	Serrasalmus	*Serrasalmus brandtii*	*Serrasalmus brandtii*	*Serrasalmus* sp.		
Cichliformes	Cichlidae	Astronotus	*Astronotus ocellatus*	*Astronotus ocellatus*			
		Australoheros	*Australoheros facetus*	*Australoeros* sp.			
		Coptodon	*Coptodon zillii*	Tilapia sp.			
		Crenicichla	*Crenicichla lacustris*	*Crenicichla* sp.			
		Geophagus	*Geophagus brasiliensis*	*Geophagus brasiliensis*			
		Oreochromis	*Oreochromis aureus*	*Oreochromis niloticus*			
Cypriniformes	Cyprinidae	Cyprinus	*Cyprinus carpio*	*Hypophthalmichthys molitrix*			
Cyprinodontiformes	Poeciliidae	Phalloceros	*Phallocero*s sp.	*Phalloceros caudimaculatus*	*Phalloceros* sp.		
		Poecilia	*Poecilia reticulata*	*Poecilia reticulata*	*Poecilia vivipara*		
Gymnotiformes	Gymnotidae	Gymnotus	*Gymnotus carapo*	*Gymnotus bahianus*	*Gymnotus carapo*	*Gymnotus pantherinus*	*Gymnotus sylvius*
Siluriformes	Auchenipteridae	Trachelyopterus	*Trachelyopterus striatulus*	*Trachelyopterus galeatus*	*Trachelyopterus striatulus*		
	Callichthyidae	Hoplosternum	*Hoplosternum littorale*	*Hoplosternum littorale*			
	Doradidae	Wertheimeria	*Wertheimeria maculata*	*Wertheimeria maculata*			
	Heptapteridae	Rhamdia	*Rhamdia quelen*	*Rhamdia quelen*	*Rhamdia jequitinhonha*		
	Loricariidae	Delturus	*Delturus carinotus*	*Delturus brevis*			
		Hypostomus	*Hypostomus gymnorhynchus*	*Hypostomus nigrolineatus*	*Hypostomus* sp.		
			Hypostomus nigromaculatus	*Hypostomus nigrolineatus*	*Hypostomus* sp.		
			Neoplecostominae gen. 2 sp. FFR-2012				
			Neoplecostomini gen.n. sp.n TEP-2017				
	Pseudopimelodidae	Lophiosilurus	*Lophiosilurus alexandri**				
	Trichomycteridae	Trichomycterus	*Trichomycterus* sp.	*Trichomycterus itacambirussu*	*Trichomycterus jequitinhonhae*	*Trichomycterus landinga*	*Trichomycterus* spp.
			Trichomycterus sp.2	*Trichomycterus itacambirussu*	*Trichomycterus jequitinhonhae*	*Trichomycterus landinga*	*Trichomycterus* spp.

* Species not previously reported for JRB

Table S2: Species reported for the Jequitinhonha River Basin.

Order	Family	Species	Status	Habitat	Reference	Current status (Catalog of Fishes)
Characiformes	Anostomidae	Hypomaticus garmani	native	freshwater	Andrade-Neto, 2009	Megaleporinus garmani
Characiformes	Anostomidae	Leporinus bahiensis	native	freshwater	Andrade-Neto, 2009	Leporinus bahiensis
Characiformes	Anostomidae	Leporinus crassilabris	native	freshwater	Andrade-Neto, 2009	Synonym of Megaleporinus elongatus
Characiformes	Anostomidae	Leporinus elongatus		freshwater	Pugedo et al., 2016	Megaleporinus elongatus
Characiformes	Anostomidae	Leporinus garmani	native	freshwater	Report CEMIG, 2005	Megaleporinus garmani
Characiformes	Anostomidae	Leporinus spp.	not described yet	freshwater	Andrade-Neto, 2009	
Characiformes	Anostomidae	Leporinus steindachneri	native	freshwater	Andrade-Neto, 2009	Leporinus steindachneri
Characiformes	Anostomidae	Leporinus taeniatus		freshwater	Pugedo et al., 2016	Leporinus taeniatus
Characiformes	Bryconidae	Brycon devillei	endangered	freshwater	Andrade-Neto, 2009	Brycon devillei
Characiformes	Bryconidae	Brycon sp.	not described yet	freshwater	Andrade-Neto, 2009	-
Characiformes	Callichthyidae	Callichthys callichthys	native	freshwater	Andrade-Neto, 2009	Callichthys callichthys
Characiformes	Characidae	Achirus lineatus	native	marine	Andrade-Neto, 2009	Achirus lineatus
Characiformes	Characidae	Acinocheirodon melanogramma	native	freshwater	Andrade-Neto, 2009	Acinocheirodon melanogramma
Characiformes	Characidae	Aphyocheirodon sp.	native	freshwater	Report CEMIG, 2007	-
Characiformes	Characidae	Astyanax bimaculatus	native	freshwater	Andrade-Neto, 2009	Astyanax bimaculatus
Characiformes	Characidae	Astyanax brevirhinus	native	freshwater	Andrade-Neto, 2009	Astyanax brevirhinus
Characiformes	Characidae	Astyanax cf. jequitinhonhae	native	freshwater	Report CEMIG, 2007	Astyanax fasciatus
Characiformes	Characidae	Astyanax fasciatus	native	freshwater	Andrade-Neto, 2009	Astyanax fasciatus
Characiformes	Characidae	Astyanax lacustris	native	freshwater	Pugedo et al., 2016	Astyanax lacustris
Characiformes	Characidae	Astyanax scabripinnis	native	freshwater	Andrade-Neto, 2009	Astyanax scabripinnis
Characiformes	Characidae	Astyanax sp.	native	freshwater	Report CEMIG, 2007	-
Characiformes	Characidae	Astyanax turmalinensis	native	freshwater	Andrade-Neto, 2009	Astyanax turmalinensis
Characiformes	Characidae	Hyphessobrycon cf. luetkeni	native	freshwater	Intertechne, 2009	Hyphessobrycon luetkenii
Characiformes	Characidae	Hyphessobrycon sp.	native	freshwater	Report CEMIG, 2007	-
Characiformes	Characidae	Knodus moenkhausii		freshwater	Pugedo et al., 2016	Knodus moenkhausii
Characiformes	Characidae	Mimagoniates sylvicola	native	freshwater	Andrade-Neto, 2009	Mimagoniates sylvicola
Characiformes	Characidae	Moenkhausia costae	Introduced	freshwater	Andrade-Neto, 2009	Moenkhausia costae
Characiformes	Characidae	Moenkhausia intermedia	native	freshwater	Intertechne, 2009	Moenkhausia intermedia
Characiformes	Characidae	Nematocharax venustus	endangered	freshwater	Andrade-Neto, 2009	Nematocharax venustus
Characiformes	Characidae	Oligosarcus hepsetus	native	freshwater	Andrade-Neto, 2009	Oligosarcus hepsetus
Characiformes	Characidae	Oligosarcus macrolepis	native	freshwater	Andrade-Neto, 2009	Oligosarcus macrolepis
Characiformes	Characidae	Serrapinnus zanatae	native	freshwater	Jerep, Camelier & Zanata, 2016	Serrapinnus zanatae
Characiformes	Crenuchidae	Characidium cf. fasciatum	native	freshwater	Intertechne, 2009	Characidium fasciatum
Characiformes	Crenuchidae	Characidium spp.	not described yet	freshwater	Andrade-Neto, 2009	-
Characiformes	Curimatidae	Cyphocharax cf gilbert	native	freshwater	Intertechne, 2009	Cyphocharax gilbert
Characiformes	Curimatidae	Cyphocharax jagunco	native	freshwater	Dutra et al., 2016	Cyphocharax jagunco
Characiformes	Curimatidae	Cyphocharax lundi	native	freshwater	Dutra et al., 2016	Cyphocharax lundi
Characiformes	Curimatidae	Cyphocharax naegelii	native	freshwater	Report CEMIG, 2005	Cyphocharax naegelii
Characiformes	Curimatidae	Steindachnerina elegans	native	freshwater	Andrade-Neto, 2009	Steindachnerina elegans

Order	Family	Species	Status	Habitat	Reference	Current status (Catalog of Fishes)
Characiformes	Erythrinidae	Hoplias brasiliensis	native	freshwater	Andrade-Neto, 2009	Hoplias brasiliensis
Characiformes	Erythrinidae	Hoplias malabaricus	native	freshwater	Andrade-Neto, 2009	Hoplias malabaricus
Characiformes	Prochilodontidae	Prochilodus argenteus	Introduced	freshwater	Sales et al., 2017	Prochilodus argenteus
Characiformes	Prochilodontidae	Prochilodus costatus	Introduced	freshwater	Godinho et al., 1999	Prochilodus costatus
Characiformes	Prochilodontidae	Prochilodus hartii	native	freshwater	Andrade-Neto, 2009	Prochilodus hartii
Characiformes	Prochilodontidae	Prochilodus lineatus	Introduced	freshwater	Sales et al., 2017	Prochilodus lineatus
Characiformes	Serrasalmidae	Colossoma macropomum		freshwater	Pugedo et al., 2016	Colossoma macropomum
Characiformes	Serrasalmidae	Serrasalmus brandtii	Introduced	freshwater	Pugedo et al., 2016	Serrasalmus brandtii
Characiformes	Serrasalmidae	Serrasalmus sp.	Introduced	freshwater	Godinho, 2008	-
Cichliformes	Cichlidae	Astronotus ocellatus	Introduced	freshwater	Bizerril & Lima, 2005	Astronotus ocellatus
Cichliformes	Cichlidae	Australoheros sp.		freshwater	Pugedo et al., 2016	-
Cichliformes	Cichlidae	Cichla kelberi	Introduced		Pugedo et al., 2016	Cichla kelberi
Cichliformes	Cichlidae	Cichla sp.	Introduced	freshwater	FADETEC, 2002	-
Cichliformes	Cichlidae	Cichlasoma facetum	native	freshwater	Report CEMIG, 2007	Valid as Australoheros facetus
Cichliformes	Cichlidae	Crenicichla sp.	native	freshwater	Intertechne, 2009	-
Cichliformes	Cichlidae	Geophagus brasiliensis	native	Freshwater; brackish	Andrade-Neto, 2009	Geophagus brasiliensis
Cichliformes	Cichlidae	Oreochromis niloticus	Introduced	freshwater	Bizerril & Lima, 2005	Oreochromis niloticus
Cichliformes	Cichlidae	Tilapia sp.	Introduced	freshwater	Godinho et al., 2001	-
Clupeiformes	Engraulidae	Anchoviella lepidentostole	native	Marine; freshwater; brackish	Andrade-Neto, 2009	Anchoviella lepidentostole
Clupeiformes	Engraulidae	Lycengraulis grossidens	native	Marine; freshwater; brackish	Andrade-Neto, 2009	Lycengraulis grossidens
Cypriniformes	Cyprinidae	Hypophthalmichthys molitrix	Introduced	freshwater	Alves et al., 2007	Hypophthalmichthys molitrix
Cyprinodontiformes	Poeciliidae	Phalloceros caudimaculatus	native	freshwater	Andrade-Neto, 2009	Phalloceros caudimaculatus
Cyprinodontiformes	Poeciliidae	Phalloceros sp.	native	freshwater	Pugedo et al., 2016	-
Cyprinodontiformes	Poeciliidae	Poecilia reticulata	Introduced	freshwater	Bizerril & Lima, 2005	Poecilia reticulata
Cyprinodontiformes	Poeciliidae	Poecilia vivipara	native	freshwater	Intertechne, 2009	Poecilia vivipara
Cyprinodontiformes	Rivulidae	Simpsonichthys espinhacensis	possibly endangered	freshwater	Nielsen, Pessali & Dutra, 2017	Simpsonichthys espinhacensis
Cyprinodontiformes	Rivulidae	Simpsonichthys ocellatus	native	freshwater	Andrade-Neto, 2009	Hypsolebias ocellatus
Cyprinodontiformes	Rivulidae	Simpsonichthys perpendicularis	endangered	freshwater	Andrade-Neto, 2009	Ophthalmolebias perpendicularis
Gymnotiformes	Gymnotidae	Gymnotus bahianus	native	freshwater	Andrade-Neto, 2009	Gymnotus bahianus
Gymnotiformes	Gymnotidae	Gymnotus carapo	native	freshwater	Andrade-Neto, 2009	Gymnotus carapo
Gymnotiformes	Gymnotidae	Gymnotus pantherinus	native	freshwater	Andrade-Neto, 2009	Gymnotus pantherinus
Gymnotiformes	Gymnotidae	Gymnotus sylvius			Pugedo et al., 2016	Gymnotus sylvius
Gymnotiformes	Sternopygidae	Eigenmania virescens	native	freshwater	Andrade-Neto, 2009	Eigenmania virescens
Mugiliformes	Mugilidae	Mugil platanus	native	Marine; freshwater; brackish	Andrade-Neto, 2009	Synonym of Mugil liza
Perciformes	Carangidae	Caranx latus	native	Marine; freshwater; brackish	Andrade-Neto, 2009	Caranx latus
Perciformes	Centropomidae	Centropomus parallelus	native	Marine; freshwater; brackish	Andrade-Neto, 2009	Centropomus parallelus
Perciformes	Centropomidae	Centropomus undecimalis	native	Marine; freshwater; brackish	Andrade-Neto, 2009	Centropomus undecimalis
Perciformes	Eleotridae	Dormitator maculatus	native	Marine; freshwater; brackish	Andrade-Neto, 2009	Dormitator maculatus
Perciformes	Eleotridae	Eleotris pisonis	native	Marine; freshwater; brackish	Andrade-Neto, 2009	Eleotris pisonis

