1. Introduction

The Cichlidae family is one of the most specious and ecologically diverse groups of freshwater fishes inhabiting the tropical zone (Winemiller et al., 1997). Recently assessed using molecular data, the phylogeny of cichlid fishes supports a monophyletic origin for the South American faunas and further suggests higher evolutionary rates among many Neotropical cichlids (Farias et al., 1999, 2000; Lundberg et al., 1998). With more than 400 species (Reis et al., 2004), the Neotropical cichlid fauna hosts the second largest species of the family: the piscivorous species of the genus *Cichla* Scheneider 1801, that reach 80 cm in standard length for more than 10 kg (Taphorn and Barbarino-Duque, 1993; Winemiller et al., 1997). Closely related to the genus *Astronotus* (Farias et al., 1999, 2000) and sharing morphological adaptation to predation similar to that of *Crenicichla* and *Teleoichla* (Stiassny, 1987, 1991; Kullander, 1998), this Neotropical genus comprises five valid species: *C. ocellaris* Block & Schneider 1801, from the coastal drainages of the Guyana shield, *C. orinocensis* Humboldt 1821, from the Orinoco and Negro Rivers to central Amazon; *C. temensis* Humboldt 1821, in clear and black water tributaries of the Orinoco, Negro and Tapajos Rivers; *C. intermedia* Machado-Allison, 1971, from the upper Negro and Orinoco Rivers; and *C. monoculus* Spix & Agassiz 1831, distributed throughout the Amazon basin as well as the rivers from the Guyana shield (Kullander, 2003; Reis et al., 2004). Due to the size reached by some *Cichla* species and their combativeness, the genus is of great economic importance for sport fishing. This has led to the translocation of various species out of their natural range (Winemiller, 2001), thus artificially increasing their geographical distribution range and the mixing of different species. However, several species are naturally found in sympatry as previously reported in the Orinoco or Amazon Rivers (Machado-Allison, 1971; Winemiller et al., 1997). Owing to its wide distribution across South America and its life history characteristics, the genus *Cichla* is of great interest for biogeographic and population genetic studies.

The Amazon hosts a great diversity of water chemistry and the hydrological typology of the Amazonian drainages currently includes three types of water (Sioli, 1984): (1) white water characterised by a great amount of dissolved solid materials and a low transparency (Andean origin); (2) clear water characterised by a low content of dissolved solids and a high transparency (Precambrian shield) and (3) black water originated from forested lowland that differs from the latter by having a higher content of humic acids and a lower pH. Owing to its location between the Southern Andes and the Brazilian shield, the upper Madera also hosts most of the diversity in water chemistry. Annual rainfall cycles further enhance habitat heterogeneity due to the
seasonal inundation of one of the largest floodplains of the Amazon (Guyot et al., 1999). Furthermore, the palaeo-geographical context of the region is of great interest since the upper watershed of the Madera River is isolated from its lower course by a large series of rapids (Lundberg et al., 1998).

A single species of the genus Cichla referred to as Cichla monoculus was reported from all the four main basins of the upper Madera (Laussanne et al., 1991) and in the Ucayali basin in Perú (Kullander, 1986). Given the palaeo-geographic and hydrologic context of the Bolivian Amazon, the populations of C. monoculus have evolved in considerable geological changes and habitat heterogeneity. Cichla species are rather sedentary (Hoeinghaus et al., 2003) ambush predators. It could then be expected that variations in water chemistry and transparency might play a role in the genetic structure and differentiation of the upper Madera populations. On the other hand, the floodplain may act as homogenising factor by making inter-basin dispersal easier within the upper Madera.

In order to better understand the structuring events and evolution of Cichla monoculus in the upper Madera (Bolivian Amazon), its phylogeography was inferred using sequences of the mtDNA control region. Specimens of C. monoculus from the Ucayali River (Perú) and C. ocellaris from the Maroni River (French Guyana) were used to infer the taxonomic position of C. monoculus in the upper Madera. The genetic variation of C. monoculus in the upper Madera is discussed in the light of the ecological and palaeo-geographical context of the region, taking into account its life history characteristics.

