CrossMark

PRIMARY RESEARCH PAPER

Hybridization and genetic introgression patterns between two South American catfish along their sympatric distribution range

Fernanda Dotti do Prado · Raquel Fernandez-Cebrián ·
Diogo Teruo Hashimoto · José Augusto Senhorini · Fausto Foresti · Paulino Martínez ·
Fábio Porto-Foresti

Abstract The presence of introgressive hybridization in the wild, especially that resulting from human interference, can have negative impacts on biodiversity. Genetic tools provide essential information for species and hybrid identification, facilitating the conservation of natural resources. Here, we tested a set of markers to precisely elucidate introgressive hybridization between *Pseudoplatystoma corruscans* and *Pseudoplatystoma reticulatum*, two South American catfishes. New microsatellites showed high interspecific genetic divergence, and simulated data demonstrated the high power of STRUCTURE and NEWHYBRIDS for hybrid identification and classification, especially when all 11 nuclear markers were used. The investigation of real populations suggested that natural hybridization is rare. Otherwise, different hybridization scenarios were observed in two wild populations: one involving advanced backcrosses and the other involving high admixture. Our data represent the first detailed evaluation of genetic introgression between these species in Parana and Paraguay Basins and suggest that genetic contamination is occurring through F1 hybrids from aquaculture facilities. The results also provide a useful set of markers for monitoring escapees to aid in the conservation of the wild population and sustainable aquaculture. Additionally, species genotypic data are freely available to be used in the future as "parental species reference" in Bayesian methods assignments.

Handling editor: Christian Sturmbauer

F. D. do Prado · F. Porto-Foresti
Universidade Estadual Paulista (UNESP),
Bauru 17033-360, Brazil

R. Fernandez-Cebrián
Geneaqua, 27002 Lugo, Spain

D. T. Hashimoto
Centro de Aquicultura da Unesp (CAUNESP),
Jaboticabal 14884-900, Brazil

J. A. Senhorini
Centro Nacional de Pesquisa e Conservação de Peixes
Continentais, Instituto Chico Mendes de Conservação
(CEPTA/ICMBio), Pirassununga 13630-970, Brazil

F. Foresti
Universidade Estadual Paulista (UNESP),
Botucatu 18618-689, Brazil

P. Martínez
Universidade de Santiago de Compostela (USC),
27002 Lugo, Spain

F. D. do Prado (✉)
Departamento de Biologia, Universidade Estadual
Paulista (UNESP), Av. Eng. Luiz Edmundo Carrijo Coube
14-01, Vargem Limpa, Bauru 17033-360, Brazil
e-mail: prado.fernandad@gmail.com

Introduction

Although it was previously considered infrequent in animals, interspecific hybridization has been found to be related to the evolution and speciation of several taxonomic groups (Mallet, 2005; Willis et al., 2012; Abbott et al., 2013). Sporadic crosses between sympatric species may lead to the formation of hybrid zones with different evolutionary outcomes, and these zones sometimes remain stable for long periods (Arnold et al., 1999; Barton, 2001; Mallet, 2005). Introgressive hybridization has been reported in non-disturbed environments (Willis et al., 2012), suggesting that gene frequencies can be modified as a result of gene flow between species (Rhymer & Simberloff, 1996).

A significant increase in the occurrence of interspecific hybrids in wild environments has been observed in recent decades due to global changes caused by human activities such as the introduction of exotic species, habitat alterations, and contact with artificial hybrids from aquaculture (Allendorf et al., 2010; Brennan et al., 2014; Hasselman et al., 2014; Hashimoto et al., 2014). This kind of hybridization is considered a genetic and ecological threat to biodiversity (Allendorf et al., 2010). Even if they are sterile, F1 hybrids (first-generation or interspecific hybrids) may directly compete with parental species and, if fertile, may modify the genetic composition of parental species through introgression (gene flow between distinct species by backcrosses) (Toledo-Filho et al., 1994, 1998; Allendorf et al., 2010). Fertile hybrids may also cause allele loss from one parental species or break apart adaptive genetic combinations that probably will never be restored (Laikre et al., 1999).

Despite the possible negative impacts of hybridization, studies on this issue as it relates to native fishes in South American hydrographic systems are still scarce. *Pseudoplatystoma reticulatum* Eigenmann & Eigenmann, 1889 (*Lato* sensu *Pseudoplatystoma fasciatum* Linnaeus, 1766) and *Pseudoplatystoma corruscans* Spix & Agassiz, 1829 (Pimelodidae, Siluriformes) are large migratory catfish widely distributed throughout

South American basins (Buitrago-Suárez & Burr, 2007). Hybrids of these species have already been captured in wild populations, especially in the Upper Parana Basin (Bignotto et al., 2009; Prado et al., 2012a; Vaini et al., 2014), indicating some level of genetic contamination. The widespread production of interspecific hybrids between the former two species in farms (Campos, 2010) and inappropriate management likely facilitate escapees and introductions of these hybrids into the natural environment (Porto-Foresti et al., 2010; Hashimoto et al., 2014). To date, ecological and genetic aspects of hybridization in wild populations remain poorly elucidated for these species.

The investigation of hybridization patterns in nature represents one of the main challenges to the conservation of wild populations. Morphological studies provide important data for hybridization surveys, but distinguishing hybrids through external morphological characteristics may be difficult, especially at advanced hybridization stages that are usually highly heterogeneous (Demandt & Bergek, 2009; Allendorf et al., 2010; Hashimoto et al., 2014). Different DNA markers have been developed and successfully applied for species and hybrid identification to circumvent the problems associated with the use of morphological traits (Machado-Schiaffino et al., 2010; Dubut et al., 2010; Bohling et al., 2013; Khosravi et al., 2013). Diagnostic molecular markers such as single nucleotide polymorphisms (SNP) or restriction-fragment-length polymorphisms (RFLP) enabled the accurate identification of parental species and their hybrids (Masaoka et al., 2012; Walter et al., 2014; Ho et al., 2015). Other types of markers, such as microsatellites (SSR—simple sequence repeats), also have been successfully applied with the same purpose (Dubut et al., 2010; Khosravi et al., 2013; Zhang et al., 2013; Kovach et al., 2015). Additionally, microsatellites provide intraspecies population genetic data useful to evaluate hybridization and introgression events, such as changes in heterozygosity levels, deviations from Hardy–Weinberg equilibrium or linkage disequilibrium (Haas et al., 2009; Allendorf et al., 2010). Thus, the combined use of diagnostic and variable markers may allow a more complete study of natural populations, integrating population genetics, and species identification, as verified in several studies (Roques et al., 1999; Dubut et al., 2010; McBride et al., 2014).

The development of clustering methods that use a Bayesian framework has allowed the assignment of individuals to their original populations depending on genetic differentiation between them (Hansen et al., 2001; Sanz et al., 2009). To study hybridization, this approach has the advantage of using a large number of markers to infer individual ancestry and classification into genetic categories (parental species, F1, F2 hybrids, or backcrosses), which is essential to studying populations with introgressive hybridization. STRUCTURE (Pritchard et al., 2000) and NEWHYBRIDS (Anderson & Thompson, 2002) are among the most popular software that use Bayesian tests to study hybridization. STRUCTURE, assuming $K = 2$, i.e., two reference parental populations, infers the fraction of an individual genome (q) inherited from each ancestral genome (Pritchard et al., 2000; Sanz et al., 2009; Vähä & Primmer, 2006), providing highly valuable information to distinguish hybrids from parental populations and even to allocate individuals into different hybrid categories depending on the scenario. On the other hand, NEWHYBRIDS returns the probability (Q) of individuals belonging to specific category, allowing distinction between hybrid classes and elucidation of the genetic introgression pattern (Anderson & Thompson, 2002).

Bayesian methods show some limitations for species and hybrid differentiation, since their precision relies on genetic differentiation between parental species as well as on the number of genetic markers available to detect introgressive hybridization. While few markers (for example, two or three nuclear markers) may simply identify species and hybrids through heterozygous patterns in admixtured individuals (hybrids presenting alleles from different species, for example), a large number of markers is needed to classify individuals beyond F1 categories, as later generations may resemble one of the parental genotypes (Allendorf et al., 2010; Prado et al., 2014). Previous studies demonstrated that the number of multiallelic markers required to correctly identify and classify hybrid categories using Bayesian approaches increases as the divergence between species diminishes (Vähä & Primmer, 2006; Sanz et al., 2009; Bohling et al., 2013; Hasselman et al., 2014), and around eight to fifteen nuclear markers could be enough to detect hybrids at Fst values between species from 0.12 to 0.4.

Currently, there are four diagnostic nuclear and one mitochondrial marker available to differentiate *P. corruscans*, *P. reticulatum,* and their hybrids (Prado et al., 2011; Hashimoto et al., 2013). Recent studies have demonstrated that microsatellite loci developed for *P. reticulatum* show important genetic divergence between the species (allelic range and frequencies) and thus have high potential for hybrid identification (Prado et al., 2014). Carvalho et al. (2013) also demonstrated that a group of eight microsatellites established by Revaldaves et al. (2005) was sufficient to distinguish hybrids and species through Bayesian inference. However, no investigation has been performed concerning the number and power of markers needed for species and hybrid identification, especially the parameters needed for precise classification among hybrid categories.

The aim of this study was to test known and novel molecular markers to (1) verify the genetic divergence between *P. reticulatum* and *P. corruscans* for a set of microsatellites developed by Prado et al. (2014) to provide additional markers for more refined hybridization studies in these species; (2) investigate introgressive hybridization occurring throughout the geographic distribution of *P. reticulatum* and *P. corruscans* species in their main overlapping area, La Plata Basin; and (3) test the power of markers for the efficient classification of parental and hybrid categories to provide a useful tool to be applied in management strategies for wild population restoration.

Materials and methods

Fin clips or blood of 706 adult *P. reticulatum* (Pr) and *P. corruscans* (Pc) were collected in 14 rivers in Paraguay (PG), Parana (PN), and Uruguay (UR) basins (La Plata Basin, South America) between 1998 and 2011 and now are part of the tissue collection of the *Laboratório de Biologia e Genética de Peixes* (UNESP, Botucatu, Brazil) and *Laboratório de Genética de Peixes* (UNESP, Bauru, Brazil) (Fig. 1; Table 1).

Ten individuals from the Mogi-Guaçu River and 14 from the Verde River previously studied by Prado et al. (2012a) were included in the analysis. The Mogi-Guaçu River represents a paradigmatic example of anthropic alterations of aquatic systems. Hybrids have
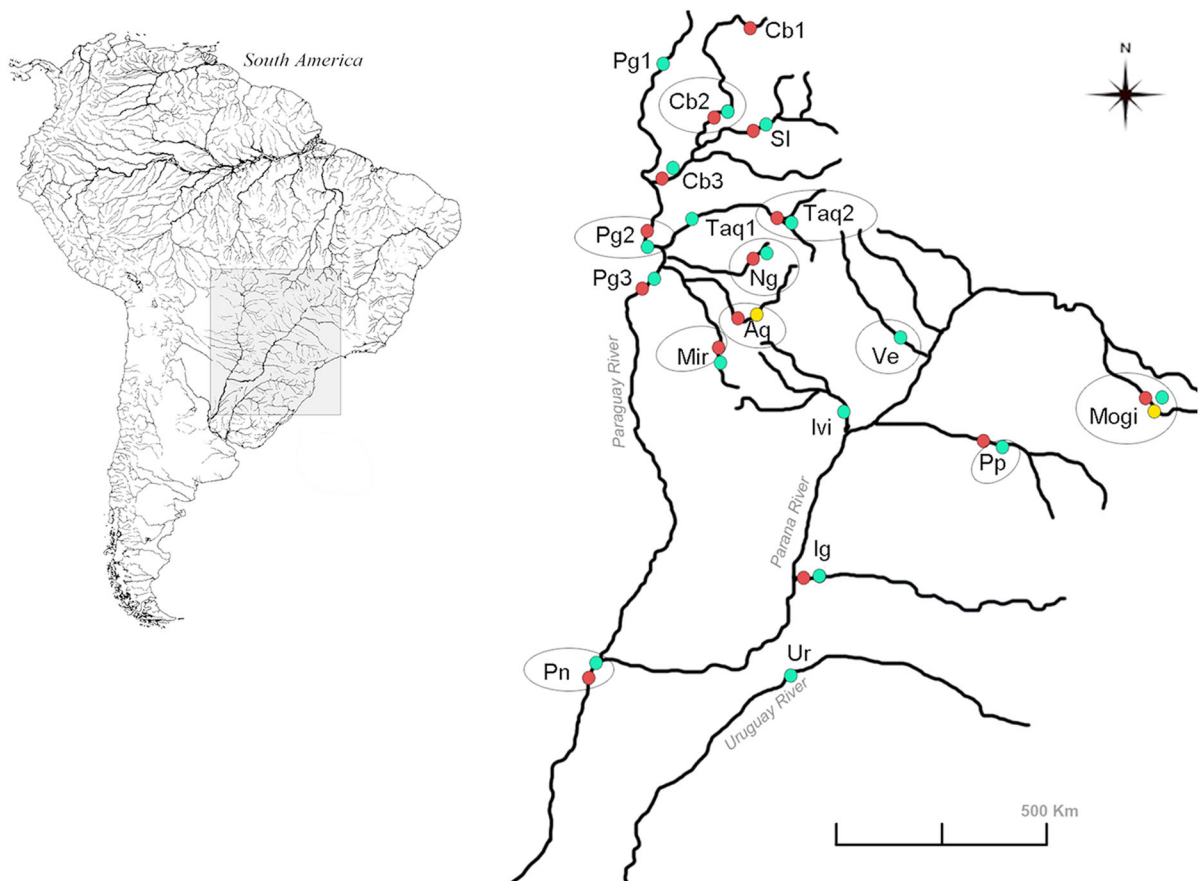
Fig. 1 Sampling sites of *P. reticulatum* (*red*), *P. corruscans* (*green*), and hybrids (*yellow*) in the Paraguay, Parana, and Uruguay River basins, classified according to nuclear SNPs. *Cb* Cuiaba, *Pg* Paraguay, *Sl* São Lourenço, *Ta* Taquari, *Ng* Negro, *Aq* Aquidauana, *Mir* Miranda, *Ve* Verde, *Mogi* Mogi-Guaçu, *Ivi* Ivinhema, *Pp* Paranapanema, *Ig* Iguaçu, *Pn* Parana, and *Ur*

