Genetic landscape clustering of a large DNA barcoding dataset reveals shared patterns of genetic divergence among freshwater fishes of the Maroni Basin

Yvan Papa[1], Pierre-Yves Le Bail[2], and Raphael Covain[1]

[1]Museum of Natural History of Geneva
[2]INRA

July 5, 2022

Abstract

Version of record in Molecular Ecology Resources: https://doi.org/10.1111/1755-0998.13402

The Maroni is one of the most speciose basins of the Guianas and hosts a megadiverse freshwater fish community. Although taxonomical references exist for both the Surinamese and Guyanese parts of the basin, these lists were mainly based on morphological identification and there are still taxonomical uncertainties concerning the status of several fish species. Here we present a barcode dataset of 1,284 COI sequences from 199 freshwater fish species (68.86% of the total number of strictly freshwater fishes from the basin) from 124 genera, 36 families, and 8 orders. DNA barcoding allowed for fast and efficient identification of all specimens studied as well as unveiling a consequent cryptic diversity, with the detection of 20 putative cryptic species and 5 species flagged for re-identification. In order to explore global genetic patterns across the basin, genetic divergence landscapes were computed for 128 species, showing a global trend of high genetic divergence between the Surinamese south-west (Tapanahony and Paloemeu), the Guianese south-east (Marouini, Litany, Tampok, Lawa...), and the river mouth in the north. This could be explained either by lower levels of connectivity between these three main parts or by the exchange of individuals with the surrounding basins. A new method of ordination of genetic landscapes successfully assigned species into cluster groups based on their respective pattern of genetic divergence across the Maroni Basin: genetically homogenous species across the basin were effectively discriminated from species showing high spatial genetic fragmentation and possible lower capacity for dispersal.

Title: Genetic landscape clustering of a large DNA barcoding dataset reveals shared patterns of genetic divergence among freshwater fishes of the Maroni Basin

Short Title: DNA barcoding of Maroni freshwater fishes

Authors: Yvan Papa[1,2]*, Pierre-Yves Le Bail[3], Raphaël Covain[1]

[1]Department of Herpetology and Ichthyology, Museum of Natural History of Geneva, Geneva, Switzerland

[2]School of Biological Sciences, Victoria University of Wellington, Wellington, New Zealand

[3]INRA, Fish Physiology and Genomics UR 1037, Rennes, France

*corresponding author (email: yvanpapa@gmail.com)

Abstract

The Maroni is one of the most speciose basins of the Guianas and hosts a megadiverse freshwater fish community. Although taxonomical references exist for both the Surinamese and Guyanese parts of the basin,

these lists were mainly based on morphological identification and there are still taxonomical uncertainties concerning the status of several fish species. Here we present a barcode dataset of 1,284 COI sequences from 199 freshwater fish species (68.86% of the total number of strictly freshwater fishes from the basin) from 124 genera, 36 families, and 8 orders. DNA barcoding allowed for fast and efficient identification of all specimens studied as well as unveiling a consequent cryptic diversity, with the detection of 20 putative cryptic species and 5 species flagged for re-identification. In order to explore global genetic patterns across the basin, genetic divergence landscapes were computed for 128 species, showing a global trend of high genetic divergence between the Surinamese south-west (Tapanahony and Paloemeu), the Guianese south-east (Marouini, Litany, Tampok, Lawa...), and the river mouth in the north. This could be explained either by lower levels of connectivity between these three main parts or by the exchange of individuals with the surrounding basins. A new method of ordination of genetic landscapes successfully assigned species into cluster groups based on their respective pattern of genetic divergence across the Maroni Basin: genetically homogenous species across the basin were effectively discriminated from species showing high spatial genetic fragmentation and possible lower capacity for dispersal.

Keywords

Cytochrome *c* oxidase subunit I, species identification, French Guiana, Suriname, genetic divergence, ichthyodiversity.

Introduction

With over 32,000 known species, fishes represent more than half of the total number of vertebrates. Although 3,900 new species have been described during the last decade (Nelson, Grande, & Wilson, 2016) and 100 are described per year in the Neotropics only (Birindelli & Sidlauskas, 2018), the ultimate goal of cataloguing all fishes is still far from being achieved. DNA barcoding using the mitochondrial cytochrome *c* oxidase I gene, or COI (Hebert, Cywinska, Ball, & DeWaard, 2003), has proven to be a powerful tool to quicken and facilitate the global effort of species identification and discovery (Barrett & Hebert, 2005; Goldstein & DeSalle, 2011; Gomes, Pessali, Sales, Pompeu, & Carvalho, 2015). This has led to the foundation of the BOLD platform, an ever-growing COI database of animal organisms (Ratnasingham & Hebert, 2007). In this context, several studies have already been carried out on the megadiverse Neotropical freshwater fish community with focuses on rivers and lakes of Brazil (de Carvalho et al., 2011; Nascimento et al., 2016; Berbel-Filho et al., 2018), Argentina (Rosso, Mabragaña, González Castro, & Díaz de Astarloa, 2012; Díaz et al., 2016), Mexico, and Guatemala (Valdez-Moreno, Ivanova, Elías-Gutiérrez, Contreras-Balderas, & Hebert, 2009). All of them have emphasised the efficacy of DNA barcoding for this model with successful species discrimination rates ranging from 90 to 100%. In this context, the Gui-BOL project by Covain et al. (http://www.boldsystems.org/index.php/MAS_Management_DataConsole?codes=GBOL) is a work group affiliated to the FishBOL campaign (Ward, Hanner, & Hebert, 2009) that aims at building a reference DNA barcode database for all fishes of the Guianas.

With a length of 400 km, a 68,700 km^2 catchment area and a mean discharge of 1780 m^3/s, the Maroni River is one of the largest rivers of the Guianas (Amatali, 1993; Négrel & Lachassagne, 2000). The Maroni and the Mana rivers, which share the same mouth estuary, constitute the Western French Guiana freshwater ecoregion characterised by its specific faunistic assemblage and its high endemism, and splits the fish fauna of Suriname from the one of French Guiana (Lemopoulos & Covain, 2019). It is also a region of faunal exchanges between the adjacent Surinamese Ecoregion to the west and the Central and Eastern French Guiana ecoregions to the east. Strong faunistic relationships with the Amazon Basin have also been highlighted, and the Maroni is supposed porous to fish dispersal from tributaries of the Amazon to the south (Cardoso & Montoya-Burgos, 2009; Fisch-Muller, Mol, & Covain, 2018; Lemopoulos & Covain, 2019). The Maroni River and its tributaries, among other watersheds of the Guiana Shield, have been extensively studied these past decades in an effort to inventory and describe its ichthyofauna. The latest complete checklists of freshwater fishes of French Guiana (Le Bail et al., 2012) and Suriname (Mol, Vari, Covain, Willink, & Fisch-muller, 2012) reported

the occurrence of 336 fish species in the Maroni Basin, including more than 250 species strictly restricted to freshwaters, from 15 orders and more than 50 families. This makes it the most speciose river basin of both countries, hosting one third of the total number of fish species of the Guiana Shield (Vari, Ferraris, Radosavljevic, & Funk, 2009). However, several species catalogued in these checklists still have an undefined status, while some other have been named with doubts on their actual taxonomical identity. Some species identified based on morphological and meristic methods display an intriguing patchwork distribution across the Guianas (e.g. *Leporinus nijsseni*described from Suriname River with occurrences reported in Nickerie River to the west of Suriname and Oyapock River to the east of French Guiana (Mol, 2012)). Furthermore, few studies have used molecular methods to better apprehend the faunistic richness of this basin, with the exceptions of some enzymatic and molecular sequencing approaches on*Leporinus* (Planquette & Renno, 1990) and Loricariidae (Covain et al., 2012, 2016; Fisch-Muller, Montoya-Burgos, Le Bail, & Covain, 2012; Fisch-Muller et al., 2018; Weber, Covain, & Fisch-Muller, 2012) or environmental DNA surveys (Cilleros et al., 2019; Murienne et al., 2019). These elements make the stability of these reference lists doubtful and incomplete, while they are essential for the management of these natural areas currently facing growing anthropic pressure.

While the accurate assessment of species richness is a fundamental prerequisite in the effective study and management of this megadiverse river system, the extensive genetic and geospatial data provided by a sampling campaign of this scale can also put into light global patterns of genetic connectivity among populations. As an example, observation of molecular data from the *Guyanancistrus* genus seems to indicate that the Maroni could be divided into a West (Suriname) and an East (Guianese) assemblage with lower genetic connectivity between them (Fisch-Muller et al., 2018). Finding evidence for high genetic divergence between these regions regardless of distance could strengthen this hypothesis. There is a growing number of methods and tools available to use genetic and spatial information in order to explore the biogeographical patterns of species and populations (Chan, Brown, & Yoder, 2011). One of them is the mapping of spatial patterns through "genetic landscapes" (Manel, Schwartz, Luikart, & Taberlet, 2003). This method can assist in identifying divergence hotspots (Wood et al., 2013) and potential barriers to gene flow (Vodă, Dapporto, Dincă, & Vila, 2015) and has already been used on genetic distances between COI sequences to provide a visual framework of genetic variation between organisms across space (Arbeláez-Cortés, Milá, & Navarro-Sigüenza, 2014; Mamos, Wattier, Burzyński, & Grabowski, 2016).

The present study makes use of a large new dataset of DNA barcodes to (1) assess the validity of the current references on the Maroni's freshwater fish species richness, (2) reveal the presence of genetic heterogeneity in order to flag potential cryptic species, and (3) investigate spatial genetic distribution within species that may reveal obstacles to connectivity across the watershed, or the presence of recent colonisation from adjacent rivers. The last point was approached with a new method of multivariate clustering of genetic landscapes.

Materials and Methods

Ethics statement

No endangered or internationally protected species at time of collection (local restrictions, IUCN or CITES listed species) were concerned by the study. Most specimens and tissue samples were obtained from Museum collections and/or by local populations or fishermen. No experimentation was conducted on live specimens. For specimens and associated tissue samples obtained from the field, specimens were collected and exported with appropriate permits: Ministry of Agriculture, Animal Husbandry and Fisheries to export fishes from Suriname in 2008. Material obtained from the Parc Amazonien de Guyane (PAG) in 2014 and 2015 was collected under the direct supervision of PAG authorities. When collecting occurred in non-protected areas of French Guiana, sampled specimens were equally declared to the French DEAL (French environmental protection ministry) before export. Immediately after collection, fish were anesthetised and sacrificed using water containing a lethal dose of eugenol (essential oils of cloves). All the work has been conducted in accordance with relevant national and international guidelines, and conforms to the legal requirements (Directive

2010/63/EU of the European Parliament and of the Council on the protection of animals used for scientific purposes, the Swiss ordinance OPAn 455.1 of OSAV, and recommendations and regulations of DT-OCAN).

Specimens collection and sampling area

All 1,284 specimens were collected on 83 sampling points across the Maroni River (1 to 18 specimens per location per species) and some of its main tributaries including Tapanahony, Paloemeu, Lawa, Litany, Tampok and Marouini rivers, as well as some more remote headwaters locations like Mitaraka Mountains or Saül uplands (Figure 1). Specimens were collected between 1997 and 2015 as part of a broader project on the ichthyological biodiversity and the ichthyofaunistic assemblages of the Guianese ecoregions (*sensu* Lemopoulos & Covain, 2019). A piece of fin or muscle tissue was collected from each specimen and stored in 80% ethanol at -20°C. To conform to the Barcoding Of Life recommendations (Ratnasingham & Hebert, 2007) and provide vouchered references, 764 specimens were fixed in 5% formaldehyde at room temperature or in ethanol at -20°C for long-term conservation and deposited in the MNHG fish collection. Six specimens were stored in the Museum of National History of Paris, six in Auburn University of Alabama, and two in the Academy of Natural Sciences of Philadelphia. The remaining 506 specimens are vouchered by tissue only, sometimes completed by a photograph. Most of these were large specimens returned to local fishermen.

