Molecular Phylogenetics and Evolution
January 2008, Volume 46, Issue 1, Pages 375-381
http://dx.doi.org/10.1016/j.ympev.2007.04.002

Species polyphyly and mtDNA introgression among three Serrasalmus sister-species

Nicolas Hubert[a, b, c, *], Juan Pablo Torrico[b, c], François Bonhomme[c] and Jean-François Renno[a, b, c]

[a]U.R. 175 Institut de Recherche pour le Développement (IRD), GAMET, BP 5095, 361 rue JF Breton, 34196 Montpellier Cedex 05, France
[b]Instituto de Biologìa Molecular y Biotecnologìa, Universidad Mayor de San Andres, La Paz, Bolivia
[c]Laboratoire Génome, Populations, Interactions, Adaptation, CNRS-IFREMER-Université Montpellier II, UMR 5171, SMEL, 1 quai de la daurade, 34200 Sète, France

*: Corresponding author : N. Hubert, email address : myloplus@excite.com

Keywords: Characidae; Neotropics; Introgression; Ancient polymorphism; mtDNA

1. Introduction

Understanding the processes that generated pattern of DNA variation in natural populations may be a difficult task. Since migration and gene flow may superimpose to genetic drift and divergence, evolutionary forces responsible of shared polymorphism may be difficult to identify (Pamilo, 1988; Nielsen & Wakely, 2001). In this context, the raise of the coalescent theory constituted a significant improvement in the comprehension of the theoretical framework behind gene genealogies (Kingman, 1982; Tajima, 1983) and its application to the analysis of DNA sequences has proven to constitute an informative approach to the problem of shared polymorphism (Chiang, 2000; Takahashi *et al*., 2001; Machado & Hey, 2002; Rokas *et al*., 2003; Bowie *et al*., 2005). The coalescent theory predicts that haplotype sharing will persist at the incipient stage of species divergence between species that founded from the same gene pool (Rosenberg, 2003). This stage of shared polymorphism without gene flow has been previously formalised as the lineage sorting period (Hoelzer *et al*., 1998). This step is characterised by the occurrence of coalescent events between alleles from isolated groups leading to erratic genealogies (Pamilo, 1988; Funk, 2003). However, recently diverging groups may still exchange genes and distinguishing between gene flow and ancestral polymorphism may be a difficult task (e.g. Nielsen & Wakeley, 2001). The piranha belongs to the characidae subfamily of Serrasalminae (Buckup 1998). Currently including 28 species ranging from 130-420 mm standard length, the piranha genera *Serrasalmus* and *Pygocentrus* constitute the most speciose group of large carnivorous Characiformes (Jégu 2003). DNA sequences from mitochondrial DNA (mtDNA) recently evidenced that these genera constitute a monophyletic group originating 9 million years ago (Ma) and that *Serrasalmus* splits into three distinct clades, all distributed throughout the Amazon, Orinoco and Paraná watersheds (Hubert *et al*., in press). The biogeography of the Amazon freshwater fish fauna has been largely influenced by the Miocene marine incursion

that happened at 5 Ma (Hubert & Renno, 2006; Nores, 1999). The analysis of mtDNA sequences within the Piranha evidenced that the colonisation of the Upper Amazon by the genera *Serrasalmus* and *Pygocentrus* occurred after the marine retreat, during the last 4 million years, from the Miocene freshwater refuges of the Brazilian and Guyana shields (Hubert & Renno, 2006; Hubert *et al*., in press).