2. Materials and methods

2.1. Hydrological context and sampling

With an area of $1.37 \times 10^6$ km$^2$, the Madera is the second largest watershed of the Amazon after the Solimões with 2.24 $\times 10^6$ km$^2$. The upper Madera, corresponding to the Bolivian Amazon, represents at least 60% of the total Madera watershed area. A marked annual cycle of rainy and dry seasons is responsible of multi-peaked floods in the Andean tributaries. The downstream pulse is stored in the Bolivian floodplain, which is one of the largest of the Amazon with a potential flood extension of 0.15 $\times 10^6$ km$^2$ (Guyot et al., 1999). Of the four main tributaries of the upper Madera River, the Madre de Dios, Beni, Mamoré and Itenez Rivers, three of these originate in the Andes and correspond to white water (Mamoré and Beni) or mixed white water–black water systems (Madre de Dios). The Itenez River, also known as the Guapore, runs through the Brazilian shield and thereby hosts clear water tributaries. However, these designations mask considerable heterogeneity, and small tributaries with mixed (white and black) or plain black waters from the forested lowland are frequently encountered along the main channel of the Madre de Dios and Mamoré Rivers. The Yata is a small tributary of the Madera River, remarkable for its central position and its black lowland waters.

The Ucayali River in the upper Amazonas (Perú) and the Madre de Dios, Mamoré, Itenez and Yata rivers in the upper Madera basin (Bolivia) were sampled (Fig. 1). In the upper Madera 41 specimens of C. monoculus were analysed: 12 in the Manuripi in the middle Madre de Dios basin, a blackwater lowland tributary; 9 in the Yata lowland river with black waters, 2 in the Securé and 4 in the Ichilo Rivers, both whitewater tributaries in the middle and upper Mamoré basin, respectively; 6 in the San Martin and 8 in the Paraguá River, both clear water tributaries in the middle and upper Itenez basin, respectively. In Perú, 4 specimens from the Iquitos region (lower Ucayali) and 2 from the Pucallpa region (upper Ucayali) were analysed. Two specimens of C. ocellaris originated from the upper Maroni River (French Guyana). The 3 specimens of C. temensis originating from Brazil (obtained through a specialised fish trader in Paris, IGUARAPE) were used as outgroup.

Specimens were caught using hook, line and gill nets with a mesh size ranging from 50 to 110 mm. For each of them, approximately 1 cm$^3$ of muscle was preserved in 96% ethanol. Vouchers were preserved in a 30% formaldehyde solution and then transferred into an 80% ethanol solution.

2.2. DNA extraction and sequencing

The mitochondrial genome is of traditional use to infer phylogenetic relationships among Neotropical fishes using either complete mtDNA (Renno et al., 1991), ribosomal genes (Alves-Gomes et al., 1995; Ortí et al., 1996; Ortí and Meyer, 1997; Farias et al., 1999, 2000; Calcagnotto et al., 2005; Hubert et al., 2005), the hypervariable control region (Montoya-Burgos, 2003; Sivasundar et al., 2001), the NADH1 and ATPase discontinuous segments (Hrbek et al., 2005), or a combined approach between mitochondrial and nuclear genes (Lopez-Fernandez et al., 2005).

In this study, a near complete segment of the mitochondrial control region, including the D-loop, was sequenced (1041 bp) for 47 specimens of Cichla monoculus and three specimens of the outgroup C. temensis. For the two specimens of C. ocellaris, it was only possible to sequence 667 bp for one sample and 762 bp for the other in the 5’DNA region, but these provided sufficient data for unambiguous alignment.