Fish preliminary identification and Maroni species coverage

Fish were morphologically identified at the species level based on literature (Planquette, Keith, & Le Bail, 1996; Keith, Le Bail, & Planquette, 2000; Le Bail, Keith, & Planquette, 2000; Mol, 2012). Fish taxonomical classification follows Le Bail et al. (2012) and Mol et al. (2012), with taxonomic updates following Fricke, Eschmeyer, & Van Der Laan (2019) (Table 1). Out of the 264 freshwater fish species certainly known to occur in the Maroni Basin according to the last checklists (Le Bail et al., 2012; Mol et al., 2012), 174 were collected, representing 65.91% of the total number of already known species (Supplementary Material 1). Additionally, five species collected during this study were not known to occur or were considered dubious in the basin in the 2012 checklists (*Charax gibbosus* , *Guyanancistrus megastictus* ,*Krobia guianensis* , *Poecilia bifurca,* and *Tomeurus gracilis*).

Extraction, PCR amplification, and DNA sequencing

Total genomic DNA was extracted with the E.Z.N.A. Tissue DNA Kit (Omega Biotek) following the instructions of the manufacturer. The PCR amplifications of COI were carried out using the Taq PCR Core Kit (Qiagen) following Covain et al. (2012). Primers and their taxonomic targets are listed in Table 2. Cycles of amplification were programmed following the profile: (1) 3 min at 94 °C, (2) 30 s at 94 °C (initial denaturing), (3) 40 s with annealing temperature ranging between 51 and 54 °C depending on primers used, (4) 40 s to 1 min at 72 °C (elongation), (5) 10 min at 72 °C (final elongation). Steps 2–4 were repeated 40 times (42 with 5COI-F / COI-R3). Some samples were amplified by Touchdown PCR following the protocol of Korbie & Mattick (2008). Purification and sequencing of PCR products were performed at Eurofins Genomics (France) and Macrogen Europe (The Netherlands) using Sanger method (Sanger, Nicklen, & Coulson, 1977). DNA sequences were edited using BioEdit 7.2.5. (Hall, 1999) and aligned with MUSCLE (Edgar, 2004). Edited sequences were deposited on BOLD with corresponding vouchers.

Fish molecular identification and barcode analysis

The barcode sequence data was used in conjunction with morphology and known geographical distribution (an integrative approach similar to Gomes et al. (2015) and Pugedo, de Andrade Neto, Pessali, Birindelli, & Carvalho (2016) to flag potential identifications errors, identify unknown specimens (e.g. juveniles), and detect putative cryptic species. Molecular identification was performed with BOLD Identification System (www.boldsystems.org), BLAST search (Altschul, Gish, Miller, Myers, & Lipman, 1990) on NCBI (http://www.ncbi.nlm.nih.gov/BLAST) and by neighbour-joining tree based identification. Specimens

showing a combination of unexpected intra-specific genetic divergence (both in our dataset and on the global BOLD database), overlooked morphological cues, and support from distribution patterns were flagged as putative cryptic species, given provisional new names, and treated as distinct species in all subsequent analyses.

All genetic distance analyses were performed under the Kimura two-parameter (K2P) substitution model (Kimura, 1980), as it is a standard metric in barcode studies (Ward, 2009; Díaz et al., 2016). Sequence divergences at the Species, Genus, and Family level were estimated using the BOLD Distance Summary Tool (BOLD Aligner, pairwise deletions). BOLD's Barcode Gap Analysis (same parameters) was used to investigate species who do not comply with barcode gap (i.e. for which distance to the Nearest Neighbour (NN) is lower than the standard barcode threshold of 2% or lower than the maximum intra-specific distance). The BIN Discordance Report tool was used to analyse the final dataset using the clustering method provided by BOLD: the Barcode Index Number (BIN, Ratnasingham & Hebert, 2013) which is the standard method in barcode studies to attribute each specimen to a new or pre-existing operational taxonomic unit. A neighbour-joining dendrogram of BOLD-aligned K2P distances was built with BOLD's Taxon ID Tree and modified with MEGA 7.0.26 (Kumar, Stecher, & Tamura, 2016) to visualise total clustering of BINs and species.

Genetic divergence landscape analysis

Patterns of genetic divergence among species were represented by genetic landscapes using the Inverse Distance Weighting (IDW) interpolation as in Vandergast, Perry, Roberto, & Hathaway (2011). We focused on intra-specific genetic divergence instead of genetic variation because of the relatively low number of specimens per species collected at each sampling point and used K2P distances as the metric to stay consistent with the barcoding approach. Genetic landscape analysis (Supplementary Material 2) was performed in R v3.5.0 (R Core Team, 2018) based on the location data (catch coordinates) of the 1,284 specimens and their respective barcode sequences. In order to take into account the intra-locality sequence variation of specimens, all equal pairs of coordinates within species were added a constant term of one or more fifth digit on their latitude value. This transformation only changed the recorded sampling point by a few meters, which conforms to the reality of the fishing area on the field. Matrices of pairwise K2P distances were computed among specimens within each species with ape 5.3 (Paradis & Schliep, 2018). The 55 species with less than three specimens and the 16 species with only one sampling point were discarded from landscape analysis.

A Mantel test was performed with 9,999 permutations using ade4 1.7-13 (Dray & Dufour, 2007) to evaluate absence of relationships between geographic and non-null K2P genetic distances for each species. All obtained p-values were corrected for multiple comparisons with the false discovery rate method (Benjamini & Hochberg, 1995). If a species showed a corrected p-value lower than 0.05, actual K2P distances were corrected by fitting a linear model between genetic and geographic distances, residuals were then used to compute the IDW interpolation to remove effect of inter-location distances in genetic distance. If the p-values were greater or equal to 0.05, IDW was performed directly on uncorrected K2P distances.

For IDW, midpoints coordinates between pairs of samples locations connected by a Delaunay triangulation (i.e. the smallest network with non-overlapping edges) were extracted for each species with the phylin package 2.0 (Tarroso, Carvalho, & Velo-Antón, 2019). Respective K2P distance values (or geographically corrected K2P distances) between pair of samples were assigned to each respective midpoint, and IWD (as implemented in phylin with default weighting method "Shepard") was used to interpolate the Z values of each coordinate in the grid within the whole species' sampling area. Real and interpolated K2P distance values or residuals were normalised to enable comparison of heterogeneous genetic divergence rates among species and then projected on basin maps using the raster package 2.9-5 (Hijmans, 2019). IDW interpolations of species for which all pairwise distances were equal to zero were calculated and projected directly using these raw zero values. The global patterns of genetic divergence across the basin ("multispecies landscape") were visualised by projecting the arithmetic mean of all previously generated normalised overlapping surfaces.

We then sought to ordinate genetic landscapes based on similar genetic divergence patterns in order to explore shared trends among species and ultimately assign each species to a broader cluster group. Ordination was

performed through Principal Component Analysis (PCA). For this, a genetic landscape data table was compiled with Z values as estimates of genetic divergence for each species listed in rows and each grid cell in the basin in columns, resulting in a table of 129 rows and 60,675 columns. Because genetic landscapes had different sizes depending of the species sampling area, many Z values were missing in the table. These missing values were replaced by the values for the corresponding coordinate pair from the multispecies landscape with MCRestimate package 2.38-0 (Johannes et al., 2018). The multispecies landscape was then treated as the "null genetic landscape" of the basin. The PCA was performed using the variance-covariance matrix of the data table using ade4 1.7-13. Since the variables consisted of tens of thousands of coordinate points, their projection on the principal components were visualised as points with two separate colour gradient codes respectively representing latitudes (from red to orange to yellow) and longitude (purple to blue to green). In order to cluster genetic landscapes into general patterns, the canonical pairwise distance matrix of the row coordinates of informative axes of the PCA was submitted to an agglomerative hierarchical clustering analysis. These axes were chosen in order to explain most of inertia without being affected by artefactual effects. The suitability of eight clustering methods were compared, and the Weighted Pair Group Method with Arithmetic Mean (WPGMA) algorithm (McQuitty, 1966) on Euclidean distances was retained for having the highest cophenetic correlation coefficient of 0.83. The package pvclust 2.0-0 (Suzuki & Shimodaira, 2006) was used to calculate two types of p-values on each cluster node via multiscale bootstrap: Approximately Unbiased (AU) (Shimodaira, 2002, 2004) and Bootstrap Probability (BP) (Efron, 1979; Felsenstein, 1985). The analysis was performed using 999 pseudoreplicates. The "mean landscapes" of the main resulting clusters were visualised by projecting the arithmetic mean of all genetic landscapes nested in said clusters.

Results

Barcode dataset

The 1,284 COI sequences collected from the Maroni Basin generated a final barcode data library of 199 species (125 Genera, 36 Families, and 8 Orders) after integrative re-identification (Figure 2). Five out of these species were flagged for being putative re-identifications from the latest Maroni checklists, with the nominal species being presumably absent from the basin (Table 3), and 20 are putative cryptic species that were revealed by our integrative approach for identification, among which 13 are already BIN-concordant on the BOLD database (Table 3). All species in the data set belong to the class Actinopterygii with the exception of *Potamotrygon marinae* from the class Elasmobranchii. The number of individuals per species ranged from 1 to 59 (mean=6.45) with 169 species represented by more than one specimen (84.92%). All final sequences were devoid of stop codons, insertions, or deletions. Mean nucleotide frequencies for the total alignment were 24.7% adenine, 27.0% cytosine, 18.4% guanine and 29.9% thymine. The mean K2P genetic divergence was 0.32% within species, 12.46% within genus, and 19.61% within family (Table 4).

Species delimitation

Barcode Gap Analysis showed a distance to the Nearest Neighbour (dNN) greater than 2% and greater than the maximum intra-specific distance for 187 species out of 199. Low dNN (<2%) was observed in only six pairs of species: *Ancistrus* cf. *leucostictus* / *Ancistrus temminckii* (dNN=1.84%), *Corydoras geoffroy* / *Corydoras* aff. *geoffroy* (dNN=0.12%), *Guyanancistrus brevispinis* / *Guyanancistrus nassauensis* (dNN=0.62%),*Hypostomus plecostomus* / *Hypostomus watwata* (dNN=1.94%),*Melanocharacidium* sp. 1 / *Melanocharacidium* sp. 2 (dNN=0.97%) and *Pimelodella geryi* / *Pimelodella* aff.*geryi* (dNN=1.47%). Among them, maximum intra-specific distance exceeded or was equal to their dNN in the following two species only:*Corydoras geoffroy* (NN=*Corydoras* aff. *geoffroy*) and*Guyanancistrus brevispinis* (NN=*Guyanancistrus nassauensis*) (Table 5).

The neighbour-joining dendrogram showed no overlap between any species, except only for *Corydoras* aff. *geoffroy* , which was nested within the *Corydoras geoffroy* cluster (Figure 3, Supplementary Material 3). Two hundred twenty nine BINs were assigned by BOLD to the total number of samples in the dataset, among which 171 were taxonomically concordant (represented by 1,195 specimens) and 55 were singletons

(i.e. were assigned to only one specimen from this dataset). The remaining three BINs (32 specimens) showed some discordance and were represented by three pairs of species that shared the same BIN in our dataset: *Guyanancistrus brevispinis* / *Guyanancistrus nassauensis* , *Corydoras geoffroy* / *Corydoras* aff.*geoffroy* and *Melanocharacidium* sp. 1 /*Melanocharacidium* sp. 2 (Figure 3). Two specimens (one*Serrasalmus rhombeus* and one *Peckoltia otali*) were not assigned any BINs, probably due to a high proportion of ambiguous bases for the latter (14Ns/652bp). However, both specimens clustered perfectly with their conspecifics, so they can be reasonably considered well identified.

Among concordant BINs, species that were assigned more than one BIN in our dataset are: *Aequidens tetramerus* , *Bryconamericus guyanensis* , *Characidium zebra* , *Cleithracara maronii* ,*Crenicichla multispinosa* , *Eigenmannia virescens* ,*Erythrinus erythrinus* , *Farlowella reticulata* ,*Gymnotus anguillaris* , *Hemiodus huraulti* , *Hoplias malabaricus* , *Hypopygus lepturus* , *Ituglanis amazonicus* ,*Moenkhausia moisae* , *Moenkhausia oligolepis* ,*Nannostomus bifasciatus* , *Pimelodella leptosoma* ,*Poptella brevispina* with two BINs; *Cetopsidium orientale* ,*Crenicichla albopunctata, Helogenes marmoratus* ,*Hyphessobrycon roseus* , *Phenacogaster wayana* ,*Pimelodella* cf. *cristata* with three BINs, and*Gasteropelecus sternicla* with four BINs (Figure 3, Supplementary Material 3). This result suggest that these 25 species could hide some unexpected diversity and that some of them may be potential cryptic species complexes. However, they were not flagged as such for the current study due to the lack of strong morphological and / or molecular evidence.

