Journal of Fish Biology (2006) **69** (Supplement A), 31–51
doi:10.1111/j.1095-8649.2006.01097.x, available online at http://www.blackwell-synergy.com

Evidence of reproductive isolation among closely related sympatric species of *Serrasalmus* (Ostariophysii, Characidae) from the Upper Madeira River, Amazon, Bolivia

N. Hubert*†§, F. Duponchelle†‡, J. Nuñez†§, R. Rivera§ and J.-F. Renno†§

†U.R. 175 Institut de Recherche pour le Développement (IRD), GAMET, BP 5095, 361 rue JF Breton, 34196 Montpellier Cedex 05, France, §Instituto de Biología Molecular y Biotecnología, Universidad Mayor de San Andrés, La Paz, Bolivia and ‡Unidad de Limnología y Recursos Acuáticos, Universidad Mayor de San Simón, Cochabamba, Bolivia

(Received 29 July 2005, Accepted 31 January 2006)

The delimitation of the *Serrasalmus* species (Characiformes, Serrasalminae) from the Bolivian Amazon (Amazonas, Madeira) was examined using intron length polymorphism assessed by EPIC-PCR. The six pairs of primers provided 16 polymorphic loci across the species of the region and the allelic diversity ranged between two and 10 alleles per locus. For each locus, the lack of departure from Hardy-Weinberg expectations in a large number of the populations analysed and the homogenous distribution of linkage disequilibrium between paralogous loci and loci belonging to different intronic systems provided strong lines of evidence that the 16 amplified loci constituted independent neutral markers. Furthermore, allelic diversity was size-dependent, thereby indicating that insertion-deletions occurred frequently but randomly in introns, and that intron length polymorphism was a valid marker for investigating the systematics of piranhas. EPIC-PCR demonstrated that eight of the nine nominal species of piranha of the Upper Madeira were reproductively isolated and identified a new species that differed from its closest morphological and genetic relatives by seven diagnostic or semi-diagnostic loci. By contrast, no diagnostic or semi-diagnostic locus was found between *S. spilopleura* and *S. eigenmanni*, nor were their allelic frequencies different, thereby questioning the validity of their biological species status, at least in the Upper Madeira. This study, which was one of the first applications of EPIC-PCR to a large-scale molecular systematic purpose, demonstrates that it is a rapid, reliable and cost-effective tool for elucidating issues pertaining to fish systematics.

Key words: *Catoprion*; EPIC-PCR; nuclear introns; piranhas; *Pygocentrus*; *Serrasalmus*.

INTRODUCTION

Piranhas are the largest of all characiform carnivores (110–410 mm in standard length L_S; Fink & Machado-Allison, 1992; Merckx *et al.*, 2000) and the most

*Author to whom correspondence should be addressed at the present address: UMR5171, Station Méditerranéenne de l'Environnement Littoral, 1 quai de la Daurade, 34200 Sète, France. Tel.: +4 67 46 33 77; fax: +4 67 46 33 99; email: myloplus@excite.com

speciose group with 34 nominal species from five genera, namely *Serrasalmus*, *Pristobrycon*, *Pygocentrus*, *Pygopristis* and *Catoprion* (Jégu, 2003). Their systematics has confounded neotropical ichtyologists for decades (Fink, 1993; Fink & Machado-Allison, 2001; Jégu & Dos Santos, 2001) and recent studies of their biology have demonstrated a great ecological diversity (Machado-Allison, 1985; Machado-Allison & Garcia, 1986; Jégu & Dos Santos, 1988; Nico & Taphorn, 1988; Jégu *et al.*, 1991; Fink & Machado-Allison, 1992; Machado-Allison & Fink, 1995, 1996). Morphological and molecular data agree that the genera *Serrasalmus*, *Pristobrycon*, *Pygocentrus*, *Pygopristis* and *Catoprion* constitute a monophyletic unit but internal relationships are still debated, in particular molecular results suggest that at least two genera (*Serrasalmus* and *Pristobrycon*) are not monophyletic (Fink & Machado-Allison, 1992; Ortí *et al.*, 1996). As a corollary, the systematics of the group is complex and still only inadequately understood since several species remain undiagnosed (Fink, 1993). These issues are exemplified within the genus *Serrasalmus*, which comprises 24 of the 34 species of piranhas and probably encompass several lineages (Ortí *et al.*, 1996).

Several centres of endemism are already known for piranhas in South America among them the Upper Madeira River (Bolivia), where extensive systematic studies only began in the early 1980s (Vari, 1982, 1984, 1989*a*,*b*; Lauzanne *et al.*, 1991; Jégu *et al.*, 1991). Among the 389 fish species found in the region (Lauzanne *et al.*, 1991), nine morphological species of piranha belonging to the genera *Serrasalmus*, *Pygocentrus* and *Catoprion* have been reported and at least three of these species, namely *Serrasalmus hollandi* Eigenmann, *Serrasalmus compressus* Jégu, Leão & Dos Santos and *Serrasalmus spilopleura* Kner, are endemic (Jégu, 2003). In the absence of detailed molecular systematic studies, the origin and biogeographic affinities of the Amazonian piranha faunas are still to be appraised. These issues have fostered the use of genetic tools to further elucidate the systematics of piranhas. Cytogenetic tools have been applied in the Lower Amazon region, where they proved of limited use to discriminate among *Serrasalmus* species, because variations in the numbers of chromosomes or nuclear organizer regions are low in this group of fishes (Porto *et al.*, 1992; Nakayama *et al.*, 2000, 2001, 2002; Centofante *et al.*, 2002). By contrast, the analysis of nuclear DNA enables the identification of non-overlapping gene pools, which permits the unequivocal identification of sympatric lineages that are reproductively isolated (De Queiroz, 1998; Sites & Marshall, 2003). Although mitochondrial DNA have smaller effective population size and higher evolutionary rates (Avise, 1994), the maternal inheritance and problems of lineage sorting of mtDNA among closely related species (Pamilo & Nei, 1988) make allelic frequency data from nuclear DNA a more suitable molecular tool for delimiting species (Zawadzki *et al.*, 2000, 2004; Fisch-Muller *et al.*, 2001; Borsa *et al.*, 2004). Various nuclear markers are available for this purpose (Avise, 1994; O'Hanlon *et al.*, 2000; Sunnucks, 2000; Zhang & Hewitt, 2003).

Pioneered by Lessa (1992), the intron-targeted PCR amplify an intron using primers designed in the contiguous sequences of the highly conserved exons. This approach, called Exon-Primed Intron-Crossing (EPIC)-PCR (Palumbi & Baker, 1994), was successfully used in several studies of population genetics (Côrte-Real *et al.*, 1994; He & Haymer, 1997, 1999; Gomulski *et al.*, 1998; Villablanca *et al.*,

1998; Daguin *et al*., 2001; Bierne *et al*., 2002, 2003; Hassan *et al*., 2003) and more recently in molecular systematics (Naeole & Haymer, 2003). EPIC-PCR enjoys several practical advantages: by using primers designed from orthologous genes, cloning and sequencing of target sequences can be avoided (Côrte-Real *et al*., 1994; Bierne *et al*., 2000); cross-species amplification is easier when primers are designed in coding sequences because exon sequences are more conserved across species, and PCR artefacts such as null alleles are expected to be less frequent for the same reason. There is now a large number of universal primers available for a broad series of loci (Chow & Hazama, 1998; Chow & Takeyama, 1998; Bierne *et al*., 2000; Hassan *et al*., 2002; Atarhouch *et al*., 2003; Chow & Nakadate, 2004), which makes the EPIC-PCR a suitable tool when prior genomic knowledge of the species understudy is scarce.

The Upper Madeira is one of the major tributaries of the Amazon and hosts a great number of endemic species (Jégu, 2003). Since the real number of piranha species remains unknown, however, the evolutionary history of the fauna is still to be elucidated. Thus, the present study aimed at delimiting piranha species among the Upper Madeira River in a molecular systematic perspective based on nuclear DNA assessed by EPIC-PCR. Since the Bolivian piranha species are morphologically very similar, species are difficult to identify at the early stages of the life cycle and this study further investigated the potential applications of EPIC-PCR in the molecular identification of the earliest ontogenetic stages.