DNA extractions were performed using the Qiagen DNAeasy Mini Kit and the primers DL20F (5′-TTAGCA AGGGCTTCTTGGGCT) and DL20R (5′-ACCCCTTAGC TCCAAAAGCTA) were used for amplification and sequencing (courtesy of Dr. J.F. Agnèse, IRD). Amplifications were carried out in a total volume of 50 μl, with 2 U Tag DNA polymerase (Promega) and 1 μl of each primer (20 pmol/μl). Cycling was performed using a touch-down profile as follows: 2 min initial denaturation at 92°C followed by 10 standard cycles with a 1°C cycle gradient starting at 66°C; finally 20 cycles with 1 min at 92°C, 1 min at 56°C and 1 min 30s at 72°C. Post-PCR extension was...
carried out for 5 min at 72 °C. Double-stranded products purification was performed using Qiagen MinElute PCR Kit and followed by direct sequencing on both strands by automatic sequencing. All templates were sequenced directly by automatic sequencing (Macrogen service). The sequences were deposited in GenBank with the following Accession numbers: DQ778661 to DQ778712.

2.3. Analysis of genetic variability

Sequences were aligned manually using the GeneDoc software (Nicholas et al., 1997). To detect potential departure from expectations of neutral molecular evolution of the mitochondrial control region, Tajima’s Test (Tajima, 1989) was performed.

Our hypothesis of the influence of habitat heterogeneity (i.e. water types) on genetic structure of *C. monoculus* in the upper Madera was tested in two stages. In a first stage using a Spatial Analysis of Molecular Variance (SAMOVA), two (k = 2), three (k = 3) and four (k = 4) groups of localities were made to explore possible genetic structure without explicit *a priori* hypotheses. The SAMOVA (Dupanloup et al., 2002) crosses the Analysis of Molecular Variance (AMOVA) (Excoffier et al., 1992) and geographical information, and involves the definition of groups of populations without the need of *a priori* delimitation by maximising the differentiation among groups of localities (i.e. delimiting the boundaries of groups such that the proportion of total genetic variance due to differences between groups is maximised). The significance of the variance components among groups (\(\phi_{CT}\)), among populations within groups (\(\phi_{SC}\)) and within populations (\(\phi_{ST}\)) was tested by 1000 permutations of individuals for each of the hierarchical levels. In the second stage, an AMOVA was performed among predefined ecological groups: a (Manuripi + Yata) group for the populations of lowland black waters, a (Securé + Ichilo) group for the populations of white waters and a (San Martín + Paraguá) group for the populations of clear waters.

To infer phylogenetic relationships among *Cichla*, we implemented a Maximum Likelihood search (ML) using PhyML (http://atgc.lirmm.fr/phyml), which uses the algorithm developed by Guindon and Gascuel (2003). The best evolutionary model and its parameters were selected from among 28 models using the Akaike Information Criterion values (Akaike, 1973) determined using the Analysis of Phylogenetics and Evolution (APE) R program (Paradis et al., 2004; Paradis, 2006). Bootstrap proportions (BP) for the maximum likelihood tree were determined with 1000 pseudoreplicates to estimate each node’s robustness. The evolutionary model for ML also was used to calculate the genetic divergence among species and among populations using PAUP* (Swofford, 2002).

Expected mismatch distribution and parameter (\(r\)) of Roger’s (1995) model of expanding population, following the coalescence method, were simulated using the approach outlined by Slatkin and Hudson (1991) and Rogers and Harpending (1992). Tests for significant fit of the model
were made using the permutation procedure implemented in Arlequin.

Parameters of nucleotide diversity, Tajima’s test of neutrality, AMOVA and mismatch were computed using Arlequin 3.0 (Excoffier et al., 2005).

3. Results

3.1. Control region sequences variations

The alignment of complete control region sequences from *C. monoculus* comprised 1041 sites. Base composition was biased markedly in favour of A and T (A + T% = 62.1%). Mean values were A = 0.32, C = 0.23, G = 0.15, T = 0.30. Very high haplotypic diversity was observed ($H = 0.98$): 39 haplotypes for 41 fishes from the upper Madera. Only one haplotype was shared by three fishes in the Manuripi River. In the Ucayali, two fishes shared one of five haplotypes, one fish from Pucallpa and the other from Iquitos. The nucleotide diversity of *C. monoculus* in the Ucayali was 0.012 ± 0.008. In the upper Madera, it ranged from 0.010 ± 0.006 in the Manuripi to 0.058 ± 0.031 in the Yata population (Table 1).