Order	Family	Species	Status	Habitat	Reference	Current status (Catalog of Fishes)
Perciformes	Gerreidae	Eugerres brasilianus	native	marine	Pugedo et al., 2016	Eugerres brasilianus
Perciformes	Gobiidae	Awaous tajasica	native	Freshwater; brackish;	Andrade-Neto, 2009	Awaous tajasica
Siluriformes	Ariidae	Genidens genidens	native	marine;brackish	Andrade-Neto, 2009	Genidens genidens
Siluriformes	Auchenipteridae	Pseudauchenipterus jequitinhonhae	native	freshwater	Andrade-Neto, 2009	Pseudauchenipterus jequitinhonhae
Siluriformes	Auchenipteridae	Trachelyopterus galeatus		freshwater	Pugedo et al., 2016	Trachelyopterus galeatus
Siluriformes	Auchenipteridae	Trachelyopterus striatulus	native	freshwater	Andrade-Neto, 2009	Trachelyopterus striatulus
Siluriformes	Callichthyidae	Aspidoras cf. rochai	native	freshwater	Report CEMIG, 2007	Aspidoras cf. rochai
Siluriformes	Callichthyidae	Corydoras sp.		freshwater	Pugedo et al., 2016	-
Siluriformes	Callichthyidae	Hoplosternum littorale		freshwater	Pugedo et al., 2016	Hoplosternum littorale
Siluriformes	Clariidae	Clarias gariepinus	Introduced	freshwater	Godinho, 2007	Clarias gariepinus
Siluriformes	Doradidae	Wertheimeria maculata	native	freshwater	Andrade-Neto, 2009	Wertheimeria maculata
Siluriformes	Heptapteridae	Imparfinis sp.	native	freshwater	Intertechne, 2009	-
Siluriformes	Heptapteridae	Pariolius sp.	native	freshwater	Report CEMIG, 2007	-
Siluriformes	Heptapteridae	Pimelodella sp.	not described yet	freshwater	Andrade-Neto, 2009	-
Siluriformes	Heptapteridae	Rhamdia jequitinhonha	endangered	freshwater	Andrade-Neto, 2009	Rhamdia jequitinhonha
Siluriformes	Heptapteridae	Rhamdia quelen	native	freshwater	Andrade-Neto, 2009	Rhamdia quelen
Siluriformes	Loricariidae	Chauliocheilos saxatilis	native	freshwater	Martins et al., 2014	Chauliocheilos saxatilis
Siluriformes	Loricariidae	Delturus brevis	native	freshwater	Andrade-Neto, 2009	Delturus brevis
Siluriformes	Loricariidae	Harttia garavelloi	native	freshwater	Andrade-Neto, 2009	Harttia garavelloi
Siluriformes	Loricariidae	Hypostomus nigrolineatus	native	freshwater	Zawadzki et al., 2016	Hypostomus nigrolineatus
Siluriformes	Loricariidae	Hypostomus sp.	not described yet	freshwater	Andrade-Neto, 2009	-
Siluriformes	Loricariidae	Microlepidogaster discus	native	freshwater	Martins et al., 2014	Microlepidogaster discus
Siluriformes	Loricariidae	Pareiorhaphis sp.	native	freshwater	Report CEMIG, 2007	-
Siluriformes	Loricariidae	Pareiorhaphis stephanus	native	freshwater	Andrade-Neto, 2009	Pareiorhaphis stephana
Siluriformes	Loricariidae	Pareiorhaphis lineata	native	freshwater	Pereira et al., 2017	Pareiorhaphis lineata
Siluriformes	Loricariidae	Parotocinclus jequi	native	freshwater	Lehman et al., 2013	Parotocinclus jequi
Siluriformes	Loricariidae	Parotocinclus sp.	native	freshwater	Report CEMIG, 2007	-
Siluriformes	Loricariidae	Pogonopoma wertheimeri	native	freshwater	Andrade-Neto, 2009	Pogonopoma wertheimeri
Siluriformes	Pimelodidae	Pseudoplatystoma spp.	Introduced	freshwater	Godinho, 2007	-
Siluriformes	Pimelodidae	Steindachneridion amblyurum	endangered	freshwater	Andrade-Neto, 2009	Steindachneridion amblyurum
Siluriformes	Trichomycteridae	Trichomycterus itacambirussu	native	freshwater	Andrade-Neto, 2009	Trichomycterus itacambirussu
Siluriformes	Trichomycteridae	Trichomycterus jequitinhonhae	native	freshwater	Andrade-Neto, 2009	Trichomycterus jequitinhonhae
Siluriformes	Trichomycteridae	Trichomycterus landinga	native	freshwater	Andrade-Neto, 2009	Trichomycterus landinga
Siluriformes	Trichomycteridae	Trichomycterus spp.	not described yet	freshwater	Report CEMIG, 2007	-
Siluriformes	Pseudopimelodidae	Microglanis cf. parahybae	native	freshwater	Intertechne, 2009	Microglanis parahybae
Synbranchiformes	Synbranchidae	Synbranchus marmoratus	native	freshwater;brackish	Andrade-Neto, 2009	Synbranchus marmoratus

Table S3: Comparison of different SWARM clustering thresholds.

	Sampling event 1			Sampling event 2		
	LIB1			LIB2		
	d=1	**d=2**	**d=3**	**d=1**	**d=2**	**d=3**
Number of swarms	107567	60259	41094	131121	70633	46339
Largest swarm	64783	135807	179901	40359	52581	91237
Max generations	28	30	31	21	24	29
	After SWARM Recount					
Clusters	107567	60259	41094	131121	70633	46339
Cluster >2 reads	5249	4563	3595	5714	4766	3823
Reads kept for calculations/Total reads = 404914	302596	349218	367415	369295	428835	452186
Alignment cached	81.44%	78.15%	75.67	79.34	74.66	70.02
Number of MOTUs	5249	4563	3595	5714	4766	3823
Number of reads	3955997	4004093	4022641	3914479	3975772	3999426
Actinopterygii-MOTUs	4821	3974	3065	5054	4128	3265
Neotropical Orders-MOTUs	3195	2534	2168	2421	2041	1691
Neotropical Families-MOTUs	2635	2058	1862	1732	1441	1007
Neotropical species-MOTUs	2579	1918	1744	1417	1212	979
Contamination removal	2056	1664	1491	967	893	820
Number of MOTUs assigned to species (minid 0.97)	155	52	25	59	42	28
Number of MOTUs assigned to genus (minid 0.97)	58	16	5	64	20	6
Number of MOTUs assigned to Family (minid 0.97)	22	7	0	25	5	2
Number of MOTUs assigned to Suborder (minid 0.97)	4	2	1	2	0	0

Table S4: Species identified applying the minimum identity of 0.97, according to each SWARM threshold.

MOTUs assigned to species (minid 0.97)											
First Sampling event						Second Sampling event					
LIB1						LIB2					
d=1		d=2		d=3		d=1		d=2		d=3	
Species	Number of MOTUs	Species	Number of MOTUs	Species	Number of MOTUs	Species	Number of MOTUs	Species	Number of MOTUs	Species	Number of MOTUs
Astronotus ocellatus	1	Astronotus ocellatus	1	Astronotus ocellatus	1	Astronotus ocellatus	1	Astronotus ocellatus	1	Astronotus ocellatus	1
Australoheros facetus	1	Australoheros facetus	1	Australoheros facetus	1	Brycon sp.	1	Brycon sp.	1	Characidium sp.	1
Brycon sp.	1	Crenicichla lacustris	6	Crenicichla lacustris	3	Characidium sp.	6	Coptodon zillii	1	Coptodon zillii	1
Characidium sp.	11	Cyphocharax gilbert	1	Cyprinus carpio	1	Coptodon zillii	1	Crenicichla lacustris	2	Crenicichla lacustris	2
Crenicichla lacustris	9	Cyprinus carpio	1	Delturus carinotus	1	Crenicichla lacustris	3	Cyphocharax gilbert	1	Cyphocharax gilbert	1
Cyphocharax gilbert	2	Delturus carinotus	2	Geophagus brasiliensis	1	Cyphocharax gilbert	1	Delturus carinotus	1	Delturus carinotus	1
Cyprinus carpio	1	Geophagus brasiliensis	1	Gymnotus carapo	1	Delturus carinotus	1	Geophagus brasiliensis	1	Geophagus brasiliensis	1
Delturus carinotus	6	Gymnotus carapo	1	Hoplias malabaricus	1	Geophagus brasiliensis	1	Gymnotus carapo	1	Gymnotus carapo	1
Geophagus brasiliensis	6	Hoplias malabaricus	1	Hoplosternum littorale	1	Gymnotus carapo	1	Hoplias intermedius	1	Hoplias intermedius	1
Gymnotus carapo	1	Hoplosternum littorale	1	Hypomasticus mormyrops	1	Hoplias intermedius	1	Hoplias malabaricus	1	Hoplias malabaricus	1
Hoplias malabaricus	1	Hypomasticus mormyrops	4	Lophiosilurus alexandri	1	Hoplias malabaricus	1	Hoplosternum littorale	1	Hoplosternum littorale	1
Hoplosternum littorale	6	Hypostomus gymnorhynchus	1		2	Hoplosternum littorale	1	Hypomasticus mormyrops	1	Hypomasticus mormyrops	1
Hypomasticus mormyrops	6	Hypostomus nigromaculatus	1	Oreochromis aureus	1	Hypomasticus mormyrops	1	Hypostomus nigromaculatus	1	Lophiosilurus alexandri	1
Hypostomus gymnorhynchus	1	Leporinus copelandii	3	Phalloceros sp.J	1	Hypostomus nigromaculatus	1	Leporinus copelandii	2	Megaleporinus garmani	1
Hypostomus nigromaculatus	1	Lophiosilurus alexandri	1	Poecilia reticulata	1	Leporinus copelandii	3	Lophiosilurus alexandri	1	Moenkhausia costae	1
Leporinus copelandii	14	Megaleporinus garmani	1	Salminus brasiliensis	1	Lophiosilurus alexandri	2	Megaleporinus garmani	2	Neoplecostominae gen. 2 sp. FFR-2012	1
Lophiosilurus alexandri	5	Moenkhausia costae	3	Trachelyopterus striatulus	1	Megaleporinus garmani	9	Moenkhausia costae	3	Neoplecostomini gen.n. sp.n TEP-2017	1
Megaleporinus garmani	14	Neoplecostominae gen. 2 sp. FFR-2012	1	Trichomycterus sp.	1	Moenkhausia costae	3	Neoplecostominae gen. 2 sp. FFR-2012	2	Oligosarcus argenteus	1
Moenkhausia costae	6	Neoplecostomini gen.n. sp.n TEP-2017	3	Trichomycterus sp.J	1	Neoplecostominae gen. 2 sp. FFR-2012	2	Neoplecostomini gen.n. sp.n TEP-2017	1	Oreochromis aureus	1
Neoplecostominae gen. 2 sp. FFR-2012	3	Oreochromis aureus	1	Wertheimeria maculata	1	Neoplecostomini gen.n. sp.n TEP-2017	1	Oligosarcus argenteus	1	Poecilia reticulata	1
Neoplecostomini gen.n. sp.n TEP-2017	6	Phalloceros sp.J	1			Oligosarcus argenteus	1	Oreochromis aureus	1	Rhamdia quelen	1
Oligosarcus argenteus	3	Poecilia reticulata	1			Oreochromis aureus	1	Phalloceros sp.J	1	Salminus brasiliensis	1
Oreochromis aureus	1	Prochilodus argenteus	2			Phalloceros sp.J	1	Poecilia reticulata	1	Serrasalmus brandtii	1
Phalloceros sp.J	2	Rhamdia quelen	1			Poecilia reticulata	1	Prochilodus argenteus	4	Trachelyopterus striatulus	1
Poecilia reticulata	1	Salminus brasiliensis	2			Prochilodus argenteus	3	Rhamdia quelen	2	Trichomycterus sp.J	1
Prochilodus argenteus	28	Trachelyopterus striatulus	1			Rhamdia quelen	2	Salminus brasiliensis	2	Wertheimeria maculata	1
Rhamdia quelen	1	Trichomycterus sp.	1			Salminus brasiliensis	1	Serrasalmus brandtii	2		
Salminus brasiliensis	5	Trichomycterus sp.J	2			Serrasalmus brandtii	3	Trachelyopterus striatulus	1		
Serrasalmus brandtii	1	Wertheimeria maculata	1			Trachelyopterus striatulus	1	Trichomycterus sp.J	1		
Trachelyopterus striatulus	5					Trichomycterus sp.J	3	Wertheimeria maculata	1		
Trichomycterus sp.	1					Wertheimeria maculata	1				
Trichomycterus sp.J	1										
Wertheimeria maculata	1										

Table S5: Taxa detected in each sampling event and sampling medium.

WATER		SEDIMENT	
Sampling event 1	**Sampling event 2**	**Sampling event 1**	**Sampling event 2**
Astronotus ocellatus	Astronotus ocellatus	Characidium sp.	Astronotus ocellatus
Australoheros facetus	Brycon sp.	Crenicichla lacustris	Brycon sp.
Brycon sp.	Characidium sp.	Cyphocharax gilbert	Characidium sp.
Characidium sp.	Coptodon zillii	Delturus carinotus	Coptodon zillii
Crenicichla lacustris	Crenicichla lacustris	Geophagus brasiliensis	Crenicichla lacustris
Cyphocharax gilbert	Cyphocharax gilbert	Hoplias malabaricus	Cyphocharax gilbert
Cyprinus carpio	Delturus carinotus	Hoplosternum littorale	Delturus carinotus
Delturus carinotus	Geophagus brasiliensis	Hypomasticus mormyrops	Geophagus brasiliensis
Geophagus brasiliensis	Gymnotus carapo	Hypostomus nigromaculatus	Gymnotus carapo
Gymnotus carapo	Hoplias intermedius	Leporinus copelandii	Hoplias intermedius
Hoplias malabaricus	Hoplias malabaricus	Lophiosilurus alexandri	Hoplosternum littorale
Hoplosternum littorale	Hoplosternum littorale	Megaleporinus garmani	Hypomasticus mormyrops
Hypomasticus mormyrops	Hypomasticus mormyrops	Moenkhausia costae	Hypostomus nigromaculatus
Hypostomus gymnorhynchus	Hypostomus nigromaculatus	Neoplecostominae gen. 2 sp. FFR-2012	Leporinus copelandii
Hypostomus nigromaculatus	Leporinus copelandii	Neoplecostomini gen.n. sp.n TEP-2017	Lophiosilurus alexandri
Leporinus copelandii	Lophiosilurus alexandri	Oligosarcus argenteus	Megaleporinus garmani
Lophiosilurus alexandri	Megaleporinus garmani	Oreochromis aureus	Moenkhausia costae
Megaleporinus garmani	Moenkhausia costae	Phalloceros sp.J	Neoplecostominae gen. 2 sp. FFR-2012
Moenkhausia costae	Neoplecostominae gen. 2 sp. FFR-2012	Prochilodus argenteus	Neoplecostomini gen.n. sp.n TEP-2017
Neoplecostominae gen. 2 sp. FFR-2012	Neoplecostomini gen.n. sp.n TEP-2017	Rhamdia quelen	Oligosarcus argenteus
Neoplecostomini gen.n. sp.n TEP-2017	Oligosarcus argenteus	Salminus brasiliensis	Oreochromis aureus
Oligosarcus argenteus	Oreochromis aureus	Serrasalmus brandtii	Phalloceros sp.
Oreochromis aureus	Phalloceros sp.	Trachelyopterus striatulus	Prochilodus argenteus
Phalloceros sp.	Poecilia reticulata	Trichomycterus sp.	Rhamdia quelen
Poecilia reticulata	Prochilodus argenteus	Wertheimeria maculata	Salminus brasiliensis
Prochilodus argenteus	Rhamdia quelen		Serrasalmus brandtii
Rhamdia quelen	Salminus brasiliensis		Trachelyopterus striatulus
Salminus brasiliensis	Serrasalmus brandtii		Trichomycterus sp.
Serrasalmus brandtii	Trachelyopterus striatulus		Wertheimeria maculata
Trachelyopterus striatulus	Trichomycterus sp.J		
Trichomycterus sp.	Wertheimeria maculata		
Trichomycterus sp.			
Wertheimeria maculata			

Table S6: Sample sites including code, city, human population* and GPS coordinates.