Uruguay Rivers. Numbers after the abbreviations indicate the different sites of sampling in the same River. *Circled sites* correspond to samples selected for subsequent hybridization and genetic diversity analysis using microsatellites (See Table 2 for numbers of individuals)

probably been deliberately introduced or have accidentally escaped from fish farms (Prado et al., 2012a). This location is currently a "closed" environment for large migrators, as it is bordered by the Emas and Marimbondo dams, which preclude fish migration to other rivers (Senhorini JA, personal communication). It thus constitutes an interesting "genetically isolated" environment in which to study introgressive hybridization in this fish group.

Identification of *P. reticulatum* and *P. corruscans* is often accomplished by fish farmers and fishermen using the skin-spot pattern. While *P. corruscans* shows dark circle spots across the whole body (excluding the head), *P. reticulatum* shows dark stripes (Buitrago-Suárez & Burr, 2007). F1 hybrids may have an intermediate pattern of dark circles and

stripes and can thus be easily recognized. One problem is that these hybrids are frequently misclassified as *P. reticulatum*, and it is nearly impossible to identify more advanced hybrids, which may be more similar to one of the parental species, depending on the admixture proportions. In this work, samples were initially identified as *P. corruscans*, *P. reticulatum,* or were not identified (possible hybrids) according to the skin pattern observed at sampling or as reported by fishermen who provided the samples (Table 1).

Analysis delimitation

This work was conducted following a series of interrelationated steps, where markers and individuals were selected according to the objectives starting from

Table 1 Total samples of *Pseudoplatystoma* from wild populations, initially identified through skin pattern, and individuals selected for further analysis

Hb	Rivers	Samples			Total
		Pc	Pr	NI	
PG	Cuiabá (Cb)	52 (**33**)	78 (**27**)		130 (**60**)
	Paraguay (Pg)	77 (**24**)	36 (**22**)		113 (**46**)
	São Lourenço (Sl)	27	11		38
	Taquari (Ta)	57 (**22**)	16 (**12**)		73 (**34**)
	Negro (Ng)	34 (**32**)	23 (**20**)		57 (**52**)
	Aquidauana (Aq)	–	30 (**30**)		30 (**30**)
	Miranda (Mir)	12	40 (**28**)		52 (**28**)
PN	Verde (Ve)	43 (**30**)	–		43 (**30**)
	Mogi-Guaçu (Mogi)	–	–	43 (**43**)	43 (**43**)
	Ivinhema (Ivi)	19	–		19
	Paranapanema (Pp)	30 (**30**)	4		34 (**30**)
	Iguaçu (Ig)	5	14		19
	Parana (Pn)	35 (**31**)	14 (**10**)		49 (**41**)
UR	Uruguay (Ur)	6	–		6
Total		397 (**202**)	263 (**119**)		706 (**394**)

Hb hydrographic basin (*PG* Paraguay, *PN* Parana and *UR* Uruguay River Basins), *Pc P. corruscans*, *Pr P. reticulatum*, *NI* not identified (populations presenting individuals with dubious morphological identification and possible interspecific hybrids). In parentheses and highlighted in bold are presented the samples selected for subsequent genetic diversity and hybridization analyses using microsatellites

a total of 706 individuals and 11 genetic markers. The four main steps are briefly summarized below:

(1) All 706 fishes (Table 1) were genotyped using four nuclear and one mitochondrial species-specific SNPs developed for *P. corruscans* and *P. reticulatum* (Prado et al., 2011; Hashimoto et al., 2013);

(2) A subsample of 394 fishes with good DNA quality (Table 1; Fig. 1) was selected for genotyping using microsatellites to assess genetic diversity and structure of the selected populations. This information was used to identify pure parental populations and evidence of hybrids in wild populations. Selection was additionally conducted by including samples covering the largest geographic area as possible of the species distribution to obtain the best representation of genetic diversity within species and to examine the genetic divergence between them. After populations including hybrids were identified, only parental species were selected to validate microsatellite markers for species differentiation using statistical tests.

(3) Next, parental individuals were used to produce simulated individuals and to test the power of a given number of markers for species and hybrid classification using different statistical methods and then to establish the most appropriate parameters.

(4) Finally, real data (394 individuals) were reanalyzed using Bayesian tests for a more refined classification of parents and hybrids at individual level.

PCR–RFLP and PCR-multiplex

DNA isolation was performed using the Wizard genomic DNA purification kit (Promega, Fitchburg, Wisconsin, USA). All collected individuals were initially classified using the previously reported diagnostic markers RAG2 (*recombination activating gene 2*), EF1 (*nuclear elongation factor 1 α*), 18S (ribosomal 18S) nuclear genes, and the 16S mitochondrial ribosomal gene using the described multiplex-PCR and PCR–RFLP protocols (Prado et al., 2011; Hashimoto et al., 2013) (Table 2). SNP alleles in the

Table 2 Primer sequences and electrophoretic fragments used for species-specific SNPs and primers and multiplex-PCR used for microsatellite genotyping

Gene	Primer sequence (5′-3′)	Diagnostic fragments (bp)		References
		P. reticulatum	*P. corruscans*	
Nuclear and mitochondrial SNPs				
16S F	ACGCCGTTTATCAAAAACAT	400	200	Palumbi (1996a)[a]
16S R	CCGGTCTGAACTCAGATCCGT			Palumbi (1996)[a]
16S PcF	TGACGATCCTATCCGGCTAT			Prado et al. (2011)[a,b]
16S PrR	TACCTGTTATTGGGTTGTA			
RAG2 Silu F	CCTGAGTGCTACCTATTCATGGA	290	330	Prado et al. (2011)[a,b]
RAG2 Silu R	CTTGGGAGGAGAGACCATC			
RAG2 PcR	AACTCCACCACTCTGAATAAATG			
RAG2 PrR	CAGAASCGGGTCTCTGTGGTT			
EF1α 835 F	ATTGGAACTTACCTGTGG	630	520	Moyer et al. (2004)[a]
EF1α 1417 R	CAGCCTTCTGTGCAGACTT			Moyer et al. (2004)[a]
EF1 PcR	CAACAATGCACGCATCTCT			Hashimoto et al. (2013)[a,b]
EF1 PrR	ATAAAGACCCAGACAAGATCG			Hashimoto et al. (2013)[a,b]
NS1 F	GTAGTCATATGCTGTCTC	350	163, 187	White et al. (1990)[a]
18S Silu R	CCATCGAAAATTGATAGGG			Hashimoto et al. (2013)[a,b]

Locus (motif)	Primer sequences (5′-3′)	References
Microsatellite markers		
Multiplex 1		
Prt3 (TG)$_{13}$	F: AGTGGCGTTAGGTCTGTGTG[NED]	Prado et al. (2014)[a]
	R: CTCTGCCATCAATACGCTCA	Present work[b]
Prt11 (GT)$_{12}$	F: TAGCAGCAGGCGATGAGAT[6-FAM]	
	R: CCTAATGTCCAGGGATTTGC	
Prt12 (CA)$_{13}$	F: AGAGCCATGCTGTTGTTGTG[PET]	
	R: GTTTGTGGGACTCGGTGACT	
Prt25 (CA)$_{15}$	F: CAAGGCGCTGTGTATCTTCTT[PET]	
	R: GATCATGCTTGGCTCAGACTT	
Prt34 (TAGA)$_{14}$	F: GGTAGACCGCAAGACAGAACA[6-FAM]	
	R: GGAACTCCTGACCTCCTATGAA	
Multiplex 2		
Prt5 (TG)$_{15}$	F: GTGCTTCCTGCTGTGAGGTA[PET]	Prado et al. (2014)[a]
	R: TGGCAACTGAGGCTTACTGA	Present work[b]
Prt27 (AC)$_{16}$	F: TGTCTCGCATCAAACTACGC[VIC]	
	R: GTCGAAACCGGGACCTTC	
Prt30 (TC)$_8$	F: CACCTGAGACACCACACGTT[6-FAM]	
	R: CGGAGGTAGGAGAGAAAGAGAG	
Prt36 (AGAT)$_{14}$	F: ACCGAGCACAGCACAGAAC[NED]	
	R: AAGGCAATGGTTGGAAGAA	
Prt39 (ATA)$_{11}$	F: GCCGCCATATTGGATCAAG[6-FAM]	
	R: GCGACTCATTATACCACCTCGT	

Pc P. corruscans, *Pr P. reticulatum*, *F* forward, *R* reverse, *bp* base pairs.

[a] Primer development

[b] Technique development: PCR reactions (for PCR–RFLP, PCR-multiplex, or microsatellites)

PCR–RFLP analysis were named according to the band length: RAG2: 290 (Pr), 330 (Pc); EF1: 630 (Pr), 520 (Pc); and 18S: 350 (Pr), 163 (Pc) (Table 2).

Accordingly, individuals were identified as parental species when they showed species-specific markers for the three SNP loci and as hybrids when they presented heterozygous genotypes (i.e., PCR-diagnostic fragments of both species *P. reticulatum* and *P. corruscans*, according to Table 2) for at least one diagnostic nuclear marker. This identification was performed through direct observation of electrophoretic diagnostic bands on gels (Table 2). In this part of the work, considering the methodology and the difficulty of classifying individuals into hybrid categories using few nuclear markers, "non-parental" fishes were simply classified as "hybrids".

The nuclear marker β-globin (GLOB) was initially used for this categorization but was ultimately discarded because 19 individuals who were soundly classified as *P. corruscans* from one population (Taquari) were heterozygous only for this marker, but entirely homozygous (species-specific, e.g., for *P. corruscans*) for the other three nuclear SNPs (RAG2, EF1, 18S). Incongruities for this marker were also observed in specific individuals when different techniques were used to amplify the same region (PCR–RFLP and multiplex-PCR) (Hashimoto et al., 2013). These differences are possibly related to the amplification of paralogous regions, since globin is a member of a multigene family (Tomoko, 1983; Goodman et al., 1988).These results introduced doubts regarding the suitability of this gene for hybrid identification, hence, being excluded for subsequent analyses.

Microsatellites

Ten microsatellite loci (Prado et al., 2014) were genotyped with two multiplex-PCRs using five primer pairs each (Table 2) with the following protocol: reactions were done in a final volume of 10 μl containing 5 μl of PCR Mix (Master mix - 100 μM of each dNTP, MgCl2 1.5 mM, Taq buffer 1X, 0.5 units of Taq polymerase) (Quiagen, Hilden, Germany), 0.4 μM of each primer, and 30 ng of genomic DNA. Cycling consisted in 15 min at 95°C, followed by 30 cycles of 30 s at 94°C, 1 min at 58°C, 1 min at 72°C, with a final extension of 60°C for 30 min on a Thermocycler Veriti™ (Applied Biosystems, Foster City, California, USA). PCR products were confirmed

in 1% agarose gels using 100 bp DNA ladder (Invitrogen, Carlsbad, California, USA) stained with SYBR Safe ™ (Applied Biosystems). Fragment size was analyzed in an ABI PRISM_3730 (Applied Biosystems) sequencer using fluorescence-labeled forward primers. Allelic length categorization was performed with Gene Mapper 3.7 software (Applied Biosystems).