The SAMOVA delimited the following groups (Table 2): for $k = 2$, $\phi_{CT} = 0.19$, the Yata, San Martín, Paraguá localities near the Brazilian shield were grouped together, and the Manuripi, Securé, and Ichilo localities farther west were grouped together, although the among-group portion of the genetic variance was not significant ($P = 0.11$). For $k = 3$, $\phi_{CT} = 0.23$, SAMOVA separated the Manuripi and

Table 1

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Percentage of variation</th>
<th>Fixation index</th>
<th>Significance tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among populations</td>
<td></td>
<td>$\phi_{CT}$</td>
<td></td>
</tr>
<tr>
<td>Manuripi (12)</td>
<td></td>
<td>0.0019 ± 0.0006</td>
<td></td>
</tr>
<tr>
<td>Yata (9)</td>
<td></td>
<td>0.020 ± 0.004</td>
<td></td>
</tr>
<tr>
<td>San Martín (6)</td>
<td></td>
<td>0.021 ± 0.006</td>
<td></td>
</tr>
<tr>
<td>Paraguá (8)</td>
<td></td>
<td>0.030 ± 0.016</td>
<td></td>
</tr>
<tr>
<td>Among species</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. monoculus (41)</td>
<td>Upper Madera</td>
<td>0.139</td>
<td></td>
</tr>
<tr>
<td>C. monoculus (6)</td>
<td>French Guyana</td>
<td>0.094 ± 0.016</td>
<td></td>
</tr>
<tr>
<td>C. ocellaris (2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. temensis (3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SC</td>
<td></td>
<td>0.137 ± 0.036</td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td></td>
<td>0.216 ± 0.019</td>
<td></td>
</tr>
<tr>
<td>AMOVA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among populations</td>
<td></td>
<td>$\phi_{CT}$</td>
<td></td>
</tr>
</tbody>
</table>
| Manuripi, Securé, Ichilo-group and (Yata, San Martín, Paraguá)-group | 19.26 | 0.23 | $P = 0.02, df = 2$
| Among populations within groups | 20.89 | 0.22 | $P = 0.00, df = 3$
| Within populations  | 59.86                   | 0.40           | $P = 0.00, df = 35$
| SAMOVA               |                         | $\phi_{CT}$   |                   |
| Among groups        |                         | 0.24           |                   |
| Among populations within groups | 12.53 | 0.16 | $P = 0.00, df = 2$
| Within populations  | 63.28                   | 0.37           | $P = 0.00, df = 35$
| AMOVA                |                         | $\phi_{CT}$   |                   |
| Among groups        |                         | 0.12           |                   |
| Among populations within groups | 25.37 | 0.29 | $P = 0.00, df = 3$
| Within populations  | 62.69                   | 0.37           | $P = 0.00, df = 35$

Pair wise genetic divergence of *Cichla* according to the F84 + G model; among populations of *C. monoculus* in the upper Madera (first semi-diagonal matrix) and among species of *C. monoculus, C. ocellaris* and *C. temensis*. The number of fish used between brackets.

Table 2

Results of the SAMOVA analysis according to $k = 2$, 3 or 4 groups *a posteriori* formed and results of the AMOVA analysis according to three water quality groups *a priori* formed.