Site	City	Population	Coordinates	
1	Mendanha	639	18° 7'15.06"S	43°30'59.16"W
2	Terra Branca	<1000	17°18'48.34''S	43°12'26.61''W
3	Jose Goncalves	4553	16°44'25.89''S	42°34'16.34''W
4	Itacambirucu	15024	16°36'24.00''S	42°49'46.00''W
5	Coronel Murta	9117	16°44'26.85''S	42°34'11.78''W
6	Aracuai	36013	16°51'10.47"S	41°51'33.53"W
7	Jequitinhonha	24131	16°26'16.74"S	41°1'1.45"W
8	Almenara	38755	16° 8'26.20"S	40°35'4.64"W
9	Salto da Divisa	6859	15°59'51.07"S	39°53'29.76"W
10	Itapebi	10495	15°56'57.69"S	39°31'27.08"W
11	Belmonte	21798	15°51'0.02"S	38°52'13.66"W

*Human population census based on IBGE, 2018

Table S7: List of samples sequenced for the custom reference database.

Sample	Species
233	*Astyanax bimaculatus*
242	*Astyanax fasciatus*
226	*Astyanax fasciatus*
6584	Australoeros sp.
6585	Australoeros sp.
1897	*Brycon aff. devillei*
4508	*Brycon ferox*
3745	*Brycon opalinus*
3679	*Brycon opalinus*
3678	*Brycon opalinus*
3062	Brycon sp.
3063	Brycon sp.
4635	Brycon sp. 2
4636	Brycon sp. 2
66	Brycon sp.
1619	Characidium sp.
NAI103	Characidium sp.
3683	*Characidium timbuiense*
294	*Cichlasoma facetum*
293	*Cichlasoma facetum*
1573	Corydoras sp.
3556	*Crenicichla lacustris*
NAI4579	*Crenicichla lacustris*
NAI3555	*Crenicichla lacustris*
3549	*Cyphocharax gilbert*
1810	*Cyphocharax gilbert*
3549	*Cyphocharax gilbert*
1830	*Delturus carinotus*
1832	*Delturus carinotus*
304	*Geophagus brasiliensis*
302	*Geophagus brasiliensis*
303	*Geophagus brasiliensis*
440	*Harttia garavelloi*
441	*Harttia garavelloi*
3703	Hasemania sp.
1164	Hisonotus sp.
104	*Hoplias brasiliensis*
3573	*Hoplias intermedius*
3714	*Hoplias intermedius*
258	*Hoplias malabaricus*
259	*Hoplias malabaricus*

Sample	Species
2644	*Hoplosternum litoralle*
2643	*Hoplosternum litoralle*
1873	*Hypomasticus garmani*
NAI3713	*Hypomasticus mormyrops*
1566_B	*Hypoptopomatinae*
1565	*Hypoptopomatinae*
1566	*Hypoptopomatinae*
4211	*Hypostomus gr. affinis*
4212	*Hypostomus gr. affinis*
3712	Hypostomus sp.
1069	Hypostomus sp.
3544	*Leporinus copelandii*
1823	*Leporinus copelandii*
NAI1823	*Leporinus copelandii*
NAI3033	*Leporinus copelandii*
296	*Leporinus crassilabris*
221	*Leporinus crassilabris*
84	*Leporinus elongatus*
85	*Leporinus elongatus*
52	*Leporinus garmani*
NAI107	*Leporinus garmani*
58	*Leporinus garmani*
93	*Leporinus steindachneri*
2672	*Loricariichthys castaneus*
BB run	*Loricariichthys castaneus*
236	*Moenkhausia costae*
NAI305	*Moenkhausia costae*
1822	Neoplecostominae
4519	*Oligosarcus argenteus*
NAI3547	*Oligosarcus argenteus*
NAI1818	*Oligosarcus argenteus*
4555	*Oligosarcus macrolepis*
1621	Pareiorhaphis sp.
3710	Pareiorhaphis sp.
NAI	Pareiorhaphis sp.
467	*Pareiorhaphis stephanus*
466	*Pareiorhaphis stephanus*
4144	*Phalloceros elachistos*
NAI4144	*Phalloceros elachistos*
NAI3704	Phalloceros sp.
3704	Phalloceros sp.

Sample	Species
1629	Pimelodella sp.
NAI116	*Prochilodus harttii*
3056	Rhamdia cf. jequitinhonha
3056A	Rhamdia cf. jequitinhonha
3055	Rhamdia cf. jequitinhonha
3056	Rhamdia cf. jequitinhonha
NAI101	*Serrasalmus brandtii*
2996	*Steindachneridion amblyurum*
2996	*Steindachneridion amblyurum*
3014	*Steindachneridion amblyurum*
232	*Steindachnerina elegans*
6556	*Trachelyopterus striatulus*
6583	*Trachelyopterus striatulus*
1833	*Trachelyopterus striatulus*
NAI1833	*Trachelyopterus striatulus*
1620	Trichomycterus sp.1
1624	Trichomycterus sp.2
NAI3708	Trichomycterus sp.
5778	Trichomycterus sp.
77	*Wertheimeria maculata*
76	*Wertheimeria maculata*

Chapter V

FINAL CONSIDERATIONS

Chapter V

Final Considerations

5.1 Thesis Summary - Main Findings

Our understanding of Neotropical fish biodiversity has greatly increased in the recent decades. The number of known freshwater fish species is 6,000, and the true estimate that the true number may be over 9,000 (Birindelli & Sidlauskas, 2018). Molecular tools have become widespread and played an important role in understanding the biodiversity in the Neotropics (Mastrochirico filho et al., 2017), and as an example, in the past ˜10 years, studies on DNA barcoding of the Brazilian ichthyofauna have thrived, clarifying previous taxonomic ambiguities and leading to the description of several new species, which expanded the confines of an already high fish diversity (Carvalho et al., 2011; Pereira et al. 2011; Nascimento et al., 2016; Jerep, Camelier & Zanata, 2016; Dutra et al., 2016, Nielsen et. al, 2017; Zawadzki et al., 2016, Pereira et al., 2017).

My studies lie in this context, by exemplifying the role of DNA approaches in improving biodiversity assessment, using both “traditional” DNA barcoding to improve accuracy of taxonomic delimitation, as well as novel metabarcoding methods for simultaneous multi-species assessment from environmental samples. I have provided valuable information for the Doce river ichthyofauna by exploring the fish biodiversity that remains hidden from

traditional survey methods. This drainage has undergone one of the worst environmental disasters ever reported (Fernandes et al., 2016; Neves et al., 2016) and despite being known as a well studied area, I have shown that many of the species inhabiting the river remain undescribed or still unkown. By highlighting the occurrence of cryptic species and providing additional genetic data for 69 species obtained prior to the dam mining burst, we expect that this data can be used as a baseline for describing new species and also for comparison purposes with data obtained following the accident and thus, provide support for management and conservation plans in this basin.

Despite the great advancements made by molecular ecologists in the past decades the knowledge gathered is still far from ideal, and a continued progress is required in the ultimate aim is to accurately describe species and protect them from anthopogenic impacts (Birindelli & Sidlauskas, 2018). Therefore, innovative methods focusing on a fast and realiable identification and description of freshwater communities could contribute to improve biodiversity assessment and speed up the information about species distribution in Brazilian catchments. To this end, I have explored the use of environmental DNA metabarcoding, which is proving to be a powerful bioassessment and biomonitoring tool in a vast array of aquatic environments (marine – Bruyn et al., 2018, Parsons et al., 2018; lakes – Handley et al., 2018; Valdez-Moreno et al., 2018; streams – Hinlo et al., 2018; Li et al., 2018; rivers – Bracken et al., 2018, Pont et al., 2018).

Still, studies on Neotropical freshwater communities are scarce (Cilleros et al., 2018) and to our knowledge, no studies focusing on fish biodiversity assessment in Brazilian waters have been conducted by means of eDNA. Given the countless future opportunities for this method in biomonitoring Neotopical freshwaters, it was important to critically assess a

number of methodological issues, prior to use eDNA data to characterise the fish community of the Jequitinhonha river basin.

First, we compared distinct preservation methods (cooling samples vs adding the cationic surfactant Benzalkonium Chloride), sampling media (water vs sediment) and sampling time (three week interval) and demonstrated that in order to achieve the best yields: i) water samples should be kept at low temperatures to reduce eDNA degradation; ii) multiple sampling collections must be conducted when trying to obtain a full picture of the whole fish community present; iii) sampling media (water vs sediment) might recover similar amount of MOTUs but still provide significant different information due to distinct temporal scales covered.

Subsequently, the evaluation of eDNA as a biodiversity assessment tool allowed the detection of over a hundred MOTUs, though only just over 30 of these could be reliably identified down to the species level. Nevertheless, I noted the occurrence of new records and flagged likely cases of recently introduced species. Furthermore, this method was shown to complement the data obtained through traditional sampling. For instance, eDNA detected in several sites the presence of a species (*Crenicichla* sp.) rarely collected by traditional sampling techniques in this area. Furthermore, two commercially important non-native species were reported for the first time in this basin (*Salminus brasiliensis*, *Lophiosilurus alexandri*), both of which are known as highly invasive species in other catchments (Alves et al., 2007). This highlights the importance of this method in both strengthening knowledge of autoctonous species distributions, as well as monitoning species introductions, with the view of limiting the loss caused by biological invasions.

Regarding the contemporary spatio-temporal variation of fish assemblages inferred by species detected in water samples, my data showed that communities can vary even within short time-frames (e.g. 3 weeks) and while few locations hosted the highest species richness during the first campaign, the distribution followed a much more homogeneous trend towards the second sampling period. Yet, despite the lack of correlation between beta diversity and longitudinal distance or presence of barriers, anthropogenic impacts might still influence fish assemblage's distribution as sites located right near the dams (Irapé and Itapebi) had the lowest species richness. The presence of impoundments impacts the environment by ecologically and physically modifying habitats through habitat loss and changes in the water flow, flood control, dynamic of nutrients, and temperature. Thus, dams are often accounted as a biodiversity reduction factor (Pelicice & Agostinho, 2007; Pompeu et al., 2012).

Species distribution depends on the interplay between ecological and spatial processes. Therefore, several factors may influence species distribution and detection. The Jequitinhonha Valley is a region of dry-climate in Brazil, as and other semi-arid and arid regions, it has an irregular precipitation, in which short rainfall periods can be followed by long dry periods (Leite et al., 2010). An extended and severe drought was recorded for this catchment until February/2017 when the first sampling was conducted, until the increase in the rainfall before the second sampling campaign (2.1-50 mm to 100-250 mm) (CPTEC/INPE, 2018). For aquatic dispersers, river network constraints can influence diversity patterns and a lower water volume might impact on the resources distribution and in the connectivity of habitats; thus, contributing to a higher dissimilarity between fish communities. An increased volume and higher connectivity of habitats might have contributed to a higher connectivity

of aquatic habitats and favoured the dispersal, leading to a higher stability amongst sample sites, regarding the contemporary species richness (inferred through water samples) (Fitzgerald et al., 2017). Thus, this result might suggest that freshwater fish assemblages in tropical habitats may vary significantly between dry and wet seasons. Furthermore, increased rainfall might also influence eDNA particle recovery, as a higher water volume might dilute the DNA in the water column and could explain the lower amount of MOTUs detected in the second campaign compared to the first one.

Regarding eDNA detection, some factors might be taken into account when inferring the data recovered. DNA molecules are known to be transported in varied rates in rivers, from short (10 km -Deiner & Altermatt, 2014) to long distances (more than 100 km - Pont et al., 2018). The transport and subsequent accumulation of eDNA molecules can cause a species richness inflation in downstream locations. However, no correlation between species richness and longitudinal distance was found for the Jequitinhonha river basin and as the sampling sites were located approximately 100 km distant and the transport distances in this catchment had negligible effects on alpha diversity. Yet, these factors (transport and accumulation) should be carefully analysed in this environment, as a species richness overestimation was detected in the last sample site located at the river mouth. Thus, suggesting a spatial accumulation of eDNA molecules and a subsequent resuspension of those.

Notably, one of the main caveats of eDNA studies is imposed by the lack of a complete reference database, hindering the identification and detection of species. However, unfiltered datasets (including the MOTUs not identified at species level) showed a positive

correlation with the filtered dataset and demonstrated that β-diversity is not influenced by the filtering criteria applied as much as sampling medium or sampling time. Thus, corroborating with the results found by Li et al. (2018) and suggesting that this data can be applied for obtaining ecological indices such as α- and β-diversity.

5.2 Challenges and limitations

Despite the significant information provided by DNA barcoding and initial results showing the potential of using eDNA for biomonitoring freshwater fish species in Brazil there is still a lot to accomplish in both fields. DNA barcoding can greatly contribute to the discovery of new species, highlight the occurrence of cryptic species, species complex, historical errors in morphological identification and introduced species. However, DNA barcoding data can only be used to provide a signal for hidden diversity and improvegspecies discovery if an integrative approach is applied. In order to provide a more robust and conclusive result DNA barcoding studies must be aligned with morphological and taxonomic analyses, which are often hindered by the lack of funding, infrastructure and shortage of trained taxonomists, known as “taxonomic impediment” (Birindelli and Sidlauskas, 2018; Dayrat, 2005; Ely et al., 2017; Taylor, 1983).

Furthermore, due to the maternal inheritance of the mitochondria, the COI fragment used for DNA barcoding studies does not allow the identification of hybrids. Thus, the hybridization between species might impose an additional challenge to DNA barcoding studies, as the introgression of mitochondrial DNA can incur in an absence of “barcoding gap” and errouneously lead to misidentifications (Ermakov et al., 2015). When the hybridization process is well documented in the literature the presence of hybrids might be suggested by

comparing mismatches between molecular and morphological identifications (Sales et al., 2018). Otherwise, the occurrence of hybrids may hinder barcode-based species identification.

Environmental DNA studies are still at an early stage and have many challenges and limitations associated with it. Some of the constraints of applying this method in Brazil has also been described for other research areas such as the lack of funding and support for scientific studies (Birindelli & Sidlauskas, 2018). Added to that are the setbacks inherent to the implementation of a relatively newly described approach which lacks standandized protocols and has many unaddressed questions, requiring empirical studies prior to its use. As an example, little is know regarding eDNA fate in neotropical lotic environments. Shedding, degradation, and transport rates are unknown and more conclusive inferences are often hampered by the lack of sufficient basic information.

Here, we found a high species richness in the most downstream sample site which is not expected to harbour higher freshwater fish diversity due to the marine influence in its waters (such as increased salinity). Yet, it is expected that transport and subsequent accumulation of eDNA downstream from its main source (Pont et al., 2018), which could potentially lead to an inflation of species detected in downstream locations. However, this question remains as a conundrum since there are no estimates for the distance travelled by eDNA molecules and its persistence and accumulation in this lotic system.

Another question that remains unanswered is regarding the correlation between the read abundance and species abundance or species biomass (Fonseca et al., 2018). Positive correlations between relative abundances of fish assemblages obtained by eDNA and traditional methods (Pont et al., 2018) and eDNA-biomass were demonstrated (Doi et al., 2015; Maruyama, et al., 2014, Takahara et al., 2012). Yet, Knudsen et al. (2019) have shown

a correlation between eDNA concentrations and and the known species distributions and abundances despite the absence of correlation between eDNA concentration and fish biomass (based on trawling data). Still many studies have questioned the correlation between eDNA and species abundances (Hansen et al., 2018). Hinlo et al. (2018) stressed the challenges of using eDNA data from flowing water to obtain density estimates as the chacteristics of the habitat might have an impact on the results obtained. Thus, using read abundance as a proxy to species abundance should be further investigated prior using read counts for obtaining biodiversity indices.