MATERIAL AND METHODS

SAMPLES LOCATION

Fishes were collected in the upper watershed of the Madeira River, which corresponds to the Bolivian Amazon. With an area of $1{\cdot}37\ 10^6$ km^2 (Guyot, 1993), the Madeira River is the second largest tributary of the Amazon and is second to the Solimões River with $2{\cdot}24\ 10^6$ km^2. The upper watershed of the Madeira River represents at least 60% of the total area of the basin. It includes the Bolivian floodplain with a potential flood extension of $0{\cdot}15\ 10^6$ km^2. The tributaries of the upper Madeira watershed exhibit a great diversity of water-chemistry depending on the relative contribution of the Andes (turbid water) or the Brazilian shield (clear water). Hence, nine sites encompassing the ecological diversity of the Madeira River were sampled between September 2002 and June 2003 (Fig. 1) using hook and line, and gillnets with a mesh-size ranging from 10 to 50 mm. The nine species described in this region (Jégu, 2003) were collected during this study (Table I). Among them, three are largely distributed in the rivers of South America [*Catoprion mento* (Cuvier), *Pygocentrus nattereri* Kner and *Serrasalmus maculatus* Kner] while three others are largely distributed in the Amazon and Orinoco rivers but are lacking in the Paraná river [*Serrasalmus eigenmanni* Norman, *Serrasalmus elongatus* Kner and *Serrasalmus rhombeus* (L.)] and three are restricted to the upper watershed of the Madeira river (*S. compressus*, *S. hollandi* and *S. spilopleura*). Samples of muscle were taken from all specimens and preserved in 90% ethanol solution. Vouchers were conserved in a 30% formaldehyde solution and then transferred into an 80% ethanol solution.

INTRON LENGTH POLYMORPHISM ASSESSMENT

The following primers were used: GPD2F/3R which amplify the intron 2 of the glyceraldehyde 3-phosphate dehydrogenase (Hassan *et al*., 2002); RPEX1F/2R which amplify the intron 1 of the S7 ribosomal protein (Chow & Takeyama, 1998); GH5F/6R

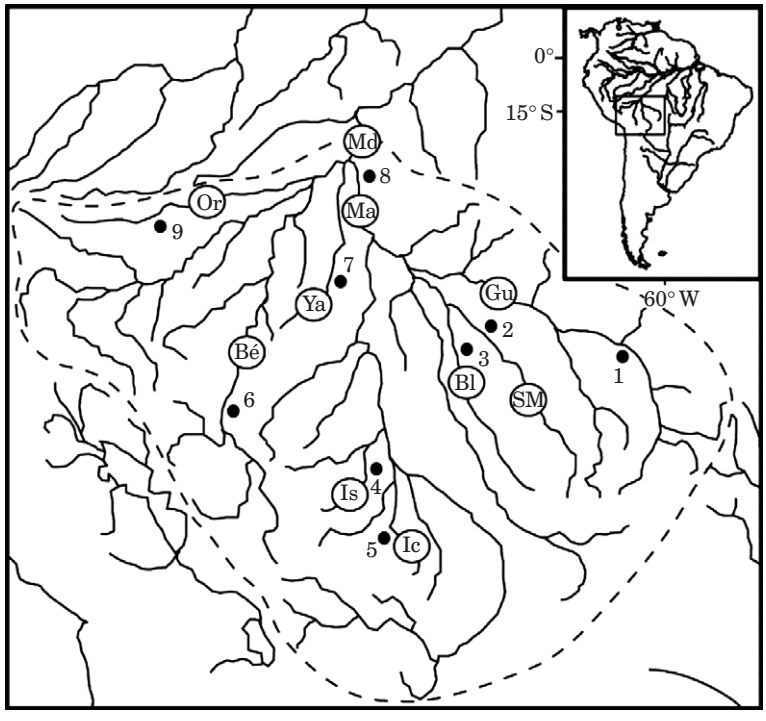

FIG. 1. Maps showing the sampling sites of the *Serrasalmus*, *Pygocentrus* and *Catoprion* specimens. Md, Madeira basin; Gu, Guaporé River (sampling site 1); SM, San Martin River (sampling site 2); Bl, Blanco River (sampling site 3); Is, Isiboro River (sampling site 4); Ic, Ichilo River (sampling site 5); Bé, Béni River (sampling site 6); Ya, Yata River (sampling site 7); Ma, Lower Mamoré River (sampling site 8); Or, Orthon River (sampling site 9).

which amplify the intron 5 of the growth hormone (Hassan *et al.*, 2002); Am2b2F/3R which amplify the intron 2 of the alpha amylase (Hassan *et al.*, 2002); PmOPSIF/R, designed from cDNA sequences of the opsin (Bierne *et al.*, 2000) and GnRH1F/R which amplify the intron 1 of the gonadotropin-releasing hormone 3 (Hassan *et al.*, 2002). For each individual, genomic DNA was extracted from muscle tissue by standard CTAB/chloroform protocol and used as a template for amplifications performed in 25 µl including 5 µl of genomic DNA, 2·5 µl of Taq buffer (×10), 62·5 pM of $MgCl_2$, 5 pM of dNTP and 20 pM of each primer. Amplifications were carried out for 30 cycles including denaturation at 95° C for 1 min, annealing at 40–50° C for 1 min and extension at 72° C for 1 min. To detect size-polymorphism, electrophoresis of non-denatured PCR products was performed in 0·8 mm-thick of 5% polyacrylamide gel. After migration at 40 W for 2 h up to 5 h, depending on the size of the alleles, the gel was fixed in a 30% ethanol solution, and then transferred to a 1% HNO_3 solution for 3 min followed by 2 min. in a 0·2% $AgNO_3$ solution and finally revealed in a 2 : 1 buffer of $NaCO_3$ (30 g L^{-1}) and formaldehyde (0·08%). Length variations were scored using the numerical procedure implemented in LabImage 3·0 b (LabImage, 2005) using a size-standard 100 bp gene ruler. When individual size assignment was ambiguous, PCR products were assayed side-by-side for a second run. This size-polymorphism assessment allowed detecting differences up to 2–3 bp but only differences of 5 bp or more were retained.

Previous studies relying on EPIC-PCR reported that various loci might be scored for a given pair of primers (France *et al.*, 1999; Atarhouch *et al.*, 2003; Hassan *et al.*, 2003; Borsa *et al.*, 2004). This phenomenon may reflect former polyploidizations, tandem

TABLE I. Allelic frequencies and intron-length polymorphism details among sampling sites (see Fig. 1) for *Serrasalmus*, *Pygocentrus* and *Catoprion* species. Sampling sites with small samples size were discarded from the analysis

	S.rho					*S.mac*					*S.hol*			*S.*sp				*S.com*	*S.spi*	*S.eig*			*P.nat*	*C.men*	*S.elo*
Sampling sites	2	3	5	7	Total	2, 3	4	6	8	Total	4	8	Total	1	7	2	Total	1, 2, 6–9	1, 2	2	7	Total	2, 3	2, 3	2, 7
N	38	10	31	58	147	16	10	14	26	71	7	9	16	12	7	22	41	13	7	12	16	28	9	9	9
RPEX1-1																									
1090	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0·07	0·04	0·16	0·11	–	–	–
1070	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0·14	0·25	0·44	0·36	–	–	–
1060	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0·79	0·71	0·40	0·53	–	–	–
1020	0·17	0·25	0·08	0·09	0·13	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1005	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0·33	–	–
995	0·68	0·75	0·82	0·89	0·80	1·00	1·00	1·00	1·00	1·00	–	–	–	–	–	–	–	–	–	–	–	–	0·67	–	1·00
955	0·15	–	0·10	0·02	0·07	–	–	–	–	–	–	–	–	–	–	0·27	0·15	–	–	–	–	–	–	–	–
950	–	–	–	–	–	–	–	–	–	–	1·00	1·00	1·00	1·00	1·00	0·73	0·85	1·00	–	–	–	–	–	1·00	–
H_E	0·49	0·40	0·31	0·22	0·34	–	–	–	–	–	–	–	–	–	–	0·41	0·25	–	0·39	0·45	0·62	0·58		–	–
H_O	0·33	0·30	0·26	0·21	0·26	–	–	–	–	–	–	–	–	–	–	0·18	0·10	–	0·14	0·33	0·31	0·32		–	–
f	0·33*	0·25	0·18	0·06	0·24***	–	–	–	–	–	–	–	–	–	–	0·56*	0·62***	–	0·65	0·27	0·52*	0·47*	0·06	–	–
RPEX1-2																									
945	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0·86	0·33	0·13	0·41	–	–	–
935	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	0·14	0·67	0·87	0·79	1·00	1·00	1·00
H_E	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0·26	0·46	0·23	0·34	–	–	–
H_O	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0·29	0·33	0·13	0·22	–	–	–
f	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0·41	0·29	0·46	0·39	–	–	–
RPEX1-3																									
790	–	–	–	–	–	1·00	1·00	0·714	1·00	0·939	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
770	–	–	–	–	–	–	–	0·286	–	0·061	–	–	–	–	–	–	–	1·00	–	–	–	–	1·00	–	–
755	0·04	–	–	–	0·01	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
745	0·96	1·00	1·00	1·00	0·99	–	–	–	–	–	1·00	1·00	1·00	–	–	–	–	–	1·00	1·00	1·00	1·00	–	–	–
735	–	–	–	–	–	–	–	–	–	–	–	–	–	1·00	1·00	1·00	1·00	–	–	–	–	–	–	–	–
725	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1·00	1·00