The Madeira is one of the major Andean tributary of the Amazon and previous phylogeographic studies evidenced that the piranha genera *Serrasalmus* and *Pygocentrus* colonised the Andean tributaries of the Amazon during only the last 2 Ma (Hubert *et al*., in press). Although the colonisation of the Upper Madeira is recent, molecular phylogenetic results suggested that speciation occurred in *Serrasalmus* within the Upper Madeira watershed (Hubert *et al*., 2006). This may be related to the existence of varied water types in the area as a function of the relative contribution of the Brazilian shield, the Tertiary sediments of the lowlands and the Andes (Sioli, 1975; Guyot *et al*., 1999). A total of seven *Serrasalmus* species genetically well differentiated and characterised by private alleles at diagnostic and semi-diagnostic nuclear loci may be found in the area (Hubert *et al*., 2006). Among this set of well-recognised species, three endemic species from the Madeira River, namely *S. compressus, S. hollandi* and a *Serrasalmus* sp (Hubert *et al*., 2006), constitute a monophyletic group suggesting that speciation occurred within the same watershed (Hubert *et al*., in press). If the three species have a recent and common origin, then they may still exhibit shared ancestral polymorphism due to a recent divergence and currently fall within the range of the lineage sorting period. In this context, poor concordance between the gene tree and species tree may be expected. Such a pattern would reinforce the hypothesis of a common geographic origin within the Madeira watershed. Hence, in order to achieve a better understanding of the structuring events and evolution of this endemic group of *Serrasalmus*

50 species in the Upper Madeira River, we explored the genealogy of the mtDNA control region
51 from samples of the three species throughout their distribution range.
52
53 **2. Materials and methods**
54 *2.1 Hydrological context and sampling*
55 The Madeira River is the second largest tributary of the Amazon (1.37×10^6 km^2) after the
56 Solimões (2.24×10^6 km^2) and is characterised by a marked annual cycle of rainy and dry
57 seasons responsible for multi-peaked floods in the Andean tributaries. The downstream pulse
58 is stored in the Bolivian floodplain, which is one of the largest of the Amazon with a potential
59 flood extension of 0.15×10^6 km^2 (Guyot *et al.*, 1999). The headwaters represent at least 60%
60 of the overall watershed area and they can be separated into four major systems with distinct
61 hydrological typology (Fig. 1). Currently, three types of water are recognised in the Amazon:
62 (1) the white waters characterised by a great amount of dissolved solid materials and a low
63 transparency (Andean origin); (2) the clear water characterised by a low content of dissolved
64 solid and a high transparency (Brazilian or Guyana shields) and (3) the black water
65 originating from the forested lowlands and differing from the latter by having a higher content
66 of humic acids and a lower pH (Sioli, 1975). Within the Upper Madeira, the Guaporé River
67 drains almost exclusively the Brazilian shield and so it is characterised by clear waters. By
68 contrast, the Mamoré and Madre de Dios Rivers originate in the Andes. Their main channels
69 are constituted by white waters and small lowland tributaries with black water are frequently
70 encountered along their main channel. Finally, the Yata is a small central tributary hosting
71 black lowland waters.
72 A total of six rivers were sampled between September 2002 and June 2003 (Fig. 1;
73 Table 1). In the Guaporé, specimens from clear water sites in the headwater (Fig. 1; 1) and the
74 lower course (Fig. 1; 2) were sampled. In the Mamoré, specimens from one white water

tributary originating in the Andean flank were sampled (Fig. 1; 3) while both a white water (Fig.1; 4) and clear water tributary (Fig. 1; 5) were prospected in the Madre de Dios. A single black water site was sampled from the Yata River (Fig. 1; 6).

2.2 DNA extraction and sequencing

Genomic DNA was isolated from ethanol-preserved tissues with the DNeasy Tissue Kit (Qiagen). The mtDNA control region was amplified using the primers CR22U: 5' TGGTTTAGTACATATTATGCAT (Hubert *et al.*, in press) and F-12R: 5' GTCAGGACCATGCCTTTGTG (Sivasundar *et al*., 2001). These primers amplify a fragment of 980 bp beginning in the position 100 of *Colossoma macropomum* control region (accession number: AF283963) and including the 3' flanking tRNA genes (tRNA Thr and tRNA Pro). PCR were performed in 50 µl volumes including 13.5-µl of template DNA (approximately 1 µg), 3 units of Taq DNA polymerase, 5 µl of Taq 10x buffer, 3 µl of $MgCl_2$ (25mM), 4 µl of dNTP (5mM) and 3 µl of each primer (10 µM). PCR conditions were as follows: 94 °C (5 min), 10 cycles of 94 °C (1 min), 66 °C to 56 °C decreasing of 1 °C per cycle (1 min 30 s), 72 °C (2 min), 25 cycles of 94 °C (1 min), 56 °C (1 min 30 s), 72 °C (2 min), followed by 72 °C (5 min). PCR products were sequenced in both directions. The consensus sequences have been deposited in GenBank and vouchers have been deposited in the Muséum National d'Histoire Naturelle, Paris (Table 1).