One of the greatest limitations also relies on the taxonomic assignment. A massive amount of data can be generated by eDNA metabarcoding studies, however, most of the information is lost during the bioinformatics filtering steps. The bottleneck caused by the incompleteness of reference database and the low taxonomic resolution of the fragment currently being used hampers the dectection of many species. If a suitable reference sequence is not present in the database, the sequence remains unassigned and thus is usually not included in the subsequent analyses. Furthermore, the lack of phylogenetic resolution results in many species being grouped together as a single one. My work shows starkly that the majority of species that were molecularly detected could not be biologically assigned to a species, which reduces significantly the impact that these efforts can have.

As previously mentioned, the risk of false positives and negatives should be considered and to obtain reliable data all steps must be meticulously conducted focused on avoiding contaminations and including as many controls as possible.

5.3 Future directions

Since studies using eDNA in the neotropics are rare and the advancement of this field in this region is still far from what has been acomplished in temperate regions of the world, endless opportunities of evaluating this methodology are available. There is still a long way to go in terms of knowledge gathering for the full implementation of eDNA metabarcoding as a biomonitoring tool in Brazil, especially comparing the results obtained by this method with traditional sampling techniques. Furthermore, a great contribution to this field would be if studies were conducted assessing the influence of shedding, degradation and transport rates on eDNA recovery in Neotropical rivers.

Perharps, most importantly, future endeavours should include a significant improvement in the reference database generation, as the majority of the fish diversity found in the Neotropics are endemic (Birindelli and Svalauskas, 2018), and we currently lack reference sequences for many. As previously discussed, and taking into consideration the advances made in high through-put sequencing associated with the reduction of its costs and adressing the problems raised by changing the markers/fragments used in DNA metabarcoding studies, the best option would be to include whole mitogenomes sequences to overcome the limitation imposed by the imcompletness of the genetic database. Furthermore, the data could contribute to phylogenetic investigations, and would be available for several eDNA metabarcoding studies applying different markers or even trying to directly recover the whole mitogenomes from eDNA samples, also avoiding the need of re-sequencing all species when changing the gene regions investigated and overcoming the

problems associated with species identification based on short amplicons (Deiner et al., 2017).

The data provided here demonstrated the potential of applying genetics and metagenetics to provide additional biodiversity information and contribute to the conservation and management of a rapidly changing neotropical freshwater fish community. This project has highlighted the occurence of potential new candidate native species and provided a baseline for further studies focused on the management and conseravation of fish species in Doce river basin, especially after the mining collapse.

Furthermore, the data and results provided herein represent a great advancement in terms of applying molecular tools for freshwater assessment and monitoring in Brazilan freshwaters. Environmental DNA is attracting a growing interest and increasing the ability to acquire data regarding communities composition, species distribution, biological invasions, and much more. However, as with any new emergent methodology it has the need for improvement in many aspects, and despite been widely used in temperate regions its application is still scarce in the Neotropics. Here, the use of eDNA metabarcoding as a biodiversisty assessment tool for Brazilian ichthyofauna was evaluated for the first time and the publication of these results will demonstrate the potential of this method estimulating the emergence of new studies in this field, aiming its optimization and applications as a biomonitoring tool in Neotropical ecosystems.

References

Birindelli, J. L. O., and Sidlauskas, B. L. (2018). Preface: How far has Neotropical Ichthyology progressed in twenty years? *Neotropical Ichthyology,* 16, 1-8. doi:10.1590/1982-0224-20180128.

Bracken, F. S. A., Rooney, S. M., Kelly-Quinn, M., King, J. J., and Carlsson, J. (2018). Identifying spawning sites and other critical habitat in lotic systems using eDNA snapshots. A case study using the sea lamprey *Petromyzon marinus*. *Ecology and Evolution*, 1–15. doi: 10.1002/ece3.4777.

Carvalho, D. C., Oliveira, D. A. A., Pompeu, P. S., Leal, C. G., Oliveira, C., and Hanner, R. (2011). Deep barcode divergence in Brazilian freshwater fishes: the case of the São Francisco River basin. *Mitochondrial DNA*, 22, 80-86. doi: 10.3109/19401736.2011.588214

Cilleros, K., Valentini, A., Allard, L., Dejean, T., Etienne, R., Grenouillet, G., et al. (2018). Unlocking biodiversity and conservation studies in high diversity environments using environmental DNA (eDNA): a test with Guianese freshwater fishes. *Molecular Ecology Resources*, 1–20. doi:10.1111/1755-0998.12900.

Dayrat, B. (2005). Towards integrative taxonomy. *Biological Journal of the Linnean Society*, 85, 407-415. doi:10.1111/j.1095-8312.2005.00503.x

Pont, D., Rocle, M., Valentini, A., Civade, R., Jean, P., Maire, A., et al. (2018). Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. *Scientific Reports*, 8, 10361. doi:10.1038/s41598-018-28424-8.

Doi, H., Uchii, K., Takahara, T., Matsuhashi, S., Yamanaka, H., and Minamoto, T. (2015). Use of droplet digital PCR for estimation of fish abundance and biomass in environmental DNA surveys. *PLoS ONE*, 10(3), 1–11. doi:10.1371/journal.pone.0122763

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

Ely, C. V., Bordignon, S. A. L., Trevisan, R., and Boldrini, I. I. (2017). Implications of poor taxonomy in conservation. *Journal of Nature Conservation,* 36, 10-13. doi: 10.1016/j.jnc.2017.01.003

Fernandes, G. W., Goulart, F. F., Ranieri, B. D., Coelho, M. S., Dales, K., Boesche, N., et al. (2016). Deep into the mud: ecological and socio-economic impacts of the dam breach in Mariana, Brazil. *Natureza & Conservação,* 14, 35-45. doi: 10.1016/j.ncon.2016.10.003

Fonseca, V. G. (2018). Pitfalls in relative abundance estimation using eDNA metabarcoding. *Molecular Ecology Resources,* 18, 923–926. doi:10.1111/1755-0998.12902.

Handley, L. L., Read, D. S., Winfield, I. J., Kimbell, H., Li, J., Hahn, C., et al. (2018). Temporal and spatial variation in distribution of fish environmental DNA in England ’s largest lake. *bioRxiv*. doi: 10.1101/376400

Hansen, B. K., Bekkevold, D., Clausen, L. W., and Nielsen, E. E. (2018). The sceptical optimist: Challenges and perspectives for the application of environmental DNA in marine fisheries. *Fish and Fisheries*, 1–18. doi:10.1111/faf.12286.

Hinlo, R., Lintermans, M., Gleeson, D., Broadhurst, B., and Furlan, E. (2018). Performance of eDNA assays to detect and quantify an elusive benthic fish in upland streams. *Biological Invasions*, 20, 1–15. doi:10.1007/s10530-018-1760-x.

Holman, L. E., de Bruyn, M., Creer, S., Carvalho, G., Robidart, J., and Rius, M. (2018). The detection of novel and resident marine non-indigenous species using environmental DNA metabarcoding of seawater and sediment. *bioRxiv*, 440768. doi: 10.1101/440768

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

Li, Y., Evans, N. T., Renshaw, M. A., Jerde, C. L., Olds, B. P., Shogren, A. J., et al. (2018). Estimating fish alpha- and beta-diversity along a small stream with environmental DNA metabarcoding. *Metabarcoding and Metagenomics*, 1–11. doi:10.3897/mbmg.2.24262.

Maruyama, A., Nakamura, K., Yamanaka, H., Kondoh, M., and Minamoto, T. (2014). The release rate of environmental DNA from juvenile and adult fish. *PLoS ONE*, 9(12), 1–13. doi: 10.1371/journal. pone.0114639

Mastrochirico-Filho, V. A., Freitas, M. V., Ariede, R. B., Lira, L. V. G., Mendes, N. J., and Hashimoto D. T. (2017). Genetic applications in the conservations of Neotropical freshwater fish. In Biological Resources of Water. IntechOpen.

Nascimento, M. H. S., Almeida, M. S., Veira, M. N. S., Limeira-Filho, D., Lima, R. C, Barros, M. C., and Fraga, E. C. (2016). DNA barcoding reveals high levels of genetic diversity in the fishes of the Itapecuru Basin in Maranhão, Brazil. *Genetics and Molecular Research* 15, 3, 1-11. doi: 10.4238/gmr.15038476

Neves, A. C. O., Nunes, F.P., Carvalho, F.A., Fernandes, G. W. (2016). Neglect of ecosystems services by mining, and the worst environmental disaster in Brazil. *Natureza & Conservação*, 14, 24-27. doi: 10.1016/j.ncon.2016.03.002

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

Parsons, K. M., Everett, M., Dahlheim, M., and Park, L. (2018). Water, water everywhere: environmental DNA can unlock population structure in elusive marine species. *Royal Society Open Science,* 5(8). doi:10.1098/rsos.180537.

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

Pereira, L. H., Pazian, M. F., Hanner, R., Foresti, F., and Oliveira, C. (2011). DNA barcoding reveals hidden diversity in the Neotropical freshwater fish *Piabina argentea* (Characiformes:

Characidae) from the Upper Parana Basin of Brazil. *Mitochondrial DNA,* 22, 87-96. doi: 10.3109/19401736.2011.588213

Pereira, E. H. L., Pessali, T. C., Andrade, F., and Reis, R. E. (2017.) Description of a new species of *Pareiorhaphis* (Loricariidae: Neoplecostominae) from the rio Jequitinhonha basin, Minas Gerais, eastern Brazil. *Neotropical Ichthyology*, 15(3). doi: 10.1590/1982-0224-20170007.

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

Pont, D., Rocle, M., Valentini, A., Civade, R., Jean, P., Maire, A., et al. (2018). Environmental DNA reveals quantitative patterns of fish biodiversity in large rivers despite its downstream transportation. *Scientific Reports*, 8, 10361. doi:10.1038/s41598-018-28424-8.

Takahara, T., Minamoto, T., Yamanaka, H., Doi, H., and Kawabata, Z. (2012). Estimation of fish biomass using environmental DNA. *PLoS ONE*, 7(4), 1–8. doi: 10.1371/journal.pone.0035868

Valdez-Moreno, M., Ivanova, N. V., Elias-Gutierrez, M., Pedersen, S. L., Bessonov, K., and Hebert, P. D. N. (2018). Using eDNA to biomonitor the fish community in a tropical oligotrophic lake. *bioRxiv*, 375089. doi:10.1101/375089.

Zawadzki, C. H., Carvalho, P. H., Birindelli, J. O., and Azevedo, F. M. (2016). *Hypostomus nigrolineatus*, a new dark-striped species from the rio Jequitinhonha and rio Pardo basins, Brazil (Siluriformes, Loricariidae). *Ichthyological Exploration of Freshwaters*, 27(3), 263-274

APPENDIX

Table A1: List of all MOTUs recovered for each treatment.

BAC 1				
rank	best_identity	order_name	family_name	taxid
species	1	Characiformes	Prochilodontidae	27695
species	1	Characiformes	Serrasalmidae	371937
species	0.976470588	Characiformes	Crenuchidae	1569729
species	1	Characiformes	Anostomidae	1925008
species	0.958823529	Characiformes	Anostomidae	1925008
species	0.976608187	Characiformes	Prochilodontidae	27695
species	1	Cyprinodontiformes	Poeciliidae	10000002
species	0.965116279	Characiformes	Crenuchidae	1569729
species	1	Characiformes	Characidae	1604128
species	0.947368421	Characiformes	Prochilodontidae	27695
species	0.936416185	Characiformes	Serrasalmidae	371937
species	0.976608187	Characiformes	Serrasalmidae	371937
species	0.947674419	Characiformes	Crenuchidae	1569729
species	0.930232558	Characiformes	Prochilodontidae	27695
species	0.976744186	Siluriformes	Loricariidae	1191588
species	0.953757225	Characiformes	Prochilodontidae	27695
species	0.970588235	Characiformes	Anostomidae	1925008
species	0.971098266	Characiformes	Anostomidae	749479
species	0.988165681	Cichliformes	Cichlidae	81339
species	0.994047619	Cichliformes	Cichlidae	47969
species	0.959302326	Characiformes	Prochilodontidae	27695
species	1	Cypriniformes	Cyprinidae	7962
suborder	0.976190476	Characiformes		1489739
species	1	Characiformes	Crenuchidae	1569729
species	0.929411765	Characiformes	Erythrinidae	27720
species	0.940828402	Characiformes	Erythrinidae	27720
species	0.988023952	Characiformes	Erythrinidae	27720
species	0.900584795	Characiformes	Characidae	1123824
species	0.936781609	Characiformes	Crenuchidae	1569729
species	0.976744186	Gymnotiformes	Gymnotidae	94172
species	0.925287356	Characiformes	Curimatidae	984680
species	0.959064328	Characiformes	Bryconidae	202829
species	0.976608187	Siluriformes	Loricariidae	168219
genus	0.959302326	Characiformes	Prochilodontidae	27694
species	0.923976608	Characiformes	Serrasalmidae	371937
species	0.941520468	Characiformes	Crenuchidae	1569729
species	0.953216374	Characiformes	Crenuchidae	1569729
species	1	Cyprinodontiformes	Poeciliidae	10000004
species	0.959537572	Characiformes	Anostomidae	749479
genus	0.947976879	Characiformes	Prochilodontidae	27694
species	0.941520468	Characiformes	Crenuchidae	1569729
species	0.959770115	Characiformes	Anostomidae	749479
species	0.935294118	Characiformes	Crenuchidae	1569729
species	0.971428571	Characiformes	Curimatidae	984680
species	0.941860465	Characiformes	Crenuchidae	1569729
superorder	0.83908046			186628
genus	0.959302326	Characiformes	Prochilodontidae	27694
species	0.970930233	Characiformes	Serrasalmidae	371937