**Table 2**

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Percentage of variation</th>
<th>Fixation index</th>
<th>Significance tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among populations</td>
<td></td>
<td>$\phi_{CT}$</td>
<td></td>
</tr>
</tbody>
</table>
| Manuripi, Securé, Ichilo-group and (Yata, San Martín, Paraguá)-group | 19.26 | 0.23 | $P = 0.02, df = 2$
| Among populations within groups | 20.89 | 0.22 | $P = 0.00, df = 3$
| Within populations  | 59.86                   | 0.40           | $P = 0.00, df = 35$
| SAMOVA               |                         | $\phi_{CT}$   |                   |
| Among groups        |                         | 0.24           |                   |
| Among populations within groups | 12.53 | 0.16 | $P = 0.00, df = 2$
| Within populations  | 63.28                   | 0.37           | $P = 0.00, df = 35$
| AMOVA                |                         | $\phi_{CT}$   |                   |
| Among groups        |                         | 0.12           |                   |
| Among populations within groups | 25.37 | 0.29 | $P = 0.00, df = 3$
| Within populations  | 62.69                   | 0.37           | $P = 0.00, df = 35$
Securé localities, which comprise two very different types of waters (black and white) from the Ichilo locality and a group including the Yata, San Martín, and Paraguá localities close to the Brazilian shield. In this analysis, the among group portion of genetic variance was found to be significant ($P = 0.02$). For $k = 4$, $\phi_{ST} = 0.24$, SAMOVA separated the Manuripi from the Securé + Ichilo, Yata, and San Martín + Paraguá groups, thereby differentiating the four sub-basins of the upper Madera with a significant among population portion of genetic variance ($P = 0.02$). In each test of hierarchical structure within-population components of genetic variation ($\phi_{ST}$) was always the highest and significant. The Yata and Manuripi populations, despite sharing the same type of lowland black water, were never grouped together, whereas the Securé and Manuripi populations (white vs. lowland waters) were grouped. The AMOVA, the second state of our test, also failed to evidence any ecological relationship between genetic structure and water quality.

3.2. D-loop haplotypic phylogeography

The largely resolved Maximum Likelihood tree (Fig. 2) was built according to the selected F84 + G model (Log (likelihood) $= -3861.05$, Gamma shape parameter $= 0.237$, transition/transversion ratio $= 2.818$). It showed that Cichla monoculus was a polyphyletic group, with C. monoculus from the Ucayali (bootstrap, boot. $= 83$) being sister group of C. ocellaris (boot. $= 100$) and C. monoculus from the upper Madera forming a separated monophyletic group (boot. $= 100$). This demonstrated that C. monoculus from the upper Madera is a species clearly differentiated from C. monoculus from Perú and from C. ocellaris, which formed a separated clade (boot. $= 99$).

Four main geographical clades, themselves organised in sub-clades, were differentiated in the upper Madera: (1) a clade including part of the Itenez samples (boot. $= 97$), sister group of the remaining haplotypes, (2) a Itenez-Yata clade (boot. $= 91$), sister group of the (3) Mamoré-Yata (boot. $= 81$) and 4) Manuripi (boot. $= 97$) clades. According to the F84 + G model the nucleotide divergence between the upper Madera populations ranged from 0.019 ± 0.006 between the Mamoré and Manuripi populations to 0.035 ± 0.020 between the Mamoré and Yata populations. At the inter-specific level, it ranged from 0.064 ± 0.005 between C. monoculus from the Ucayali and C. ocellaris to 0.34 ± 0.032 between C. ocellaris and C. temensis (Table 1). The nucleotide divergence between C. monoculus from the Ucayali and C. monoculus from the upper Madera (0.094 ± 0.016) was higher than that between C. monoculus from the Ucayali and C. ocellaris (0.064 ± 0.005).

The mismatch distribution among haplotypes from the upper Madera displayed a unique mode at about 20 mutational steps (Fig. 2). The observed pattern matched almost perfectly the simulated one for an ancient expansion model ($P = 1; r = 0.002$). The neutrality was rejected by the Tajima’s $D$ test only for the Manuripi population (Table 1).