Genetic divergence landscape analysis

A genetic landscape was produced for 128 species out of 199 (Supplementary Material 3). Twenty-one of them showed no genetic divergence across the basin, while 107 displayed various patterns. Fifteen out of these 107 species showed a significant relationship between genetic and geographic distances after false discovery rate correction (*Bryconops affinis* , *Bryconamericus guyanensis* ,*Corydoras geoffroy* , *Crenicichla multispinosa* ,*Cteniloricaria platystoma* , *Gasteropelecus sternicla* ,*Geophagus harreri* , *Harttia guianensis* , *Lithoxus*aff. *planquettei* , *Metaloricaria paucidens* ,*Moenkhausia grandisquamis* , *Myloplus ternetzi* ,*Poptella brevispina* , *Pseudancistrus barbatus* and*Rineloricaria* aff. *stewarti* 3). Accordingly, Z values for these species were computed on K2P distances between residuals of the linear model instead of the raw K2P distances. Z values of the multispecies landscape ranged from 9.43E-06 to 0.81 and number of species contributing to the calculation of each cell ranged from one to 95 (Figure 4). The lowest and highest mean genetic divergences were located at specific points of the basin with a low number of species sampled, i.e. the mouth of the Maroni River and the Saül uplands. The remaining of the map had a better species coverage and showed that the highest mean genetic divergences were observed between the West Upper Maroni (Tapanahony and Paloemeu rivers) and the East Upper Maroni (Lawa, Litany, Tampok and Marouini rivers), as well as between the West Upper Maroni and the Lower Maroni and Nassau Mountains. Relatively high divergence was also observed between the East Upper Maroni and the Lower Maroni, as well as between the Tampok and the Marouini. As a whole, the basin was divided into three large regions displaying high genetic divergences between each other without influence of geographic distances: the West Upper Maroni, the East Upper Maroni, and the Lower Maroni, with West Upper Maroni being the most divergent of the three.

Ordination of genetic landscapes patterns

Most of the genetic landscape structures were explained by the first three axes of the PCA, which accounted for 40.61%, 16.8%, and 7.70% of total inertia respectively (Figure 5). The multispecies landscape computed on the mean Z values of all analysed species was always projected at the centre of axes and effectively acted as a "null landscape" from which other landscapes were ordinated. Axis 1 aligned species with globally low genetic divergence across the basin in negative values (*Pimelodus ornatus* , *Hypostomus gymnorhynchus* , *Ageneiosus inermis* , *Serrasalmus rhombeus* ...) with species with high genetic divergences in positive values (*Helogenes marmoratus* , *Metaloricaria paucidens* , *Bryconamericus guyanensis* , *Curculionichthys*sp. Maroni...). Examination of variables revealed that most loadings were positive with only few of them around zero or weakly negative, revealing potential size effect in the ordination along the first axis (Figure 6). Accordingly, this axis was discarded from further clustering analyses. Axis 2 mostly expressed a latitudinal

influence in landscape ordinations. Landscapes projected on negative values of this axis had lower genetic divergence in the north. They included high coverage landscapes with global genetic divergence that was either mostly low (*Auchenipterus nuchalis* , *Serrasalmus rhombeus* ...) or mostly high (*Platydoras costatus, Myloplus ternetzi...*). Conversely, landscapes with higher genetic divergences in the north were projected in positive values of axis 2, including high coverage landscapes like *Gasteropelecus sternicla*or *Leporinus fasciatus* and low coverage ones like*Prochilodus rubrotaeniatus* or *Leporinus granti* (Figure 5). Projection of variables supported the split between the south (red) in negative values and the north (yellow) in positive values (Figure 6). Axis 3 displayed a pattern of opposition between higher genetic divergence in the west (*Pimelodella leptosoma* ,*Guyanancistrus brevispinis* , *Cteniloricaria platystoma* ,*Curculionichthys* sp. Maroni...) in negative values and higher genetic divergence in the east (*Triportheus brachipomus* ,*Jupiaba keithi* , *Semaprochilodus varii* ...) in positive values (Figure 5). This pattern was supported by variable projections with most western longitudes (green) projected in negative values and eastern longitudes (purple) in positive values (Figure 6).

The WPGMA of axes 2 and 3 of the PCA clustered the 129 genetic landscapes (including the multispecies) into nine main cluster groups noted A to I (Figure 7). Twenty-eight of the clusters, comprising 106 landscapes, were supported by internal nodes all having an AU greater than 95%, but nodes supports were strong overall with most of the nodes having an AU greater than 80%. Group A included two small strongly supported clusters. The first cluster included *Guyanancistrus brevispinis* and *Pimelodella leptosoma* , the two species with the lowest loadings on axis 3. As stated above, both landscapes displayed a strong genetic isolation between the Upper West and the rest of the basin. The second cluster included the four landscapes with the lowest loadings on axis 2, i.e. landscapes where the global genetic divergence across the basin is mostly low, but the highest divergence is in the south (*Auchenipterus nuchalis* , *Serrasalmus rhombeus* ,*Hypostomus gymnorhynchus* ...). Group B was a strongly supported cluster of the four species with the greatest genetic divergence in the north, which accordingly had the highest loadings on axis 2 (*Gasteropelecus sternicla, Helogenes marmoratus, Eigenmannia virescens,* and *Leporinus fasciatus*). Group C included 15 species that also projected on positive values of axis 2 and displayed particularly high genetic homogeneity in the south, with the notable exceptions of *Curculionichthys* sp. Maroni and*Cteniloricaria platystoma* , which should probably be part of their own separate cluster (cf. genetic landscape patterns in Supplementary Material 4). Group D grouped together the species that displayed the highest genetic heterogeneity among southeast locations compared to the rest of the basin, and had the highest loadings on axis 3 (e.g. *Ageneiosus inermis* , *Semaprochilodus varii* ...). On the opposite, species in group E all showed the highest genetic divergence between the south-west and the rest of the basin (e.g. *Moenkhausia oligolepis* , *Bryconamericus guyanensis* , *Anostomus brevior* ...) with negative loadings on both axes 2 and 3. Group F contained landscapes which all projected on negative values of axis 2, with patterns very similar to group E in that they all showed a high Upper West / Upper East divergence (e.g.*Hypopomus artedi* , *Pimelodella* cf. *cristata...*) but also included more patchy patterns (e.g. *Moenkhausia intermedia* , *Nannostomus bifasciatus* ...). Group G is a small strongly supported cluster characterised by a high genetic homogeneity in the Upper East, as shown by low values on axis 2.

Groups H and I consist of 63 landscapes (almost half of the total) that were either too small in area, displayed unusual patterns, or consisted of species that were from under-sampled parts of the basin (i.e. not the mouth, the Tapanahony or the Saut Wayo / Langa Sula region). As expected, the multispecies landscape was part of one of these groups (group H), being the null landscape from which all others where compared in the multivariate analysis. Accordingly, the six species with the smallest sampling areas were grouped with it in a strongly supported cluster. Landscapes from groups H and I were all projected close to the centre of axes two and three (but not always on axis 1), and nodes within these groups were generally less well supported than in the rest of the tree.

Discussion

Checklists updates: Re-identified species

Five species that were thought to be present in the Maroni according to the current checklists are now strongly suspected to be different species after examination of their respective COI haplotype compared to con-specific individuals from other basins. We thus flagged the specimens caught in the Maroni with the following provisional names:*Imparfinis* aff. *pijpersi* , *Lithoxus* aff.*planquettei* , *Loricaria* aff. *nickeriensis* ,*Pachypops* aff. *fourcroi* , and *Pimelodella* aff.*cristata* (Table 3). *Pachypops* aff. *fourcroi* has black spots on the back that are not present in the nominal species.*Pimelodella* aff. *cristata* is one of the two putative species that was probably misidentified as *Pimelodella cristata*in the Maroni (the other being *P* . cf. *cristata* that is here considered as a cryptic species of *P.* aff. *cristata* , see below). Tree based identification on the global database shows that this cluster of specimens is widely distributed (several rivers of Suriname and French Guiana) and actually branches far from *P. cristata* and *P.* cf. *cristata* , with its sister barcode species being *Pimelodella macturki.* However, although all*P* . aff. *cristata* on BOLD share the same BIN, it also includes two specimens identified as *P. vittata* and seven*Pimelodella* sp., making it currently discordant.

Checklists updates: Putative new cryptic species

The 13 putative new species that were BIN concordant on the BOLD database (Table 3) are as follows: (1) *Ancistrus* sp. is represented by one specimen from the remote Mitaraka Mountains that had a similar morphology to *Ancistrus temminckii* . However, Tree Based Identification analysis (Figure 3, Supplementary Material 3) placed it as a distinct sister group to a clade composed of *Ancistrus* cf.*leucostictus* and *A. temminckii.* (2) *Bryconops* aff.*melanurus* : two specimens from Langa Sula and Saut Wayo (Marouini) showed a dNN greater than 12% with the four *Bryconops melanurus*that were caught in the same two locations. They share the same BIN with four specimens from Sinnamary, Suriname and Mana rivers that were also re-identified as *B.* aff. *melanurus* on BOLD database. (3) The eight specimens of *Corydoras* aff. *guianensis* , differing from the nominal species by a faint black margin along the dorsal-fin spine, constituted the sister group of the blunt snouted*Corydoras* aff. *breei* (lineage 9 in Alexandrou & Taylor (2011)) and the nominal species *C. guianensis* in our dataset with a minimum K2P distance of respectively 5.49% and 5.13%. The five*C. guianensis* specimens available were caught in the Tapanahony and Paloemeu rivers in Suriname (type locality of *C. guianensis*being Nickerie River in Suriname) while the eight *C.* aff.*guianensis* came from the east tributaries in French Guiana. (4)*Cyphocharax* aff. *spilurus* 2 was first morphologically identified as *C. spilurus* but showed a dNN of 13.84% with all other *C. spilurus* and 8.59% with *C* . cf. *spilurus*(see below), while the shortest dNN was 7.40% with *C. biocellatus.* This specimen was caught in the same location as seven*C. spilurus* in the Tapanahony, while the two other *C. spilurus* were caught in the east tributaries. It displays a large dark spot on the caudal peduncle, and shares a BIN on the BOLD database with another *Cyphocharax* aff. *spilurus* 2 from Sipaliwini River (Suriname). Although not present in the two checklists, *C. aff. spilurus* 2 was already known to occur in the Maroni and suspected to live in sympatry with *C. spilurus* (Planquette et al., 1996). This putative species is different from *C.* aff. *spilurus sensu*Le Bail et al. (2012) described from Approuague and Iracoubo. (5) The three juvenile specimens of *Guianacara* sp. Tapanahony showed a minimum dNN of 3.26% with *G. owroewefi* from our dataset. They were all caught in the Paloemeu River in sympatry with *G. owroewefi* . None of them display the diagnostic pattern of *G. oelemariensis* described from upper Marowijne River in Suriname and were thus treated as an unknown species. (6) *Hemigrammus* aff.*guyanensis* from Litany River likely corresponds to the form illustrated for *H. guyanensis* in Planquette et al. (1996). However, the true *H. guyanensis* is illustrated in Mol (2012) as*H* . aff. *ocellifer* . Specimens from Tapanahony River correspond to this form and are here re-identified as *H. guyanensis* . They are characterised by a longitudinal black line, and a distinct iridescent spot on upper part of caudal peduncle. Specimens from Litany lack this last characteristic. (7) *Leptocharacidium*sp.: This non-identified specimen caught in Wayu Camp (Paloemeu) and firstly assigned to *Melanocharacidium* was peculiar in showing a very high genetic distance from all other closest species in our dataset (dNN=17.71% with *Parodon guyanensis* instead of another Crenuchidae species) and on the BOLD database (dNN=14.04% with a*Characidium sp.* from Bolivia). It also has an unusual morphology, i.e.