BAC 1				
rank	best_identity	order_name	family_name	taxid
genus	0.936781609	Characiformes	Prochilodontidae	27694
family	0.941860465	Siluriformes	Loricariidae	31002
species	1	Characiformes	Characidae	1123824
superorder	0.830508475			186628
species	0.959537572	Characiformes	Prochilodontidae	27695
species	0.976744186	Characiformes	Prochilodontidae	27695
species	0.970760234	Characiformes	Anostomidae	1925008
species	0.954022989	Characiformes	Curimatidae	984680
species	0.931428571	Siluriformes	Trichomycteridae	245746
species	0.935672515	Characiformes	Erythrinidae	1053573
species	0.965317919	Characiformes	Curimatidae	984680
species	0.913294798	Characiformes	Crenuchidae	1569729
species	0.948571429	Characiformes	Curimatidae	984680
genus	0.906432749	Characiformes	Anostomidae	1925018
species	0.970588235	Characiformes	Serrasalmidae	371937
species	0.971098266	Characiformes	Curimatidae	984680
species	0.953216374	Characiformes	Prochilodontidae	27695
genus	0.923076923	Characiformes	Bryconidae	42585
species	0.970588235	Characiformes	Anostomidae	1925008
genus	0.947368421	Characiformes	Prochilodontidae	27694
family	0.884393064	Siluriformes	Loricariidae	31002
species	0.965517241	Characiformes	Curimatidae	984680
genus	0.919075145	Characiformes	Anostomidae	1925018
genus	0.964705882	Characiformes	Anostomidae	1925018
species	0.964705882	Characiformes	Serrasalmidae	371937
species	0.941176471	Characiformes	Serrasalmidae	371937
species	1	Cichliformes	Cichlidae	8128
genus	0.953757225	Characiformes	Prochilodontidae	27694
species	0.907514451	Characiformes	Erythrinidae	27720
species	1	Cyprinodontiformes	Poeciliidae	8081
family	0.936416185	Siluriformes	Loricariidae	31002
species	0.959537572	Characiformes	Anostomidae	749479
species	0.931428571	Characiformes	Prochilodontidae	27695
species	0.953757225	Characiformes	Crenuchidae	1569729
genus	0.954285714	Characiformes	Prochilodontidae	27694
species	0.941520468	Characiformes	Anostomidae	1925003
species	0.936781609	Siluriformes	Pseudopimelodidae	490146
species	0.900584795	Characiformes	Erythrinidae	27720
species	0.930232558	Characiformes	Crenuchidae	1569729
species	0.970238095	Cyprinodontiformes	Poeciliidae	10000004
species	0.942528736	Characiformes	Crenuchidae	1569729
species	0.959064328	Characiformes	Serrasalmidae	371937
genus	0.930635838	Characiformes	Prochilodontidae	27694
species	0.888888889	Characiformes	Serrasalmidae	371937
species	0.948863636	Siluriformes	Pseudopimelodidae	490146
species	0.964497041	Characiformes	Bryconidae	202829
family	0.959537572	Siluriformes	Loricariidae	31002
family	0.897142857	Characiformes	Curimatidae	30721

BAC 1				
rank	best_identity	order_name	family_name	taxid
suborder	0.92	Siluriformes		1489793
genus	0.965116279	Characiformes	Prochilodontidae	27694
species	0.925287356	Siluriformes	Pimelodidae	162147
genus	0.919075145	Characiformes	Prochilodontidae	27694
species	0.96	Characiformes	Prochilodontidae	27695
species	0.959770115	Characiformes	Prochilodontidae	27695
genus	0.953757225	Characiformes	Prochilodontidae	27694
species	0.948275862	Characiformes	Prochilodontidae	27695
family	0.901162791	Characiformes	Serrasalmidae	42495
genus	0.970930233	Characiformes	Prochilodontidae	27694
species	0.948863636	Characiformes	Serrasalmidae	371937
family	0.913294798	Siluriformes	Loricariidae	31002
genus	0.942857143	Characiformes	Prochilodontidae	27694
family	0.947058824	Characiformes	Characidae	7992
family	0.876470588	Characiformes	Anostomidae	7986
species	0.976608187	Siluriformes	Loricariidae	163993
species	0.948571429	Siluriformes	Pimelodidae	162147
genus	0.924418605	Characiformes	Prochilodontidae	27694
family	0.935294118	Characiformes	Characidae	7992
species	0.885245902	Characiformes	Crenuchidae	1569729
species	0.965116279	Characiformes	Anostomidae	1925003
species	0.929411765	Characiformes	Crenuchidae	1569729
species	0.913294798	Characiformes	Prochilodontidae	27695
species	0.942196532	Characiformes	Anostomidae	1925003
species	0.948275862	Siluriformes	Pseudopimelodidae	490146
species	0.895953757	Characiformes	Crenuchidae	1569729
species	0.900584795	Characiformes	Crenuchidae	1569729
species	0.959064328	Characiformes	Crenuchidae	1569729
family	0.936046512	Siluriformes	Loricariidae	31002
species	0.926136364	Characiformes	Prochilodontidae	27695
genus	0.906976744	Characiformes	Prochilodontidae	27694
species	0.954022989	Siluriformes	Pseudopimelodidae	490146
species	0.964912281	Characiformes	Bryconidae	930266
species	0.920454546	Characiformes	Crenuchidae	1569729
species	0.971098266	Characiformes	Anostomidae	1925003
species	0.948275862	Characiformes	Curimatidae	984680
genus	0.959302326	Characiformes	Prochilodontidae	27694
genus	0.959302326	Characiformes	Prochilodontidae	27694
species	0.959064328	Characiformes	Anostomidae	1925003
species	0.971264368	Characiformes	Curimatidae	984680
species	0.925287356	Siluriformes	Pimelodidae	162147
genus	0.954545455	Siluriformes	Clariidae	13012
species	0.902857143	Characiformes	Crenuchidae	1569729
family	0.952380952	Characiformes	Characidae	7992
species	0.942857143	Siluriformes	Loricariidae	745493
species	0.935672515	Characiformes	Crenuchidae	1569729
species	0.802139037	Characiformes	Serrasalmidae	371937
species	0.942196532	Characiformes	Prochilodontidae	27695

BAC 1				
rank	best_identity	order_name	family_name	taxid
species	0.954022989	Characiformes	Bryconidae	202829
species	0.880681818	Characiformes	Erythrinidae	27720
species	0.912790698	Characiformes	Crenuchidae	1569729
genus	0.941176471	Characiformes	Anostomidae	1925018
species	1	Cichliformes	Cichlidae	50733
species	0.954022989	Siluriformes	Pimelodidae	162147
species	0.953488372	Characiformes	Crenuchidae	1569729
genus	0.912790698	Characiformes	Anostomidae	1925018
species	0.948275862	Characiformes	Prochilodontidae	27695
species	0.959537572	Characiformes	Prochilodontidae	27695
suborder	0.959064328	Characiformes		1489739
species	0.965714286	Characiformes	Prochilodontidae	27695
species	0.954022989	Characiformes	Prochilodontidae	27695
species	0.954285714	Characiformes	Prochilodontidae	27695
species	0.947368421	Characiformes	Crenuchidae	1569729
suborder	0.9375	Siluriformes		1489793
genus	0.925714286	Characiformes	Characidae	681908
suborder	0.934911243	Characiformes		1489739
species	0.913294798	Characiformes	Crenuchidae	1569729
species	0.918128655	Characiformes	Crenuchidae	1569729
species	0.954285714	Characiformes	Curimatidae	984680
species	0.947674419	Characiformes	Serrasalmidae	371937
genus	0.959302326	Characiformes	Prochilodontidae	27694
species	0.882681564	Characiformes	Erythrinidae	27720
species	0.891428571	Characiformes	Crenuchidae	1569729
genus	0.941860465	Characiformes	Prochilodontidae	27694
family	0.941176471	Characiformes	Characidae	7992
species	0.936046512	Characiformes	Prochilodontidae	27695
family	0.931034483	Siluriformes	Pseudopimelodidae	255574
species	0.96	Characiformes	Prochilodontidae	27695
species	0.952662722	Characiformes	Bryconidae	202829
species	0.970588235	Characiformes	Bryconidae	202829
species	0.971098266	Characiformes	Anostomidae	749479
species	0.959064328	Characiformes	Anostomidae	1925003
genus	0.947976879	Characiformes	Prochilodontidae	27694
species	0.949152542	Siluriformes	Pimelodidae	162147
species	0.941176471	Siluriformes	Trichomycteridae	10000003
genus	0.959064328	Characiformes	Anostomidae	1925018
family	0.936046512	Siluriformes	Loricariidae	31002
suborder	0.893854749	Siluriformes		1489793
species	0.947976879	Characiformes	Prochilodontidae	27695
species	0.96	Siluriformes	Loricariidae	745493
species	0.964705882	Characiformes	Erythrinidae	1053573
species	0.912280702	Characiformes	Serrasalmidae	371937
species	0.918128655	Characiformes	Erythrinidae	27720
species	0.970414201	Cichliformes	Cichlidae	81339
family	0.892045455	Siluriformes	Loricariidae	31002
species	0.901734104	Characiformes	Erythrinidae	27720

BAC 2				
rank	best_identity	order_name	family_name	taxid
species	0.970588235	Characiformes	Anostomidae	1925008
species	0.959064328	Characiformes	Bryconidae	202829
species	0.970930233	Characiformes	Serrasalmidae	371937
genus	0.906432749	Characiformes	Anostomidae	1925018
suborder	0.959064328	Characiformes		1489739
species	0.965517241	Characiformes	Curimatidae	984680
genus	0.919075145	Characiformes	Anostomidae	1925018
genus	0.953488372	Characiformes	Anostomidae	1925018
genus	0.953757225	Characiformes	Prochilodontidae	27694
species	0.959064328	Characiformes	Crenuchidae	1569729
species	0.948275862	Characiformes	Prochilodontidae	27695
species	0.947674419	Characiformes	Serrasalmidae	371937
species	0.829411765	Characiformes	Serrasalmidae	371937
species	0.930232558	Characiformes	Crenuchidae	1569729
genus	0.903409091	Characiformes	Characidae	681908
genus	0.965116279	Characiformes	Prochilodontidae	27694
species	0.959770115	Characiformes	Prochilodontidae	27695
species	0.970238095	Characiformes	Bryconidae	126315
species	0.941860465	Characiformes	Crenuchidae	1569729
species	0.970588235	Characiformes	Anostomidae	1925008
genus	0.947368421	Characiformes	Prochilodontidae	27694
genus	0.919075145	Characiformes	Anostomidae	1925018
species	0.913294798	Characiformes	Crenuchidae	1569729
genus	0.924855491	Characiformes	Characidae	681908
species	0.941176471	Characiformes	Serrasalmidae	371937
genus	0.906976744	Characiformes	Prochilodontidae	27694
genus	0.959302326	Characiformes	Prochilodontidae	27694
species	0.936046512	Characiformes	Prochilodontidae	27695
genus	0.952941177	Characiformes	Anostomidae	1925018
species	0.953216374	Characiformes	Prochilodontidae	27695
species	0.931428571	Characiformes	Prochilodontidae	27695
species	0.965714286	Characiformes	Prochilodontidae	27695
genus	0.919075145	Characiformes	Characidae	681908
species	0.947674419	Characiformes	Crenuchidae	1569729
species	0.959302326	Characiformes	Prochilodontidae	27695
genus	0.953757225	Characiformes	Anostomidae	1925018
species	0.914285714	Characiformes	Prochilodontidae	27695
species	0.953757225	Characiformes	Prochilodontidae	27695
genus	0.959302326	Characiformes	Prochilodontidae	27694
species	1	Cypriniformes	Cyprinidae	7962
species	0.953216374	Characiformes	Crenuchidae	1569729
species	0.891428571	Characiformes	Crenuchidae	1569729
species	1	Siluriformes	Auchenipteridae	984686
family	0.919075145	Siluriformes	Loricariidae	31002
species	0.976608187	Characiformes	Serrasalmidae	371937
species	0.941520468	Characiformes	Crenuchidae	1569729
species	0.994252874	Siluriformes	Pimelodidae	162147
species	0.941520468	Characiformes	Crenuchidae	1569729

BAC 2				
rank	best_identity	order_name	family_name	taxid
superorder	0.83908046			186628
species	1	Siluriformes	Callichthyidae	114109
species	0.947368421	Characiformes	Prochilodontidae	27695
species	0.976608187	Characiformes	Prochilodontidae	27695
suborder	0.976190476	Characiformes		1489739
species	1	Cichliformes	Cichlidae	870598
superorder	0.831460674			186628
species	0.994011976	Characiformes	Erythrinidae	1053573
species	0.988165681	Cichliformes	Cichlidae	81339
species	0.965116279	Characiformes	Crenuchidae	1569729
species	0.994047619	Cichliformes	Cichlidae	47969
species	0.976744186	Siluriformes	Loricariidae	1191588
family	0.959302326	Siluriformes	Loricariidae	31002
genus	0.971264368	Siluriformes	Loricariidae	503149
species	1	Cyprinodontiformes	Poeciliidae	10000002
species	0.988372093	Siluriformes	Loricariidae	1191588
species	1	Characiformes	Characidae	1604128
family	0.942196532	Siluriformes	Loricariidae	31002
species	0.931818182	Siluriformes	Trichomycteridae	245746
suborder	1	Characiformes		1489739
species	0.976744186	Gymnotiformes	Gymnotidae	94172
species	0.976608187	Siluriformes	Loricariidae	163993
species	0.988023952	Characiformes	Curimatidae	984680
species	1	Characiformes	Characidae	1123824
species	1	Characiformes	Bryconidae	202829
species	0.900584795	Characiformes	Characidae	1123824
species	1	Characiformes	Anostomidae	1925008
species	1	Siluriformes	Trichomycteridae	10000003
species	1	Characiformes	Crenuchidae	1569729
genus	0.942196532	Characiformes	Characidae	681908
species	1	Characiformes	Serrasalmidae	371937

ICE 1				
rank	best_identity	order_name	family_name	taxid
species	1	Characiformes	Characidae	1123824
species	0.970588235	Characiformes	Anostomidae	1925008
species	0.971264368	Characiformes	Prochilodontidae	27695
genus	0.930635838	Characiformes	Prochilodontidae	27694
genus	0.947368421	Characiformes	Prochilodontidae	27694
species	0.942196532	Characiformes	Anostomidae	1925003
species	0.941520468	Characiformes	Anostomidae	1925003
species	0.959064328	Characiformes	Bryconidae	202829
species	0.900584795	Characiformes	Erythrinidae	27720
species	0.959064328	Characiformes	Crenuchidae	1569729
species	0.953757225	Characiformes	Prochilodontidae	27695
genus	0.959302326	Characiformes	Prochilodontidae	27694
species	0.918128655	Characiformes	Erythrinidae	27720
genus	0.959302326	Characiformes	Prochilodontidae	27694
genus	0.959302326	Characiformes	Prochilodontidae	27694
genus	0.941860465	Characiformes	Prochilodontidae	27694
species	0.925287356	Siluriformes	Pimelodidae	162147
species	0.958823529	Characiformes	Serrasalmidae	371937
species	0.90960452	Characiformes	Crenuchidae	1569729
species	0.942528736	Characiformes	Crenuchidae	1569729
genus	0.906976744	Characiformes	Prochilodontidae	27694
genus	0.954545455	Siluriformes	Clariidae	13012
species	0.949152542	Characiformes	Prochilodontidae	27695
species	0.942857143	Siluriformes	Loricariidae	745493
suborder	0.893854749	Siluriformes		1489793
species	0.802139037	Characiformes	Serrasalmidae	371937
genus	0.965116279	Characiformes	Prochilodontidae	27694
species	0.914285714	Siluriformes	Loricariidae	745493
species	0.964705882	Characiformes	Erythrinidae	1053573
family	0.947058824	Characiformes	Characidae	7992
species	0.912280702	Characiformes	Serrasalmidae	371937
species	0.918128655	Characiformes	Erythrinidae	27720
genus	0.924418605	Characiformes	Anostomidae	1925018
family	0.876470588	Characiformes	Anostomidae	7986
species	0.880681818	Characiformes	Erythrinidae	27720
subfamily	0.925287356	Siluriformes	Loricariidae	503143
genus	0.941176471	Characiformes	Anostomidae	1925018
genus	0.964705882	Characiformes	Anostomidae	1925018
species	0.953216374	Characiformes	Prochilodontidae	27695
species	0.931428571	Characiformes	Prochilodontidae	27695
species	0.913294798	Characiformes	Crenuchidae	1569729
genus	0.923076923	Characiformes	Bryconidae	42585
species	0.954285714	Characiformes	Prochilodontidae	27695
species	0.925287356	Siluriformes	Pimelodidae	162147
species	0.953216374	Characiformes	Bryconidae	930266
species	0.959064328	Characiformes	Prochilodontidae	27695
genus	0.902298851	Characiformes	Bryconidae	42585
species	0.920454546	Characiformes	Crenuchidae	1569729