Genetic diversity and structure within and between species

To verify the actual group number (K) in the selected sample from microsatellites, multilocus genotypes were analyzed using the Bayesian clustering procedure implemented in STRUCTURE 2.3.4 software (Pritchard et al., 2000) with no prior information about species identification. It was assumed admixture ancestry, correlated alleles, K value tested from 1 to 10, 20 replicates for each K value, 500 000 MCMC chains and 200 000 burn-in generations. Online software STRUCTURE HARVESTER 0.3 (Earl & Von Holdt, 2012) was used for estimating the best-fitted number of K based on the ΔK method described by Evanno et al. (2005). CLUMPP 1.1.2 (Jakobsson & Rosenberg, 2007) software was used to verify the most probable cluster membership coefficient among the 20 runs. These two programs were used for all posterior STRUCTURE evaluations. Discriminant analyses of principal components (DAPC) (Jombart et al., 2010) were also obtained using the R platform (R Development Core Team, 2014). Data were transformed using PCA (principal component analysis) and retained an appropriate number of PCs and discriminant functions. DAPC was loaded using ADEGENET package (Jombart & Ahmed, 2011) for the R software 3.2.3 (http://www.r-project.org).

To elucidate the species-specific allelic variants at microsatellites and to ascertain genetic differences between them at each locus, we used only parental individuals. The presence of null alleles and genotyping errors at each locus was tested with MICRO-CHECKER 2.2.3 (Van Oosterhout et al., 2004). Genetic divergence was estimated by Fst (Weir & Cockerham, 1984) and Rst (Rousset, 1996; Goodman, 1997) using GENEPOP 3.4 (Raymond & Rousset, 1995). Analysis of molecular variance (AMOVA) to split genetic variation into its components between species, among populations within species, and within populations was performed using ARLEQUIN 3.11

(Excoffier et al., 2005). Significance was tested with 1000 permutations. Allelic frequencies for parental species were obtained by FSTAT 2.9.3.2 (Goudet, 2002), and exclusive, species-specific (private) alleles (Ap) were manually scored.

Genetic diversity for microsatellites was obtained for each population, considering both all parental populations and populations including hybrids. The number of individuals analyzed per locus (N), allele number per locus (A), allelic size range (A range), and observed (Ho) and expected (He) heterozygosities were obtained using FSTAT 2.9.3.2 and GENEPOP 3.4. Deviations from Hardy–Weinberg Equilibrium (HWE) and the sense and magnitude of deviation (F_{IS}) were tested using exact probability tests (Haldane, 1954; Guo & Thompson, 1992) implemented in GENEPOP 3.4. The null hypothesis of linkage equilibrium between all pairs of loci was tested for each sample using the same program. Significance values ($P < 0.05$) for multiple tests were adjusted using sequential Bonferroni correction (Rice, 1989). Allelic richness was first estimated by Fstat 2.9.3.2 by comparing the two parental populations, and then by considering all parental and hybrid samples.

Power of markers determined using Bayesian methods

We considered three sets of genetic markers to evaluate their performance for hybrid categorization: (i) only microsatellites, (ii) all nuclear markers (microsatellites and SNPs), and (iii) only the diagnostic markers. We estimated the power of each group of markers to differentiate between species and hybrid categories and assessed the performance of two Bayesian methods: STRUCTURE and NEWHYBRIDS. The q (admixture degree) range for parental and hybrid categories obtained by STRUCTURE and the Q (posterior probability of belonging to each parental or hybrid category) threshold obtained by NEWHYBRIDS were tested for the best classification of individuals as parental 1 *P. reticulatum* (Pr), parental 2 *P. corruscans* (Pc), F1, F2, backcross to Pr (BPr), or backcross to Pc (BPc). A simulation was performed starting from SNP and microsatellite frequencies of the parental species to test the power and suitability of each statistical approach and the best set of markers using HYBRIDLAB 1.0 software (Nielsen et al., 2006). The genotypes of 200 *P. corruscans* and

200 *P. reticulatum* individuals were simulated and subsequently used to simulate 100 genotypes of each hybrid category (F1, F2, and backcrosses).

STRUCTURE was used to obtain individual membership coefficients q (q1: *P. reticulatum* and q2: *P. corruscans*) considering admixture ancestry, correlated alleles, assumed $K = 2$, with 20 runs for each K value, and the same parameters applied as before. NEWHYBRIDS was used to identify the posterior probability (Q) of an individual belonging to each of the six aforementioned categories. The parameters were set to 100,000 sweeps after a *burn-in* period of 100,000 sweeps.

Testing strategies on real data

We used the information obtained from simulation (genetic markers and statistical inference) to obtain the most accurate classification of individuals by which to assess hybridization and introgression in wild populations. The real data were analyzed in STRUCTURE using the option USEPOPINFO model to specify the species of origin for the original pure samples (i.e., POPFLAG = 1), while the ancestry was estimated for the remaining samples (i.e., POPFLAG = 0 for populations with hybridization). This option uses information from the populations (species) of origin as references to infer the ancestry of other samples with unknown origin (individuals with admixtured genotypes). This approach relies on the assumption that the genetic composition of the populations of origin is known; in our data, this assumption is valid for the species studied. NEWHYBRIDS was also used for category allocation starting from simulation data.

Final classification of individuals was performed using Bayesian analyses considering the most efficient set of markers and the appropriate q range and Q threshold for STRUCTURE and NEWHYBRIDS obtained from simulated data, respectively (see results), and by considering previous works on the field (Sanz et al., 2009; Trigo et al., 2014). The maternal species origin for each individual obtained using the mitochondrial SNP marker was finally added to the nuclear data to identify the maternal origin of F1 or further hybridization events. For advanced hybrids such as backcrosses, mitochondrial information helped to elucidate the direction of hybridization, i.e., from which kind of crosses introgression occurs (which reciprocal hybrids are capable of reproduction

and which parental species are involved in these crosses, for example).

Results

Initial identification of hybrids in the wild

Classification of all individuals ($N = 706$) using nuclear SNPs (RAG2, EF1 and 18S) demonstrated the absence of hybrids and only pure individuals of *P. reticulatum* and *P. corruscans* in the following populations: Paraguay (Pg), Cuiaba (Cb), São Lourenço (Sl), Taquari (Ta), Negro (Ng), Miranda (Mir) (Paraguay Basin); Verde (Ve), Ivinhema (Ivi), Parana-panema (Pp), Iguaçu (Ig), Parana (Pn) (Parana Basin); and Uruguay(Ur) (Uruguay Basin) rivers (data not shown). Both species were detected in most populations excluding Ve, Ivi, and Ur, where only *P. corruscans* was identified. Hybrids were observed in two populations: 15 in the Aquidauana River (Paraguay Basin) and 17 in the Mogi-Guaçu River (Parana Basin). The presence of individuals presenting heterozygote genotypes for at least one diagnostic locus was the primary evidence of the presence of hybrids in these populations.

Genetic diversity and differentiation within and between species using microsatellites

In this step, 394 individuals (119 individuals identified as *P. reticulatum* from six populations, 202 identified as *P. corruscans* from seven populations, and all samples from the Aquidauana and Mogi-Guaçu Rivers) (Fig. 1; Table 1) were further analyzed with microsatellites. Eight loci out of ten evaluated were selected to assess the power of microsatellites for hybridization analysis: Prt3, Prt5, Prt12, Prt25, Prt26, Prt 27, Prt30, and Prt39. Prt11 and Prt34 were discarded due to technical problems.

The first genetic assignment with STRUCTURE supported K = 2 as two highly differentiated groups corresponding to the parental species. For *P. reticulatum*, q values ranged from 0.946 to 0.998 (average $q = 0.995$), except for one individual ($q = 0.860$), while for *P. corruscans*, the values ranged from $q = 0.924$ to 0.998 (average $q = 0.995$). Genetic admixture was observed between species in the Aquidauana and Mogi-Guaçu Rivers, with individuals

showing a wide range of intermediate q values between parental clusters (Fig. 2). Concordant results were verified with DAPC (Fig. 3), which showed distribution of the microsatellite genotypes into two main clusters (*P. corruscans* and other *P. reticulatum*). The Mogi-Guaçu population was distributed into both species clusters and between them (due to the presence of F1 hybrids), while Aquidauana individuals clustered closest to *P. reticulatum*.

This result along with the first genetic identification using the SNPs, strongly suggested that all selected parental individuals (Table 1) were "pure", except for the Aquidauana and Mogi-Guaçu Rivers, where hybrids were present.

Parental populations (red and green clusters—Fig. 2) were used to assess genetic differentiation between species at each locus. Exact tests demonstrated genotypic frequencies in accordance with HWE for *P. reticulatum* and *P. corruscans* for most loci ($P > 0.05$ after sequential Bonferroni correction). Heterozygote deficiency ($P < 0.006$) was detected for three loci (Prt5, Prt 27, and Prt 39) in three populations of *P. reticulatum* and for one locus (Prt 5) in two populations of *P. corruscans*. MICROCHECKER did not suggest specific genotyping errors and the presence of null alleles at the microsatellite loci analyzed can be considered negligible, since it was only observed for isolated parental populations and loci.

A high degree of differentiation was observed for most microsatellites between parental species (Fst from 0.147 to 0.678; average Fst = 0.292; Rst from 0.702 to 0.997; average Rst = 0.899) (Table 3). This differentiation was especially high for two loci (Prt3 and Prt25) and very low for Prt5. Allelic richness was higher in *P. corruscans* (mean Ar = 17.609) than in *P. reticulatum* (mean Ar = 11.659). A total of 71 alleles out of 94 were private of *P. reticulatum*, and 124 out of 150 were private of *P. corruscans*. Prt3, Prt12, Prt30, and Prt36 did not share alleles between species, being considered diagnostic for species identification (Table 3; Supplementary. Table 1). Prt36 showed the highest differences concerning allelic size between species (232-272 vs. 326-433 for *P. reticulatum* and *P. corruscans*, respectively) and Prt25 for allele number (17 vs. 2 for *P. reticulatum* and *P. corruscans*, respectively). In spite of showing low interspecific differentiation (Fst = 0.080), Prt5 was retained for further analyses, as several alleles were private of parental species (Supplementary. Table 1).
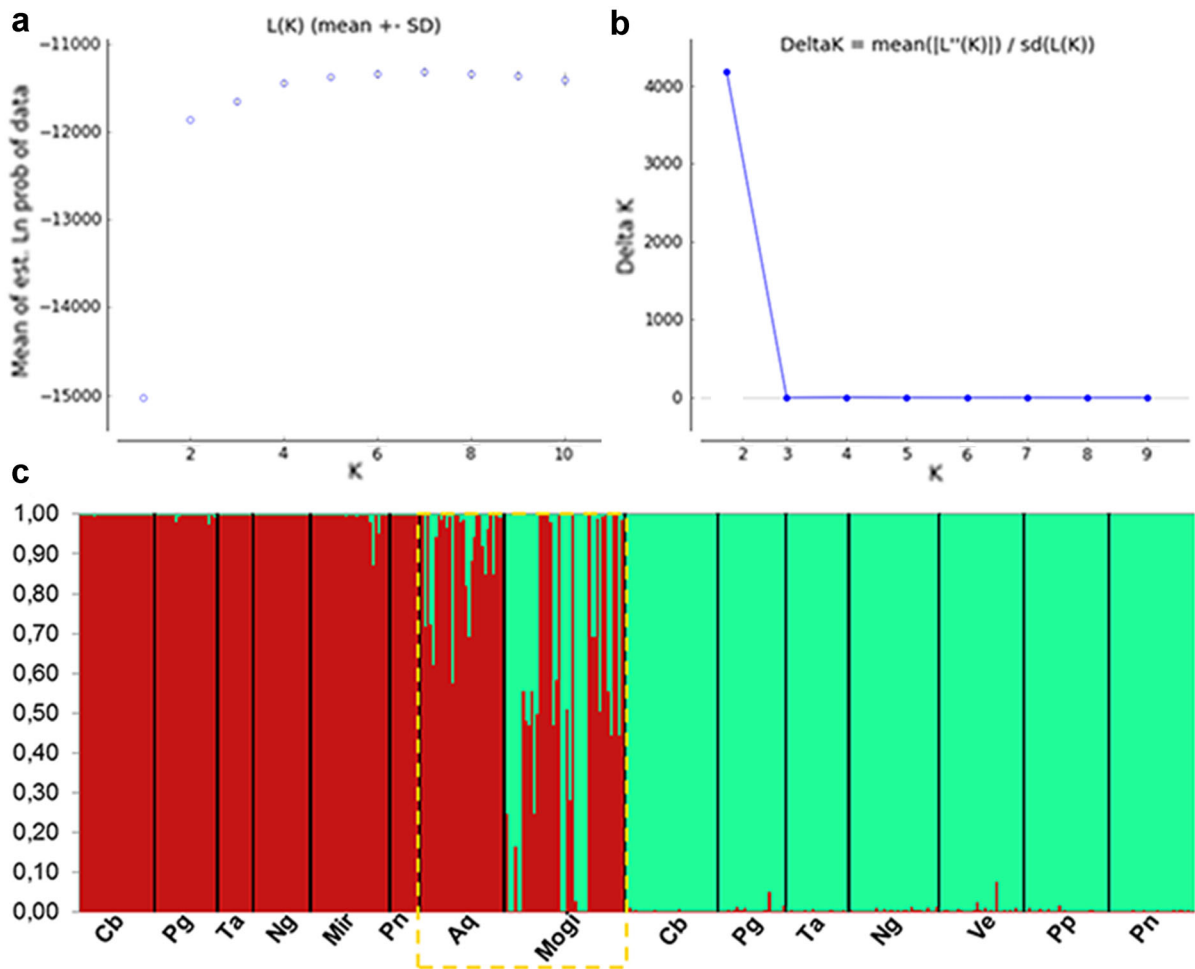
Fig. 2 STRUCTURE analysis of microsatellite genotypes for all samples indicating $K = 2$ for *P. reticulatum* (*red*) and *P. corruscans* (*green*) genotypes assignment. A: LnP(D) and B: delta(k) for $K = 2$ estimated by HARVESTER; C: STRUCTURE bar plot according to replicates averaged by CLUMPP.