a small operculary spot and a longitudinal line, but lacks the diagnostic two unbranched rays in pelvic fins. It is tentatively placed in *Leptocharacidium* , awaiting further analyses. (8) Six specimens first identified as *Moenkhausia collettii* from Litany and Sector Apsik were renamed *Moenkhausia* aff. *collettii* after showing a very high dNN of 15.33% with previously recorded *M. collettii* from French Guiana in our dataset. (9) *Nannacara* sp. Litany was represented by a single specimen in his own BIN. It is the first occurrence of a *Nannacara* found in the Maroni basin and showed a minimum dNN of 10.49% with other *Nannacara* from the BOLD database, including the two *Nannacara* species that occur in French Guiana: *N. aureocephalus* and *N. anomala* . (10) The ten *Pimelodella geryi* specimens clustered into two clear groups in the neighbour-joining tree (Supplementary Material 3). The first group included specimens from the west (Tapanahony) and the east of the basin, while the second cluster included only specimens from the west. Although the dNN between these two groups is relatively low (1.47), the second cluster exclusive to the west basin has its own concordant BIN and was flagged as a potential *P.* aff. *geryi* . (11)*Pimelodella* cf. *cristata* : Although *Pimelodella cristata* is present in the Maroni checklists, tree based identification of available close specimens from BOLD database coupled with observation of catch localities seem to indicate that the nominal species, described from Takutu River in Guyana, is present in the Corantijn River in Suriname but not in any basin of French Guiana. Instead, two potentially undescribed species formerly identified as *P. cristata* occur in the Maroni: *Pimelodella* cf. *cristata* and*Pimelodella* aff. *cristata* . According to Tree Based Identification, *Pimelodella* cf. *cristata* , which is present in several rivers of Suriname as well as the Maroni, is the closer of the two from the nominal *P* . *cristata* (although its NN is actually *Pimelodella* sp.). It is represented by five specimens in our dataset that are split in three concordant BINs. (12) One unidentified *Pimelodella* sp. from Paloemeu shows a high distance with its NN (*P* . *geryi* , 8.59) and its sister group (*P* . *megalops* , 9.58) in our dataset. Tree based identification on BOLD database showed that this specimen is most closely related to another group of unidentified *Pimelodella* from another basin in Sipaliwini. (13) Three unidentified*Rineloricaria* specimens were caught in the Paloemeu River and were named *Rineloricaria* aff. *stewarti* 2. They constitute a sister species of an undescribed species from western Suriname. Their dNN in our dataset was 9.95% with *Rineloricaria* aff.*stewarti* 3 (*Rineloricaria* aff. *stewarti sensu* Le Bail et al. (2012)), a species broadly distributed in French Guiana and eastern Suriname.

Four putative cryptic species were BIN concordant in our dataset but not on the BOLD database (Table 3): (1) *Cyphocharax* cf.*spilurus* sister group was a group composed of *C.biocellatus* specimens in our dataset, as opposed to the *C .spilurus* cluster that branched further. However, this specimen showed morphological differences (e.g. absence of the lateral spot) with*C. biocellatus* and a high dNN of 7.27. On the BOLD database, it shares a BIN with five other *C.* cf. *spilurus*from other rivers in Suriname and seven specimens from Guyana identified as *C. spilurus* by other authors. (2) While all *Eigenmannia*cf. *limbata* specimens shared a BIN with specimens identified as*E. limbata* and *E* . *nigra* , tree based identification strongly implies that *E. limbata* is a Surinamese species while *E.* cf. *limbata* is a French Guianese species. Moreover, all *E* . cf. *limbata* specimens display a black humeral spot that is not present in *E. limbata.*(3) *Gymnotus* aff. *carapo:* one specimen from Wawapsi Creek (Paloemeu) showed a dNN of 7.30% with the *G. carapo* specimens in our dataset. It shares the same BIN with three specimens from Brazil identified as *G. carapo* on BOLD database. However, tree based identification shows that *G.* aff. *carapo* and these three Brazilian specimens constitute a group situated very far from all others*G. carapo* from several locations (Suriname, Brazil, Guyana and French Guiana) on BOLD database and that their sister group is composed of several *G* . *pantana* and unidentified *Gymnotus* . G*ymnotus carapo* is known to be a widespread species in South America and contains several sub-species (Craig, Crampton, & Albert, 2017). (4) Despite a similar colour pattern with the nominal species, one *Sternopygus* cf. *macrurus* from Tampok shows a high dNN of 13.96% with other *S. macrurus* specimens from our dataset. Tree based identification shows that K2P distances among specimens identified as *S. macrurus* on BOLD database are particularly high and that this species seems to consist of a complex of several operational taxonomic units. The two closest *S. macrurus* to our*S.* cf. *macrurus* came from Suriname and Ecuador.

Three putative cryptic species showed BIN discordance in our dataset after integrative identification (Table 3). (1) The two *Corydoras*aff. *geoffroy* specimens shared the same BIN as *C. geoffroy* . They are also nested

in the *C. geoffroy* complex in the neighbour-joining dendrogram. However, there is still strong assumptions that it could be a cryptic species. The main hint is that some unidentified *Corydoras* that live in sympatry with *C .geoffroy* in the Litany River (Sector Apsik) display a dark lateral stripe that is not present on *C. geoffroy* . Although these striped specimens share the same mitochondrial DNA sequence as *C. geoffroy* , we decided to flag these specimens as *C.* aff.*geoffroy* waiting for further evidence on the status of these morphs. (2) *Melanocharacidium sp. 1* (two specimens from Wawapsi Creek, Paloemeu) *and sp. 2* (one specimen from Wayu Camp, Paloemeu and two from Sector Apsik, Litany) share the same BIN and are a different group than other *Melanocharacidium* species on BOLD database with a dNN of 16.30% with *M. blennioides* . They were kept split as *sp. 1* and *sp. 2* as the Wayu Camp specimen in the Surinamese Paloemeu River displays a much shorter genetic distance with the two Litany specimens in French Guiana than with the other two specimens from Paloemeu, which could indicate a potential genetic flow barrier between sp.1 and sp.2 although they can occur in the same river. Species from the Crenuchidae family are still under-studied, as «undescribed Crenuchidae may be present in Suriname, especially in the Interior of the country» (Mol, 2012).

Barcode Gap

Apart from the new putative species cited above, only three pairs of species from the final dataset were below the 2% barcode gap (Table 5). Among them, the only ones that shared the same BIN in our dataset and on BOLD were *Guyanancistrus brevispinis* and *G. nassauensis*(dNN=0.62%). However, *G. nassauensis* is morphologically very distinct from *G. brevispinis* and has been formally described in detail (Fisch-Muller et al., 2018). Moreover, mitochondrial introgression between the two species has been reported, which explains this shared COI haplotype. The remaining two pairs, *Ancistrus* cf.*leucostictus* / *Ancistrus temminckii* (dNN=1.84%) and*Hypostomus plecostomus* / *Hypostomus watwata* (dNN=1.94%) still bear little doubt on their species delimitation. Each of the four species have their own associated BIN, and their dNN is still relatively high and close to the classical 2%. Moreover, *Hypostomus* is known to be a genera with low inter-specific COI divergence (de Queiroz et al., 2020). Overall, results from our dataset conform to the now well-established notion that the 2% barcode gap is a good start to flag low inter-species divergence (only six pairs of species show a dNN lower than <2% in our dataset), but that it does not always apply to some groups of fish (Pereira, Hanner, Foresti, & Oliveira, 2013; de Queiroz et al., 2020) and thus should not be interpreted as a stand-alone metric.

Genetic divergence landscape analysis

The multispecies landscape (Figure 4) computed using all available species, provided a mean reference pattern of genetic divergences across the Maroni basin. This pattern supports evidence for the hypothesis of lower levels of genetic connectivity between the west and the east of the basin. Especially, the Tapanahony / Paloemeu rivers in Suriname showed the highest mean divergence with both the east and the north with three divergence hotspots (average Z values between 0.5 and 0.6 in yellow in Figure 4). To assess the extent of this West-East bipartition, we briefly compared intra-basin and inter-basin genetic divergence of 31 species with enough sampling coverage across the Guianas on BOLD (Supplementary Material 5). Two groups were detected: a western group that includes the West Maroni (Tapanahony) and nine basins from Suriname and Guyana, and an eastern group that includes the East and North Maroni areas as well as six basins of French Guiana. The Mana (east) and the Corantijn (west) were the only rivers that were not assigned to their respective expected groups, but both are highly connected to the Maroni, either through the mouth (Mana) or through putative shared waters in the upstream Surinamese areas (Corantijn). This global bi-partition could be a consequence of multiple entries of species from the western and eastern rivers in the Maroni. However, although the trend is detectable, very few individuals among the species investigated displayed a lower genetic divergence with individuals from surrounding basins than with individuals from the Maroni (Supplementary Material 6), implying that faunal exchanges were rather ancient. The hypothesis of past migrations out of the Maroni to colonise surrounding basins, as it may have happened several time in the *Guyanancistrus brevispinis* complex (Fisch-Muller et al., 2018), might fit better with the general lack of

highly divergent haplotypes in the basin (but see *Ageneiosus inermis* , *Gymnotus* aff. *carapo* or *Sternopygus* cf.*macrurus* above for examples of recent entries from headwaters of Amazonian tributaries).

Ordination of individual patterns of the different species relative to the mean pattern reveals individual trends, and their respective contribution to the consensus (Figure 5). The method of genetic landscape reconstruction is sensitive to the sampling coverage of each species (all species not being collected in every place) given that genetic landscapes are interpolations from point data. Accordingly, resulting ordination patterns are highly dependent on the number and distribution of specimen captures across the landscape. However, if species displaying a broader coverage are better represented further from the centre of the axes, similarity in the grouping of patterns of genetic divergences along axes seems preserved regardless of the species sampling coverage. This result is reinforced by the hierarchical classification that created clusters of similar trends expressed by the different species independently from sampling coverage, providing a relatively good confidence in the various reconstructed cluster groups (Figure 7). For instance, cluster group D grouped species displaying higher divergence in the south-eastern part of the basin, a pattern contrasting with other results, including both species with broad sampling (*Ageneiosus inermis* and *Triportheusbrachipomus*) and species with smaller coverage (*Rineloricaria* aff. *stewarti* 3, *Hypomasticusdespaxi* , *Caenotropus maculosus* , *Jupiaba keithi* , or *Semaprochilodus varii*). An example of pattern revealed by the method is the higher genetic divergence in the west of the basin observed in several species (*G. brevispinis* and*P. leptosoma* from group A and all species from group E (Figure 7)). This observed pattern could result from the progressive establishment of hotspots of mutual exclusion among populations linked to more ancient and favoured dispersal routes from the Maroni Basin, something that has also been suggested for *G. brevispinis* , which first dispersed from the lower Maroni to the west toward the Suriname River and then toward Upper Corantijn River (Fisch-Muller et al., 2018).

A rapid examination of the clusters showed that they contain most taxonomic groups, including Siluriformes, Characiformes, Gymnotiformes, Cichliformes, etc. (Figure 7). In addition, comparison of local communities revealed equally likely distributed patterns among faunistic assemblages (Supplementary Material 5). The community of Langa Sula in the mainstream of Marouini River in French Guiana and the community of Wawapsi Creek, a small forest creek tributary of the Paloemeu River in Suriname comprise around 60 species each, but only 16 of them are shared between both. In both communities, including among shared species, all patterns of genetic divergence were present. Confrontation of pattern distribution relative to the taxonomy (which could be considered a proxy of the phylogeny) reveals weakly overdispersed patterns (i.e. weak negative autocorrelation of divergence patterns with the taxonomy) with absence of taxonomic structure among patterns (i.e. non-significant autocorrelation). Overdispersion of traits in regard of a phylogeny is often considered as a signature of competitive interactions driving community assemblage rules, meaning that dissimilar traits are expected to co-occur (Cavender-Bares, Ackerly, Baum, & Bazzaz, 2004; Cavender-Bares, Kozak, Fine, & Kembel, 2009; Pausas & Verdu, 2010).

Present faunistic assemblages result from interaction of multifactorial processes acting at different scales and operating over a wide range of time and space (Lowe & McPeek, 2014; Mittelbach & Schemske, 2015; Lemopoulos & Covain, 2019). In this context, dispersal abilities seem to be a key factor in species distributions, contributing to colonisation of new areas, range shifts and gene mixing among populations (Jonsson et al., 2016). Accordingly, observed genetic divergences among species likely result from limitations to dispersal, possibly related to competition, within the basin, responsible to limitation to gene flows and driving individual species response in assembly processes.