2.3 Analysis of mtDNA variability

Multiple alignments of the control region were performed using CLUSTAL W (Thompson *et al.*, 1993). Sequences were aligned with 3 different schemes of gap opening and extending costs as follow, opening cost = 5 and extending cost = 4; opening cost = 15 and extending cost = 6 (default setting); opening cost = 20 and extending cost = 8, in order to detect

potential alignment ambiguous sites defined as positions with gap assignment differing among alternatives cost functions (Gatesy *et al.*, 1994). Phylogenetic relationships among the control region haplotypes sampled were constructed using Maximum Likelihood (ML) as implemented in PhyML (http://atgc.lirmm.fr/phyml) following the algorithm developed by Guindon & Gascuel (2003). The Akaike Information Criterion (AIC) identified the optimal model as implemented in Modeltest 3.7 (Posada & Crandall, 1998), and was further used for tree searches and bootstrap analyses based on 1000 replicates in PhyML. Within each mtDNA clades identified, genealogies of the control region haplotypes were constructed following the statistical parsimony method of Templeton *et al.* (1992) as implemented in the TCS software (Clement *et al*., 2000). Alternative ambiguous connections resulting from homoplastic mutations were resolved by comparison with the ML tree. Finally, the analysis of molecular variance (AMOVA; Excoffier *et al*., 1992) provided an estimate of the distribution of nucleotide diversity at three levels of subdivision: among species (CT); among watersheds, within species (SC) and among individuals, within watersheds (ST). The correlation of alleles at each of the three hierarchical levels was assessed using the Φ-statistics (Excoffier *et al*., 1992) tested by 1000 permutations of individuals as implemented in Arlequin 2.0 (Schneider *et al*., 2000).

3. Results and discussion

A total of 957 bp were sequenced in 70 specimens including 23 *S. compressus*, 22 *S. hollandi* and 25 *S*. sp (Table 1). Together with nine sequences of *S. compressus*, *S. hollandi* and *S*. sp previously published (Hubert *et al*., in press), control region sequences from 79 individuals were analysed here. *Serrasalmus marginatus* is the sister species of the clade including *S. compressus*, *S. hollandi* and *S.* sp (Hubert *et al*., in press) and two sequences of *S. marginatus* previously published were used as outgroup for subsequent analyses (Table 1).

The three alignments schemes provided the same alignment indicating that no alignment ambiguous sites were present in this data set. Within the 957 sites analysed, 89 sites were variable among which 66 were informative, and a single insertion-deletion of 1 bp was observed. The AIC indicated that the HKY+I+Γ model fitted the present data set better than others and was used for subsequent ML searches (Fig. 2; -lnL = 2239.58). A poor correspondence between the gene tree and the species tree was observed and three clusters of sequences were identified in the ML tree, namely cluster I, II and III (Fig. 2). In general, internal branches were short and deep nodes were statistically poorly supported (Fig. 2). As no alignment ambiguous sites were detected, the lack of statistical support seems to be better explained by a fast differentiation of the mtDNA lineages rather than character conflict due to molecular saturation and homoplasy. The latter hypothesis is consistent with previous phylogenetic results arguing for a fast differentiation of the *Serrasalmus* lineages (Hubert *et al*., in press).

Cluster I is further subdivided into two distinct clades, the first represented only by sequences from individuals of *S. compressus* and the second by sequences from individuals of *S*. sp (Fig. 2). Likewise, cluster II is further subdivided into two distinct clades, the first including seven sequences from *S. compressus* and the second including 18 sequences from *S.* sp in addition to one from *S. compressus*. The parsimony network inferred for cluster II indicates that haplotype sharing occurs between these two species and hybridisation and introgression cannot be rejected. Finally, cluster III harbours no subdivision. This clade consists of a poorly supported polytomy represented by sequences from both *S. hollandi* and *S*. sp. Once again, the parsimony network evidences some haplotype sharing between these two species, which cannot be explained by the retention of ancestral polymorphism alone. In this case, introgression through hybridisation is likely. The AMOVA evidenced that most of the nucleotide variability was found within watershed rather than species as 50% of the

variability in the control region sequences was explained by variation within watershed while only 33% of the variability was explained by differences between species (Table 2). However, the variation between species was found significant indicating that drift shaped species genealogy for long enough to imprint a significant differentiation of the mtDNA lineages.