4. Discussion

The ecological complexity of the upper Madera region is the result of a long evolutionary history that may be traced back to the Miocene after which its current configuration was preserved from later marine incursions (Lundberg et al., 1998; Montoya-Burgos, 2003). Since then, this region is partially isolated from the rest of the Amazon watershed by large series of rapids representing a total or partial upstream barrier for many fish species (e.g. Arapaima gigas, Hrbek et al., 2005).

The high nucleotidic divergence observed between the haplotypes of C. monoculus in the upper Madera and the Ucayali (around 9%) is much higher than in several South American characiform species such as the curimatá or coporo Prochilodus lineatus (0.3% to 3.6% in the Paraná, Sivasundar et al., 2001) or between several species of piranhas (ranging from 1.5% to 10% in Serrasalmus and 3.9% to 10% in Pygocentrus, Hubert, 2005). The largely resolved Maximum Likelihood tree evidenced that Cichla monoculus is a polyphyletic group with the C. monoculus from the Ucayali (hereafter C. monoculus sensu stricto) being sister group of C. ocellaris from the Maroni in French Guiana. As the Cichla species from the Ucayali was described as C. monoculus by Kullander (1986), the Cichla from the upper Madera, which form a separated monophyletic group, is a species clearly differentiated from C. monoculus sensu stricto. The partial geographical isolation of the upper Madera with a total and long-lasting genetic isolation of Cichla in this region might have led to the formation of an endemic species of Cichla by allopatric speciation. The Cichla species of the Bolivian Amazon is very likely a new undescribed species, although it would need to be compared to the other described species for confirmation.

In the upper Madera, the hypothesised relationship between biogeographic structure and water quality was not confirmed by the Spatial Analysis of Molecular Variance (SAMOVA) nor by the AMOVA.

The ML tree indicated that each clade is restricted to a particular river basin, except in the Yata where secondary colonisation may have occurred due to its central location (Fig. 2). This was consistent with the SAMOVA results, as the more strongly supported structure ($k = 4$) separated each river basin within the Bolivian Amazon. This suggests that the floodplain does not act as a dispersal corridor for unrestricted gene flow among populations in different sub-basins.

In spite of a high number of samples analysed there, the nucleotide diversity was low in the Manuripi population compared with the other populations, which suggested a founder event. The high haplotypic diversity in the Itenez basin, which included two clades (one being the sister group of all the others), suggest that the Itenez basin may have acted as an aquatic refuge in the upper Madera for this Cichla species. This result is consistent with the mismatch distribution which emphasised an ancient demographic expansion of C. monoculus in the upper Madera, while the
non neutral Tajima’s $D$ test in the Manuripi population corroborates a founder effect.

Given the remarkably strong between-river structure of *Cichla* in some portions of the upper Madera, it is interesting to consider what implications this may have for species cohesion. This structure, if reflected by other DNA loci, might indicate incipient speciation associated with the restricted dispersal behaviour of *Cichla* (Hoeinghaus et al., 2003). The robustness of this genetic structure could be investigated using nuclear intron length variation.
(Hubert et al., 2006). Additionally, an experimental study in controlled environment would be useful to test the levels of reproductive isolation between the geographical populations, as well as investigations of the phenotypic differentiation among these populations.

Acknowledgments

The authors thank Dr. I. Farias (UFM, Manaus—Brazil), Dr. G. Ortí (University of Nebraska, USA), Dr. F. Bonhomme (CNRS, Montpellier—France) and Dr. E. Paradis (IRD, Montpellier—France) for fruitful discussions, and two anonymous reviewers for comments on an earlier version of this manuscript. This study was part of the CAV-IAR program developed by IRD in partnership with the “Universidad Mayor de San Andrés” (UMSA) in Bolivia and the University of Montpellier II (France). The authors also thank N. Lovejoy and Y. Fermon for assistance obtaining data for *Cichla ocellaris*.

References


Hubert, N., 2005. Evolution de l’ichtyofaune en Amérique du Sud: Biogéographie des Characiformes et phylogéographie des piranhas *Serrasal-