Conclusion

The 1,284 COI sequences introduced in this study represent the first extensive fish barcode dataset of the Maroni and the first dataset of this scale for the Guiana Shield. Using barcoding as a secondary tool to identify specimens during this study has been a success. Although the majority of specimens could be identified on the field with morphologic and meristic methods only, this traditional way of identification has its limits: post-larvae or small damaged specimens are sometimes impossible to classify based on morphological

keys (Ward et al., 2009). Species for which identification relies heavily on pigmentation (e.g. *Moenkhausia* and *Jupiaba*) are also highly problematic when identification is performed long after sampling because conservation in ethanol and formaldehyde results in the loss of pigments from the specimens. By building a DNA barcode library and relying on other sequences deposited by the scientific community, it was possible to easily verify some dubious identifications and pinpoint some overlooked diversity in our data. More than two thirds of the freshwater fishes of the Maroni now have at least one reference sequence available to facilitate identification. Biodiversity assessment in rivers have recently shown a trend of shifting from traditional "sampling by catch" methods to the less destructive use of environmental DNA metabarcoding (Taberlet, Coissac, Hajibabaei, & Rieseberg, 2012), a method that is currently being trialled in French Guiana rivers (Cilleros et al., 2019; Murienne et al., 2019). This method relies on solid reference databases to identify environmental DNA sequences to the species level or below, and our dataset is a step forward towards this goal.

The present study also updated the 2012 checklists in several ways. Five species absent or dubious in these checklists have been confirmed to occur in the Maroni. Five more species have been flagged as being in need of formal re-identification, and 20 new putative cryptic species have been added to the Maroni inventory. They all require further investigations to better assess their taxonomical status. Other updates include the change of taxonomical status of three recently described species (*Distocyclus guchereauae* Meunier, Jegu, & Keith 2014, *Guyanancistrus nassauensis* Mol, Fisch-Muller & Covain 2018, and *Mastiglanis durantoni* De Pinna & Keith 2019) and the new assignation of "doubtful" status to 13 species, based on the authors' extensive observations of the basin since 2012 (Supplementary Material 1).

We showed that the use of genetic landscape divergence is an efficient way to explore shared patterns of genetic connectivity when sampling is spatially extensive but number of specimens per locations is relatively low. Using a multivariate analysis (PCA) coupled with a hierarchical clustering method (WPGMA) proved to be an efficient way to classify these landscape patterns in order to facilitate the construction of species-specific genetic connectivity hypotheses. While this methodology is in no way restricted to river networks (on the opposite, land and sea may even be better suited to it), using this methodology on different basins and at different times could assist in the detection of habitat fragmentation, one of the biggest threats that freshwater ecosystems are currently facing. More extensive sampling, i.e. by using environmental DNA, would allow the application of this clustering landscape approach to compare levels of genetic diversity (as opposed to divergence) across the basin, another important indicator of the resilience capacity of species in an ecosystem.

Acknowledgements

We are grateful to Francois Meunier, Yves Fermon, and Philippe Keith, MNHN; Mark Sabaj and John Lundberg, ANSP; Philippe Gaucher, CNRS Guyane; Regis Vigouroux, Philippe Cerdan, and Sebastien Le Reun, Hydreco Guyane; Chrystelle Delord, Marie Le Noc, Marie Nevoux and Jean-Marc Roussel, INRA, Rennes; Michel Jegu, IRD; Mael Dewynter, Benjamin Adam and Antoine Baglan, Biotope Guyane; Frederic Melki, Biotope France; Jan Mol, University of Suriname; Kenneth Wan Tong You and Paul Ouboter, NZCS; Juan Montoya-Burgos, UNIGE; Sonia Fisch-Muller and Claude Weber, MHNG; Raphaelle Rinaldo and Guillaume Longin, Parc Amazonien de Guyane; Gregory Quartarollo and the Guyane Wild Fish Association ; Olivier Tostain, Ecobios Cayenne; Jonathan W. Armbruster, Auburn University; Sebastien Brosse, University of Toulouse ; as well as all our friends from Maroni River communities, Wayanas, Tekos, and Bushinengues for their contribution to field collection of specimens and logistic assistance. The French Guiana DEAL, PAG, and Prefecture; and the Surinamese Ministry of Agriculture, Animal Husbandry and Fisheries provided the necessary authorisations and collecting permits. Oceane Leclercq, Sabrina Grillard, and Alexandre Lemopoulos, MHNG are acknowledged for laboratory assistance. Oceane Leclercq benefited from a grant of the Biotope Foundation for Biodiversity, Meze, France.

References

Alexandrou, M. A., & Taylor, M. I. (2011). Evolution, ecology and taxonomy of the Corydoradinae revisited. In I. A. M. Fuller & H. G. Evers (Eds.), *Identifying Corydoradinae catfish: Aspidoras, Brochis, Corydoras, Scleromystax, C-numbers & CW-numbers: supplement 1.* (pp. 101–114). Kidderminster, England: Ian Fuller Enterprises.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* , *215* (3), 403–410. http://doi.org/10.1016/S0022-2836(05)80360-2

Amatali, M. (1993). Climate and surface water hydrology. In P. E. Ouboter (Ed.), *The Freshwater Ecosystems of Suriname* (Vol. 70, pp. 29–51). Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-94-011-2070-8

Arbelaez-Cortes, E., Mila, B., & Navarro-Siguenza, A. G. (2014). Multilocus analysis of intraspecific differentiation in three endemic bird species from the northern Neotropical dry forest. *Molecular Phylogenetics and Evolution* , *70* (1), 362–377. http://doi.org/10.1016/j.ympev.2013.10.006

Barrett, R. D. ., & Hebert, P. D. . (2005). Identifying spiders through DNA barcodes. *Canadian Journal of Zoology* , *83* (3), 481–491. http://doi.org/10.1139/z05-024

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing.*Journal of the Royal Statistical Society: Series B (Methodological)* , *57* (1), 289–300. http://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Berbel-Filho, W. M., Ramos, T. P. A., Jacobina, U. P., Maia, D. J. G., Torres, R. A., & Lima, S. M. Q. (2018). Updated checklist and DNA barcode-based species delimitations reveal taxonomic uncertainties among freshwater fishes from the mid-north-eastern Caatinga ecoregion, north-eastern Brazil. *Journal of Fish Biology* , *93* (2), 311–323. http://doi.org/10.1111/jfb.13758

Birindelli, J. L. O., & Sidlauskas, B. L. (2018). Preface: How far has Neotropical Ichthyology progressed in twenty years? *Neotropical Ichthyology* , *16* (3). http://doi.org/10.1590/1982-0224-20180128

Bivand, R., Keitt, T., & Rowlingson, B. (2019). rgdal: Bindings for the "Geospatial" Data Abstraction Library. Retrieved from https://cran.r-project.org/package=rgdal

Cardoso, Y. P., & Montoya-Burgos, J. I. (2009). Unexpected diversity in the catfish Pseudancistrus brevispinis reveals dispersal routes in a Neotropical center of endemism: the Guyanas Region. *Molecular Ecology* , *18* (5), 947–964. http://doi.org/10.1111/j.1365-294X.2008.04068.x

Cavender-Bares, J., Kozak, K. H., Fine, P. V. A., & Kembel, S. W. (2009). The merging of community ecology and phylogenetic biology.*Ecology Letters* , *12* (7), 693–715. http://doi.org/10.1111/j.1461-0248.2009.01314.x

Cavender-Bares, J., Ackerly, D. D., Baum, D. A., & Bazzaz, F. A. (2004). Phylogenetic Overdispersion in Floridian Oak Communities.*The American Naturalist* , *163* (6), 823–843. http://doi.org/10.1086/386375

Chan, L. M., Brown, J. L., & Yoder, A. D. (2011). Integrating statistical genetic and geospatial methods brings new power to phylogeography. *Molecular Phylogenetics and Evolution* ,*59* (2), 523–537. http://doi.org/10.1016/j.ympev.2011.01.020

Cilleros, K., Valentini, A., Allard, L., Dejean, T., Etienne, R., Grenouillet, G., ... Brosse, S. (2019). Unlocking biodiversity and conservation studies in high-diversity environments using environmental DNA (eDNA): A test with Guianese freshwater fishes. *Molecular Ecology Resources* , *19* (1), 27–46. http://doi.org/10.1111/1755-0998.12900

Covain, R., Fisch-Muller, S., Montoya-Burgos, J. I., Mol, J. H., Le Bail, P.-Y., & Dray, S. (2012). The Harttiini (Siluriformes, Loricariidae) from the Guianas: a multi-table approach to assess their diversity, evolution, and distribution. *Cybium* , *36* (1), 115–161.

Covain, R., Fisch-Muller, S., Oliveira, C., Mol, J. H., Montoya-Burgos, J. I., & Dray, S. (2016). Molecular phylogeny of the highly diversified catfish subfamily Loricariinae (Siluriformes, Loricariidae) reveals incongruences with morphological classification. *Molecular Phylogenetics and Evolution* , *94* , 492–517. http://doi.org/10.1016/j.ympev.2015.10.018

Craig, J. M., Crampton, W. G. R., & Albert, J. S. (2017). Revision of the polytypic electric fish Gymnotus carapo (Gymnotiformes, Teleostei), with descriptions of seven subspecies. *Zootaxa* , *4318* (3), 401–438. http://doi.org/10.11646/zootaxa.4318.3.1

de Carvalho, D. C., Oliveira, D. A., Pompeu, P. S., Leal, C. G., Oliveira, C., & Hanner, R. (2011). Deep barcode divergence in Brazilian freshwater fishes: the case of the Sao Francisco River basin.*Mitochondrial DNA* , *22* (sup1), 80–86. http://doi.org/10.3109/19401736.2011.588214

de Queiroz, L. J., Cardoso, Y., Jacot-des-Combes, C., Bahechar, I. A., Lucena, C. A., Rapp Py-Daniel, L., ... Montoya-Burgos, J. I. (2020). Evolutionary units delimitation and continental multilocus phylogeny of the hyperdiverse catfish genus Hypostomus. *Molecular Phylogenetics and Evolution* , *145* , 106711. http://doi.org/10.1016/j.ympev.2019.106711

Diaz, J., Villanova, G. V., Brancolini, F., del Pazo, F., Posner, V. M., Grimberg, A., & Arranz, S. E. (2016). First DNA Barcode Reference Library for the Identification of South American Freshwater Fish from the Lower Parana River. *PLOS ONE* , *11* (7), e0157419. http://doi.org/10.1371/journal.pone.0157419

Dray, S., & Dufour, A.-B. (2007). The ade4 Package: Implementing the Duality Diagram for Ecologists. *Journal of Statistical Software* ,*22* (4), 1–20. http://doi.org/10.18637/jss.v022.i04

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* ,*32* (5), 1792–1797. http://doi.org/10.1093/nar/gkh340

Efron, B. (1979). Bootstrap methods: Another look at the jackknife.*Annals of Statitics* , *7* (1), 1–26. http://doi.org/10.1214/aos/1176344552

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* , *39* (4), 783–791. http://doi.org/10.1111/j.1558-5646.1985.tb00420.x

Fisch-Muller, S., Mol, J. H. A., & Covain, R. (2018). An integrative framework to reevaluate the Neotropical catfish genus Guyanancistrus (Siluriformes: Loricariidae) with particular emphasis on the Guyanancistrus brevispinis complex. *PLOS ONE* , *13* (1), e0189789. http://doi.org/10.1371/journal.pone.0189789

Fisch-Muller, S., Montoya-Burgos, J. I., Le Bail, P.-Y., & Covain, R. (2012). Diversity of the Ancistrini (Siluriformes: Loricariidae) from the Guianas: the Panaque group, a molecular appraisal with descriptions of new species. *Cybium* , *36* (1), 163–193.

Fricke, R., Eschmeyer, W. N., & Van Der Laan, R. (2019). Eschmeyer's Catalog of Fishes: Genera, Species, References. Retrieved May 24, 2019, from http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp

Gerritsen, H. (2018). mapplots: Data Visualisation on Maps. Retrieved from https://cran.r-project.org/package=mapplots

Goldstein, P. Z., & DeSalle, R. (2011). Integrating DNA barcode data and taxonomic practice: Determination, discovery, and description.*BioEssays* , *33* (2), 135–147. http://doi.org/10.1002/bies.201000036

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

Hall, T. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* , *41* , 95–98.