The maintenance of ancestral polymorphism from a common ancestor may be expected to result in a distinct distribution of the coalescent events between species when compared with hybridisation and gene flow. Recent isolation and ancient polymorphism is likely to relate species through coalescent events generally older than the speciation event as homogamy tend to increase the proportion of young coalescent events within species (Pamilo & Nei, 1988). By contrast, hybridisation and gene flow will relate species polymorphism through coalescent events from varied ages (Wakeley, 1996). In this context, distributions of pairwise differences between species are likely to be distinct when considering isolation and ancestral polymorphism or gene flow through hybridisation, the latter leading to haplotype sharing of recently derived haplotypes and young coalescent events between species.

Distribution of pairwise differences within species and within clusters confirmed that the clusters poorly matched the species limits as sequences were more closely related within clusters than within species (Fig. 2). Likewise, the distribution of pairwise differences between species exhibited a complex trimodal distribution very similar to the distribution of pairwise differences within species. A major mode is found around 15-17 differences and two minor modes, the first around two differences and the second around 33-35 differences (Fig. 2D). The superposition of the modes around 15-17 and 33-35 differences in the within species and between species distributions is characteristic of recent isolation and ancient polymorphism with an excess of old coalescent events within species. By contrast, the mode around 2 differences between species is characteristic of young coalescent events within species rather than between species (Fig. 2D). If introgression through past hybridisation

created this mode between sympatric species, comparisons with an allopatric and physically isolated outgroup should differ by lacking it. The distribution of pairwise differences between *S. marginatus* from the Paraná and *S. compressus*, *S. hollandi* and *S.* sp from the Madeira lacks this mode at two differences and further supports that the excess of recent coalescent events between sympatric species from the Madeira originated from introgression through past hybridisation (Fig. 2E).

The present pattern of mixed mtDNA lineages between species has several implications. The distributions of pairwise differences between sympatric (*S. compressus*, *S. hollandi*, *S*. sp) or allopatric species (with *S. marginatus*) indicate that recent isolation and ancestral polymorphism alone is unlikely to produce haplotype sharing and account for the occurrence of recent coalescent events between sympatric species. The present result makes the hypothesis of mtDNA introgression through past hybridisation very likely. This contrast with the well differentiation of allelic pools from nuclear DNA (nDNA) previously described between *Serrasalmus compressus*, *S. hollandi* and *S*. sp (Hubert *et al*., 2006). Actually, several causes may be account to this apparent discrepancy between mtDNA and nDNA. Only size differences between alleles were previously assessed for nDNA and pattern of coalescence between alleles has not been considered (Hubert *et al*., 2006). Hence, recent coalescent events between species in the nDNA may have not been previously detected through the analyses of length differences due to insertion-deletion events. However, this artefact seems unlikely in front of the number of nuclear loci previously analysed (Hubert *et al*., 2006) Alternatively, the occurrence of mtDNA introgression through maternal lineages cannot be discarded and seems very likely.

Another implication from the present study concerns the geography and ecology of the speciation events at the origin of the three sympatric species from the Upper Madeira, namely *S. compressus*, *S. hollandi* and *S*. sp. The genealogy of the control region haplotypes argues

that this group of sympatric species still falls in the range of the lineage sorting period. The three species are tightly restricted to the Madeira River and the present pattern supports a common and recent origin in the same watershed rather than more complex scenarios involving allopatric divergence in different watersheds, secondary contacts and extirpations. Also, the abundance of each of the three species in the different tributaries of the Upper Madeira was not properly addressed here, as this was not the focus of the present study, some trends seems to emerge from the present sampling (Table 1). The two species, *Serrasalmus hollandi* and *S.* sp seems to be alternatively distributed as the former was more frequently sampled in white- to mixed-water tributaries (Béni and Mamoré river) while the latter was almost exclusively observed in clear- to black-water tributaries (Yata, Itenez and Manuripi rivers). Cytogenetic studies of *Serrasalmus* in the central Amazon previously detected cryptic reproductive units distributed alternatively in white or black waters (Centofante *et al.*, 2002). The present pattern supports a recent and common geographic origin and suggests that adaptive divergence to the variety of water type in the headwaters of the Madeira River may have been an important factor in shaping reproductive isolation between these endemic species (Schluter, 2001).