ICE 1				
rank	best_identity	order_name	family_name	taxid
family	0.88	Siluriformes	Loricariidae	31002
species	0.970238095	Cyprinodontiformes	Poeciliidae	10000004
species	0.952662722	Characiformes	Bryconidae	202829
species	0.970588235	Characiformes	Bryconidae	202829
family	0.913294798	Siluriformes	Loricariidae	31002
species	0.959064328	Characiformes	Anostomidae	1925003
family	0.952380952	Characiformes	Characidae	7992
species	0.949152542	Siluriformes	Pimelodidae	162147
species	0.941176471	Siluriformes	Trichomycteridae	10000003
family	0.942196532	Siluriformes	Loricariidae	31002
species	0.96	Siluriformes	Loricariidae	745493
species	0.943181818	Characiformes	Crenuchidae	1569729
species	0.970414201	Cichliformes	Cichlidae	81339
species	0.942196532	Characiformes	Prochilodontidae	27695
family	0.892045455	Siluriformes	Loricariidae	31002
genus	0.960227273	Characiformes	Prochilodontidae	27694
species	0.954022989	Characiformes	Bryconidae	202829
family	0.919075145	Siluriformes	Loricariidae	31002
species	0.941176471	Characiformes	Erythrinidae	27720
species	0.959770115	Characiformes	Prochilodontidae	27695
species	0.941860465	Characiformes	Crenuchidae	1569729
species	0.907514451	Characiformes	Erythrinidae	27720
species	0.971098266	Characiformes	Curimatidae	984680
family	0.897142857	Characiformes	Curimatidae	30721
genus	0.954285714	Characiformes	Prochilodontidae	27694
suborder	0.92	Siluriformes		1489793
species	0.96	Characiformes	Prochilodontidae	27695
species	0.959770115	Characiformes	Prochilodontidae	27695
genus	0.953757225	Characiformes	Prochilodontidae	27694
species	0.918604651	Characiformes	Bryconidae	930266
family	0.936046512	Siluriformes	Loricariidae	31002
species	0.954285714	Characiformes	Curimatidae	984680
family	0.901162791	Characiformes	Serrasalmidae	42495
species	0.947674419	Characiformes	Serrasalmidae	371937
species	0.954022989	Siluriformes	Pseudopimelodidae	490146
species	0.930232558	Characiformes	Crenuchidae	1569729
species	0.959064328	Characiformes	Anostomidae	1925003
species	0.947976879	Characiformes	Anostomidae	749479
family	0.931034483	Siluriformes	Pseudopimelodidae	255574
species	0.965517241	Characiformes	Curimatidae	984680
genus	0.944751381	Characiformes	Prochilodontidae	27694
species	0.902857143	Characiformes	Crenuchidae	1569729
family	0.936046512	Siluriformes	Loricariidae	31002
family	0.913793103	Siluriformes	Loricariidae	31002
species	0.918604651	Characiformes	Crenuchidae	1569729
subfamily	0.925287356	Siluriformes	Loricariidae	503143
species	0.959537572	Characiformes	Anostomidae	1925003
species	0.924855491	Characiformes	Crenuchidae	1569729

ICE 1				
rank	best_identity	order_name	family_name	taxid
species	0.923529412	Characiformes	Erythrinidae	1053573
family	0.884393064	Siluriformes	Loricariidae	31002
species	0.964912281	Characiformes	Bryconidae	202829
species	0.931428571	Siluriformes	Pimelodidae	162147
species	0.948571429	Siluriformes	Pseudopimelodidae	490146
species	0.926136364	Characiformes	Curimatidae	984680
species	0.935294118	Characiformes	Crenuchidae	1569729
species	0.970760234	Characiformes	Anostomidae	1925008
species	0.959770115	Characiformes	Anostomidae	749479
species	0.885245902	Characiformes	Crenuchidae	1569729
species	0.964497041	Characiformes	Bryconidae	202829
species	0.965116279	Characiformes	Anostomidae	1925003
species	0.914285714	Characiformes	Prochilodontidae	27695
species	0.913294798	Characiformes	Prochilodontidae	27695
species	0.953757225	Characiformes	Crenuchidae	1569729
genus	0.919075145	Characiformes	Prochilodontidae	27694
suborder	0.9375	Siluriformes		1489793
species	0.948275862	Siluriformes	Pseudopimelodidae	490146
species	0.900584795	Characiformes	Crenuchidae	1569729
species	0.964912281	Characiformes	Bryconidae	930266
species	0.948863636	Characiformes	Serrasalmidae	371937
species	0.971098266	Characiformes	Anostomidae	1925003
species	0.96	Characiformes	Prochilodontidae	27695
genus	0.947976879	Characiformes	Prochilodontidae	27694
species	0.925287356	Characiformes	Crenuchidae	1569729
species	0.953488372	Characiformes	Anostomidae	749479
species	0.965909091	Cichliformes	Cichlidae	870598
species	0.994047619	Cichliformes	Cichlidae	47969
species	0.988023952	Characiformes	Erythrinidae	27720
genus	0.919075145	Characiformes	Anostomidae	1925018
species	0.929411765	Characiformes	Crenuchidae	1569729
genus	0.965116279	Characiformes	Prochilodontidae	27694
species	0.895953757	Characiformes	Crenuchidae	1569729
species	0.942528736	Siluriformes	Pimelodidae	162147
species	0.936781609	Siluriformes	Pseudopimelodidae	490146
species	0.913294798	Characiformes	Crenuchidae	1569729
species	0.926136364	Characiformes	Prochilodontidae	27695
genus	0.906976744	Characiformes	Prochilodontidae	27694
species	0.882681564	Characiformes	Erythrinidae	27720
species	0.948275862	Characiformes	Curimatidae	984680
family	0.925287356	Siluriformes	Loricariidae	31002
species	0.947674419	Characiformes	Prochilodontidae	27695
species	0.959302326	Characiformes	Prochilodontidae	27695
species	0.941520468	Characiformes	Crenuchidae	1569729
species	0.964705882	Characiformes	Serrasalmidae	371937
genus	0.897142857	Characiformes	Prochilodontidae	27694
species	0.935294118	Characiformes	Erythrinidae	1053573
suborder	0.964705882	Characiformes		1489739

ICE 1				
rank	best_identity	order_name	family_name	taxid
species	0.971264368	Characiformes	Curimatidae	984680
family	0.941176471	Characiformes	Characidae	7992
species	0.970930233	Characiformes	Anostomidae	749479
species	0.959537572	Characiformes	Curimatidae	984680
species	0.941520468	Characiformes	Crenuchidae	1569729
species	0.948275862	Characiformes	Prochilodontidae	27695
suborder	0.959064328	Characiformes		1489739
species	1	Cypriniformes	Cyprinidae	7962
species	0.934911243	Characiformes	Erythrinidae	27720
suborder	0.959064328	Characiformes		1489739
species	0.888888889	Characiformes	Serrasalmidae	371937
species	0.954022989	Siluriformes	Pimelodidae	162147
species	0.947368421	Characiformes	Crenuchidae	1569729
species	0.976744186	Gymnotiformes	Gymnotidae	94172
genus	0.959302326	Characiformes	Prochilodontidae	27694
species	0.959537572	Characiformes	Prochilodontidae	27695
species	1	Cichliformes	Cichlidae	50733
species	0.953488372	Characiformes	Crenuchidae	1569729
species	0.954022989	Characiformes	Curimatidae	984680
species	0.959537572	Characiformes	Prochilodontidae	27695
genus	0.906432749	Characiformes	Anostomidae	1925018
genus	0.953757225	Characiformes	Prochilodontidae	27694
species	0.965517241	Characiformes	Curimatidae	984680
species	0.935672515	Characiformes	Erythrinidae	1053573
species	0.965317919	Characiformes	Curimatidae	984680
species	0.948863636	Siluriformes	Pseudopimelodidae	490146
suborder	0.907514451	Siluriformes		1489793
species	0.976744186	Characiformes	Prochilodontidae	27695
family	0.959537572	Siluriformes	Loricariidae	31002
family	0.936416185	Siluriformes	Loricariidae	31002
species	0.959537572	Characiformes	Anostomidae	749479
family	0.935294118	Characiformes	Characidae	7992
species	0.970930233	Characiformes	Serrasalmidae	371937
species	0.953216374	Characiformes	Crenuchidae	1569729
family	0.884393064	Siluriformes	Loricariidae	31002
species	0.959537572	Characiformes	Anostomidae	749479
species	0.948571429	Characiformes	Curimatidae	984680
genus	0.936781609	Characiformes	Prochilodontidae	27694
species	0.948571429	Siluriformes	Pimelodidae	162147
species	0.970588235	Characiformes	Serrasalmidae	371937
species	0.970588235	Characiformes	Crenuchidae	1569729
species	0.923976608	Characiformes	Serrasalmidae	371937
family	0.941860465	Siluriformes	Loricariidae	31002
species	0.936781609	Characiformes	Crenuchidae	1569729
species	0.959064328	Characiformes	Bryconidae	202829
species	0.970588235	Characiformes	Anostomidae	1925008
genus	0.959302326	Characiformes	Prochilodontidae	27694
species	0.971098266	Characiformes	Anostomidae	749479

ICE 1				
rank	best_identity	order_name	family_name	taxid
species	0.971428571	Characiformes	Curimatidae	984680
species	0.953757225	Characiformes	Prochilodontidae	27695
species	0.976608187	Siluriformes	Loricariidae	168219
species	0.976608187	Siluriformes	Loricariidae	163993
suborder	0.976190476	Characiformes		1489739
species	0.947674419	Characiformes	Crenuchidae	1569729
species	0.930232558	Characiformes	Prochilodontidae	27695
species	1	Siluriformes	Loricariidae	745493
species	1	Siluriformes	Callichthyidae	114109
species	0.947368421	Characiformes	Prochilodontidae	27695
species	0.882022472	Siluriformes	Loricariidae	745493
species	0.925287356	Characiformes	Curimatidae	984680
species	0.976608187	Characiformes	Serrasalmidae	371937
species	0.936416185	Characiformes	Serrasalmidae	371937
species	1	Characiformes	Crenuchidae	1569729
species	0.965116279	Characiformes	Crenuchidae	1569729
genus	0.970760234	Siluriformes	Loricariidae	36713
species	0.976608187	Characiformes	Prochilodontidae	27695
species	1	Cichliformes	Cichlidae	870598
family	0.919075145	Siluriformes	Loricariidae	31002
species	1	Characiformes	Bryconidae	202829
genus	0.994252874	Siluriformes	Clariidae	13012
species	0.994252874	Siluriformes	Pimelodidae	162147
species	1	Siluriformes	Pseudopimelodidae	490146
genus	0.971264368	Siluriformes	Loricariidae	503149
species	1	Characiformes	Anostomidae	749479
species	0.988372093	Siluriformes	Loricariidae	1191588
species	1	Characiformes	Bryconidae	930266
species	1	Siluriformes	Trichomycteridae	10000003
species	0.940828402	Characiformes	Erythrinidae	27720
species	1	Siluriformes	Auchenipteridae	984686
species	1	Characiformes	Characidae	1604128
species	0.994011976	Characiformes	Erythrinidae	1053573
species	0.976744186	Siluriformes	Loricariidae	1191588
species	0.977011494	Siluriformes	Pimelodidae	162147
species	1	Cichliformes	Cichlidae	870598
species	1	Cyprinodontiformes	Poeciliidae	10000004
species	0.929411765	Characiformes	Erythrinidae	27720
species	1	Characiformes	Anostomidae	1925008
species	0.988023952	Characiformes	Curimatidae	984680
species	0.958823529	Characiformes	Anostomidae	1925008
species	0.988165681	Cichliformes	Cichlidae	81339
species	1	Characiformes	Characidae	1604128
species	1	Cyprinodontiformes	Poeciliidae	10000002
species	0.976470588	Characiformes	Crenuchidae	1569729
species	1	Characiformes	Serrasalmidae	371937
species	1	Characiformes	Prochilodontidae	27695

ICE 2				
rank	best_identity	order_name	family_name	taxid
species	0.971428571	Characiformes	Curimatidae	984680
genus	0.959302326	Characiformes	Prochilodontidae	27694
suborder	0.959064328	Characiformes		1489739
species	0.907514451	Characiformes	Erythrinidae	27720
species	0.964497041	Characiformes	Bryconidae	202829
suborder	0.959064328	Characiformes		1489739
genus	0.919075145	Characiformes	Prochilodontidae	27694
species	0.953216374	Characiformes	Bryconidae	930266
species	0.900584795	Characiformes	Crenuchidae	1569729
species	0.948275862	Characiformes	Prochilodontidae	27695
family	0.901162791	Characiformes	Serrasalmidae	42495
species	0.941176471	Characiformes	Serrasalmidae	371937
genus	0.906976744	Characiformes	Prochilodontidae	27694
genus	0.902298851	Characiformes	Bryconidae	42585
species	0.930232558	Characiformes	Crenuchidae	1569729
genus	0.959302326	Characiformes	Prochilodontidae	27694
species	0.891428571	Characiformes	Crenuchidae	1569729
family	0.88	Siluriformes	Loricariidae	31002
species	0.958823529	Characiformes	Serrasalmidae	371937
family	0.952380952	Characiformes	Characidae	7992
species	0.942857143	Siluriformes	Loricariidae	745493
subfamily	0.925287356	Siluriformes	Loricariidae	503143
species	0.925287356	Characiformes	Crenuchidae	1569729
species	0.959770115	Characiformes	Prochilodontidae	27695
genus	0.959302326	Characiformes	Prochilodontidae	27694
species	0.959537572	Characiformes	Prochilodontidae	27695
species	0.923976608	Characiformes	Serrasalmidae	371937
species	0.948571429	Characiformes	Curimatidae	984680
species	1	Cichliformes	Cichlidae	50733
species	0.971264368	Characiformes	Prochilodontidae	27695
genus	0.924418605	Characiformes	Prochilodontidae	27694
species	0.959770115	Characiformes	Anostomidae	749479
species	0.885245902	Characiformes	Crenuchidae	1569729
genus	0.965116279	Characiformes	Prochilodontidae	27694
genus	0.897142857	Characiformes	Prochilodontidae	27694
species	0.959064328	Characiformes	Crenuchidae	1569729
species	0.947674419	Characiformes	Serrasalmidae	371937
genus	0.959302326	Characiformes	Prochilodontidae	27694
genus	0.959302326	Characiformes	Prochilodontidae	27694
genus	0.906976744	Characiformes	Prochilodontidae	27694
species	0.971098266	Characiformes	Anostomidae	749479
genus	0.965116279	Characiformes	Prochilodontidae	27694
genus	0.924418605	Characiformes	Anostomidae	1925018
species	0.888888889	Characiformes	Serrasalmidae	371937
species	0.970760234	Characiformes	Anostomidae	1925008
genus	0.964705882	Characiformes	Anostomidae	1925018
genus	0.953757225	Characiformes	Prochilodontidae	27694
species	0.953757225	Characiformes	Prochilodontidae	27695