TABLE I. Continued

	S.rho					*S.mac*					*S.hol*			*S.*sp				*S.com*	*S.spi*	*S.eig*			*P.nat*	*C.men*	*S.elo*
Sampling sites	2	3	5	7	Total	2, 3	4	6	8	Total	4	8	Total	1	7	2	Total	1, 2, 6–9	1, 2	2	7	Total	2, 3	2, 3	2, 7
H_E	0·08	–	–	–	0·02	–	–	0·42	–	0·11	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
H_O	0·03	–	–	–	0·01	–	–	0·28	–	0·06	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
f	0·66	–	–	–	0·67**	–	–	0·33	–	0·47*	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
GH5-1																									
1150	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0·28	–	–
1140	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0·72	1·00	–
1135	–	–	–	–	–	1·00	1·00	1·00	1·00	1·00	–	–	–	–	–	–	–	–	1·00	1·00	1·00	1·00	–	–	–
1130	–	–	–	–	–	–	–	–	–	–	–	–	–	1·00	1·00	1·00	1·00	–	–	–	–	–	–	–	1·00
1120	0·24	0·20	0·21	0·33	0·27	–	–	–	–	–	1·00	1·00	1·00	–	–	–	–	1·00	–	–	–	–	–	–	–
1105	0·76	0·80	0·79	0·67	0·73	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
H_E	0·37	0·34	0·34	0·44	0·39	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
H_O	0·33	0·40	0·23	0·41	0·35	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
f	0·11	−0·20	0·33	0·07	0·12	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0·23	–	–
GH5-2																									
510	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1·00	–
490	–	–	–	–	–	–	–	–	–	–	–	–	–	1·00	1·00	1·00	1·00	–	–	–	–	–	–	–	–
480	0·01	–	0·07	0·08	0·05	–	–	–	–	–	1·00	1·00	1·00	–	–	–	–	1·00	–	–	–	–	–	–	1·00
475	0·99	1·00	0·93	0·92	0·95	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1·00	–	–
465	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0·57	0·75	0·66	0·69	–	–	–
455	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0·43	0·25	0·34	0·31	–	–	–
450	–	–	–	–	–	0·28	0·80	0·54	0·42	0·47	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
440	–	–	–	–	–	0·72	0·20	0·46	0·58	0·53	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
H_E	0·03	–	0·12	0·14	0·10	0·42	0·34	0·52	0·50	0·50	–	–	–	–	–	–	–	–	0·53	0·39	0·47	0·43	–	–	–
H_O	0·03	–	0·13	0·05	0·06	0·31	0·00	0·22	0·39	0·27	–	–	–	–	–	–	–	–	0·29	0·33	0·44	0·39	–	–	–
f	0·000	–	−0·05	0·64***	0·4***	0·26	1***	0·59	0·23	0·46***	–	–	–	–	–	–	–	–	0·48	0·15	0·06	0·1	–	–	–
GPD2-1																									
875	1·00	1·00	1·00	1·00	1·00	–	–	–	–	–	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	–	–	1·00

860	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1·00	–
830	–	–	–	–	–	0·09	0·20	0·07	0·39	0·78	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
805	–	–	–	–	–	0·91	0·80	0·93	0·61	0·22	–	–	–	–	–	–	–	–	–	–	–	–	1·00	–	–
H_E	–	–	–	–	–	0·18	0·34	0·14	0·48	0·35	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
H_O	–	–	–	–	–	0·19	0·40	0·14	0·31	0·26	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
f	–	–	–	–	–	−0·07	−0·2	−0·04	0·37	0·26	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
GPD2-2																									
720	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1·00	1·00	–
705	1·00	1·00	1·00	1·00	1·00	–	–	–	–	–	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	–	–	–
690	–	–	–	–	–	1·00	1·00	1·00	1·00	1·00	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1·00
H_E	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
H_O	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
f	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
GPD2-3																									
515	1·00	1·00	1·00	1·00	1·00	–	–	–	–	–	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00
505	–	–	–	–	–	1·00	1·00	1·00	1·00	1·00	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
H_E	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
H_O	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
f	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
GPD2-4																									
360	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0·61		–
350	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	0·39	1·00	–
340	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1·00
H_E	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0·50	–	
H_O	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0·56	–	
f	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	−0·11	–	–
Am2b2-1																									
585	–	–	–	–	–	0·09	0·05	0·04	0·02	0·05	–	–	–	1·00	1·00	1·00	1·00	1·00	–	–	–	–	–	–	–
570	–	–	–	–	–	0·19	0·15	0·18	0·12	0·15	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
560	1·00	1·00	1·00	1·00	1·00	0·72	0·80	0·78	0·86	0·80	1·00	1·00	1·00	–	–	–	–	–	1·00	1·00	1·00	1·00	–	–	1·00
545	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
540	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1·00	–

TABLE I. Continued

	S.rho					*S.mac*					*S.hol*			*S.*sp				*S.com*	*S.spi*	*S.eig*			*P.nat*	*C.men*	*S.elo*
Sampling sites	2	3	5	7	Total	2, 3	4	6	8	Total	4	8	Total	1	7	2	Total	1, 2, 6–9	1, 2	2	7	Total	2, 3	2, 3	2, 7
530	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1·00	–	–
H_E	–	–	–	–	–	0·45	0·35	0·36	0·24	0·33	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
H_O	–	–	–	–	–	0·31	0·20	0·36	0·27	0·29	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
f	–	–	–	–	–	0·32	0·45	0·02	−0·12	0·14	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Am2b2-2																									
525	1·00	1·00	1·00	1·00	1·00	0·19	0·05	0·04	0·52		–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
520	–	–	–	–	–	0·63	0·35	0·60	0·27		1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	–	–	–	–	–	–	1·00
510	–	–	–	–	–	0·18	0·60	0·36	0·21		–	–	–	–	–	–	–	–	1·00	1·00	1·00	1·00	–	–	–
500	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1·00	–	–
H_E	–	–	–	–	–	0·56	0·54	0·52	0·63		–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
H_O	–	–	–	–	–	0·63	0·60	0·36	0·50		–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
f	–	–	–	–	–	−0·13	−0·12	0·32	0·2	0·21*	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Am2b2-3																									
440	0·65	0·75	0·81	0·73	0·73	–	–	–	–	–	1·00	1·00	1·00	0·21	0·07	0·18	0·17	–	1·00	1·00	1·00	1·00	–	–	–
435	0·35	0·25	0·19	0·27	0·27	1·00	1·00	1·00	1·00	1·00	–	–	–	0·80	0·93	0·82	0·83	1·00	–	–	–	–	1·00	–	–
425	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1·00
H_E	0·46	0·40	0·32	0·40	0·40	–	–	–	–	–	–	–	–	0·34	0·14	0·30	0·29	–	–	–	–	–	–	–	–
H_O	0·33	0·10	0·26	0·32	0·30	–	–	–	–	–	–	–	–	0·08	0·14	0·09	0·10	–	–	–	–	–	–	–	–
f	0·28	0·28	0·19	0·17	0·25***	–	–	–	–	–	–	–	–	0·77	0·00	0·71***	0·66***	–	–	–	–	–	–	–	–
PmOPSI-1																									
440	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1·00	–	–
420	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1·00	1·00	1·00	1·00	–	–	–
410	0·47	0·40	0·86	0·46	0·55	–	–	–	–	–	0·29	0·22	0·43	1·00	1·00	1·00	1·00	0·35	–	–	–	–	–	–	–
395	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0·19	–	–	–	–	–	–	–
390	0·53	0·60	0·15	0·54	0·45	–	–	–	–	–	0·71	0·78	0·57	–	–	–	–	0·46	–	–	–	–	–	1·00	1·00
380	–	–	–	–	–	1·00	1·00	1·00	1·00	1·00	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–