Acknowledgments

This work was part of the PhD of Nicolas Hubert on the evolution of the piranha. This research was supported by Institut de Recherche pour le Développement (IRD, France); Instituto de Biología Molecular y Biotechnología, La Paz (IBM y B, Bolivia), Instituto de Limnología, La Paz (Bolivia), and the laboratory GPIA, Montpellier (France). We thank N. Bierne, B. Guinand and E. Lambert from the GPIA laboratory; G. Rodriguo, N. Mamani and V. Iñiguez from the IBMB for laboratory supports and facilities. We wish to thank F. Carvajal, A. Parada, L. Torres, T. Yunoki for their help during field sampling, J. Pinto, R.

Marin and M. Legendre for their support. We thank P. Pruvost, L. Nandrin and R. Causse from the MNHN for providing facilities in the ichthyological collection.

References

Bowie, R.C., Fjeldså, J., Hackett, S.J., Bates, J.M., Crowe, T.M., 2005. Coalescent models reveal the relative roles of ancestral polymorphism, vicariance, and dispersal in shaping phylogeographical structure of an African montane forest robin. Molecular Phylogenetics and Evolution 38, 171-188.

Buckup, P.A., 1998. Relationships of the Characidiinae and phylogeny of Characiform fishes (Teleostei: Ostariophysi). In: Malabarba, L.R., Reis, R.E., Vari, R.P., Lucena, Z.M., Lucena, C.A.S. (Eds.), Phylogeny and classification of Neotropical fishes, Universidade Católica do Rio Grande do Sul, Porto Alegre (EDIPUCRS), pp. 251-260.

Centofante, L., Porto, J.I.R., Feldberg, E., 2002. Chromosomal polymorphism in *Serrasalmus spilopleura* Kner, 1858 (Characidae, Serrasalminae) from central Amazon Basin. Caryologia 55, 37-45.

Chiang, T.Y., 2000. Lineage sorting accounting for dissociation between chloroplast and mitochondrial lineages in oaks of southern france. Genome 43, 1090-1094.

Clement, M., Posada, D., Crandall, K.A., 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9, 1657-1659.

Excoffier, L., Smouse, P., Quattro, J.M., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: an application to human mitochondrial DNA restriction data. Genetics 131, 479-491.

Funk, D.J., Omland, K.E., 2003. Species-level paraphylyl and polyphyly: frequency, causes and consequences, with insights from animal mitochondria DNA. Annula Review of Ecology, Evolution and Systematics 34, 397-423.

250 Gatesy, J., DeSalle, R., Wheeler, W., 1994. Alignment ambiguous nucleotide sites and the
251 exclusion of systematic data. Molecular Phylogenetics and Evolution 2, 152-157.
252 Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large
253 phylogenies by Maximum Likelihood. Systematic Biology 52, 696-704.
254 Guyot, J.L., Jouanneau, J.M., Wasson, J.G., 1999. Characterisation of the river bed and
255 suspended sediments in the Rio Madeira drainage basin (Bolivian Amazonia). Journal of
256 South American Sciences 12, 401-410.
257 Hoelzer, G.A., Wallman, J., Melnick, D.J., 1998. The effects of social structure, geographical
258 structure, and population size on the evolution of mitochondrial DNA: II. Molecular clocks
259 and the lineage sorting period. Journal of Molecular Evolution 47, 21-31.
260 Hubert, N., Duponchelle, F., Nuñez, J., Riveira, R., Renno, J.F., 2006. Evidence of
261 reproductive isolation among sympatric closely related species of Serrasalmus
262 (Ostariophysii, Characidae) from the Upper Madeira River. Journal of Fish Biology 69A,
263 31-51.
264 Hubert, N., Renno, J.F., 2006. Historical Biogeography of South American Freshwater fishes.
265 Journal of Biogeography 33, 1414-1436.
266 Hubert, N., Duponchelle, F., Nuñez, J., Garcia-Davila, C., Paugy, D., Renno, J.F. (in press)
267 Phylogeography of the piranha genera *Serrasalmus* and *Pygocentrus*: implications for the
268 diversification of the Neotropical Ichthyofauna. Molecular Ecology.
269 Jégu, M., 2003. Serrasalminae. In: Reis, R.E., Kullander, S.O., Ferraris, C.J. (Eds.), Check
270 List of freshwater fishes of South and Central America. Universidade Católica do Rio
271 Grande do Sul, Porto Alegre (EDIPUCRS), pp. 182-196.
272 Kingman, J.F.C., 1982. The coalescent. Stochastic Process and their Applications 13, 245-
273 248.