Cb Cuiabá, *Pg* Paraguay, *Ta* Taquari, *Ng* Negro, *Mir* Miranda, *Pn* Parana, *Aq* Aquidauana, *Mogi* Mogi-Guaçu, *Ve* Verde, *Pp* Paranapanema. *Yellow dashed lines* indicate the sites with the presence of hybrids and genetic admixture between the two species

The distribution of genetic diversity, as examined using AMOVA (Table 4), showed a significant genetic divergence between species (33.7%; $P = 0$) that was much higher than that between populations within species (1.94%; $P < 0.05$), suggesting a high potential of microsatellites for species identification.

Genetic diversity was assessed for all selected populations, including the Aquidauana and Mogi-Guaçu Rivers (Table 1). Multilocus deviations of HWE ($P < 0.006$) for parental populations were verified in the Negro and Miranda populations for *P. reticulatum* ($F_{IS} = 0.111$ and 0.146, respectively) and for the Verde River in *P. corruscans* ($F_{IS} = -0.024$)

(Table 5). Genetic diversity was higher in the Mogi-Guaçu (He = 0.864; Ar = 8.932) than in the Aquidauana River (He = 0.764; Ar = 5.702) (Table 5). Mogi-Guaçu showed higher genetic diversity than the parental species (Table 5), as a consequence of both parental genomes being present at similar proportions (Figs. 2, 3). In contrast, although hybrids were detected in the Aquidauana River, *P. reticulatum* was the dominant genome (Figs. 2, 3) and genetic diversity was similar to that of neighboring populations of this species within the same river basin (Mann–Whitney test Aquidauana vs. Paraguay, Taquari and Negro: $P > 0.05$ both for Ar and He)

Fig. 3 DAPC-based clustering of microsatellite genotypes for all samples organized per species and both populations with hybrids (4 a priori groups). *Pr P. reticulatum*, *Pc P. corruscans*, *Aq* Aquidauana River, *Mogi* Mogi-Guaçu River
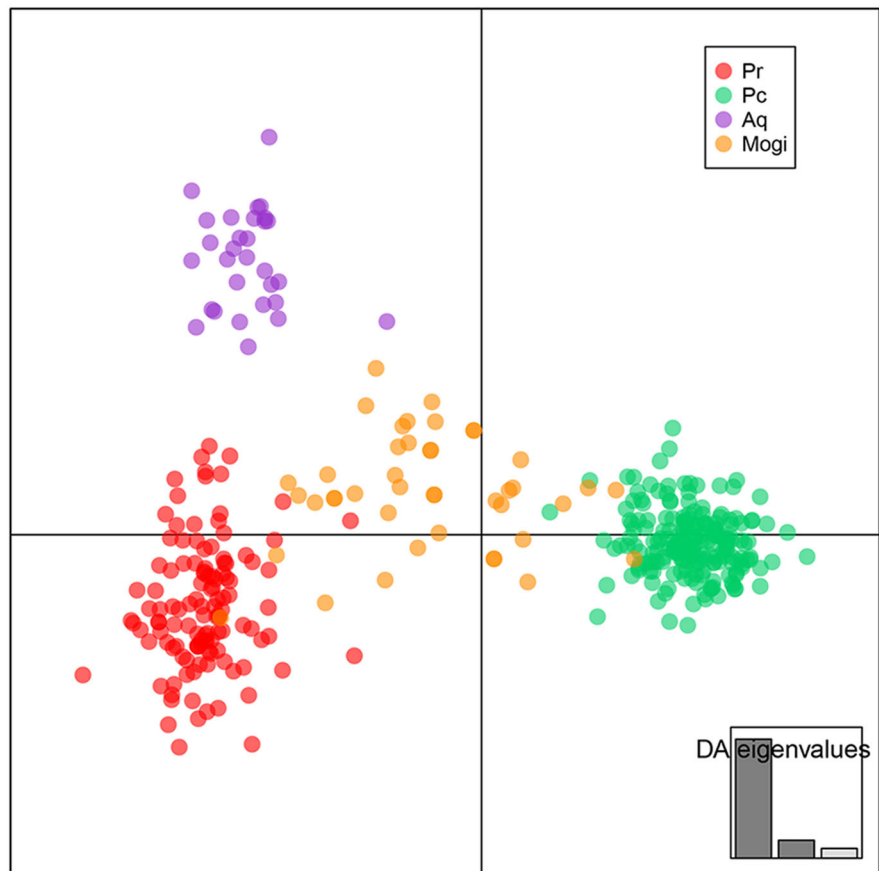
(Table 5). F_{IS} values were consistently negative in Aquidauana, denoting a significant heterozygote excess ($F_{IS} = -0.126$; $P < 0.006$), while in the Mogi-Guaçu River, a significant heterozygote deficit was detected ($F_{IS} = 0.101$; $P < 0.006$) (Table 5). Finally, no consistent linkage disequilibrium between loci was detected in parental populations (28 pairwise loci comparisons per site; Bonferroni: $P < 0.0018$), except for two populations in *P. reticulatum* and *P. corruscans* (the Cuiaba and Verde Rivers, respectively) (Table 5). Conversely, 11% and 71% linkage disequilibriums were detected in Aquidauana and Mogi-Guaçu populations, respectively, denoting different degrees of admixture of parental species.

Power of markers: simulated data

Using simulated data, STRUCTURE assignment of parental species was high and showed slight differences among the three sets of markers considered (Table 3). The mean admixture coefficient q was always higher for *P. reticulatum* ($q \geq 0.921$) than for *P. corruscans* ($q \geq 0.892$) (Table 6; Fig. 4). The best performance was achieved with all markers (q mean = 0.973 for *P. reticulatum* and 0.969 for *P. corruscans*), but the narrowest q range was obtained when only using diagnostic markers (0.959-0.963 for *P. reticulatum* and q from 0.938 to 0.963 for *P. corruscans*). Using a threshold value <0.9 to distinguish parental individuals from hybrids, STRUCTURE achieved 100% classification success for *P. reticulatum* and 99% success for *P. corruscans* when all combinations of markers (microsatellites, all markers or only diagnostic loci) were considered, excluding a single individual that was misclassified when only microsatellites were used. Using a more stringent threshold value of $q < 0.95$, STRUCTURE correctly identified most hybrid individuals, especially when only diagnostic markers were used (100% for *P. reticulatum* and 99% for *P. corruscans*). Although STRUCTURE showed high potential to identify F1 hybrids (q mean ~ 0.5 and narrow q range, especially

Table 3 Allelic variability and genetic differentiation (F_{ST} and R_{ST}) per locus between parental species

Locus	Prt 3	Prt 5	Prt 12	Prt 25	Prt 27	Prt 30	Prt 36	Prt 39	All
P. reticulatum									
N	117	115	117	116	119	117	110	112	115.4
A (Ap)	7 (7)	16 (7)	7 (7)	17 (15)	9 (6)	14 (14)	10 (10)	14 (5)	94 (74)
Ar	6.937	15.910	6.940	16.891	8.914	13.701	10.000	13.981	11.659
A range (bp)	184–197	219–280	276–292	167–207	211–233	99–133	232–272	268–306	99–306
P. corruscans									
N	197	196	187	198	196	197	178	197	193.3
A (Ap)	3 (3)	24 (13)	13 (13)	2 (0)	31 (28)	13 (13)	44 (44)	20 (10)	150 (124)
Ar	2.806	22.448	12.941	1.913	29.571	12.510	39.782	18.899	17.609
A range (bp)	159–161	221–297	296–320	169–171	225–293	69–95	326–433	255–288	69–433
Interspecific									
Fst	0.678	0.083	0.223	0.664	0.229	0.202	0.147	0.154	0.292
Rst	0.997	0.000	0.958	0.833	0.838	0.911	0.935	0.702	0.899

N number of individuals, *A* number of alleles, *A range* allelic range, *bp* base pairs, *Ap* private alleles, *Ar* allelic richness. Underlined: fully diagnostic loci between species without shared alleles

Table 4 Multilocus analysis of molecular variance (AMOVA) between *P. reticulatum* and *P. corruscans*

Variance	Variance components	Percentage of variance (%)
Between species	1.315	33.69**
Among populations within species	0.076	1.94*
Within populations	2.514	64.39**

* $P < 0.05$; ** $P = 0.000$

with diagnostic markers), these hybrids could not be distinguished from F2 individuals unless additional information was included (e.g., F1 should be heterozygous for all diagnostic markers, while F2 is only heterozygous for 50% as on average).

When applying NEWHYBRIDS (Table 6, Fig. 5), we used two threshold-assignment probabilities ($Q \geq 0.9$ or ≥ 0.6), considering the balance between correct classification of the parental and hybrid classes. NEWHYBRIDS performed very similarly to STRUCTURE in distinguishing parental species; most individuals were correctly classified using a threshold of 0.9 with all markers considered, and 100% success was achieved when $Q \geq 0.6$. Diagnostic markers showed even better performance than all markers in classifying parental classes and F1 hybrids at the 0.9 threshold. However, when the remaining hybrid classes were considered, the best results were achieved using all markers. F2 was the class with the poorest assignment success, and backcrosses showed

intermediate figures. Globally, all markers performed best at the 0.6 threshold, with greater than 90% assignment success for most classes, excluding F2 (80%).

Power of markers: real data

Considering the previously observed power of the different sets of markers, the assignment of individuals in wild populations was conducted using all the 11 nuclear genetic markers. Threshold values of $q < 0.9$ and $Q < 0.6$ were used to identify and classify hybrids in STRUCTURE and NEWHYBRIDS, respectively. Additionally, the mtDNA marker was used to identify the maternal origin of hybrids (Table 7).

The use of specific genotypes of known origin as reference increases the correct assignment of individuals with unknown origin (Pritchard et al., 2000), and consequently increases the probability of hybrids identification in admixed populations (Sanz et al.,

Table 5 Multilocus genetic diversity per site using eight microsatellites

Site	N	A	Ar	EHW (Exact tests)			% Significant LD*
				Ho	He	Fis	
P. reticulatum							
Cb	27	7.5	5.626	0.681	0.715	0.048	14
Pg	22	8.4	6.182	0.704	0.730	0.035	0
Ta	12	6.6	5.901	0.604	0.734	0.177	0
Ng	20	7.6	5.906	0.635	0.715	**<u>0.111</u>**	0
Mir	28	8.5	6.055	0.638	0.746	**<u>0.146</u>**	0
Pn	10	5.1	5.082	0.697	0.735	0.051	0
Hybrid occurrence							
Aq	30	7.3	5.702	0.860	0.764	**<u>−0.126</u>**	11
Mogi	43	15.4	8.932	0.777	0.864	**<u>0.101</u>**	71
P. corruscans							
Cb	33	10.9	6.679	0.631	0.650	0.029	0
Pg	24	10.6	7.176	0.667	0.665	−0.003	0
Ta	22	10.8	7.286	0.667	0.686	**0.027**	0
Ng	32	13.1	7.538	0.667	0.679	0.018	0
Ve	30	9.5	6.034	0.634	0.619	**<u>−0.024</u>**	32
Pp	30	10.0	6.355	0.631	0.633	0.003	0
Pn	31	13.0	7.459	0.638	0.674	**0.053**	0

Aq the Aquidauana River, *Mogi* the Mogi-Guaçu River, *N* number of individuals, *A* number of alleles, *Ar* allelic richness, *He* expected heterozygosity, *Fis* Fixation index; Bold: $P < 0.05$; Bold and underlined: $P < 0.006$ after Bonferroni correction (8 loci), *LD* linkage disequilibrium; *Percent of locus x locus pairwise comparisons presenting significant evidence of linkage disequilibrium ($P < 0.006$ after Bonferroni correction—total of 28 comparisons locus by locus). Sampling sites—*Cb* Cuiabá, *Pg* Paraguay, *Ta* Taquari, *Ng* Negro, *Mir* Miranda, *Aq* Aquidauana, *Mogi* Mogi-Guaçu, *Ve* Verde, *Pp* Paranapanema, *Pn* Parana (correspondent to Fig. 1)

2009). As representatives of "pure parental species" of reference for future hybridization studies involving *P. corruscans* and *P. reticulatum* species, parental populations genotyped at all eight microsatellite markers were included in this work (Suppl. Table 2).