Hebert, P. D. N., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences* ,*270* (1512), 313–321. http://doi.org/10.1098/rspb.2002.2218

Hijmans, R. J. (2019). raster: Geographic Data Analysis and Modeling. Retrieved from https://cran.r-project.org/package=raster

Johannes, M., Ruschhaupt, M., Froehlich, H., Mansmann, U., Buness, A., Warnat, P., ... Beissbarth, T. (2018). MCRestimate: Misclassification error estimation with cross-validation.

Jonsson, K. A., Tottrup, A. P., Borregaard, M. K., Keith, S. A., Rahbek, C., & Thorup, K. (2016). Tracking Animal Dispersal: From Individual Movement to Community Assembly and Global Range Dynamics. *Trends in Ecology & Evolution* , *31* (3), 204–214. http://doi.org/10.1016/j.tree.2016.01.003

Keith, P., Le Bail, P.-Y., & Planquette, P. (2000). *Atlas des poissons d'eau douce de Guyane. Tome 2, Fascicule I: Batrachoidiformes, Mugiliformes, Beloniformes, Cyprinodontiformes, Synbranchiformes, Perciformes, Pleuronectiformes, Tetraodontiformes.* Paris, France: Museum national d'histoire naturelle.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* , *16* (2), 111–120. http://doi.org/10.1007/BF01731581

Korbie, D. J., & Mattick, J. S. (2008). Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nature Protocols* , *3* (9), 1452–1456. http://doi.org/10.1038/nprot.2008.133

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution* , *33* (7), 1870–1874. http://doi.org/10.1093/molbev/msw054

Le Bail, P.-Y., Covain, R., Jégu, M., Fisch-Muller, S., Vigouroux, R., & Keith, P. (2012). Updated checklist of the freshwater and estuarine fishes of French Guiana. *Cybium* , *36* (1), 293–319.

Le Bail, P.-Y., Keith, P., & Planquette, P. (2000). *Atlas des poissons d'eau douce de Guyane. Tome 2, Fascicule II - Siluriformes* . Paris, France: Museum national d'histoire naturelle.

Lehner, B., Verdin, K., & Jarvis, A. (2008). New global hydrography derived from spaceborne elevation data. *Eos, Transactions American Geophysical Union* , *89* (10), 93–94. http://doi.org/10.1029/2008EO100001

Lemopoulos, A., & Covain, R. (2019). Biogeography of the freshwater fishes of the Guianas using a partitioned parsimony analysis of endemicity with reappraisal of ecoregional boundaries. *Cladistics* , *35* (1), 106–124. http://doi.org/10.1111/cla.12341

Lowe, W. H., & McPeek, M. A. (2014). Is dispersal neutral? *Trends in Ecology & Evolution* , *29* (8), 444–450. http://doi.org/10.1016/j.tree.2014.05.009

Mamos, T., Wattier, R., Burzyński, A., & Grabowski, M. (2016). The legacy of a vanished sea: a high level of diversification within a European freshwater amphipod species complex driven by 15 My of Paratethys regression. *Molecular Ecology* , *25* (3), 795–810. http://doi.org/10.1111/mec.13499

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution* , *18* (4), 189–197. http://doi.org/10.1016/S0169-5347(03)00008-9

McQuitty, L. L. (1966). Similarity Analysis by Reciprocal Pairs for Discrete and Continuous Data. *Educational and Psychological Measurement* , *26* (4), 825–831. http://doi.org/10.1177/001316446602600402

Mittelbach, G. G., & Schemske, D. W. (2015). Ecological and evolutionary perspectives on community assembly. *Trends in Ecology & Evolution* , *30* (5), 241–247. http://doi.org/10.1016/j.tree.2015.02.008

Mol, J. H. A. (2012). *The Freshwater Fishes of Suriname* . Leiden, Netherlands: Brill Academic Pub.

Mol, J. H. A., Vari, R. P., Covain, R., Willink, P. W., & Fisch-muller, S. (2012). Annotated checklist of the freshwater fishes of Suriname. *Cybium* , *36* (1), 263–292.

Murienne, J., Cantera, I., Cerdan, A., Cilleros, K., Decotte, J., Dejean, T., ... Brosse, S. (2019). Aquatic eDNA for monitoring French Guiana biodiversity. *Biodiversity Data Journal* , *7* , 1–9. http://doi.org/10.3897/BDJ.7.e37518

Nascimento, M. H. S., Almeida, M. S., Veira, M. N. S., Limeira Filho, D., Lima, R. C., Barros, M. C., & 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). http://doi.org/10.4238/gmr.15038476

Négrel, P., & Lachassagne, P. (2000). Geochemistry of the Maroni River (French Guiana) during the low water stage: implications for water–rock interaction and groundwater characteristics. *Journal of Hydrology* , *237* (3–4), 212–233. http://doi.org/10.1016/S0022-1694(00)00308-5

Nelson, J. S., Grande, T. C., & Wilson, M. V. H. (2016). *Fishes of the World* (5th ed.). Hoboken, NJ: John Wiley & Sons.

Paradis, E., & Schliep, K. (2018). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* ,*35* (3), 526–528. http://doi.org/10.1093/bioinformatics/bty633

Pausas, J. G., & Verdú, M. (2010). The jungle of methods for evaluating phenotypic and phylogenetic structure of communities. *BioScience* ,*60* (8), 614–625. http://doi.org/10.1525/bio.2010.60.8.7

Pereira, L. H. G., Hanner, R., Foresti, F., & Oliveira, C. (2013). Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? *BMC Genetics* , *14* (1), 20. http://doi.org/10.1186/1471-2156-14-20

Planquette, P., Keith, P., & Le Bail, P.-Y. (1996). *Atlas des poissons d'eau douce de Guyane: Tome 1* . Paris: Muséum national d'Histoire naturelle.

Planquette, P., & Renno, J.-F. (1990). Les Leporinus de la Guyane française (Pisces , Characiformes , Anostomidae), avec une note sur les techniques d'identification des espèces. *Revue Française d'Aquariologie* , *17* (2), 33–40.

Pugedo, M. L., de Andrade Neto, F. R., Pessali, T. C., Birindelli, J. L. O., & 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. http://doi.org/10.1007/s10709-016-9903-4

R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Ratnasingham, S., & Hebert, P. D. N. (2007). bold: The Barcode of Life Data System (http://www.barcodinglife.org). *Molecular Ecology Notes* , *7* (3), 355–364. http://doi.org/10.1111/j.1471-8286.2007.01678.x

Ratnasingham, S., & Hebert, P. D. N. (2013). A DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System. *PLoS ONE* , *8* (7), e66213. http://doi.org/10.1371/journal.pone.0066213

Rosso, J. J., Mabragaña, E., González Castro, M., & Díaz de Astarloa, J. M. (2012). DNA barcoding Neotropical fishes: recent advances from the Pampa Plain, Argentina. *Molecular Ecology Resources* ,*12* (6), 999–1011. http://doi.org/10.1111/1755-0998.12010

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* , *74* (12), 5463–5467. http://doi.org/10.1073/pnas.74.12.5463

Shimodaira, H. (2002). An approximately unbiased test of phylogenetic tree selection. *Systematic Biology* , *51* (3), 492–508. http://doi.org/10.1080/10635150290069913

Shimodaira, H. (2004). Approximately unbiased tests of regions using multistep-multiscale bootstrap resampling. *Annals of Statitics* ,*32* (6), 2616–2641. http://doi.org/10.1214/009053604000000823

Stabler, B. (2013). shapefiles: Read and Write ESRI Shapefiles. Retrieved from https://cran.r-project.org/package=shapefiles

Suzuki, R., & Shimodaira, H. (2006). Pvclust: an R package for assessing the uncertainty in hierarchical clustering.*Bioinformatics* , *22* (12), 1540–1542. http://doi.org/10.1093/bioinformatics/btl117

Taberlet, P., Coissac, E., Hajibabaei, M., & Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology* , *21* (8), 1789–1793. http://doi.org/10.1111/j.1365-294X.2012.05542.x

Tarroso, P., Carvalho, S. B., & Velo-Antón, G. (2019). Phylin 2.0: Extending the phylogeographical interpolation method to include uncertainty and user-defined distance metrics. *Molecular Ecology Resources* , (February), 1–14. http://doi.org/10.1111/1755-0998.13010

Valdez-Moreno, M., Ivanova, N. V., Elías-Gutiérrez, M., Contreras-Balderas, S., & Hebert, P. D. N. (2009). Probing diversity in freshwater fishes from Mexico and Guatemala with DNA barcodes.*Journal of Fish Biology* , *74* (2), 377–402. http://doi.org/10.1111/j.1095-8649.2008.02077.x

Vandergast, A. G., Perry, W. M., Roberto, L. V., & Hathaway, S. A. (2011). Genetic landscapes GIS Toolbox: tools to map patterns of genetic divergence and diversity. *Molecular Ecology Resources* ,*11* (1), 158–161. http://doi.org/10.1111/j.1755-0998.2010.02904.x

Vari, R. P., Ferraris, C. J., Radosavljevic, A., & Funk, V. A. (2009).*Checklist of the Freshwater Fishes of the Guiana Shield .Biological Society of Washington* . Biological Society of Washington. http://doi.org/10.2988/0097-0298-17.1.i

Vodă, R., Dapporto, L., Dincă, V., & Vila, R. (2015). Why Do Cryptic Species Tend Not to Co-Occur? A Case Study on Two Cryptic Pairs of Butterflies. *PLOS ONE* , *10* (2), e0117802. http://doi.org/10.1371/journal.pone.0117802

Ward, R. D. (2009). DNA barcode divergence among species and genera of birds and fishes. *Molecular Ecology Resources* , *9* (4), 1077–1085. http://doi.org/10.1111/j.1755-0998.2009.02541.x

Ward, R. D., Hanner, R., & Hebert, P. D. N. (2009). The campaign to DNA barcode all fishes, FISH-BOL. *Journal of Fish Biology* ,*74* (2), 329–356. http://doi.org/10.1111/j.1095-8649.2008.02080.x

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

Weber, C., Covain, R., & Fisch-Muller, S. (2012). Identity of Hypostomus plecostomus (Linnaeus, 1758), with an overview of Hypostomus species from the Guianas (Teleostei: Siluriformes: Loricariidae).*Cybium* , *36* (1), 195–227.

Wood, D. A., Vandergast, A. G., Barr, K. R., Inman, R. D., Esque, T. C., Nussear, K. E., & Fisher, R. N. (2013). Comparative phylogeography reveals deep lineages and regional evolutionary hotspots in the Mojave and Sonoran Deserts. *Diversity and Distributions* , *19* (7), 722–737. http://doi.org/10.1111/ddi.12022

Data Accessibility

All collection information and sequence data are available on the Barcode of life Data System (BOLD) in the project "Gui-BOL Barcoding Guianese fishes". All BOLD accession numbers for this dataset are available in Supplementary Material 6. R scripts used in the analyses are available on GitHub (https://github.com/yvanpapa/Maroni_Barcode).

Author Contributions

Y.P. contributed to molecular data collection, designed and ran the analyses, and wrote the manuscript. P.Y.L.B contributed to specimen collection and reviewed the manuscript. R.C. conceptualised and supervised the research, contributed to specimen and molecular data collection, contributed to design of analyses and reviewed the manuscript.

Tables

Table 1 . Species that have been taxonomically updated since 2012 Maroni checklists (Le Bail et al., 2012; Mol et al., 2012) to follow current nomenclature. (+) Currently still referred as *Glanidium leopardus* in BOLD dataset. (++) Currently still referred as *Mastiglanis* cf. *asopos* in BOLD dataset.