H_E	0·51	0·51	0·25	0·50	0·50	–	–	–	–	0·44	0·37	0·51	–	–	–	0·66	–	–	–	–	–	–	–	–	–	–	–	–	–	–
H_O	0·54	0·40	0·23	0·33	0·37	–	–	–	–	0·29	0·44	0·38	–	–	–	0·39	–	–	–	–	–	–	–	–	–	–	–	–	–	–
f	−0·01	0·22	0·11	0·35**	0·26***	–	–	–	–	0·37	−0·23	0·27	–	–	–	0·42	–	–	–	–	–	–	–	–	–	–	–	–	–	–
PmOPSL-2																														
345	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	0·43	0·44	0·44	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00
335	–	–	–	–	–	–	–	–	–	0·57	0·56	0·56	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
H_E	–	–	–	–	–	–	–	–	–	0·53	0·52	0·51	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
H_O	–	–	–	–	–	–	–	–	–	0·29	0·67	0·50	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
f	–	–	–	–	–	–	–	–	–	0·48	−0·3	0·02	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
GnRH1-1																														
1145	–	–	–	–	–	–	–	–	–	0·14	0·17	0·16	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1130	–	–	–	–	–	–	–	–	–	0·21	0·28	0·25	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1100	–	–	–	–	–	0·13	–	0·04	0·15	0·10	0·29	0·44	0·38	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1095	–	–	–	–	–	0·72	0·60	0·54	0·73	0·67	–	–	–	–	–	–	1·00	1·00	1·00	1·00	1·00	–	1·00	1·00	1·00	–	1·00	1·00	–	1·00
1085	0·27	0·30	0·16	0·36	0·29	–	–	–	–	0·36	0·11	0·22	1·00	1·00	1·00	1·00	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1075	–	–	–	–	–	0·15	0·40	0·42	0·12	0·23	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1065	0·13	0·15	0·32	0·16	0·19	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1045	0·21	0·25	0·37	0·22	0·25	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
1035	0·39	0·30	0·15	0·27	0·27	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
995	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	1·00	–	–	–	–
H_E	0·70	0·774	0·72	0·73	0·75	0·46	0·51	0·55	0·44	0·49	0·78	0·73	0·75	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
H_O	0·56	0·500	0·74	0·67	0·65	0·44	0·40	0·36	0·39	0·39	0·86	0·89	0·88	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
f	0·22*	0·37*	−0·03	0·08	0·14***	0·05	0·22	0·36	0·12	0·21	−0·11	−0·24	−0·18	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
GnRH1-2																														
550	–	–	–	–	–	0·13	–	–	0·12	0·08	–	–	–	–	–	–	1·00	–	–	–	–	–	–	–	–	1·00	–	1·00	–	1·00
540	–	–	–	–	–	0·56	0·80	0·25	0·61	0·55	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
530	1·00	1·00	1·00	1·00	1·00	0·31	0·20	0·75	0·27	0·37	1·00	1·00	1·00	1·00	1·00	1·00	–	1·00	1·00	1·00	1·00	1·00	1·00	1·00	1·00	–	1·00	–	–	–

TABLE I. Continued

	S.rho					*S.mac*					*S.hol*			*S.*sp				*S.com*	*S.spi*	*S.eig*			*P.nat*	*C.men*	*S.elo*
Sampling sites	2	3	5	7	Total	2, 3	4	6	8	Total	4	8	Total	1	7	2	Total	1, 2, 6–9	1, 2	2	7	Total	2, 3	2, 3	2, 7
H_E	–	–	–	–	–	0·59	0·34	0·39	0·55	0·56	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
H_O	–	–	–	–	–	0·25	0·20	0·21	0·23	0·23	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
f	–	–	–	–	–	0·58***	0·42	0·46	0·58***	0·59***	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
H_E (*total*)	0·19	0·17	0·15	0·17	0·18	0·19	0·17	0·21	0·22	0·21	0·13	0·12	0·13	0·02	0·01	0·05	0·04	0·05	0·08	0·09	0·10	0·09	0·10	–	–
H_O (*total*)	0·15	0·12	0·13	0·14	0·14	0·15	0·13	0·14	0·15	0·14	0·10	0·14	0·13	0·01	0·01	0·02	0·01	0·03	0·05	0·07	0·06	0·07	0·09	–	–
f (*total*)	0·19***	0·31*	0·11	0·18***	0·21***	0·2**	0·26*	0·34***	0·27***	0·33***	0·2	−0·26	0·01	0·77	0	0·62***	0·64***	0·42	0·41	0·24	0·35*	0·33*	0·05	–	–

f, Weir & Cockerham's (1984) estimate of the correlation of alleles within individuals relative to the populations of the *Serrasalmus*, *Pygocentrus* and *Catoprion* species; *N*, sample size; H_O, observed heterozigozity; *H*e, unbiased expected heterozigosity (Nei, 1987).
S.rho, *Serrasalmus rhombeus; S.hol*, *Serrasalmus hollandi; S.elo*, *Serrasalmus elongatus; S.*sp, *Serrasalmus* sp.; *S.com*, *Serrasalmus compressus; S.spi*, *Serrasalmus spilopleura; S.eig*, *Serrasalmus eigenmanni; S.mac*, *Serrasalmus maculatus; P.nat*, *Pygocentrus nattereri; C.men*, *Catoprion mento*.

duplications and other phenomena occurring during lineage evolution (Atarhouch *et al.*, 2003). Since the species understudies are diploïd (Porto *et al.*, 1992; Nakayama *et al.*, 2000, 2001, 2002; Centofante *et al.*, 2002), gene duplication processes are more likely involved. Hence, it is important to distinguish between orthologous and paralogous genes given that, two homologous genes are termed orthologous if their most recent common ancestor did not undergo gene duplication, otherwise they are termed paralogous (Page & Holmes, 1998). Since the present study aimed at delimiting species among closely related taxa, homologies are to be warranted as EPIC-PCR relies on universal primers. When several paralogous genes were amplified for a given intronic system, however, the following criteria for the delimitation of the putative orthologous loci were used: (1) when intraspecific allelic variants were observed, a locus was evidenced by the presence of all the allelic combinations following a Mendelian inheritance (homozygotes and heterozygotes); (2) if not, in order to circumvent erroneous allelic assignment, only loci where the allelic size-variation was lower between orthologous than paralogous loci, were retained.

ANALYSIS OF INTRON LENGTH POLYMORPHISM

The expectations of the Hardy–Weinberg (HW) genotypic proportions were first checked given that, if the putative loci analysed constitute neutral markers following a Mendelian inheritance, HW expectations should be met for each, at least at the population level. Thus, the observed heterozygosity (H_O), the unbiased estimate of expected heterozygosity (H_E) (Nei, 1987) and the correlation of alleles within individual assessed by the estimator f (Weir & Cockerham, 1984) were estimated for each species and populations when samples were available. To give the expected allelic distribution for each locus under the null hypothesis of HW proportions ($f = 0$), 1000 pseudo-matrices of individual genotypes were generated from the original matrix by random permutations of alleles at a locus. The null hypothesis of $f = 0$ was rejected when $P < 0{\cdot}025$ following a two tailed test. Then, if the loci analysed constituted independent markers, the proportions of linkage disequilibrium among paralogous loci might be expected to not exceed those found among loci from different intronic systems. The correlation of alleles among locus was estimated using the coefficient of genotypic disequilibrium D (Black & Krafsur, 1985). For each pair of alleles, the coefficient of correlation R_{ij} and the genotypic disequilibrium D_{ij} (i, allele at locus 1; j, allele at locus 2) are calculated and the overall genotypic disequilibrium D is computed by a weighted average procedure. The significance of D was further assessed by a permutation procedure of alleles among species and the null hypothesis of $D = 0$ was rejected when $P < 0{\cdot}05$ (one tailed test). The estimations of f and D and permutations tests were performed using the procedures *F*stats and LinkDis in the Genetix software (Belkhir *et al.*, 2004). Potential differences in the proportions of linkage disequilibrium among putative paralogous loci or loci from different intronic systems by the χ^2 test of homogeneity were further cheeked. Testing for correlation between locus size and allelic diversity also assessed the allelic variation across loci. Assuming that insertion-deletion events (INDELs) appear at random, their occurrence would be greater in large than small loci, thereby producing size-dependent allelic diversity. Correlation tests were performed using a covariance analysis as implemented in Statgraphics (Stat Graphics, 1999).

Estimating the amount of diagnostic and semi-diagnostic loci checked the genetic differentiation of the nine species. A locus is termed diagnostic between two species when both species exhibit private and alternative alleles; the term 'semi-diagnostic' applies to a locus for which some of the alleles are displayed by both species, while the remaining alleles are private. When species are examined, the reproductive isolation between distinct evolutionary lineages is detected from the absence of heterozygote for the private alleles at the diagnostic loci. A neighbour-joining (NJ) phenogram based on allelic frequencies was constructed by the mean of the Cavalli-Sforza & Edwards's (1967) chord distance using Phylip 3·57 c (Felsenstein, 1993), to depict the pattern of genetic relationships among the populations of each species. *Catoprion mento*, which is the sister

group of the genera *Serrasalmus* and *Pygocentrus* (Ortí *et al.*, 1996), was used for rooting the phenogram. Statistical support of the topology as well as gene pool differentiation were estimated using bootstrap proportion (BP) computed with 1000 pseudo-replicates. Levels of genetic differentiation among population and species are finally compared in a mismatch distribution of the pair-wise genetic distance used for the NJ phenogram.