Machado, C.A., Hey, J., 2002. The causes of phylogenetic conflict in a classic Drosophila species group. Proceedings of the Royal Scoiety of London, Series B 270, 1193-1202.

Nielsen, R., Wakely, J.,2001. Distinguishing migration from isolation: a markov chain monte carlo approach. Genetics 158, 885-896.

Nores, M., 1999. An alternative hypothesis to the origin of Amazonian bird diversity. Journal of Biogeography 26, 475-485.

Pamilo, P., Nei, M., 1988. Relationships between gene trees and species trees. Molecular Biology and Evolution 5, 568-581.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817-818.

Rokas, A., Melika, G., Abe, Y., Nieves-Aldrey, J.L., Cook, J.M., Stone, G.N., 2003. Lifecycle closure, lineage sorting, and hybridization revealed in a phylogenetic analysis of european oak gall wapsps (Hymenoptera: Cynipidae: Cynipini) using mitochondrial sequence data. Molecular Phylogenetics and Evolution 26, 36-45.

Rosenberg, N.A., 2003. The shapes of neutral gene genealogies in two species: probabilities of monophyly, paraphyly, and polyphyly in a coalescent model. Evolution 57, 1465-1477.

Schluter, D., 2001. Ecology and the origin of species. Trends in Ecology and Evolution 16, 372-380.

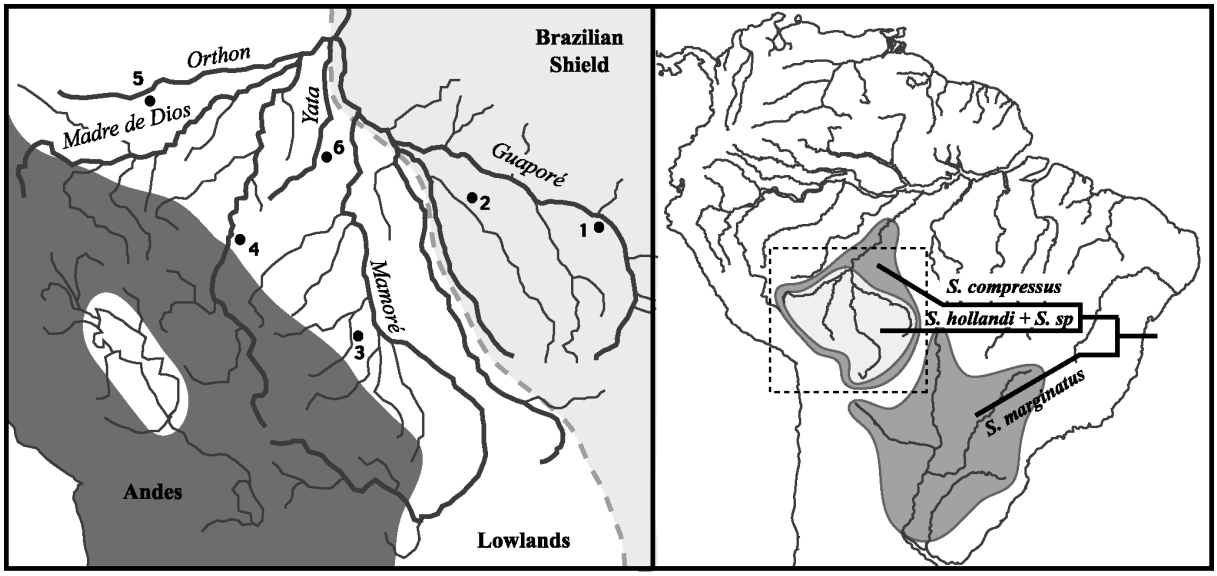
Schneider, S., Roessli, D., Excoffier, L., 2000. Arlequin version 2.0: a software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva. Geneva, Switzerland.