Current update	**Identification in 2012 checklists**	**Barco**
Anablesoides aff. *holmiae* Eigenmann, 1909	*Rivulus* aff. *holmiae* Eigenmann, 1909	
Anablepsoides amphoreus (Huber, 1979)	*Rivulus amphoreus* Huber, 1979	
Ananablepsoides gaucheri (Keith, Nandrin & Le Bail, 2006)	*Rivulus gaucheri* Keith, Nandrin & Le Bail, 2006	
Anablepsoides holmiae (Eigenmann, 1909)	*Rivulus holmiae* Eigenmann, 1909	
Anablepsoides igneus (Huber, 1991)	*Rivulus igneus* Huber, 1991	
Anablepsoides lungi (Berkenkamp 1984)	*Rivulus lungi* Berkenkamp, 1984	
Copella carsevennensis (Regan, 1912)	*Copella arnoldi* (Regan, 1912)	Yes
Curculionichthys sp. Maroni	Hypoptopomatinae Gen. nov. aff. *Parotocinclus*	Yes
Distocyclus guchereauae Meunier, Jégu & Keith 2014	*Distocyclus* sp.	Yes
Gephyromochlus leopardus (Hoedeman, 1961)$^{+}$	*Glanidium leopardum* (Hoedeman, 1961)	Yes
Guyanancistrus nassauensis Mol, Fisch-Muller & Covain 2018	*Guyanancistrus* sp. (Nassau Mountains)	Yes
Knodus heteresthes (Eigenmann 1908)	*Bryconamericus heteresthes* Eigenmann, 1908	Yes
Laimosemion agilae (Hoedeman 1954)	*Rivulus agilae* Hoedeman, 1954	Yes
Laimosemion breviceps (Eigenmann, 1909)	*Rivulus breviceps* Eigenmann, 1909	
Laimosemion frenatus (Eigenmann, 1912)	*Rivulus frenatus* Eigenmann, 1912	
Laimosemion geayi (Vaillant, 1899)	*Rivulus geayi* Vaillant, 1899	
Laimosemion xiphidius (Huber 1979)	*Rivulus xiphidius* Huber, 1979	
Mastiglanis durantoni de Pinna & Keith, 2019^{++}	*Mastiglanis* aff. *asopos* Bockmann, 1994	Yes
Pseudoqolus koko (Fisch-Muller & Covain 2012)	*Panaqolus koko* Fisch-Muller & Covain, 2012	Yes
Serrasalmus eigenmanni Norman 1929	*Pristobrycon eigenmanni* (Norman, 1929)	Yes
Serrapinnus gracilis (Gáry 1960)	*Odontostilbe gracilis* (Géry, 1960)	Yes
Tetragonopterus georgiae (Géry 1965)	*Moenkhausia georgiae* Géry, 1965	Yes

Table 2. Primers used for targeting and amplification of the COI region. Six primers were designed from this study and from Fisch-Muller et al. (2018) to obtain longer sequences of >800 bp. Other primers are from Ward, Zemlak, Innes, Last, & Hebert (2005) and amplify a region of approximately 650 bp. Sil = Siluriformes, Char = Characiformes.

Primer name	**Primer sequence**	**Targeted taxa**	**Reading direction**	**Annealing Temp. (C°)**	**Design**
5COI-F	5'-CTC GGC CAT CCT ACC TGT G-3'	Sil, Char, Cichlidae	Forward	53 - 54	Fisch-Muller et al. (2018)

Primer name	Primer sequence	Targeted taxa	Reading direction	Annealing Temp. (C°)	Design
5COI-gym-F	5'-CTC RGC CAT TCT ACC TGT GG-3'	Gymnotiformes	Forward	51	This study
FishF1	5'-TCA ACC AAC CAC AAA GAC ATT GGC AC-3'	Others	Forward	52	Ward et al. (2005)
FishF2	5'-TCG ACT AAT CAT AAA GAT ATC GGC AC-3'	Others	Forward	54	Ward et al. (2005)
RCOI-R2	5'-CGG GTG TCT ACG TCC ATT CCA ACT G-3'	Siluriformes	Reverse	53	Fisch-Muller et al. (2018)
5COI-R3	5'-AAC TGT AAA YAT GTG RTG GGC YCA-3'	Characiformes	Reverse	54	This study
RCOI-cic-R	5'-CCA ACA GTA AAT ATG TGA TGG GCT CA-3'	Cichlidae	Reverse	54	This study
5COI-gym-R	5'-CRT AGT GRA AGT GGG CAA CT-3'	Gymnotiformes	Reverse	51	This study
FishR1	5'-TAG ACT TCT GGG TGG CCA AAG AAT CA-3'	Others	Reverse	52	Ward et al. (2005)
FishR2	5'-ACT TCA GGG TGA CCG AAG AAT CAG AA-3'	Others	Reverse	54	Ward et al. (2005)

Table 3 . Twenty-five species for which status has been updated during this study using COI barcodes, morphology, and geographic distribution. Species referred as "cryptic" are additions to the Maroni checklist and share occurrence with their nominal species. Species that have been flagged as putative re-identification are referred as "New Taxonomical Combination". For these, the nominal species are suspected not to be present in the Maroni basin. For eight putative species that are BIN-discordant on the BOLD database,

records sharing the same BIN are indicated in column 2. (+) Putative species BIN-discordant in our dataset. Some species are still awaiting update on BOLD and are currently still referred with their former names on the platform, which are (++) *Cyphocharax* aff. *spilurus,* (§)*Hemigrammus guyanensis,* (¶) *Imparfinis* sp.

Current Update	**Status**	**BIN concordant with BOLD database?**	**Identification in 2012 checklists**
Ancistrus sp.	Cryptic	Yes	*Ancistrus temminckii* (Valenciennes, 1840)
Bryconops aff. *melanurus* (Bloch, 1794)	Cryptic	Yes	*Bryconops melanurus* (Bloch, 1794)
Corydoras aff. *geoffroy*$^{+}$ Lacepède, 1803	Cryptic	No: *Corydoras geoffroy*	*Corydoras geoffroy* Lacepède, 1803
Corydoras aff. *guianensis* Nijssen, 1970	Cryptic	Yes	*Corydoras guianensis* Nijssen, 1970
Cyphocharax aff. *spilurus* 2 (Günther 1864)$^{++}$	Cryptic	Yes	*Cyphocharax spilurus* (Günther, 1864)
Cyphocharax cf. *spilurus* (Günther 1864)	Cryptic	No: *Cyphocharax spilurus*	*Cyphocharax spilurus* (Günther, 1864)
Eigenmannia cf. *limbata* (Schreiner & Miranda Ribeiro 1903)	Cryptic	No: *E. limbata* / *E. nigra* / *E.* sp.	*Eigenmannia humboldtii* (Steindachner 1878)
Guianacara sp. Tapanahony	Cryptic	Yes	*Guianacara owroewefi* Kullander & Nijssen, 1989
Gymnotus aff. *carapo* Linnaeus, 1758	Cryptic	No: *G. pantanal* / *G. carapo* / *G. chaviro*	*Gymnotus carapo* Linnaeus, 1758
Hemigrammus aff. *guyanensis* Géry, 1959§	Cryptic	Yes	*Hemigrammus guyanensis* Géry, 1959
Imparfinis aff. *pijpersi* (Hoedeman, 1961)¶	New Taxonomical Combination	Yes	*Imparfinis pijpersi* (Hoedeman, 1961)
Leptocharacidium sp.	Cryptic	Yes	
Lithoxus aff. *planquettei* Boeseman, 1982	New Taxonomical Combination	Yes	*Lithoxus planquettei* Boeseman, 1982
Loricaria aff. *nickeriensis* Isbrücker, 1979	New Taxonomical Combination	Yes	*Loricaria nickeriensis* Isbrücker, 1979
Melanocharacidium sp. 1^{+}	Cryptic	No: *Melanocharacidium* sp. 2	
Melanocharacidium sp. 2^{+}	Cryptic	No: *Melanocharacidium* sp. 1	
Moenkhausia aff. *collettii* (Steindachner 1882)	Cryptic	Yes	*Moenkhausia collettii* (Steindachner, 1882)
Nannacara sp. Litany	Cryptic	Yes	

Current Update	Status	BIN concordant with BOLD database?	Identification in 2012 checklists
Pachypops aff. *fourcroi* (Lacepède, 1802)	New Taxonomical Combination	Yes	*Pachypops fourcroi* (Lacepède, 1802)
Pimelodella aff. *cristata* (Müller & Troschel, 1849)	New Taxonomical Combination	No: *P.* sp. / *P. vittata*	*Pimelodella cristata* (Müller & Troschel, 1849)
Pimelodella aff. *geryi* Hoedeman, 1961	Cryptic	Yes	*Pimelodella geryi* Hoedeman, 1961
Pimelodella cf. *cristata* (Müller & Troschel, 1849)	Cryptic	Yes	*Pimelodella cristata* (Müller & Troschel, 1849)
Pimelodella sp.	Cryptic	Yes	
Rineloricaria aff. *stewarti* 2 (Eigenmann, 1909)	Cryptic	Yes	*Rineloricaria* sp.2
Sternopygus cf. macrurus (Bloch & Schneider, 1801)	Cryptic	No: *S. macrurus*	*Sternopygus macrurus* (Bloch & Schneider, 1801)

Table 4. K2P genetic divergence values within different taxonomic levels from 1,284 specimens of Maroni Basin analysed.

K2P genetic divergence (%)	Taxa	Minimum	Mean	Maximum
Within Species	169	0.00	0.32	14.37
Within Genus	36	0.12	12.46	25.52
Within Family	25	3.57	19.61	30.70

Table 5. Table of twelve species showing a potential absence of barcode gap due to a maximum intra-specific K2P distance equal or higher than the distance to their Nearest Neighbour (in bold), or a distance to Nearest Neighbour (dNN) <2%, meaning that these species share particularly close COI sequences.

Family	Species (Pairs of Nearest Neighbours)	Max. Intra- Specific Distance	Pairwise distance (dNN)
Loricariidae	*Ancistrus* cf. *leucostictus* / *Ancistrus temminckii*	0.37 / 0.32	1.84
Callichthyidae	*Corydoras* aff. *geoffroy* / *Corydoras geoffroy*	0 / **0.62**	0.12
Loricariidae	*Guyanancistrus brevispinis* / *Guyanancistrus nassauensis*	**1.19** / 0.13	0.62
Loricariidae	*Hypostomus plecostomus* / *Hypostomus watwata*	0.26 / 0	1.94
Crenuchidae	*Melanocharacidium* sp. 1 / *Melanocharacidium* sp. 2	0.12 / 0.31	0.97

Family	Species (Pairs of Nearest Neighbours)	Max. Intra- Specific Distance	Pairwise distance (dNN)
Heptapteridae	*Pimelodella* aff. *geryi* / *Pimelodella geryi*	0 / 0.17	1.47

Figures

Figure 1. Distribution map of sample sites and effort. Relief map of the Maroni Basin (delimited in black) showing 83 sampling points (red dots) of specimens collected. Names of main tributaries are indicated in blue and main sampling locations are in black. Red dot size refer to sampling effort in each site. The map was reconstructed using raster images and shapefiles obtained from HydroSHEDS (Lehner, Verdin, & Jarvis, 2008) project website (http://www.worldwildlife.org/pages/hydrosheds) and the R packages raster v2.9.5 (Hijmans, 2019), rgdal v1.4-8 (Bivand, Keitt, & Rowlingson, 2019), shapefiles v0.7 (Stabler, 2013), and mapplots v1.5-1 (Gerritsen, 2018).

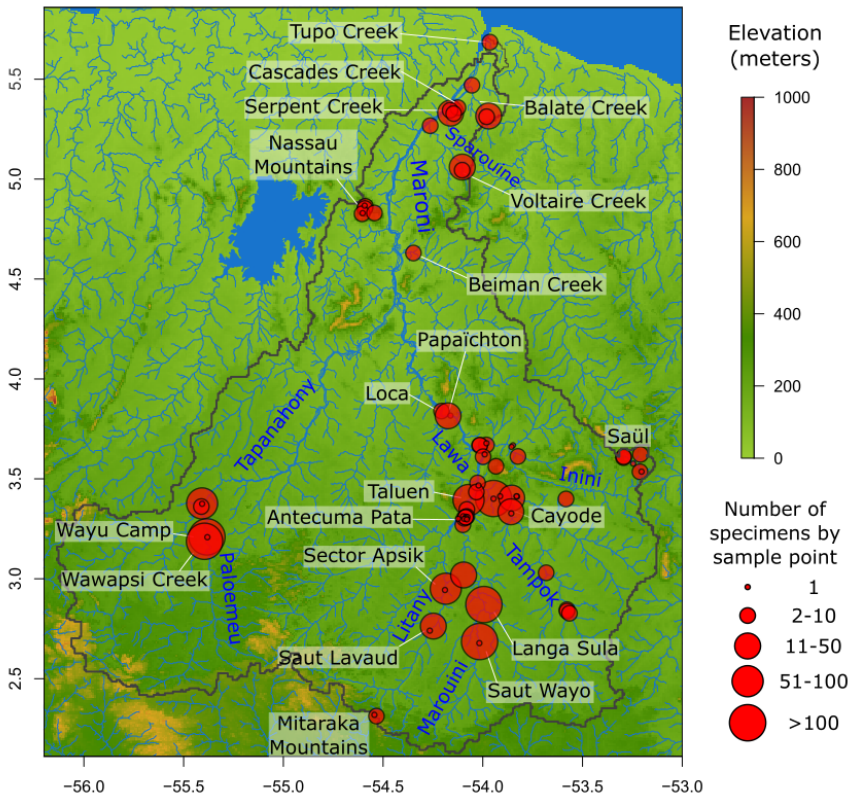

Figure 2. Number of Species, Genera, and Families by Order in the total data set, following taxonomic classification from BOLD. Acanthuriformes include only Sciaenidae (Perciformes *sedis mutabilis sensu* Fricke et al. (2019)).