RESULTS

Of the 345 fish analysed in this study, 41 individuals could not be assigned unequivocally to any of the nine nominal species based on morphological criteria. Their morphological characteristics differed substantially from those of genera *Catoprion*, *Pristobrycon*, *Pygocentrus* and *Pygopristis*, so they were assigned to the genus *Serrasalmus*. In all analyses, these fishes were grouped separately then analysed as if they were a different morphological species, and termed *Serrasalmus* sp.

INTRON LENGTH VARIABILITY

The following nuclear loci were amplified: *GPD2-1*, *GPD2-2*, *GPD2-3*, *GPD2-4* with the primers *GPD2F*/*3R*; *RPEX1-1*, *RPEX1-2*, *RPEX1-3* with the primers *RPEX1F*/*2R*; *GH5-1*, *GH5-2* with the primers *GH5F*/*6R*; *Am2b2-1*, *Am2b2-2*, *Am2b2-3* with the primers Am2b2F/3R; *PmOPSI-1*, *PmOPSI-2* with the primers *PmOPSIF*/*R* and *GnRH1-1*, *GnRH1-2* with the primers *GnRH1F*/*R*. The allelic diversity varied greatly between loci ranging from two alleles for *RPEX1-2* up to 10 for *GH5-2* and *GnRH1-1* in the whole data set (Table I). It is worth noting that no variation was observed in the number of amplified paralogous loci across species. Among this set of 16 putative loci, two exhibited intraspecific monomorphism (*GPD2-2* and *GPD2-3*; Table I). For these two loci, size-variation was lower between orthologous than paralogous loci. Among the remaining 14 loci, allelic size-variation was greater between orthologous than paralogous loci in three cases (between *RPEX1-1* and *RPEX1-2*; *Am2b2-1* and *Am2b2-2*; *PmOPSI-1* and *PmOPSI-2*; Table I). Intraspecific polymorphism, however, was detected for this set of 14 loci and all the allelic combinations were observed (Table I). For *RPEX1-1* and *RPEX1-2*, *S. eigenmanni* exhibited allelic variations at both loci and all the allelic combinations of the three and two alleles of *RPEX1-1* and *RPEX1-2*, respectively, were observed. Additionally, no heterozygous combinations between alleles of *RPEX1-1* and *RPEX1-2* were observed in *S. eigenmanni*. In the same way, all the allelic combinations of the three alleles of *RPEX1-1* of *S. rhombeus* were observed and no heterozygous genotypes including the allele *RPEX1-2*935* were observed, thereby indicating that this allele belong to the loci *RPEX1-2* (Table I). Furthermore, all the orthologous loci shared similar alleles, at least in part. The same observations were made for *Am2b2-1* and *Am2b2-2* in *S. maculatus* or *PmOPSI1-1* and *PmOPSI1-2* in *S. hollandi* where all the allelic combinations were observed for each locus and orthologous loci shared similar alleles (Table I).

Three lines of evidence suggested that the 16 loci from the present data set constitute independent neutral markers. For polymorphic loci, permutation test of genotypic HW proportions (*f*) were found largely significant among

populations when they were frequently not within populations (*RPEX1-1*, *GH5-1*, *GH5-2*, *Am2b2-3*, *PmOPSI-1* and *GnRH1-1* for *S. rhombeus*; *RPEX1-1*, *RPEX1-2* and *GH5-2* in *S. eigenmanni*; *GH5-2*, *GPD2-1*, *Am2b2-1*, *Am2b2-2*, *GnRH1-1* and *GnRH1-2* in *S. maculatus*; *Am2b2-3* in *S.* sp; *PmOPSI-1*, *PmOPSI-2* and *GnRH1-1* in *S. hollandi*; Table I), indicating that most of the populations sampled locally fit HW proportions. Hence, inferred genotype frequencies were compatible with HW expectation drawn from a Mendelian inheritance for each locus. Then, distribution of linkage disequilibrium between putative paralogous loci or loci belonging to different intronic systems was homogenous (χ^2, d.f. = 1, $P = 0{\cdot}40$) with 9/15 and 75/106 cases, respectively, where D was significant (permutation test, $P < 0{\cdot}05$). Finally, a significant positive relationship was observed between locus size and allelic diversity ($F_{4,\ 67}$, $P = 0{\cdot}003$) indicating that INDELs occurred at random in the present data set.

DELIMITING PIRANHA SPECIES

The genetic differentiation between nine of the 10 species was large (Table II): the number of diagnostic and semi-diagnostic loci ranges from seven (*S. compressus v. S. hollandi*) to 14 (*P. nattereri v. S. eigenmanni*) against a maximum of 16 possible loci. Furthermore the respective number of diagnostic and semi-diagnostic loci greatly fluctuated between pairs of species: there were seven semi-diagnostic loci but no more than a single diagnostic locus between *S. hollandi* and *S. rhombeus*, and up to 12 diagnostic loci and two semi-diagnostic loci between *P. nattereri* and *S. eigenmanni*. The unidentified *Serrasalmus* sp. was genetically well differentiated from all other sympatric species, with the total number of diagnostic and semi-diagnostic loci ranging from seven to 12 depending on comparisons. By contrast, *S. eigenmanni* and *S. spilopleura*, although described as distinct morphological species, shared the same alleles for 16 loci and both lack alternative alleles.

TABLE II. Genetic differentiation between the 10 species of the genera *Serrasalmus*, *Pygocentrus* and *Catoprion* (see Table I). Upper matrix, Cavalli-Sforza & Edwards' (1967) chord distance; lower matrix, diagnostic *v.* semi-diagnostic (in parentheses) loci

	S.rho	*S.mac*	*S.hol*	*S.elo*	*S.*sp	*S.com*	*S.spi*	*S.eig*	*P.nat*	*C.men*
S.rho	–	0·665	0·315	0·610	0·425	0·550	0·475	0·469	0·681	0·735
S.mac	**7**(4)	–	0·749	0·648	0·719	0·674	0·667	0·661	0·731	0·815
S.hol	**1**(7)	**9**(4)	–	0·498	0·266	0·376	0·480	0·474	0·877	0·590
S.elo	**6**(1)	**7**(4)	**7**(1)	–	0·621	0·436	0·733	0·727	0·797	0·690
*S.*sp	**4**(3)	**9**(3)	**3**(5)	**6**(1)	–	0·247	0·635	0·629	0·791	0·666
S.com	**5**(3)	**7**(4)	**4**(3)	**8**(1)	**5**(2)	–	0·664	0·658	0·647	0·643
S.spi	**6**(3)	**8**(4)	**6**(2)	**9**(1)	**8**(2)	**9**(1)	–	0·002	0·896	0·733
S.eig	**6**(3)	**8**(4)	**6**(2)	**9**(1)	**8**(2)	**9**(1)	0(0)	–	0·890	0·727
P.nat	**9**(3)	**4**(6)	**11**(2)	**9**(2)	**10**(2)	**9**(1)	**12**(2)	**12**(2)	–	0·769
C.me	**9**(2)	**7**(3)	**7**(1)	**9**(0)	**9**(1)	**9**(2)	**10**(1)	**10**(1)	**12**(1)	–

Cavalli-Sforza & Edwards's (1967) chord distance ranged from 0·247 (*S. spilopleura v. S. compressus*) to 0·896 (*S. spilopleura v. P. nattereri*) (Table II). The NJ phenogram confirmed the great differentiation of gene pools among species [Fig. 2(a) and Appendix]. The high BP values among populations of each species indicate that gene flow is high within species, which contrasts the strong differentiation between species. The NJ phenogram identified at least seven genetic units. Branch lengths indicate long-term barriers to gene flow between these sympatric taxa except for *S. spilopleura* and *S. eigenmanni* whose populations are nested within each other with high statistical confidence. *Serrasalmus maculatus* showed the most divergent gene pool of the *Serrasalmus* species and the new *Serrasalmus* sp. was most closely related to *S. compressus* [Fig. 2(a)]. Mismatch distribution of pair-wise genetic distance confirmed the great divergence of species gene pools [Fig. 2(b)]. Genetic distance between populations ranged between 0·001 and 0·090 and exhibited a unimodal distribution related to the lack of alternative alleles between populations. By contrast, the pair-wise genetic distance between species harboured a plurimodal distribution related to the three genetic cluster of the NJ phenogram [cluster 5 including *S. rhombeus*, *S. eigenmanni*, *S. spilopleura* and *S. hollandi*; cluster 4 including *S.* sp and *S. compressus*; cluster 1 including *S. maculatus* and *P. nattereri*; Fig. 2(a)].