ICE 2				
rank	best_identity	order_name	family_name	taxid
genus	0.942196532	Characiformes	Prochilodontidae	27694
species	0.947976879	Characiformes	Anostomidae	749479
family	0.923976608	Characiformes	Serrasalmidae	42495
suborder	0.907514451	Siluriformes		1489793
species	0.970588235	Characiformes	Serrasalmidae	371937
genus	0.919075145	Characiformes	Anostomidae	1925018
species	0.931428571	Characiformes	Prochilodontidae	27695
species	0.913294798	Characiformes	Crenuchidae	1569729
species	0.829411765	Characiformes	Serrasalmidae	371937
species	1	Characiformes	Crenuchidae	1569729
species	0.953216374	Characiformes	Crenuchidae	1569729
species	0.941520468	Characiformes	Crenuchidae	1569729
species	0.936046512	Characiformes	Prochilodontidae	27695
genus	0.970760234	Siluriformes	Loricariidae	36713
species	0.970930233	Characiformes	Serrasalmidae	371937
species	0.970760234	Characiformes	Anostomidae	1925008
species	1	Siluriformes	Auchenipteridae	984686
species	0.936781609	Characiformes	Crenuchidae	1569729
genus	0.941176471	Characiformes	Anostomidae	1925018
species	0.914285714	Characiformes	Prochilodontidae	27695
species	0.959537572	Characiformes	Anostomidae	749479
species	0.935294118	Characiformes	Crenuchidae	1569729
genus	0.906432749	Characiformes	Anostomidae	1925018
species	0.918128655	Characiformes	Crenuchidae	1569729
species	0.942196532	Characiformes	Anostomidae	1925003
species	0.965714286	Characiformes	Prochilodontidae	27695
genus	0.953757225	Characiformes	Prochilodontidae	27694
genus	0.930635838	Characiformes	Characidae	681908
genus	0.925714286	Characiformes	Characidae	681908
species	0.970930233	Siluriformes	Loricariidae	1191587
suborder	0.935672515	Characiformes		1489739
species	0.882022472	Siluriformes	Loricariidae	745493
species	0.965116279	Characiformes	Anostomidae	1925008
species	0.970588235	Characiformes	Anostomidae	1925008
species	0.954022989	Characiformes	Prochilodontidae	27695
genus	0.947368421	Characiformes	Prochilodontidae	27694
genus	0.930635838	Characiformes	Prochilodontidae	27694
species	1	Cyprinodontiformes	Poeciliidae	8081
species	1	Characiformes	Bryconidae	202829
species	1	Cyprinodontiformes	Poeciliidae	10000004
genus	0.994252874	Siluriformes	Clariidae	13012
species	0.925287356	Characiformes	Curimatidae	984680
species	1	Siluriformes	Pseudopimelodidae	490146
species	0.941520468	Characiformes	Crenuchidae	1569729
species	1	Siluriformes	Loricariidae	745493
species	0.941860465	Characiformes	Crenuchidae	1569729
species	0.970238095	Characiformes	Bryconidae	126315
species	0.929411765	Characiformes	Erythrinidae	27720

ICE 2				
rank	best_identity	order_name	family_name	taxid
species	0.959302326	Characiformes	Prochilodontidae	27695
species	0.947674419	Characiformes	Crenuchidae	1569729
species	0.976744186	Siluriformes	Loricariidae	1191588
species	0.970588235	Characiformes	Anostomidae	1925008
superorder	0.83908046			186628
species	0.988023952	Characiformes	Erythrinidae	27720
species	0.988023952	Characiformes	Curimatidae	984680
species	0.926553672	Siluriformes	Pimelodidae	162147
suborder	1	Characiformes		1489739
species	1	Characiformes	Anostomidae	749479
species	0.953757225	Characiformes	Prochilodontidae	27695
species	0.988165681	Cichliformes	Cichlidae	81339
species	1	Cypriniformes	Cyprinidae	7962
species	1	Characiformes	Bryconidae	930266
species	0.977011494	Siluriformes	Pimelodidae	162147
species	1	Cichliformes	Cichlidae	870598
family	0.959302326	Siluriformes	Loricariidae	31002
family	0.919075145	Siluriformes	Loricariidae	31002
species	0.994252874	Siluriformes	Pimelodidae	162147
species	0.947368421	Characiformes	Prochilodontidae	27695
genus	0.971264368	Siluriformes	Loricariidae	503149
superorder	0.830508475			186628
species	0.976608187	Characiformes	Serrasalmidae	371937
species	0.931818182	Siluriformes	Trichomycteridae	245746
species	1	Cyprinodontiformes	Poeciliidae	10000002
species	0.936416185	Characiformes	Serrasalmidae	371937
species	0.976047904	Cichliformes	Cichlidae	8130
species	0.994011976	Characiformes	Erythrinidae	1053573
species	0.930232558	Characiformes	Prochilodontidae	27695
species	1	Characiformes	Characidae	1604128
family	0.942196532	Siluriformes	Loricariidae	31002
species	0.976608187	Characiformes	Prochilodontidae	27695
species	0.976608187	Siluriformes	Loricariidae	163993
species	0.965116279	Characiformes	Crenuchidae	1569729
suborder	0.976190476	Characiformes		1489739
species	0.971098266	Characiformes	Anostomidae	749479
species	0.900584795	Characiformes	Characidae	1123824
species	0.994047619	Cichliformes	Cichlidae	47969
species	1	Characiformes	Characidae	1123824
species	1	Characiformes	Characidae	1604128
species	0.976744186	Gymnotiformes	Gymnotidae	94172
species	0.958823529	Characiformes	Anostomidae	1925008
species	1	Characiformes	Anostomidae	1925008
species	0.976470588	Characiformes	Crenuchidae	1569729
species	1	Characiformes	Serrasalmidae	371937
species	1	Characiformes	Prochilodontidae	27695

SED 1				
rank	best_identity	order_name	family_name	taxid
species	0.976608187	Siluriformes	Loricariidae	163993
species	0.988023952	Characiformes	Erythrinidae	27720
species	1	Cichliformes	Cichlidae	50733
species	0.941520468	Characiformes	Crenuchidae	1569729
genus	0.919075145	Characiformes	Anostomidae	1925018
species	0.913294798	Characiformes	Crenuchidae	1569729
genus	0.947976879	Characiformes	Prochilodontidae	27694
species	0.942196532	Characiformes	Anostomidae	1925003
genus	0.953757225	Characiformes	Prochilodontidae	27694
species	0.959064328	Characiformes	Crenuchidae	1569729
species	0.941176471	Characiformes	Serrasalmidae	371937
species	0.948275862	Characiformes	Curimatidae	984680
genus	0.959302326	Characiformes	Prochilodontidae	27694
genus	0.959302326	Characiformes	Prochilodontidae	27694
family	0.941176471	Characiformes	Characidae	7992
species	0.958823529	Characiformes	Serrasalmidae	371937
species	0.902857143	Characiformes	Crenuchidae	1569729
species	0.971098266	Characiformes	Anostomidae	749479
genus	0.947976879	Characiformes	Prochilodontidae	27694
family	0.952380952	Characiformes	Characidae	7992
species	0.96	Siluriformes	Loricariidae	745493
family	0.947058824	Characiformes	Characidae	7992
species	0.953488372	Characiformes	Anostomidae	749479
family	0.884393064	Siluriformes	Loricariidae	31002
species	0.964912281	Characiformes	Bryconidae	202829
species	0.970414201	Cichliformes	Cichlidae	81339
species	0.942196532	Characiformes	Prochilodontidae	27695
family	0.876470588	Characiformes	Anostomidae	7986
species	0.931428571	Siluriformes	Pimelodidae	162147
species	0.954022989	Characiformes	Bryconidae	202829
species	0.948571429	Siluriformes	Pseudopimelodidae	490146
species	0.912790698	Characiformes	Crenuchidae	1569729
species	0.965909091	Cichliformes	Cichlidae	870598
species	0.965317919	Characiformes	Serrasalmidae	371937
species	0.953216374	Characiformes	Crenuchidae	1569729
genus	0.947368421	Characiformes	Prochilodontidae	27694
species	0.959770115	Characiformes	Anostomidae	749479
species	0.959537572	Characiformes	Anostomidae	749479
species	0.931428571	Characiformes	Prochilodontidae	27695
species	0.936781609	Siluriformes	Pseudopimelodidae	490146
family	0.901162791	Characiformes	Serrasalmidae	42495
species	0.948863636	Characiformes	Serrasalmidae	371937
species	0.970238095	Cyprinodontiformes	Poeciliidae	10000004
species	0.970930233	Characiformes	Anostomidae	749479
species	0.96	Characiformes	Prochilodontidae	27695
species	0.970588235	Characiformes	Bryconidae	202829
family	0.913294798	Siluriformes	Loricariidae	31002
species	0.942528736	Characiformes	Crenuchidae	1569729

SED 1				
rank	best_identity	order_name	family_name	taxid
species	0.959064328	Characiformes	Serrasalmidae	371937
genus	0.906976744	Characiformes	Prochilodontidae	27694
genus	0.944751381	Characiformes	Prochilodontidae	27694
species	0.949152542	Siluriformes	Pimelodidae	162147
species	0.942857143	Siluriformes	Loricariidae	745493
family	0.936046512	Siluriformes	Loricariidae	31002
species	0.918604651	Characiformes	Crenuchidae	1569729
species	0.943181818	Characiformes	Crenuchidae	1569729
species	0.959537572	Characiformes	Anostomidae	1925003
species	0.924855491	Characiformes	Crenuchidae	1569729
species	0.923529412	Characiformes	Erythrinidae	1053573
genus	0.924418605	Characiformes	Anostomidae	1925018
family	0.892045455	Siluriformes	Loricariidae	31002
species	0.880681818	Characiformes	Erythrinidae	27720
species	0.926136364	Characiformes	Curimatidae	984680
suborder	0.907514451	Siluriformes		1489793
family	0.936416185	Siluriformes	Loricariidae	31002
species	0.947368421	Characiformes	Crenuchidae	1569729
genus	0.919075145	Characiformes	Prochilodontidae	27694
species	0.959770115	Characiformes	Prochilodontidae	27695
species	0.900584795	Characiformes	Crenuchidae	1569729
species	0.926136364	Characiformes	Prochilodontidae	27695
species	0.882681564	Characiformes	Erythrinidae	27720
species	0.964912281	Characiformes	Bryconidae	930266
species	0.971098266	Characiformes	Anostomidae	1925003
family	0.931034483	Siluriformes	Pseudopimelodidae	255574
species	0.965517241	Characiformes	Curimatidae	984680
species	0.949152542	Characiformes	Prochilodontidae	27695
species	0.895953757	Characiformes	Erythrinidae	27720
species	0.959064328	Characiformes	Anostomidae	1925003
genus	0.942857143	Characiformes	Prochilodontidae	27694
species	0.941176471	Siluriformes	Trichomycteridae	10000003
species	0.802139037	Characiformes	Serrasalmidae	371937
family	0.913793103	Siluriformes	Loricariidae	31002
species	0.964705882	Characiformes	Erythrinidae	1053573
species	0.912280702	Characiformes	Serrasalmidae	371937
species	0.918128655	Characiformes	Erythrinidae	27720
genus	0.960227273	Characiformes	Prochilodontidae	27694
species	0.901734104	Characiformes	Erythrinidae	27720
family	0.919075145	Siluriformes	Loricariidae	31002
species	0.941176471	Characiformes	Erythrinidae	27720
subfamily	0.925287356	Siluriformes	Loricariidae	503143
species	0.953216374	Characiformes	Prochilodontidae	27695
species	0.929411765	Characiformes	Crenuchidae	1569729
species	0.964705882	Characiformes	Serrasalmidae	371937
genus	0.965116279	Characiformes	Prochilodontidae	27694
genus	0.897142857	Characiformes	Prochilodontidae	27694
species	0.913294798	Characiformes	Crenuchidae	1569729

SED 1				
rank	best_identity	order_name	family_name	taxid
species	0.900584795	Characiformes	Erythrinidae	27720
species	0.935294118	Characiformes	Erythrinidae	1053573
species	0.918604651	Characiformes	Bryconidae	930266
suborder	0.964705882	Characiformes		1489739
species	0.954022989	Siluriformes	Pseudopimelodidae	490146
species	0.959064328	Characiformes	Anostomidae	1925003
species	0.925287356	Siluriformes	Pimelodidae	162147
species	0.952662722	Characiformes	Bryconidae	202829
genus	0.954545455	Siluriformes	Clariidae	13012
species	0.947674419	Characiformes	Prochilodontidae	27695
suborder	0.893854749	Siluriformes		1489793
family	0.942196532	Siluriformes	Loricariidae	31002
species	0.935672515	Characiformes	Crenuchidae	1569729
species	0.947976879	Characiformes	Prochilodontidae	27695
family	0.959537572	Siluriformes	Loricariidae	31002
species	0.96	Characiformes	Prochilodontidae	27695
suborder	0.9375	Siluriformes		1489793
species	0.895953757	Characiformes	Crenuchidae	1569729
species	0.948275862	Characiformes	Prochilodontidae	27695
family	0.936046512	Siluriformes	Loricariidae	31002
genus	0.970930233	Characiformes	Prochilodontidae	27694
species	0.920454546	Characiformes	Crenuchidae	1569729
family	0.88	Siluriformes	Loricariidae	31002
genus	0.947674419	Characiformes	Prochilodontidae	27694
genus	0.959064328	Characiformes	Anostomidae	1925018
species	0.914285714	Siluriformes	Loricariidae	745493
species	0.964912281	Characiformes	Bryconidae	930266
genus	0.930635838	Characiformes	Prochilodontidae	27694
species	0.885245902	Characiformes	Crenuchidae	1569729
species	0.953757225	Characiformes	Crenuchidae	1569729
species	0.948275862	Siluriformes	Pseudopimelodidae	490146
species	0.942528736	Siluriformes	Pimelodidae	162147
suborder	0.934911243	Characiformes		1489739
species	0.954285714	Characiformes	Curimatidae	984680
species	0.918128655	Characiformes	Erythrinidae	27720
genus	0.902298851	Characiformes	Bryconidae	42585
species	0.90960452	Characiformes	Crenuchidae	1569729
species	0.959537572	Characiformes	Anostomidae	749479
species	0.965317919	Characiformes	Curimatidae	984680
species	0.948863636	Siluriformes	Pseudopimelodidae	490146
genus	0.964705882	Characiformes	Anostomidae	1925018
genus	0.954285714	Characiformes	Prochilodontidae	27694
species	0.954285714	Characiformes	Prochilodontidae	27695
species	0.925287356	Siluriformes	Pimelodidae	162147
species	0.959064328	Characiformes	Prochilodontidae	27695
genus	0.941860465	Characiformes	Prochilodontidae	27694
genus	0.941176471	Characiformes	Anostomidae	1925018
species	0.959537572	Characiformes	Prochilodontidae	27695