Bayesian results for the Aquidauana River showed an overall genomic composition very close to that of *P. reticulatum* (mean $q1 = 0.861$) and identified a total of 13 *P. reticulatum* and 17 hybrids (one F1 hybrid, 11 backcrosses BPr and five other hybrids with uncertain classification, probably advanced categories). Assignment was mostly concordant between programs, excluding an individual (number 8) classified as *P. reticulatum* with STRUCTURE ($q1 = 0.900$) but as a backcross (BPr; $Q = 0.674$) with NEWHYBRIDS and individuals 15, 19, 20, and 29, classified as a hybrid with STRUCTURE but showing low certainty of hybrid categorization with NEWHYBRIDS. These fishes probably correspond to hybrids beyond the first backcross and were identified

only as "hybrid" (H) individuals, without a specific hybrid categorization (Table 7). The most frequent hybrid class (BPr), backcrosses, was slightly above ($q = 0.754$) the expected value ($q = 0.750$). In all hybrids, the mtDNA was of *P. reticulatum* origin.

In the Mogi-Guaçu River, 14 *P. reticulatum*, 11 *P. corruscans,* and a total of 18 hybrids (10 F1, 4 BPr, and 4 BPc backcrosses) were identified using STRUCTURE and NEWHYBRIDS (Table 7). Assignment in this population was fully concordant between programs even though this setting was more complex, with both parental genomes contributing in similar proportions to the sample analyzed (mean $q1 = 0.521$ and $q2 = 0.479$). Additionally, the probability of assignment of individuals to the correct class was very high in NEWHYBRIDS (mean $Q = 0.951$). In this population, the mtDNA genomes of hybrids came from both species, although *P. reticulatum* markers were more highly represented (13 out of 18 hybrids). The mean genomic composition (q) of hybrid classes

Table 6 Results of the Bayesian analysis with STRUCTURE (K = 2) and NEWHYBRIDS on the simulated genotypes considering different molecular marker sets

Markers (N loci)	Category	STRUCTURE				NEWHYBRIDS		
		Average proportion of assignment (range)		Correct assignment*		Average probability (range)	Correct classification**	
		q1	q2	0.90	0.95	Q	0.90	0.60
Micros. (8)	Pr	**0.959** (0.921–0.965)	**0.041** (0.035–0.079)	100%	93%	**0.991** (0.951–0.995)	100%	100%
	Pc	**0.046** (0.035–0.108)	**0.954** (0.892–0.965)	99%	78%	**0.978** (0.751–0.996)	97%	100%
	F1	**0.496** (0.401–0.596)	**0.504** (0.404–0.599)	–	–	**0.949** (0.723–0.983)	88%	100%
	F2	**0.516** (0.265–0.831)	**0.484** (0.169–0.735)	–	–	**0.603** (0.032–0.999)	30%	57%
	BPr	**0.725** (0.499–0.964)	**0.275** (0.036–0.501)	–	–	**0.721** (0.005–0.957)	30%	75%
	BPc	**0.258** (0.073–0.495)	**0.742** (0.505–0.927)	–	–	**0.742** (0.012–0.946)	24%	80%
All (11)	Pr	**0.973** (0.934–0.976)	**0.027** (0.024–0.067)	<u>100%</u>	95%	**0.998** (0.963–0.999)	100%	<u>100%</u>
	Pc	**0.031** (0.023–0.099)	**0.969** (0.901–0.977)	<u>100%</u>	92%	**0.996** (0.963–0.999)	100%	<u>100%</u>
	F1	**0.498** (0.433–0.589)	**0.502** (0.411–0.567)	–	–	**0.988** (0.881–0.997)	98%	<u>100%</u>
	F2	**0.496** (0.172–0.763)	**0.504** (0.238–0.828)	–	–	**0.785** (0.007–1.000)	56%	<u>80%</u>
	BPr	**0.743** (0.575–0.931)	**0.257** (0.070–0.425)	–	–	**0.921** (0.050–0.989)	85%	<u>96%</u>
	BPc	**0.255** (0.057–0.471)	**0.745** (0.529–0.943)	–	–	**0.870** (0.051–0.991)	79%	<u>92%</u>
Diagnostic (7)	Pr	**0.963** (0.959–0.963)	**0.037** (0.037–0.041)	100%	100%	**0.992** (0.989–0.992)	100%	100%
	Pc	**0.038** (0.037–0.062)	**0.962** (0.938–0.963)	100%	99%	**0.991** (0.985–0.992)	100%	100%
	F1	**0.498** (0.483–0.553)	**0.502** (0.447–0.517)	–	–	**0.970** (0.912–0.975)	100%	100%
	F2	**0.504** (0.092–0.773)	**0.496** (0.227–0.908)	–	–	**0.686** (0.006–0.999)	38%	70%
	BPr	**0.740** (0.565–0.959)	**0.260** (0.041–0.435)	–	–	**0.806** (0.014–0.949)	35%	86%
	BPc	**0.253** (0.037–0.432)	**0.747** (0.568–0.963)	–	–	**0.801** (0.010–0.944)	33%	87%

The STRUCTURE q values consist in the average proportion of assignment of each simulated parental and hybrid category considering two clusters. Micros: eight microsatellite markers (Prt3, 5, 12, 25, 27, 30, 36, and 39); All: eight microsatellite markers plus RAG2, EF1 and 18S nuclear SNPs; Diagnostic: four diagnostic microsatellite markers (Prt3, 12, 30, and 36) plus RAG2, EF1, and 18S nuclear SNPs; *Pr P. reticulatum, Pc P. corruscans, BPr* Backcross F1 × Pr; *BPc* backcross F1 × Pc, *Percent of individuals correctly assignment to their genetic category using STRUCTURE at 0.9 and 0.95 threshold values; **Percent of individuals correctly classified as their own genetic category using NEWHYBRIDS at 0.9 and 0.6 threshold values. Underlined: Parameters that allowed the greatest correct assignment and classification of parental and hybrid categories, considering both STRUCTURE and NEWHYBRIDS programs

showed a value slightly higher than expected for F1 (0.514) and BPc (0.784) individuals, but a markedly lower value for BPr (0.640) individuals. In most cases, the type of backcross was concordant with the mtDNA genome (i.e., BPc was associated with Pc mtDNA, while BPr was associated with Pr mtDNA).

Discussion

Genetic markers

The initial identification of hybrids in wild populations using nuclear gene markers (RAG2, EF1, and 18S) suggested the presence of 15 hybrids in the

Aquidauana River and 17 in the Mogi-Guaçu River, which nearly corroborated the results of the Bayesian tests, even when the number of markers increased (17 and 18 hybrids in each river, respectively). However, while the simple identification of hybrids occurrence in wild stocks represents important data for species conservation (Prado et al., 2012a; Vaini et al., 2014), a precise evaluation of introgression patterns in the wild, as well as ecological and reproductive inferences, requires the distinction among hybrid categories, which is only possible using appropriate statistical parameters and a higher number of markers. The use of a small number of nuclear markers to analyze hybridization in wild populations may be effective when hybridization is recent and the introgression rate
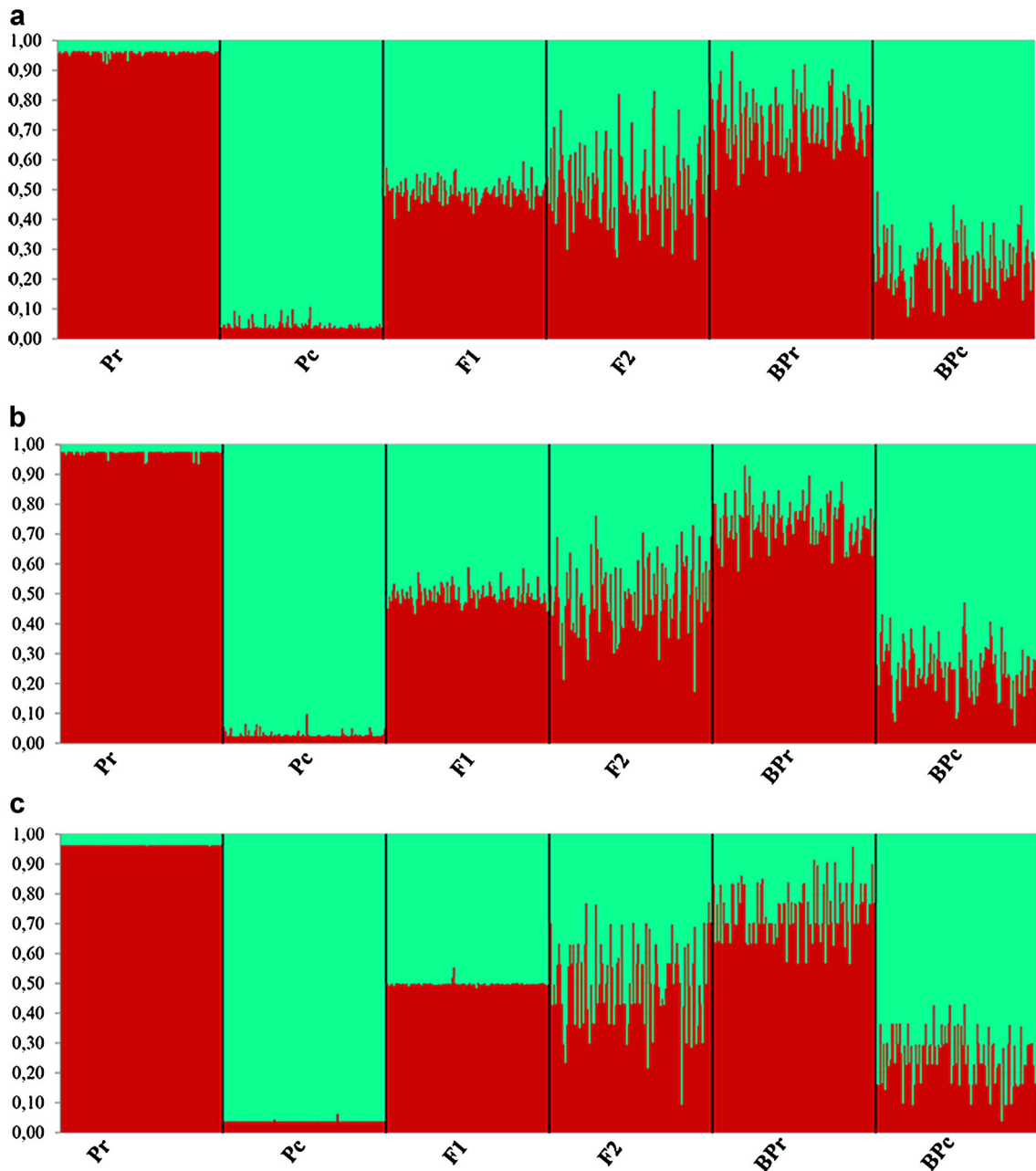
Fig. 4 Bar plot of Bayesian clustering results obtained with STRUCTURE ($K = 2$) for simulated genotypes using only microsatellite data (**a**), all markers (**b**) and diagnostic markers (**c**). *Pr P. reticulatum*, *Pc P. corruscans*, *F1* first hybrid generation, *F2* second hybrid generation, *BPr* Backcrosses with *P. reticulatum*, *BPc* Backcrosses with *P. corruscans*