DISCUSSION

Naeole & Haymer (2003) pioneered EPIC-PCR to discriminate between species in a species complex of fruit flies. The present study was the first large-scale application of EPIC-PCR to the systematics of fish species. The lack of significant departures for each loci from HW expectations in populations, the homogeneity of genotypic disequilibrium between putative paralogous loci or loci belonging to different intronic systems and the significant correlation between allelic diversity and locus size constituted strong lines of evidence that the 16 loci of the present study constitute independent neutral nuclear markers following a Mendelian inheritance. The high level of polymorphism validated the relevance of EPIC-PCR for investigating issues pertaining to systematics. EPIC-PCR yielded unambiguous results about barriers to gene flow between eight of the nine nominal species of piranha in the Upper Madeira basin, while questioning the difference between *S. eigenmanni* and *S. spilopleura*, and revealing a new species of *Serrasalmus*. These clear-cut results contrast with the numerous difficulties or ambiguities generally encountered in piranha studies that relied on morphological (Fink & Machado-Allison, 2001; Jégu & Dos Santos, 2001) or cytogenetic criteria (Nakayama *et al.*, 2002).

At the intraspecific level, EPIC-PCR showed that the mean level of heterozygosity [H_O (total); Table I] among species of genus *Serrasalmus* ranged from 2·6 to 12·6%. These values fall in the range of H_O-values that were described from isozyme data for other Characiformes: 2·7% in *Schizodon intermedius* Garavello & Britski and 10·9% in *Leporinus friderici* (Bloch) from the central Amazon (Chiari & Sodré, 1999), and 13·2% in *Prochilodus lineatus* (Valenciennes) from the Upper Paraná (Revaldaves *et al.*, 1997). The multilocus departures from Hardy-Weinberg expectations [f (total); Table I] for both species and sampling sites indicates that, piranhas, at least *S. rhombeus*,

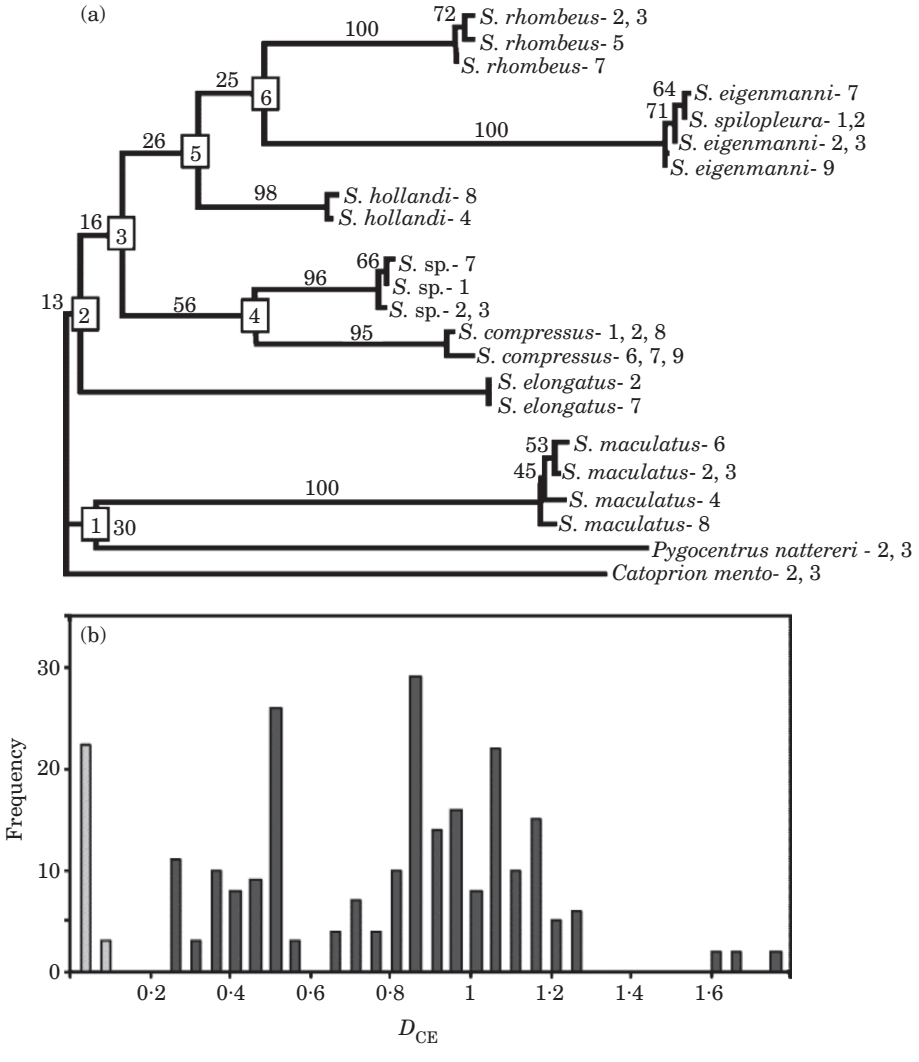

FIG. 2. (a) Neighbour-joining phenogram based on Cavalli-Sforza and Edwards's (1967) chord distance (D_{CE}) and rooted with *Catoprion mento*. Confidence values for nodes are percentages over 100 replicates. Nodes numbers referred to the allelic key to species in Appendix. Note that bootstrap proportions strongly support each species (see Table I) except for the differentiation of *Serrasalmus eigenmanni* (*S.eig*) and *Serrasalmus spilopleura* (*S.spi*). (b) Mistmatch distribution of pair-wise distances between populations (□) and species (■).

S. eigenmanni and *S. maculatus* for which sample size is sufficiently large, might be structured in genetic populations, without congruence with sample locations, among the Bolivian floodplain. It is worth noting that cytogenetic studies already documented slight differences in chromosome numbers of populations of *S. maculatus* in the central Amazon (Nakayama *et al.*, 2000, 2001; Centofante *et al.*, 2002). This result suggests that some complex population process

(*e.g.* Wahlund effect, kinship, population expansion) might be acting and need further sampling that will have to be analysed by EPIC-PCR.

No single diagnostic or semi-diagnostic locus was found between *S. eigenmanni* and *S. spilopleura* in this study. It cannot be ruled out that they might be different species yet, since possible differences might be found for loci other than those examined here. Nevertheless, this is quite unlikely in view of the high ($\geq$43%) proportion of diagnostic and semi-diagnostic loci between other species of piranhas from the Upper Madeira. Furthermore, no difference in the allelic frequencies of these two species was found, which further questions the validity of their status, at least in the Upper Madeira. Jégu & Dos Santos (2001) had previously emphasized that morphological differences between these two taxa were tenuous. Furthermore, no sympatric occurrence of *S. eigenmanni* and *S. spilopleura* has ever been reported, neither in the study of Jégu & Dos Santos (2001) nor in this study: *S. eigenmanni* is largely distributed in the Upper Madeira basin, but not in the Upper Guaporé River, where only *S. spilopleura* has been found until now. The allopatric distribution patterns of these two taxa and the absence of genetic differences between them suggest that their taxonomy has been misunderstood.

While EPIC-PCR questioned the validity of these two taxa in the Bolivian Amazon, it provided strong lines of evidence of a new species of *Serrasalmus*, which differed by no fewer than seven diagnostic and semi-diagnostic loci from its genetically closest (*S. compressus*) and morphologically closest (*S. rhombeus*) sympatric relatives in the Upper Madeira (see Fig. 2). This finding is of importance for understanding evolutionary relationships among piranhas in the Upper Madeira and to define relevant policies for conservation purposes. Currently, there is no single information on whether this new species of *Serrasalmus* is endemic to the Upper Madeira or present in other regions of the Amazon basin. Certainly, the thorough morphological and genetic diagnosis of this new species, which is currently on its way, will permit a finer revision of the preserved specimens, so as to determine other possible sites of occurrence, which will have to be sampled and analysed by EPIC-PCR to address this question.

That a high number of alternative alleles at diagnostic loci were found between closely related and morphologically very similar *Serrasalmus* species, hints that the present study opens heuristic perspectives about the molecular identification of piranha species, as exemplified by the allelic key to the species of the Upper Madeira in the Appendix. This key can be used to diagnose species at all stages of the life cycle, from eggs to adults, opening new perspectives for elucidating the ontogenetic and seasonal variations in resource use and habitat partitioning among sympatric species of piranhas.