Sioli, H., 1975. Tropical rivers as expressions of their terrestrial environments. In: Golley, F.B., Medina, E. (Eds.), Tropical Ecological Systems: Trends in Terrestrial and Aquatic Research, Springer Verlag, Berlin, pp. 275–288.

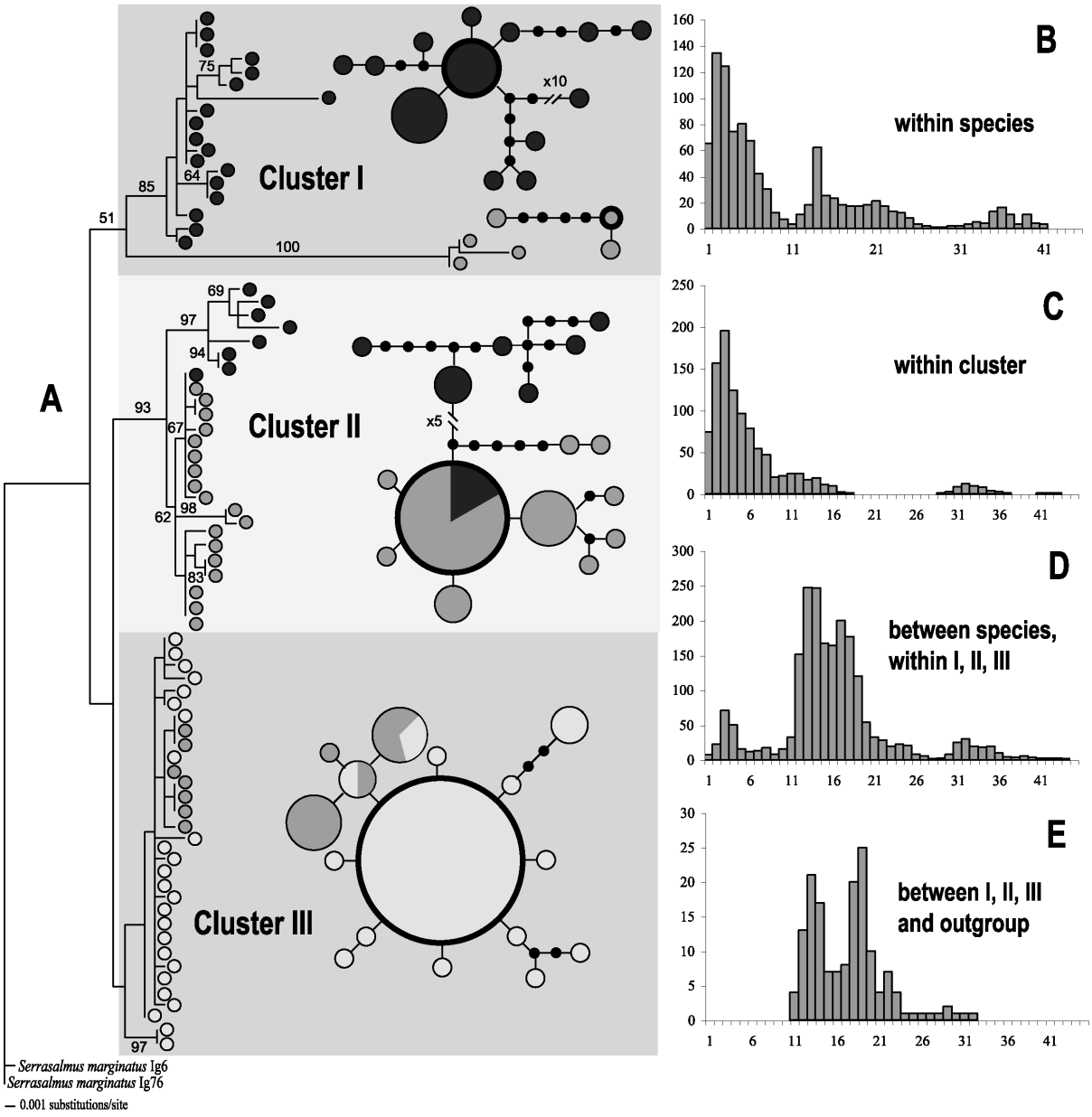
298 Sivasundar, A., Bermingham, E., Ortí, G., 2001. Population structure and biogeography of
299 migratory freshwater fishes (*Prochilodus*: Characiformes) in major South American rivers.
300 Molecular Ecology 10, 407-417.
301 Tajima, F., 1983. Evolutionary relationships of DNA sequences in finite populations.
302 Genetics 105, 437-460.
303 Takahashi, K., Terai, Y., Nishida, M., Okada, N., 2001. Phylogenetic relationships and
304 ancient incomplete lineage sortin among cichlid fishes in lake tanganyika as revealed by
305 analysis of the insertion of retroposons. Molecular Biology and Evolution 18, 2057-2066.
306 Templeton, A.R., Crandall, K., Sing, C.F., 1992. A cladistic analyses of phenotypic
307 associations with haplotypes inferred from restriction endonuclease mapping and DNA
308 sequence data. III. Cladogram estimation. Genetics 132, 619-633.
309 Thompson, J.D., Higgins, D.G., Gibson, T.J., 1993. CLUSTAL W: improving the sensitivity
310 of progressive multiple sequence alignment through sequence weighting, position-specific
311 gap penalties and weight matrix choice. Nucleic Acids Research 22, 4673-4680.
312 Wakeley, J., 1996. The variance of pairwise nucleotide differences in two populations with
313 migration. Theoritical Populations Biology 49, 39-57.