is low (Vähä & Primmer, 2006). As fertility has been demonstrated for *P. corruscans* and *P. reticulatum* hybrids (Prado et al., 2012b), hybrid classification must use a higher number of markers and statistical methodologies, as previously suggested for more complex scenarios (Toledo-Filho et al., 1994; Boecklen & Howard, 1997; Vähä & Primmer, 2006; Sanz et al., 2009). The application of nuclear SNPs in this study confirmed their diagnostic utility (Prado et al., 2011; Hashimoto et al., 2013) using a large and
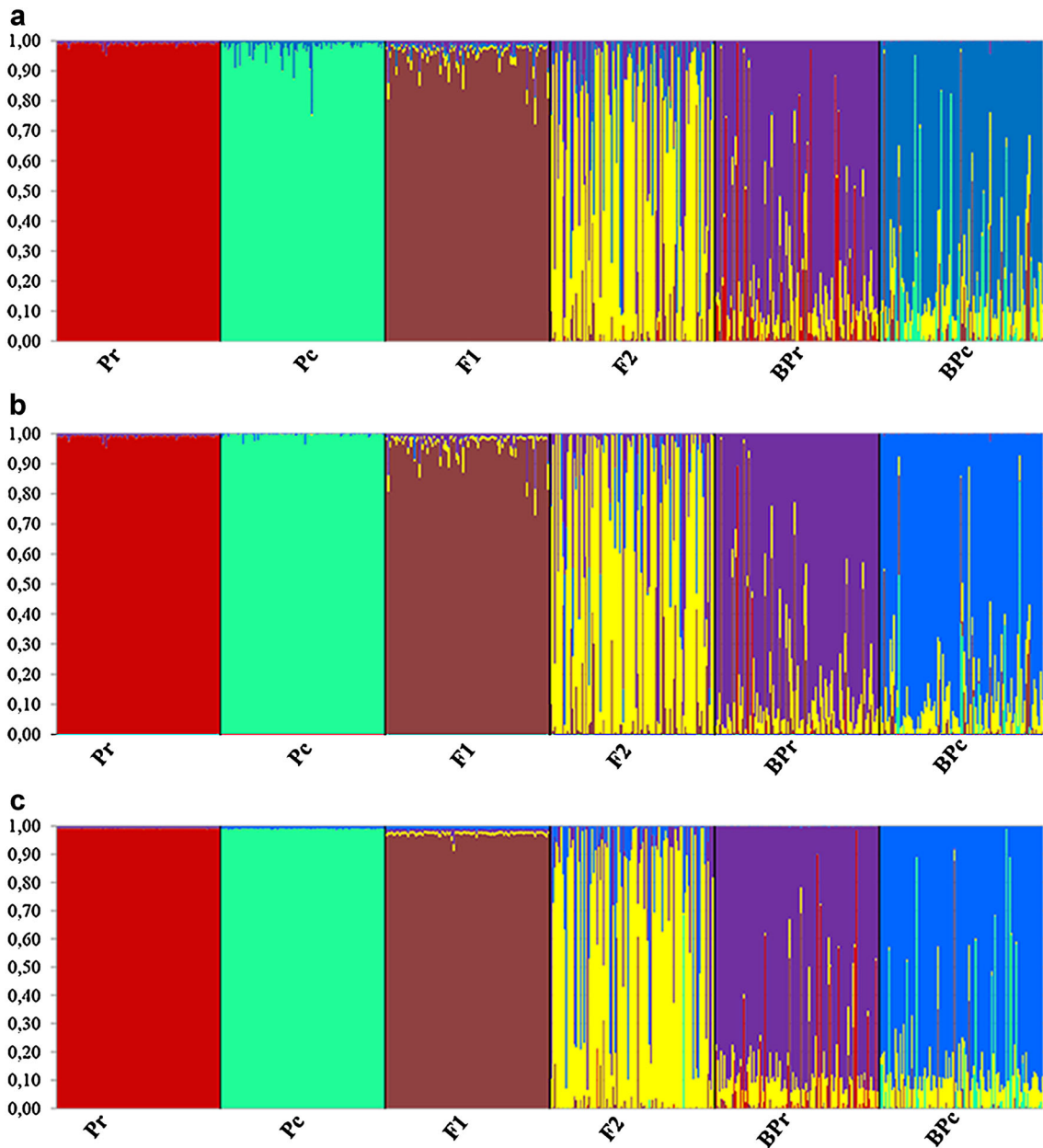
Fig. 5 Bar plot of posterior classification probabilities obtained with NEWHYBRIDS for each simulated genotype using only microsatellite data (**a**) and all markers (**b**) and diagnostic markers (**c**). *Pr P. reticulatum*, *Pc P. corruscans*, *F1* first hybrid generation, *F2* second hybrid generation, *BPr* Backcrosses with *P. reticulatum*, *BPc* Backcrosses with *P. corruscans*

geographically representative sample of both species, and these markers were included in subsequent analyses.

A first step in the use of multiallelic genetic markers such as microsatellites for species identification is to determinate the degree of allelic divergence between

species (Roques et al., 1999). In this study, the eight microsatellites tested demonstrated much higher genetic divergence between species (33.69%) than within species (1.94%) and accordingly were used to allocate all individuals analyzed into two groups corresponding to *P. corruscans* and *P. reticulatum* in

Table 7 Classification of individuals from the Aquidauana and Mogi-Guaçu Rivers using 11 nuclear markers and Bayesian methods implemented in STRUCTURE ($K = 2$, using parental species as reference), NEWHYBRIDS, and an additional mitochondrial SNP

Pop	STRUCTURE*		NEWHYBRIDS*	Mt SNP	Genetic category
	q1 (Pr)	*q2* (Pc)	Q		
Aq					
1	0.983	0.017	0.997 (Pr)	Pr	Pr
2	**0.798**	**0.202**	**0.983 (BPr)**	**Pr**	**BPr**
3	0.986	0.014	0.999 (Pr)	Pr	Pr
4	**0.702**	**0.298**	**0.975 (BPr)**	**Pr**	**BPr**
5	**0.602**	**0.398**	**0.971 (BPr)**	**Pr**	**BPr**
6	**0.759**	**0.241**	**0.937 (BPr)**	**Pr**	**BPr**
7	0.985	0.015	0.999 (Pr)	Pr	Pr
8	**0.900**	**0.100**	**0.674 (BPr)**	**Pr**	**H**
9	0.984	0.016	0.999 (Pr)	Pr	Pr
10	0.901	0.099	0.961 (Pr)	Pr	Pr
11	0.952	0.048	0.998 (Pr)	Pr	Pr
12	**0.542**	**0.458**	**0.971 (F1)**	**Pr**	**♀Pr x ♂Pc**
13	0.986	0.014	1.000 (Pr)	Pr	Pr
14	0.986	0.014	0.997 (Pr)	Pr	Pr
15	**0.866**	**0.134**	**0.787 (Pr)**	**Pr**	**H**
16	0.947	0.053	0.944 (Pr)	Pr	Pr
17	**0.791**	**0.209**	**0.975 (BPr)**	**Pr**	**BPr**
18	**0.719**	**0.281**	**0.974 (BPr)**	**Pr**	**BPr**
19	**0.898**	**0.102**	**0.976 (Pr)**	**Pr**	**H**
20	**0.804**	**0.196**	**0.517 (Pr)**	**Pr**	**H**
21	0.985	0.015	0.999 (Pr)	Pr	Pr
22	0.986	0.014	0.999 (Pr)	Pr	Pr
23	**0.846**	**0.154**	**0.981 (BPr)**	**Pr**	**BPr**
24	**0.805**	**0.195**	**0.913 (BPr)**	**Pr**	**BPr**
25	**0.748**	**0.252**	**0.989 (BPr)**	**Pr**	**BPr**
26	0.986	0.014	0.999 (Pr)	Pr	Pr
27	**0.713**	**0.287**	**0.984 (BPr)**	**Pr**	**BPr**
28	0.984	0.016	0.999 (Pr)	Pr	Pr
29	**0.863**	**0.137**	**0.777 (Pr)**	**Pr**	**H**
30	**0.813**	**0.187**	**0.676 (BPr)**	**Pr**	**BPr**
Total	Pr:14, H:16		Pr:17, H:13		17 Hybrids (H: 5, F1:1, BPr:11)
Mogi					
1	**0.180**	**0.820**	**0.800 (BPc)**	**Pr**	**BPc**
2	0.014	0.986	1.000 (Pc)	Pc	Pc
3	0.014	0.986	1.000 (Pc)	Pc	Pc
4	**0.199**	**0.801**	**0.969 (BPc)**	**Pc**	**BPc**
5	0.014	0.986	0.999 (Pc)	Pc	Pc
6	0.014	0.986	0.997 (Pc)	Pc	Pc
7	**0.537**	**0.463**	**0.864 (F1)**	**Pr**	**♀Pr x ♂Pc**
8	**0.489**	**0.511**	**0.998 (F1)**	**Pc**	**♀Pc x ♂Pr**
9	**0.478**	**0.522**	**0.998 (F1)**	**Pc**	**♀Pc x ♂Pr**

Table 7 continued

Pop	STRUCTURE*		NEWHYBRIDS*	Mt SNP	Genetic category
	q1 (Pr)	*q2* (Pc)	*Q*		
10	**0.530**	**0.470**	**0.875 (F1)**	**Pr**	**♀Pr x ♂Pc**
11	**0.180**	**0.820**	**0.800 (BPc)**	**Pc**	**BPc**
12	**0.548**	**0.452**	**0.899 (F1)**	**Pr**	**♀Pr x ♂Pc**
13	0.986	0.014	1.000 (Pr)	Pr	Pr
14	0.985	0.015	0.999 (Pr)	Pr	Pr
15	0.981	0.019	0.957 (Pr)	Pr	Pr
16	0.978	0.022	0.997 (Pr)	Pr	Pr
17	0.904	0.096	0.927 (Pr)	Pr	Pr
18	**0.575**	**0.425**	**0.723 (BPr)**	**Pr**	**BPr**
19	**0.557**	**0.443**	**0.918 (F1)**	**Pr**	**♀Pr x ♂Pc**
20	0.983	0.017	0.997 (Pr)	Pr	Pr
21	0.014	0.986	1.000 (Pc)	Pc	Pc
22	0.014	0.986	1.000 (Pc)	Pc	Pc
23	**0.458**	**0.542**	**0.973 (F1)**	**Pr**	**♀Pr x ♂Pc**
24	**0.305**	**0.695**	**0.720 (BPc)**	**Pc**	**BPc**
25	0.956	0.044	0.971 (Pr)	Pr	Pr
26	0.083	0.917	0.963 (Pc)	Pc	Pc
27	0.014	0.986	1.000 (Pc)	Pc	Pc
28	0.014	0.986	0.999 (Pc)	Pc	Pc
29	0.023	0.977	0.999 (Pc)	Pc	Pc
30	0.014	0.986	1.000 (Pc)	Pc	Pc
31	0.974	0.026	0.997 (Pr)	Pr	Pr
32	**0.717**	**0.283**	**0.994 (BPr)**	**Pr**	**BPr**
33	**0.717**	**0.283**	**0.994 (BPr)**	**Pr**	**BPr**
34	0.951	0.049	0.998 (Pr)	Pr	Pr
35	**0.552**	**0.448**	**0.923 (BPr)**	**Pr**	**BPr**
36	0.960	0.040	0.994 (Pr)	Pr	Pr
37	0.983	0.017	0.999 (Pr)	Pr	Pr
38	**0.527**	**0.473**	**0.982 (F1)**	**Pr**	**♀Pr x ♂Pc**
39	**0.509**	**0.491**	**0.850 (F1)**	**Pr**	**♀Pr x ♂Pc**
40	0.986	0.014	0.999 (Pr)	Pr	Pr
41	0.986	0.014	0.999 (Pr)	Pr	Pr
42	**0.509**	**0.491**	**0.850 (F1)**	**Pr**	**♀Pr x ♂Pc**
43	0.970	0.030	0.970 (Pr)	Pr	Pr
Total	Pr: 14, Pc: 11, H: 18		Pr: 14, Pc: 11, H: 18		18 Hybrids (F1: 10, BPr: 4, BPc: 4)

Pr P. reticulatum, Pc P. corruscans; *H* individual identified as hybrid without definition on hybrid category, *BPr* Backcross F1 × Pr; *BPc* backcross F1 × Pc; ♀Pr x ♂Pc and ♀Pc x ♂Pr F1 hybrids with the sex of parentals established using the mtDNA SNP; *Mt* mitochondrial; *Threshold value of 0.90 for STRUCTURE and 0.6 for NEWHYBRIDS; highlighted in bold individuals identified as hybrids

the STRUCTURE Bayesian test. We cannot totally exclude the possibility that the only individual assigned to Pr with low confidence could be a hybrid with a high genomic component of *P. reticulatum*. However, this individual's genotype contained all diagnostic markers consistent with *P. reticulatum*;

doubts about its origin may be due to background noise arising from the use of markers that are not entirely diagnostic between species. The DAPC analyses indicate high genetic variations between parental clusters, corroborating the STRUCTURE results.

Other statistics (Fst and Rst) also revealed high interspecific divergence. Fst is based on the relative component of the divergence and therefore underestimates genetic divergence when the mean He (expected heterozygosity within each parental species population) values are high, a typical feature of hypervariable markers such as microsatellites. In our study, Fst showed moderate-to-high genetic divergence for most microsatellites (between 0.147 and 0.678 excluding Prt5 that presented low interspecific divergence), even for diagnostic microsatellites that did not share any allele (for example Prt36: Fst = 0.147), confirming Fst drawbacks for this assessment. Conversely, Rst (between 0.702 and 0.997 excluding Prt5), which includes allelic size in its computation, provided a much more confident estimate on the genetic divergence and on the utility of microsatellites for species identification. Similar results were obtained by Roques et al. (1999), who reported an average Fst = 0.292 and Rst = 0.899. Our data showed that high genetic divergence has accumulated between *P. corruscans* and *P. reticulatum*, two closely related species that can hybridize and produce fertile hybrids. Other studies have demonstrated that the accumulation of allelic differences is common even between closely related species (Roques et al., 1999; Demandt & Bergek, 2009).