EPIC-PCR challenges rapid techniques of nuclear polymorphism assessment as it allows a rapid and large screening without prior genomic knowledge. Compared with other DNA-based techniques such as microsatellites, EPIC-PCR is the only technique based on universal primers that allows a fast screening even for cross-species studies. Furthermore, since gene duplications could produce more than one locus, this characteristic is useful because it potentially provides a number of polymorphisms with only one pair of primers. By contrast with the great morphological and cytogenetic similarity of the *Serrasalmus* species, allelic data from intron length polymorphism evidence large level of

genetic divergences which in turn suggest some great eco-ethological differentiation and highly specialized mating systems. A result illustrating potential applications of EPIC-PCR is the constitution of the allelic key to *Serrasalmus* species from the Bolivian Amazon. It must be emphasized that such rapid and cost-effective technique of genetic characterization should find strong applications in evolutionary studies of neotropical fishes.

This work is part of the Ph.D of N. Hubert about the biogeography and population genetics of the Bolivian piranha fauna developed by IRD and the Instituto de Biología Molecular y Biotecnología (Universidad Mayor de San Andres, La Paz). We thank these institutions for their support and especially V. Iñiguez and N. Mamani from the La Paz University, T. Yunoki and A. Parada from Trinidad University. We would like to thank F. Bonhomme, E. and D. Barras and the two anonymous referees for their helpful comments on the earlier versions of the manuscript.

References

Atarhouch, T., Rami, M., Cattaneo-Berrebi, G., Ibanez, C., Augros, S., Boissin, E., Dakkak, A. & Berrebi, P. (2003). Primers for EPIC amplification of intron sequences for fish and other vertebrate population genetic studies. *Biotechniques* **35,** 676–682.

Avise, J. C. (1994). *Molecular Markers, Natural History and Evolution*. New York: Champman & Hall.

Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. & Bonhomme, F. (2004). *GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations*. Montpellier: Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II.

Bierne, N., Lehnert, S. A., Bédier, E., Bonhomme, F. & Moore, S. S. (2000). Screening for intron-length polymorphism in paenid shrimps using exon-primed intron-crossing (EPIC)-PCR. *Molecular Ecology* **9,** 233–235.

Bierne, N., David, P., Langlade, A. & Bonhomme, F. (2002). Can habitat specialization maintain a mosaic hybrid zone in marine bivalves? *Marine Ecology Progress Series* **245,** 157–170.

Bierne, N., Borsa, P., Daguin, C., Jollivet, D., Viard, D., Bonhomme, F. & David, P. (2003). Introgression patterns in the mosaic hybrid zone between *Mytilus edulis* and *Mytilus galloprovoncialis*. *Molecular Ecology* **12,** 447–461.

Black, W. C. & Krafsur, E. S. (1985). A FORTRAN program for the calculation and analysis of two-locus linkage disequilibrium coefficients. *Theoretical and Applied Genetics* **70,** 491–496.

Borsa, P., Collet, A. & Durand, J.-D. (2004). Nuclear-DNA markers confirm the presence of two anchovy species in the Mediterranean. *C. R. Biologies* **327,** 1113–1123.

Cavalli-Sforza, L. L. & Edwards, A. W. F. (1967). Phylogenetic analysis: models and estimation procedures. *Evolution* **32,** 550–570.

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.

Chiari, L. & Sodré, L. M. K. (1999). Genetic variability in five species of anostomidae (Ostariophysi – Characiformes). *Genetics and Molecular Biology* **22,** 517–523.

Chow, S. & Hazama, K. (1998). Universal PCR primers for S7 ribosomal protein gene in introns in fish. *Molecular Ecology* **7,** 1247–1263.

Chow, S. & Nakadate, M. (2004). PCR primers for fish G6PD gene intron and characterization of intron length variation in the albacore *Thunnus alalunga*. *Molecular Ecology Notes* **4,** 391–395.

Chow, S. & Takeyama, H. (1998). Intron length variation observed in the creatine kinase and ribosomal protein genes of Swordfish *Xiphias gladius*. *Fisheries Science* **64,** 397–402.

Côrte-Real, H. B. S. M., Dixon, D. R. & Holland, P. W. H (1994). Intron-targeted PCR: a new approach to survey neutral DNA polymorphism in bivalve populations. *Marine Biology* **120,** 407–413.

Daguin, C., Bonhomme, F. & Borsa, P. (2001). The zone of sympatry and hybridization of *Mytilus edulis* and *M. galloprovincialis*, as described by intron length polymorphism at locus *mac-1*. *Heredity* **86,** 342–354.

De Queiroz, K. (1998). The general lineage concept of species, species criteria, and the process of speciation. In *Endless Forms: Species and Speciation* (Howard, D. J. & Berlocher, S. H., eds), pp. 57–75. New York: Oxford University Press.

Felsenstein, J. (1993). *PHYLIP (Phylogeny Inference Package)*, *Version 3.5c*. Seattle, WA: Department of Genetics, University of Washington.

Fink, W. L. (1993). Revision of the piranha genus *Pygocentrus* (Teleostei, Characiformes). *Copeia* **1993,** 665–687.

Fink, W. L. & Machado-Allison, A. (1992). Three new species of piranhas from Brazil and Venezuela (Teleostei: Characiformes). *Ichtyological Exploration of Freshwaters* **3,** 55–71.

Fink, W. L. & Machado-Allison, A. (2001). *Serrasalmus hastatus*, a new species of piranha from Brasil, with comments on *Serrasalmus altuvei* and *Serrasalmus compressus* (Teleostei: Characiformes). *Occasional papers of the Museum of Zoology, the University of Michigan* **730,** 1–18.

Fisch-Muller, S., Mazzoni, S. R. & Weber, C. (2001). Genetic and morphological evidences for two new sibling species of *Ancistrus* (Siluriformes: Loricariidae) in upper Rio Tocantins drainage, Brazil. *Ichthyological Exploration of Freshwaters* **12,** 289–304.

France, S. C., Tachino, N., Duda, T. F., Schleser, R. A. & Palumbi, S. R. (1999). Intraspecific genetic divessity in the marine shrimp *Penaeus Vannamei*: multiple polymorphic elongation factor-1α loci revealed by intron sequencing. *Marine Biotechnology* **1,** 261–268.

Gomulski, L. M., Bourtzis, K., Brogna, S., Morandi, P. A., Bonvicini, C., Sebastiani, F., Torti, C., Guglielmino, C. R., Savakis, C., Gasperi, G. & Malacrida, A. R. (1998). Intron size polymorphism of the *Adh1* gene parallels the worldwide colonization history of the Mediterranean fruit fly, *Ceratitis capitata*. *Molecular Ecology* **7,** 1729–1741.

Guyot, J. L. (1993). *Hydrogéochimie des fleuves de l'Amazonie bolivienne*. Paris: Editions de l'ORSTOM, Collection Etudes et Thèses.

Hassan, M., Lemaire, C., Fauvelot, C. & Bonhomme, F. (2002). Seventeen new EPIC-PCR amplifiable introns in fish. *Molecular Ecology* **2,** 334–340.

Hassan, M., Harmelin-Vivien, M. & Bonhomme, F. (2003). Lessepsian invasion without bottleneck: example of two rabbitfish species (*Siganus rivulatus* and *Siganus luridus*). *Journal of Experimental Marine Biology and Ecology* **291,** 219–232.

He, M. & Haymer, D. S. (1997). Polymorphic intron sequences detected within and between populations of the Oriental fruit fly (Diptera: Tephritidae). *Annals of the Entomological Society of America* **90,** 825–831.

He, M. & Haymer, D. S. (1999). Genetic relationships of populations and the origins of new infestations of the Mediterranean fruit fly. *Molecular Ecology* **8,** 1247–1257.

Jégu, M. (2003). Characiformes: Characidae, Serrasalminae. In *Check-List of Freshwater Fishes from South and Central America* (Kullander, S. O., Ferraris, C. J. Jr & Reis, R. E., eds), pp. 182–196. Porto Alegre: EDIPUCRS.

Jégu, M. & Dos Santos, G. M. (1988). Le genre *Serrasalmus* (Pisces, Serrasalmidae) dans le bas Tocantins (Brésil, Parà), avec la description d'une espèce nouvelle, *S. geryi*, du bassin Aragaia-Tocantins. *Revue d'Hydrobiologie Tropicale* **21,** 239–274.

Jégu, M. & Dos Santos, G. M. (2001). Mise au point à propos de *Serrasalmus spilopleura* Kner, 1858 et réhabilitation de *S. maculatus* Kner, 1858 (Characidae: Serrasalminae). *Cybium* **25,** 119–143.