Fig. 1. Distribution range of *Serrasalmus marginatus*, *S. compressus, S. hollandi* and known sampling area of *S*. sp, and sampling sites of *S. compressus, S. hollandi* and *S*. sp within the Upper Madeira watershed (each point may represent more than one locality). The Brazilian shield is represented in light grey while the Andes are represented in dark grey. 1, upper Guaporé; 2, lower Guaporé in the San Martin River; 3, lower Mamoré in the Isiboro River; 4, Béni River in the Madré de Dios watershed; 5, Orthon River in the Manuripi tributary; 6, Yata River.

Fig. 2. Phylogenetic relationships among control regions sequences of *Serrasalmus compressus, S. hollandi* and *S*. sp. A. ML tree inferred using the model HKY+I+Γ with the following parameters: base frequencies A = 0.31, G = 0.22, C = 0.17, T = 0.30, transition/transversion ratio = 11.98, proportion of invariable sites = 0.76, gamma shape parameter = 0.66, number of categories = 4. For each cluster identified, the corresponding genealogy inferred using the statistical parsimony framework of Templeton *et al*., 1992 is provided. Ancestral haplotypes inferred are indicated with bold lines. B, mismatch distribution of pairwise differences within the three species *S. compressus, S. hollandi* and *S*. sp. C, mismatch distribution of pairwise differences within the three clusters I, II and III. D, mismatch distribution of pairwise differences between species within the clade including cluster I, II and III. E, mismatch distribution of pairwise differences between the outgroup and the species from the clade including cluster I, II and III.


Brazilian
Shield
Orthon
Madre de Dios
Yata
Guaporé
Mamoré
5
6
4
2
1
3
Andes
Lowlands
S. compressus
S. hollandi + S. sp
S. marginatus

334


A
Cluster I
Cluster II
Cluster III
x10
x5
75
64
85
51
100
69
97
94
93
67
98
62
83
97
Serrasalmus marginatus Ig6
Serrasalmus marginatus Ig76
0.001 substitutions/site
B
within species
C
within cluster
D
between species, within I, II, III
E
between I, II, III and outgroup

335