Power of markers using Bayesian methods

Roques et al. (1999), using from eight to ten loci with moderate Fst values (0.09–0.20) in *Sebastes* spp., achieved between 87 and 100% species assignment accuracy; using only the four microsatellites with the highest divergence, 95% assignment was reached. Vähä & Primmer (2006), simulating the performance of different Bayesian methods verified that 48 loci with moderate-to-low Fst values (approximately 0.12) achieved 90% hybrid identification, while the 12 loci with the highest Fst values (0.21) provided similar accuracy.

In our study, a high success rate was achieved in classifying simulated individuals as parental or hybrid when all groups of markers were considered (8

microsatellites, 7 diagnostic loci, and all the 11 markers tested) using STRUCTURE at a 0.9 threshold (100% in most cases). This outcome demonstrates the high genetic divergence between parental species and highlights the value of tested markers in analyzing hybridization between *P. corruscans* and *P. reticulatum*. This threshold is in accordance with previous studies using the same program to distinguish between species and hybrids (Sanz et al., 2009; Aboim et al., 2010; Trigo et al., 2014).

Although NEWHYBRIDS correctly classified most parental species and F1 hybrids (100% with diagnostic markers), it misclassified a large number of F2 and backcross hybrids using a Q threshold value of 0.9. The results improved significantly with a Q = 0.6 threshold, and, in this case, all 11 markers performed significantly better than did the diagnostic markers; a relatively small fraction of F2 and backcross hybrids were misclassified (20% of F2, 4% of backcrosses with *P. reticulatum* and 8% of backcrosses with *P. corruscans*). The same threshold was used by Trigo et al. (2014) to achieve the best performance, but higher (0.870; Hasselman et al., 2014) and lower (0.5; Vähä & Primmer et al., 2006; Vilas et al., 2010) thresholds have also been reported. This variability suggests that the Q cut-off should be adjusted to each scenario of divergence between species or populations. Despite its limitations, the Bayesian method implemented in NEWHYBRIDS was effective for hybrid classification in our study, as previously reported (Aboim et al., 2010; Devitt et al., 2011).

Bayesian approaches evaluated in our study had some limitations and our data suggest that a combination of them may be useful in analyzing real scenarios. While STRUCTURE performed best in distinguishing hybrids from parental species, NEWHYBRIDS performed well in hybrid classification in the simplest categories (F1, F2, backcrosses BC) of hybridization. It is clear that more complex categories resulting from full hybrid fertility and from older hybridization events would require a greater number of markers, especially diagnostic ones if they are available. Even in this case, however, the complementary of both Bayesian approaches would be very useful. Moreover, while diagnostic markers perform best at classification, our study demonstrated that a small number of additional non-diagnostic but highly differentiated markers can aid in hybrid categorization.

Assessing genetic admixture in wild

Analysis of hybridization in the Aquidauana and Mogi-Guaçu populations was carried out following the best strategy devised from the simulations; Bayesian approaches implemented in STRUCTURE ($q < 0.9$) and NEWHYBRIDS ($Q > 0.6$) were combined for a more robust categorization of hybrids, and all 11 nuclear markers were ultimately used. The two programs yielded similar results, confirming the consistency of the hybrid classification, enabling the detection of F1 hybrids and backcrosses in the Aquidauana and Mogi-Guaçu Rivers. Only five individuals showed discrepancies between the programs, but in the Mogi-Guaçu, which represents the most complex scenario, STRUCTURE and NEWHY-BRIDS provided largely concordant results. The discordances detected in the Aquidauana River are very likely related to limitations in the characterization of more advanced hybridization stages (BPr x Pr crosses), as Pr and Bpr were the most frequent classes at this site (43.3% and 39.2%, respectively). When few F1 hybrids (1%, for example) and high introgression occur, it is difficult to classify individuals because they represent advanced backcrosses (Vähä & Primmer, 2006; Sanz et al., 2009).

Our study consistently showed the absence of hybrids from most wild populations, even though *P. corruscans* and *P. reticulatum* are sympatric in most populations, which suggests that intercrossing does not occur in the wild or is very rare. Previous studies reported differences in body size at sexual maturity in these species (Resende et al., 1996), which may prevent natural hybridization. Furthermore, new microsatellite data from our study along with previous information (Bignotto et al., 2009; Torrico et al., 2009; Carvalho et al., 2013; Hashimoto et al., 2013) show the high genetic divergence between species, which indicates no or very low gene flow between them. It is estimated that *P. corruscans* diverged from *P. reticulatum* approximately 11.8 to 10 Myr (Torrico et al., 2009). Based on the results of this study, pre-zygotic barriers between these species are related to ecological and reproductive behavior, leading to isolation between them. While more studies are necessary, post-zygotic barriers could also play a role in partially restricting gene flow between species due to some level of reduced hybrid fertility or low fitness.

We confirmed the occurrence of *P. reticulatum–P. corruscans* hybrids in the Aquidauana and Mogi-Guaçu Rivers. Such hybridization has been related to escapees or releases from fish farms or research facilities in these areas (Prado et al., 2012a). To analyze these populations, it is important to bear in mind that both parental species are large catfishes (>1 m of body length) with a long time to reach sexual maturity (approximately, 4.5 years) (Resende et al., 1996; Mateus & Penha, 2007). Although there are no precise data on the viability, fitness, and fertility of hybrids in nature, it has been shown that F1 individuals reproduce in captivity, generating F2 individuals and backcrosses (Prado et al., 2012b). As escapees from fish farms are usually fingerlings or juveniles (Fernandes et al., 2003), the identification of adult hybrids in our study suggests that hybrids are viable and capable of growth in the wild. Additionally, the large number of backcross hybrids identified in our work indicates that F1 hybrids very likely reproduce in nature, as these hybrids most likely escaped from aquaculture facilities.

Hypotheses explaining the differences in hybridization between Aquidauana and Mogi-Guaçu Rivers are relevant not only to understanding the impact of introgression but also to understanding the reproductive behavior of hybrids in the wild and thus to proposing appropriate management guidelines. According to the available data, the most likely origin of hybrids in Aquidauana is farms in the vicinity of this river where F1 hybrids have been produced since the 90 s (Toledo-Filho et al., 1994, 1998). The preferred cross on farms is ♀Pr x ♂Pc due to its better growth performance (Campos, 2010). Most hybrids in this river were backcrosses with *P. reticulatum*; only one F1 were detected. Additionally, the genetic diversity of Aquidauana was very similar to that of non-affected populations in the same area, and a significant heterozygote excess ($F_{IS} = -0.121$) and moderate linkage disequilibrium (11% pairwise comparisons) were observed. Deviations from HW and linkage equilibrium are useful in evaluating hybridization (Allendorf et al., 2010) and may indicate that crosses in the wild are not at random, that the hybridization event is of recent origin, or that there is constant flow of some hybrid classes (Haas et al., 2009). Additionally, all hybrids detected in Aquidauana were of *P. reticulatum* maternal origin, in agreement with the pure species in this population and the preferred

crosses at farms, which employ females of this species. Genetic data from the Aquidauana population and farm practices suggest that F1 hybrids from farms have been escaping at moderate rates (suggested by the low frequency of F1 hybrids) since approximately the 90 s, when this practice began. If escapees are infrequent but constant over time, F1 hybrids (3.3%) would mostly cross with the parental species (Pr), the most frequent class (43.3%), to produce BPrs (39.3%), and increase their proportion depending on how long this scenario has persisted. Additionally, BPr x Pr crosses should occur and produce more advanced hybrids, as our data suggest. Apparently, F1 and BPr hybrids are viable and can reproduce in the wild, supporting previous findings, and they likely cause active introgression and genetic alteration of the recipient population.

Conversely, in the Mogi-Guaçu population, both parental species are present. Fish farms producing F1 hybrids are also present in this river and, as in the Aquidauana River, escapees or releases are likely not infrequent. Additionally, it should be noted that the sampled area is a closed system where no gene flow is possible from adjacent areas of the hydrographic basin. In Mogi-Guaçu, all types of hybrids were detected excluding F2, and F1 individuals were much more frequent (23.3%) than were backcrosses (BCs) (9.3% each class), the reverse of the pattern observed in Aquidauana. The proportions of observed BCs agree well with the expected frequencies of crosses between F1 and both parental classes (Pr: 32.6%; Pc: 25.6%). Additionally, crosses between BCs and the most frequent classes (Pr, Pc, and F1) are likely occurring given the variable q values of some BCs (BPr: from 0.552 to 0.717; BPc: from 0.695 to 0.820). As expected, genetic diversity in Mogi-Guaçu was highest among those in all populations analyzed because of the admixture of two equally frequent and divergent gene pools. Additionally, a significant heterozygote deficit ($F_{IS} = 0.101$) and high linkage disequilibrium (71% pairwise comparisons) was detected, likely as a consequence of reproductive isolation between individuals, especially individuals of parental species. Finally, mtDNA analysis, an effective tool for investigating hybridization events in the wild (Gunnel et al., 2008; Metcalf et al., 2008; Broughton et al., 2011), showed that the maternal genome in some F1 hybrids was Pc. All in all, in this population introgression is apparently occurring

between species due to hybrid fertility. The high proportion of F1 individuals suggests that there are many more escapees from farms in this region than in Aquidauana, although we cannot fully exclude the possibility that parental species could intercross in such disturbed population. Mogi-Guaçu does not represent a panmictic population due to isolation barriers between parental species, but hybrids act as bridges that facilitate introgression.

Implications for conservation

The extinction of unique parental alleles and the loss of local adaptations established during evolutionary history are probably irreversible (Allendorf et al., 2010). The "dilution" of parental genomes via genetic introgression may result in genetic extinction, i.e., genetically distinct species cease to exist and give rise to a single hybrid population (Huxel, 1999; Epifanio & Philipp, 2001; Muhlfeld et al., 2009). In Salmonidae, introgressions between native and non-native species have given rise to hybrid populations in extensive geographic areas, resulting in genomic extinction (Muhlfeld et al., 2009). These authors indicate that even small amounts of hybridization sharply reduce the ability of male and female salmon to reproduce, and breeding success has declined approximately 50% with only 20% genetic mixing. Data from the present study evidenced genetic introgression between species in the Aquidauana and Mogi-Guaçu Rivers that will have unknown consequences regarding adaptation.

Regarding management, some authors have suggested removing hybrids from wild environments to minimize genetic damage to native species and to prevent general introgression (Allendorf et al., 2004; 2010). Recent genetic monitoring of *Pseudoplatystoma* hybrids in nature (Vaini et al., 2014) indicates that hybrids constituted more than 50% of sampled individuals in some rivers of the Parana Basin. Another difficulty for species conservation is that, considering the long-distance migration of parental species (Resende et al., 1996) and hybrid fertility, it is possible that hybrids are able to migrate and reproduce in areas far from the region of escape and introduction. It is highly probable that genetic contamination of wild stocks is occurring at higher rates than have been detected in recent surveys. In this scenario, hybrid fertility and ongoing escapes from farms in Brazil reduce the efficiency of the hybrid-removal strategy,

as many wild specimens are admixed to different degrees. Hybridization practices at farms represent an important threat to the future of these species if they spread given the increasing importance of these species in Brazilian aquaculture. The set of markers and statistical approach developed in this work to assess hybridization between species will be very useful in monitoring escapes and evaluating the degree of introgression in wild populations near farms. This information is essential to establishing controls at farms to move toward more sustainable aquaculture practices that help to protect the rich biodiversity of Brazilian basins. Before more appropriate tools become available, genetic methods should be routinely applied in programs to monitor aquaculture broodstock and juvenile hybrids, to support sustainable development, i.e., integration of the production of fish and conservation measures to reduce environmental impacts from aquaculture (Hashimoto et al., 2014). Policies that regulate and supervise hybrid production in Brazilian fish farms are also necessary to prevent future escapes.