Jégu, M., Leão, E. L. & Dos Santos, G. M. (1991). *Serrasalmus compressus*, une espèce nouvelle du Rio Madeira, Amazonie (Pisces: Serrasalmidae). *Ichtyological Exploration of Freshwaters* **2,** 97–108.
LabImage (2005). *Gel Analysis Software*. Halle: Kapelan GmbH.
Lauzanne, L., Loubens, G. & Le Guennec, B. (1991). Liste commentée des poissons de l'Amazonie bolivienne. *Revue d'Hydrobiologie Tropicale* **24,** 61–76.
Lessa, E. P. (1992). Rapid survey of DNA sequence variation in natural populations. *Molecular Biology and Evolution* **9,** 323–330.
Machado-Allison, A. (1985). Estudios sobre la subfamilia Serrasalminae. Parte III. Sobre el estatus generico y relaciones filogeneticas de los generos *Pygopristis*, *Pygocentrus*, *Pristobrycon* y *Serrasalmus* (Teleostei, Characidae, Serrasalminae). *Acta Biologica Venezuela* **12,** 19–42.
Machado-Allison, A. & Fink, W. L. (1995). *Sinopsis de las especies de la subfamilia serrasalminae presentes en la cuenca del orinoco: claves, diagnosis e illustraciones*. Caracas: Universidad central de Venezuela.
Machado-Allison, A. & Fink, W. L. (1996). *Los peces caribes de Venezuela. Diagnosis, claves, aspectos ecologicos y evolutivos*. Caracas: Universidad central de Venezuela.
Machado-Allison, A. & Garcia, C. (1986). Food habits and morphological changes during ontogeny in three Serrasalmins fish species of the Venezuelan floodplains. *Copeia* **1986,** 193–196.
Merckx, A., Jégu, M. & Dos Santos, G. M. (2000). Une nouvelle espèce de *Serrasalmus* (Teleostei: Characidae: Serrasalminae), *Serrasalmus altispinnis n. sp.*, décrite du rio Uatumã (Amazonas, Brésil) avec une description complémentaire de *Serrasalmus rhombeus* (Linnaeus, 1766) du plateau Guyanais. *Cybium* **24,** 181–201.
Naeole, C. K. & Haymer, D. S. (2003). Use of oligonucleotide arrays for molecular taxonomic studies of closely related species in the Oriental fruit fly (*Bactrocera dorsalis*) complex. *Molecular Ecology Notes* **3,** 662–665.
Nakayama, C. M., Porto, J. I. R. & Feldberg, E. (2000). Ocorrência de dois citótipos em *Serrasalmus spilopleura* Kner, 1858 (Characiformes, Serrasalmidae) de região de confluência dos rios negro e solimões, Amazonas, Brasil. *Acta Amazonica* **30,** 149–154.
Nakayama, C., Jégu, M., Porto, J. I. R. & Feldberg, E. (2001). Karyological evidence for a cryptic species of piranha within *Serrasalmus rhombeus* (Characidae, Serrasalminae) in the Amazon. *Copeia* **2001,** 866–869.
Nakayama, C., Porto, J. I. R. & Feldberg, E. (2002). A comparative cytogenetic study of five piranha species (*Serrasalmus*, Serrasalminae) from the Amazon basin. *Genetica* **114,** 231–236.
Nei, M. (1987). *Molecular Evolutionary Genetics*. New York: Columbia University Press.
Nico, L. G. & Taphorn, D. C. (1988). Food habits of piranhas in the low llanos of Venezuela. *Biotropica* **20,** 311–321.
O'Hanlon, P. C., Peakall, R. & Briese, D. T. (2000). A review of new PCR-based markers and their utility to weed ecology. *Weed Research* **40,** 239–254.
Ortí, G., Petry, P., Porto, J. I. R., Jégu, M. & Meyer, A. (1996). Patterns of nucleotide change in the mitochondrial ribosomal RNA genes and the phylogeny of piranhas. *Journal of Molecular Evolution* **42,** 169–182.
Page, R. D. M. & Holmes, E. C. (1998). *Molecular Evolution, a Phylogenetic Approach*. Oxford: Blackwell Science.
Palumbi, S. R. & Baker, C. S. (1994). Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. *Molecular Biology and Evolution* **11,** 426–435.
Pamilo, P. & Nei, M. (1988). Relationships between gene trees and species trees. *Molecular Biology and Evolution* **5,** 568–581.
Porto, J. I. R., Feldberg, E., Nakayama, C. M. & Falcão, J. N. (1992). A checklist of chromosome numbers and karyotypes of Amazonian freshwater fishes. *Revue d'Hydrobiologie Tropicale* **25,** 287–299.

Revaldaves, E., Renesto, E. & Machado, M. F. P. S. (1997). Genetic variability of *Prochilodus lineatus* (Characiformes, Prochilodontidae) in the upper Paranà River. *Brazilian Journal of Genetic*. **20,** 381–388.
Sites, J. W. & Marshall, J. C. (2003). Delimiting species: a Renaissance issue in systematic biology. *Trends in Ecology and Evolution* **18,** 462–470.
StatGraphic (1999). Statistical Package. Los Angeles, CA: Statistical Graphics Corporation.
Sunnucks, P. (2000). Efficient genetic markers for population biology. *Trends in Ecology and Evolution* **15,** 199–203.
Vari, R. P. (1982). Systematics of the Neotropical Characoid genus *Curimatopsis* (Pisces: Characoidei). *Smithsonian Contributions to Zoology* **373**.
Vari, R. P. (1984). Systematics of the neotropical characiform genus *Potamorhina* (Pisces: Characiformes). *Smithsonian Contributions to Zoology* **400**.
Vari, R. P. (1989*a*). Systematics of the neotropical characiform genus *Curimata* Bosc (Pisces: Characiformes). *Smithsonian Contributions to Zoology* **474**.
Vari, R. P. (1989*b*). Systematics of the neotropical characiform genus *Psectrogaster* Eigenmann and Eigenmann (Pisces: Characiformes). *Smithsonian Contributions to Zoology* **481**.
Villablanca, F. X., Roderick, G. K. & Palumbi, S. R. (1998). Invasion genetics of the Mediterranean fruit fly: variation in multiple nuclear introns. *Molecular Ecology* **7,** 547–560.
Weir, B. S. & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38,** 1358–1370.
Zawadzki, C. H., Reis R. E. & Renesto, E. (2000). Allozyme descrimination of three species of *Loricariichthys* (Siluriformes: Loricariidae) from Southern Brazil. *Revue Suisse de Zoologie* **107,** 1–12.
Zawadzki, C. H., Renesto, E., de Paiva, S. & Lara-Kamei, M. C. S. (2004). Allozyme differentiation of four populations of *Hypostomus* (Teleostei: Loricariidae) from Ribeirão Keller, a small stream in the upper Rio Parana basin, Brazil. *Genetica* **121,** 251–257.
Zhang, D. X. & Hewitt, G. M. (2003). Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology* **12,** 563–584.

APPENDIX. Allelic key to the Upper Madeira species belonging to the genera *Serrasalmus* and *Pygocentrus*

1	a.- Always the alleles *Am2b2-1*530*, *Am2b2-2*500* and often the alleles *RPEX1-1*1005*, *GPD2-4*360*, *PmOPSI-1*440*, *GnRH1-1*995*: *Pygocentrus nattereri* b.- Always the alleles *GH5-2*450*, *GH5-2*440*, *GPD2-3*505* and often the alleles *RPEX1-2*790*, *Am2b2-1*570*, *GnRH1-1*1075*: *Serrasalmus maculatus* c.- 2
2	a.- Always the alleles *Am2b2-3*425* and always the allele *RPEX1-1*995* associated with the alleles *RPEX1-3*725*, *GH5-1*1130*, *GPD2-2*690* and *PmOPSI-1*390*: *Serrasalmus elongatus* b.- 3
3	a.- Always the allele *Am2b2-1*585*: 4 b.- 5
4	a.- Always the alleles *RPEX1-3*735* and *GH5-2*490*: *Serrasalmus* sp. b.- Often the allele *PmOPSI-1*395* and always the allele *RPEX1-3*770* associated with the alleles *GH5-1*1120*, *GH5-2*480*, *Am2b2-1*585*: *Serrasalmus compressus*
5	a.- Always the alleles *RPEX1-1*950*, *Am2b2-2*520* and often the alleles *PmOPSI-1*335*, *GnRH1-1*1145* and *GnRH1-1*1130*: *Serrasalmus hollandi* b.- 6
6	a.- Always the alleles *RPEX1-1*1090*, *RPEX1-1*1070*, *RPEX1-1*1060*, *GH5-2*465*, *GH5-2*455*, *PmOPSI-1*420*, *RPEX1-2*945*: *Serrasalmus eigenmanni* or *Serrasalmus spilopleura* b.- Always the alleles *RPEX1-1*1020*, *RPEX1-1*1005*, *RPEX1-1*955*, *Am2b2-2*525* and often the alleles *RPEX1-3*755*, *GH5-2*475*, *GnRH1-1*1065*, *GnRH1-1*1045* and *GnRH1-1*1035*: *Serrasalmus rhombeus*