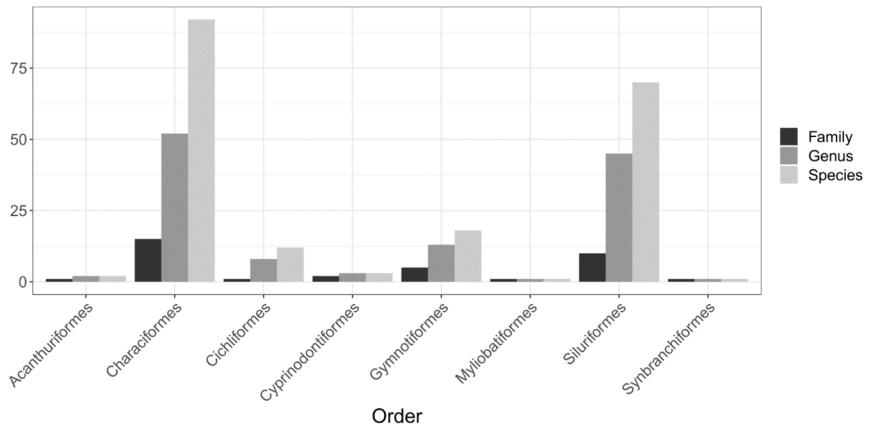

Figure 3. Neighbour joining tree computed on Kimura two-parameter distances of 1,284 COI sequences for 199 freshwater fish species from the Maroni Basin. Sequences sharing the same species identification and BIN are clustered on the same branch. Tip labels: species name and number of specimens. Labels of species with multiple BINs are coloured in blue, labels of species sharing the same BIN are coloured in red. (*) indicates the two clades that contain a specimen for which a BIN has not been assigned yet.

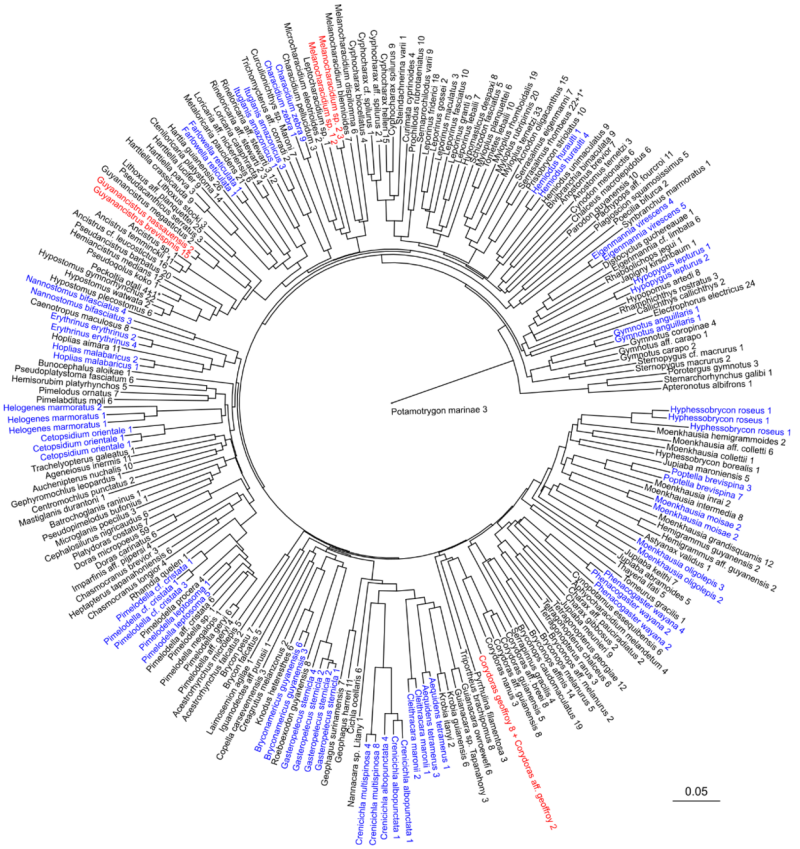

Figure 4. Average "multispecies" genetic divergence landscape. Mean of all overlaid genetic divergence landscapes from 128 fish species (Supplementary Material 4). White dots are sampling points. Upper left

inset: number of overlapping species landscapes in each grid cell, showing the sampling coverage for the calculation of the multispecies landscape.

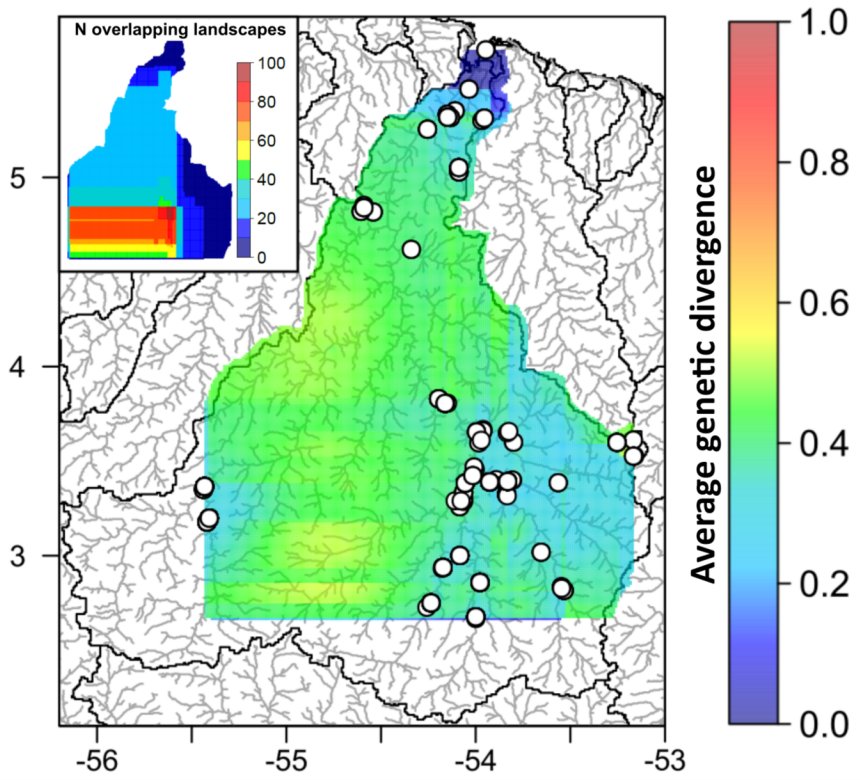

Figure 5: Principal Components Analysis of genetic landscapes of 128 species and the average "multispecies" landscape. Projection of the genetic landscapes on the first two axes (**A**) and on axes 2 and 3 (**B**); same projection with the species names on the first two axes (**C**) and on axes 2 and 3 (**D**) of the PCA. Species are labelled as first letter of genera + three first letter of species as detailed in Supplementary Material 7. Letters (N, S, W, and E) refer to geographic directions from the centre of Maroni Basin. Signs (+ and -) correspond to higher or lower genetic divergences associated to these directions. Combinations of signs and letters describe types of ordinated patterns along axes (see individual projected landscapes, Supplementary Material 3).

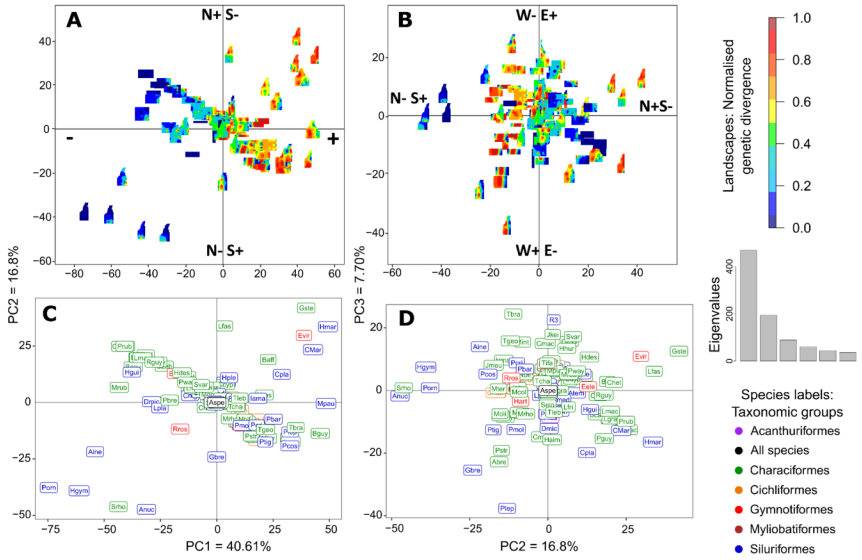

Figure 6. Projection of PCA loadings (scores) onto axes. Colour gradients from red to yellow and from green to purple correspond to latitude and to longitude of grid cells respectively (as shown in "colour scale").

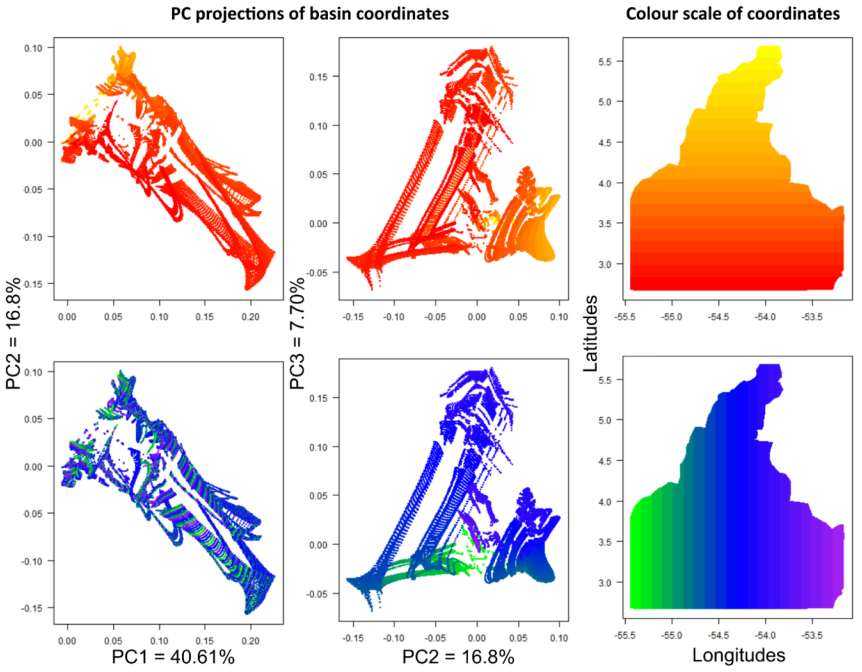

Figure 7. WPGMA tree performed on axes 2 and 3 of the genetic landscape PCA. Only clusters supports [?] 95% are reported above nodes. Approximately Unbiased (AU) are in red (left) and Bootstrap Probability (BP) are in green (right). Red squares represent clusters supported by an AU $>$ 95%. Species labelled as in Supplementary Material 7. Mean genetic landscapes were computed for the main cluster groups (large, noted A to I) and for the clusters supported by AU $>$ 95% as well as the sister nodes of main cluster groups (small).

They were calculated as the mean of all genetic landscapes in their respective cluster (minus the multispecies landscape if present). The root corresponds to the multispecies landscape. Upper right: projection of the species labels from the genetic landscapes PCA axes 2 and 3, coloured according to WPGMA main cluster groups.