SED 1				
rank	best_identity	order_name	family_name	taxid
suborder	0.959064328	Characiformes		1489739
species	0.913294798	Characiformes	Prochilodontidae	27695
genus	0.923076923	Characiformes	Bryconidae	42585
suborder	0.92	Siluriformes		1489793
species	0.941520468	Characiformes	Anostomidae	1925003
species	0.953216374	Characiformes	Bryconidae	930266
genus	0.906432749	Characiformes	Anostomidae	1925018
species	0.954022989	Characiformes	Curimatidae	984680
species	0.935672515	Characiformes	Erythrinidae	1053573
species	0.948275862	Characiformes	Prochilodontidae	27695
species	0.971098266	Characiformes	Curimatidae	984680
species	0.959064328	Characiformes	Bryconidae	202829
family	0.941860465	Siluriformes	Loricariidae	31002
family	0.897142857	Characiformes	Curimatidae	30721
family	0.935294118	Characiformes	Characidae	7992
species	0.965517241	Characiformes	Curimatidae	984680
family	0.884393064	Siluriformes	Loricariidae	31002
species	0.953488372	Characiformes	Crenuchidae	1569729
species	0.964497041	Characiformes	Bryconidae	202829
species	0.965116279	Characiformes	Anostomidae	1925003
species	0.994047619	Cichliformes	Cichlidae	47969
species	0.970760234	Characiformes	Anostomidae	1925008
species	0.888888889	Characiformes	Serrasalmidae	371937
species	0.907514451	Characiformes	Erythrinidae	27720
species	0.970588235	Characiformes	Anostomidae	1925008
genus	0.959302326	Characiformes	Prochilodontidae	27694
species	0.948571429	Siluriformes	Pimelodidae	162147
species	0.923976608	Characiformes	Serrasalmidae	371937
genus	0.953757225	Characiformes	Prochilodontidae	27694
species	0.970930233	Characiformes	Serrasalmidae	371937
suborder	0.959064328	Characiformes		1489739
species	0.941860465	Characiformes	Crenuchidae	1569729
genus	0.959302326	Characiformes	Prochilodontidae	27694
species	0.954022989	Siluriformes	Pimelodidae	162147
genus	0.912790698	Characiformes	Anostomidae	1925018
species	0.976744186	Characiformes	Prochilodontidae	27695
genus	0.924418605	Characiformes	Prochilodontidae	27694
genus	0.936781609	Characiformes	Prochilodontidae	27694
species	0.934911243	Characiformes	Erythrinidae	27720
species	0.935294118	Characiformes	Crenuchidae	1569729
species	0.936781609	Characiformes	Crenuchidae	1569729
species	0.948571429	Characiformes	Curimatidae	984680
species	0.971264368	Characiformes	Prochilodontidae	27695
species	0.925287356	Characiformes	Curimatidae	984680
species	0.959537572	Characiformes	Prochilodontidae	27695
species	0.971098266	Characiformes	Anostomidae	749479
species	0.941520468	Characiformes	Crenuchidae	1569729
species	0.994252874	Siluriformes	Pimelodidae	162147

SED 1				
rank	best_identity	order_name	family_name	taxid
species	0.959064328	Characiformes	Bryconidae	202829
species	0.970588235	Characiformes	Anostomidae	1925008
species	1	Characiformes	Anostomidae	749479
species	0.971428571	Characiformes	Curimatidae	984680
species	0.936416185	Clupeiformes	Engraulidae	458582
species	0.959302326	Characiformes	Prochilodontidae	27695
species	0.953757225	Characiformes	Prochilodontidae	27695
species	1	Siluriformes	Loricariidae	745493
species	1	Siluriformes	Callichthyidae	114109
species	1	Siluriformes	Pseudopimelodidae	490146
species	0.947674419	Characiformes	Crenuchidae	1569729
species	0.976608187	Characiformes	Serrasalmidae	371937
species	0.930232558	Characiformes	Prochilodontidae	27695
species	0.936416185	Characiformes	Serrasalmidae	371937
species	0.947368421	Characiformes	Prochilodontidae	27695
species	0.965116279	Characiformes	Crenuchidae	1569729
species	0.976608187	Characiformes	Prochilodontidae	27695
species	0.988372093	Siluriformes	Loricariidae	1191588
genus	0.971264368	Siluriformes	Loricariidae	503149
species	1	Characiformes	Bryconidae	930266
species	1	Characiformes	Crenuchidae	1569729
species	1	Siluriformes	Trichomycteridae	10000003
species	0.994011976	Characiformes	Erythrinidae	1053573
species	1	Cichliformes	Cichlidae	870598
suborder	1	Characiformes		1489739
species	1	Characiformes	Characidae	1604128
species	0.988165681	Cichliformes	Cichlidae	81339
genus	0.970760234	Siluriformes	Loricariidae	36713
family	0.919075145	Siluriformes	Loricariidae	31002
species	0.940828402	Characiformes	Erythrinidae	27720
species	1	Cyprinodontiformes	Poeciliidae	10000004
species	0.977011494	Siluriformes	Pimelodidae	162147
species	1	Cichliformes	Cichlidae	870598
species	0.929411765	Characiformes	Erythrinidae	27720
species	1	Siluriformes	Auchenipteridae	984686
species	0.976744186	Siluriformes	Loricariidae	1191588
species	0.988023952	Characiformes	Curimatidae	984680
species	0.958823529	Characiformes	Anostomidae	1925008
species	1	Characiformes	Bryconidae	202829
species	1	Characiformes	Characidae	1604128
species	1	Characiformes	Anostomidae	1925008
species	1	Cyprinodontiformes	Poeciliidae	10000002
species	0.976470588	Characiformes	Crenuchidae	1569729
species	1	Characiformes	Serrasalmidae	371937
species	1	Characiformes	Prochilodontidae	27695

SED 2				
rank	best_identity	order_name	family_name	taxid
species	1	Characiformes	Bryconidae	930266
species	0.994047619	Cichliformes	Cichlidae	47969
suborder	0.976190476	Characiformes		1489739
species	0.976047904	Cichliformes	Cichlidae	8130
species	0.925287356	Characiformes	Curimatidae	984680
species	1	Cypriniformes	Cyprinidae	7962
species	0.926553672	Siluriformes	Pimelodidae	162147
genus	0.936781609	Characiformes	Prochilodontidae	27694
genus	0.953757225	Characiformes	Prochilodontidae	27694
species	0.954022989	Siluriformes	Pimelodidae	162147
species	0.954022989	Characiformes	Curimatidae	984680
species	0.959537572	Characiformes	Prochilodontidae	27695
suborder	0.959064328	Characiformes		1489739
genus	0.964705882	Characiformes	Anostomidae	1925018
species	0.953216374	Characiformes	Prochilodontidae	27695
species	0.964705882	Characiformes	Serrasalmidae	371937
species	0.954022989	Characiformes	Prochilodontidae	27695
species	0.954285714	Characiformes	Prochilodontidae	27695
species	0.925287356	Siluriformes	Pimelodidae	162147
genus	0.897142857	Characiformes	Prochilodontidae	27694
species	0.947368421	Characiformes	Crenuchidae	1569729
species	0.953216374	Characiformes	Bryconidae	930266
genus	0.930635838	Characiformes	Characidae	681908
species	0.913294798	Characiformes	Crenuchidae	1569729
species	0.900584795	Characiformes	Crenuchidae	1569729
species	0.953757225	Characiformes	Prochilodontidae	27695
family	0.901162791	Characiformes	Serrasalmidae	42495
species	0.941176471	Characiformes	Serrasalmidae	371937
genus	0.970930233	Characiformes	Prochilodontidae	27694
species	0.948275862	Characiformes	Curimatidae	984680
genus	0.959302326	Characiformes	Prochilodontidae	27694
species	0.965517241	Characiformes	Curimatidae	984680
species	0.959064328	Characiformes	Serrasalmidae	371937
species	0.971098266	Characiformes	Anostomidae	749479
species	0.947674419	Characiformes	Prochilodontidae	27695
genus	0.947674419	Characiformes	Prochilodontidae	27694
species	0.918604651	Characiformes	Crenuchidae	1569729
species	0.96	Siluriformes	Loricariidae	745493
subfamily	0.925287356	Siluriformes	Loricariidae	503143
species	0.943181818	Characiformes	Crenuchidae	1569729
species	0.901734104	Characiformes	Erythrinidae	27720
family	0.942196532	Siluriformes	Loricariidae	31002
species	0.971428571	Characiformes	Curimatidae	984680
species	0.941520468	Characiformes	Crenuchidae	1569729
species	0.970760234	Characiformes	Anostomidae	1925008
family	0.935294118	Characiformes	Characidae	7992
genus	0.919075145	Characiformes	Anostomidae	1925018
species	0.914285714	Characiformes	Prochilodontidae	27695

SED 2				
rank	best_identity	order_name	family_name	taxid
species	0.965714286	Characiformes	Prochilodontidae	27695
species	0.913294798	Characiformes	Prochilodontidae	27695
genus	0.953757225	Characiformes	Prochilodontidae	27694
genus	0.919075145	Characiformes	Prochilodontidae	27694
suborder	0.9375	Siluriformes		1489793
genus	0.953757225	Characiformes	Prochilodontidae	27694
species	0.900584795	Characiformes	Erythrinidae	27720
genus	0.959302326	Characiformes	Prochilodontidae	27694
species	0.918128655	Characiformes	Erythrinidae	27720
species	0.891428571	Characiformes	Crenuchidae	1569729
genus	0.906976744	Characiformes	Prochilodontidae	27694
genus	0.944751381	Characiformes	Prochilodontidae	27694
species	0.949152542	Characiformes	Prochilodontidae	27695
genus	0.965116279	Characiformes	Prochilodontidae	27694
species	0.971098266	Characiformes	Anostomidae	749479
species	0.959064328	Characiformes	Bryconidae	202829
species	0.948571429	Characiformes	Curimatidae	984680
species	0.971264368	Characiformes	Prochilodontidae	27695
species	0.953488372	Characiformes	Crenuchidae	1569729
species	0.931428571	Characiformes	Prochilodontidae	27695
species	0.959770115	Characiformes	Prochilodontidae	27695
species	0.918604651	Characiformes	Bryconidae	930266
species	0.947674419	Characiformes	Serrasalmidae	371937
species	0.958823529	Characiformes	Serrasalmidae	371937
species	0.895953757	Characiformes	Erythrinidae	27720
species	1	Characiformes	Characidae	1123824
genus	0.959302326	Characiformes	Prochilodontidae	27694
species	0.935294118	Characiformes	Crenuchidae	1569729
species	0.953216374	Characiformes	Crenuchidae	1569729
species	0.934911243	Characiformes	Erythrinidae	27720
genus	0.906432749	Characiformes	Anostomidae	1925018
species	0.959770115	Characiformes	Anostomidae	749479
suborder	0.934911243	Characiformes		1489739
species	0.959064328	Characiformes	Crenuchidae	1569729
species	0.829411765	Characiformes	Serrasalmidae	371937
genus	0.959302326	Characiformes	Prochilodontidae	27694
family	0.923976608	Characiformes	Serrasalmidae	42495
species	0.965317919	Characiformes	Serrasalmidae	371937
species	0.959537572	Characiformes	Prochilodontidae	27695
species	0.976744186	Characiformes	Prochilodontidae	27695
genus	0.942196532	Characiformes	Prochilodontidae	27694
species	0.976744186	Gymnotiformes	Gymnotidae	94172
species	0.923976608	Characiformes	Serrasalmidae	371937
species	0.888888889	Characiformes	Serrasalmidae	371937
species	0.970588235	Characiformes	Serrasalmidae	371937
genus	0.959302326	Characiformes	Prochilodontidae	27694
species	0.907514451	Characiformes	Erythrinidae	27720
species	0.885245902	Characiformes	Crenuchidae	1569729

SED 2				
rank	best_identity	order_name	family_name	taxid
species	0.929411765	Characiformes	Crenuchidae	1569729
species	0.976608187	Siluriformes	Loricariidae	163993
suborder	0.959064328	Characiformes		1489739
genus	0.930635838	Characiformes	Prochilodontidae	27694
genus	0.947368421	Characiformes	Prochilodontidae	27694
species	0.964705882	Characiformes	Serrasalmidae	371937
species	0.936781609	Characiformes	Crenuchidae	1569729
species	0.959537572	Characiformes	Anostomidae	749479
species	0.941860465	Characiformes	Crenuchidae	1569729
species	0.970588235	Characiformes	Anostomidae	1925008
species	1	Cichliformes	Cichlidae	50733
species	0.900584795	Characiformes	Characidae	1123824
species	0.970930233	Characiformes	Serrasalmidae	371937
genus	0.941176471	Characiformes	Anostomidae	1925018
species	0.941520468	Characiformes	Crenuchidae	1569729
species	0.970588235	Characiformes	Anostomidae	1925008
species	0.959302326	Characiformes	Prochilodontidae	27695
species	0.994252874	Siluriformes	Pimelodidae	162147
species	0.953757225	Characiformes	Prochilodontidae	27695
species	1	Siluriformes	Auchenipteridae	984686
species	0.977011494	Siluriformes	Pimelodidae	162147
species	0.947674419	Characiformes	Crenuchidae	1569729
family	0.919075145	Siluriformes	Loricariidae	31002
species	1	Siluriformes	Trichomycteridae	10000003
species	1	Cichliformes	Cichlidae	870598
species	0.976744186	Siluriformes	Loricariidae	1191588
species	0.958823529	Characiformes	Serrasalmidae	371937
genus	0.994252874	Siluriformes	Clariidae	13012
species	0.947368421	Characiformes	Prochilodontidae	27695
species	1	Characiformes	Crenuchidae	1569729
species	0.976608187	Characiformes	Serrasalmidae	371937
species	0.936416185	Characiformes	Serrasalmidae	371937
species	1	Characiformes	Anostomidae	749479
species	0.930232558	Characiformes	Prochilodontidae	27695
genus	0.971264368	Siluriformes	Loricariidae	503149
species	0.965116279	Characiformes	Crenuchidae	1569729
species	0.976608187	Characiformes	Prochilodontidae	27695
genus	0.970760234	Siluriformes	Loricariidae	36713
species	1	Characiformes	Characidae	1604128
species	0.929411765	Characiformes	Erythrinidae	27720
species	1	Cyprinodontiformes	Poeciliidae	10000002
species	0.988372093	Siluriformes	Loricariidae	1191588
species	0.994011976	Characiformes	Erythrinidae	1053573
species	1	Cichliformes	Cichlidae	870598
species	0.988165681	Cichliformes	Cichlidae	81339
species	1	Siluriformes	Loricariidae	745493
species	0.988023952	Characiformes	Curimatidae	984680
species	1	Siluriformes	Pseudopimelodidae	490146

SED 2				
rank	best_identity	order_name	family_name	taxid
suborder	0.907514451	Siluriformes		1489793
species	1	Characiformes	Bryconidae	202829
suborder	1	Characiformes		1489739
species	1	Characiformes	Characidae	1604128
species	1	Characiformes	Anostomidae	1925008
species	0.958823529	Characiformes	Anostomidae	1925008
species	0.976470588	Characiformes	Crenuchidae	1569729
species	1	Characiformes	Prochilodontidae	27695
species	1	Characiformes	Serrasalmidae	371937