Conclusions

Considering the level of genetic divergence between *P. corruscans* and *P. reticulatum*, the set of loci used in this work performed well for Bayesian identification and classification of species and hybrids compared to methods used in previous studies (Roques et al., 1999; Vähä & Primmer, 2006; Sanz et al., 2009; Bohling et al., 2013; Hasselman et al., 2014). Additionally, the heterogeneity of the hybridization scenarios indicates that the use of the full set of markers through both the STRUCTURE and NEWHYBRIDS Bayesian methods would be the most comprehensive methodology, since they complement the results interpretation. Genetic results also provide information on parental microsatellites (allelic frequencies per locus and the complete microsatellite genotypes of all parental populations) for future use as reference parental datasets in Bayesian analyses. Highly divergent loci could also result in high categorization success using the same methods. Our data also showed that pre-zygotic reproductive barriers exist between *P. reticulatum* and *P. corruscans*, preventing natural hybridization. Considering the low probability of spontaneous interspecific crosses in the wild, a well-supported hypothesis explaining the high

amount of hybrids of these species identified in the wild involves aquaculture escapes and the introduction of hybrids from aquaculture (Bignotto et al., 2009; Prado et al., 2012a; Hashimoto et al., 2014; Vaini et al., 2014; present work). Although advanced hybrids may also come from fish farms, our data suggest that hybrids are reproducing and causing genetic introgression in the wild.

References

Abbott, R., D. Albach, S. Ansell, J. W. Arntzen, J. W. Baird, N. Bierne & J. Boughman, 2013. Hybridization and speciation. Journal of Evolutionary Biology 26: 229–246.

Aboim, M. A., J. Mavárez, L. Bernatchez & M. M. Coelho, 2010. Introgressive hybridization between two Iberian endemic cyprinid fish: a comparison between two independent hybrid zones. Journal of Evolutionary Biology 23: 817–828.

Allendorf, F. W., P. A. Hohenlohe & G. Luikart, 2010. Genomics and the future of conservation genetics. Nature 11: 697–709.

Allendorf, F. W., R. F. Leary, N. P. Hitt, K. L. Knudsen, L. L. Lundquist & P. Spruell, 2004. Intercrosses and the U.S. Endangered Species Act: should Hybridized Populations be Included as Westslope Cutthroat Trout? Conservation Biology 18: 1203–1213.

Anderson, E. & E. A. Thompson, 2002. A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160: 1217–1229.

Arnold, M. L., M. R. Bulger, J. M. Burke, A. L. Hempel & J. H. Williams, 1999. Natural hybridization: how low can you go and still be important? Ecology 80: 371–381.

Barton, N. H., 2001. The role of hybridization in evolution. Molecular Ecology 10: 551–568.

Bignotto, T. S., A. J. Prioli, S. M. A. P. Prioli, T. C. Maniglia, T. A. Boni, L. C. Lucio, V. N. Gomes, R. A. Prioli, A. V. Oliveira, H. F. Julio & L. M. Prioli, 2009. Genetic divergence between *Pseudoplatystoma corruscans* and *Pseudoplatystoma reticulatum* (Siluriformes: pimelodidae) in the Parana River basin. Brazilian Journal of Biology 69: 681–689.

Boecklen, W. J. & D. J. Howard, 1997. Genetic analysis of hybrid zones: numbers of markers and power of resolution. Ecology 78: 2611–2616.

Bohling, J. H., J. R. Adams & L. P. Waits, 2013. Evaluating the ability of Bayesian clustering methods to detect hybridization and introgression using an empirical red wolf data set. Molecular Ecology 22: 74–86.

Brennan, A. C., G. Woodward, O. Seehausen, V. Muñoz-fuentes, C. Moritz, A. Guelmami, R. J. Abbott & P. Edelaar, 2014. Hybridization due to changing species

distributions: adding problems or solutions to conservation of biodiversity during global change? Evolutionary Ecology Research 16: 475–491.

Broughton, R. E., K. C. Vedala, T. M. Crowl & L. L. Ritterhouse, 2011. Current and historical hybridization with differential introgression among three species of cyprinid fishes (genus *Cyprinella*). Genetica 139: 699–707.

Buitrago-Suárez, U. A. & B. M. Burr, 2007. Taxonomy of the catfish genus *Pseudoplatystoma* Bleeker (Siluriformes: pimelodidae) with recognition of eight species. Zootaxa 1512: 1–38.

Campos, J. L., 2010. O cultivo do pintado (*Pseudoplatystoma corruscans*, Spix; Agassiz, 1829), outras espécies do genero *Pseudoplatystoma* e seus híbridos. In Baldisserotto, B. B. & L. C. Gomes (eds.), Espécies Nativas Para a Piscicultura no Brasil. Universidade Federal de Santa Maria, Santa Maria: 335–361.

Carvalho, D. C., A. S. Seerig, B. S. A. F. Brasil, D. V. Crepaldi & D. A. A. Oliveira, 2013. Molecular identification of the hybrid between the catfish species *Pseudoplatystoma corruscans* and *Pseudoplatystoma reticulatum* using a set of eight microsatellite markers. Journal of Fish Biology 83: 671–676.

Demandt, M. H. & S. Bergek, 2009. Identification of cyprinid hybrids by using geometric morphometrics and microsatellites. Journal of Applied Ichthyology 25: 695–701.

Devitt, T. J., S. J. E. Baird & C. Moritz, 2011. Asymmetric reproductive isolation between terminal forms of the salamander ring species *Ensatina eschscholtzii* revealed by fine-scale genetic analysis of a hybrid zone. BMC Evolutionary Biology 11: 245.

Dubut, V., M. Sinama, J.-F. Martin, E. Meglécz, J. Fernandez, R. Chappaz, A. Gilles & C. Costedoat, 2010. Cross-species amplification of 41 microsatellites in European cyprinids: a tool for evolutionary, population genetics and hybridization studies. BMC Research Notes 3: 135.

Earl, D. A. & B. M. VonHoldt, 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4: 359–361.

Epifanio, J. & D. Philipp, 2001. Simulating the extinction of parental lineages from introgressive hybridization: the effects of fitness, initial proportions of parental taxa, and mate choice. Reviews in Fish Biology and Fisheries 10: 339–354.

Evanno, G., S. Regnaut & J. Goudet, 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14: 2611–2620.

Excoffier, L., G. Laval & S. Schneider, 2005. ARLEQUIN ver. 3. 0: an integrated software package for population genetics data analysis. Evolutionary Bioinformatics 1: 47–50.

Fernandes, R., L. C. Gomes & A. A. Agostinho, 2003. Pesquepague: negócio ou fonte de dispersão de espécies exóticas? Acta Scientiarum Biological Sciences 25: 115–120.

Goodman, M., J. Pedwaydon, J. Czelusniak, J. Suzuki, T. Gotoh, L. Moens, F. Shishikura, et al., 1988. An evolutionary tree for invertebrate globin sequences. Journal of Molecular Evolution 7: 236–249.

Goodman, S. J., 1997. RST CALC: a collection of computer programs for calculating unbiased estimates of genetic differentiation and determining their significance for microsatellite data. Molecular Ecology 6: 881–885.

Goudet, J., 2002. FSTAT, a program to estimate and test gene diversities and fixation indices version 2.9.3. http://www.unil.ch/izea/softwares/Fstat.html.

Gunnel, K., M. K. Tada, F. A. Hawthorne, E. R. Keeley & M. B. Ptacek, 2008. Geographic patterns of introgressive hybridization between native Yellowstone cutthroat trout (*Oncorhynchus clarkii bouvieri*) and introduced rainbow trout (*O. mykiss*) in the South Fork of the Snake River watershed. Idaho. Conservation Genetics 9: 49–64.

Guo, S. W. & E. A. Thompson, 1992. Performing the exact test of Hardy–Weinberg proportion for multiple alleles. Biometrics 48: 361–372.

Haas, F., M. A. Pointer, N. Saino, A. Brodin, N. I. Mundy & B. Hansson, 2009. A analysis of population genetic differentiation and genotype-phenotype association across the hybrid zone of carrion and hooded crows using microsatellites and MC1R. Molecular Ecology 18: 294–305.

Haldane, J. B. S., 1954. An exact test for randomness of mating. Journal of Genetics 52: 631–635.

Hansen, M. M., E. Kenchington & E. E. Nielsen, 2001. Assigning individual fish to populations using microsatellite DNA markers. Fish and Fisheries 2: 93–112.

Hashimoto, D. T., F. D. Prado, J. A. Senhorini, F. Foresti & F. Porto-Foresti, 2013. Detection of post-F1 fish hybrids in broodstock using molecular markers: approaches for genetic management in aquaculture. Aquaculture Research 44: 876–884.

Hashimoto, D. T., F. D. Prado, J. A. Senhorini, F. Foresti & F. Porto-Foresti, 2014. Aquaculture of Neotropical Catfish Hybrids: Genetic Strategies for Conservation and Management. In Bradley, R. (ed.), Carp and Catfish: Biology, Behavior and Conservation Strategies. Nova Science Publishers, New York: 1–10.

Hasselman, D. J., E. E. Argo, M. C. McBride, P. Bentzen, T. F. Schultz, A. A. Perez-Umphrey & E. P. Palkovacs, 2014. Human disturbance causes the formation of a hybrid swarm between two naturally sympatric fish species. Molecular Ecology 23: 1137–1152.

Ho, N. C. K., A. L. F. C. H. Ho, G. D. Underwood, A. Underwood, D. Zhang & J. Lin, 2015. A simple molecular protocol for the identification of hybrid Western Atlantic seahorses, *Hippocampus erectus* × *H. reidi*, and potential consequences of hybrids for conservation. Journal of Zoo and Aquarium Research 3: 11–20.

Huxel, G. R., 1999. Rapid displacement of native species by invasive species: effects of hybridization. Biological Conservation 89: 143–152.

Jakobsson, M. & N. A. Rosenberg, 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23: 1801–1806.

Jombart, T. & I. Ahmed, 2011. Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27: 3070–3071.

Jombart, T., S. Devillard & F. Balloux, 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics 11: 94.

Khosravi, R., H. R. Rezaei & M. Kaboli, 2013. Detecting hybridization between Iranian wild wolf (*Canis lupus pallipes*) and free-ranging domestic dog (*Canis familiaris*) by analysis of microsatellite markers. Zoological Sciences 30: 27–34.

Kovach, A. I., J. Walsh, J. Ramsdell & W. Kelley Thomas, 2015. Development of diagnostic microsatellite markers from whole-genome sequences of Ammodramus sparrows for assessing admixture in a hybrid zone. Ecology and Evolution 5: 2267–2283.

Laikre, L., A. Antunes, A. Apostolidis, P. Berrebi, A. Duguid, A. Ferguson J. L. García-Marín, R. Guyomard, M. M. Hansen, K. Hindar, M. L. Koljonen, C. Largiader, P. Martínez, E. E. Nielsen, S. Palm, D. Ruzzante, N. Ryman & C.Triantaphyllidis, 1999. Conservation Genetic Management of Brown Trout (Salmo trutta) in Europe. ("TROUTCONCERT": EU FAIR CT97-3882). Stockholm University, Sweden.

Machado-Schiaffino, G., F. Juanes & E. Garcia-Vazquez, 2010. Introgressive hybridization in North American hakes after secondary contact. Molecular Phylogenetics and Evolution 55: 552–558.

Mallet, J., 2005. Hybridization as an invasion of the genome. Trends in Ecology and Evolution 20: 229–237.

Masaoka, T., H. Okamoto, K. Arakia, H. Nagoya, A. Fujiwarab & T. Kobayashic, 2012. Identification of the hybrid between *Oryzias latipes* and *Oryzias curvinotus* using nuclear genes and mitochondrial gene region. Marine Genomics 7: 37–41.

Mateus, L. A. F. & J. M. F. Penha, 2007. Dinâmica populacional de quatro espécies de grandes bagres na bacia do rio Cuiabá, Pantanal norte, Brasil (Siluriformes, Pimelodidae). Revista Brasileira de Zoologia 24: 87–98.

McBride, M. C., T. V. Willis, R. G. Bradford & P. Bentzen, 2014. Genetic diversity and structure of two hybridizing anadromous fishes (*Alosa pseudoharengus*, *Alosa aestivalis*) across the northern portion of their ranges. Conservation Genetics 15: 1281–1298.

Metcalf, J. L., M. R. Siegle & A. P. Martin, 2008. Hybridization dynamics between Colorado's native cutthroat trout and introduced rainbow trout. Journal of Heredity 99: 149–156.

Moyer, G. R., B. M. Burr & C. Krajewski, 2004. Phylogenetic relationships of thorny catfishes (Siluriformes, Doradidae) inferred from molecular and morphological data. Zoological Journal of Linnean Society 140: 551–575.

Muhlfeld, C. C., C. Clint, T. E. McMahon, D. Belcer & J. L. Kershner, 2009. Spatial and temporal spawning dynamics of native westslope cutthroat trout, Oncorhynchus clarkii lewisi, introduced rainbow trout, *Oncorhynchus mykiss*, and their hybrids. Canadian Journal of Fisheries and Aquatic Sciences 66: 1153–1168.

Nielsen, E. E., L. A. Bach & P. Kotlicki, 2006. Hybridlab (version 1.0): a program for generating simulated hybrids from population samples. Molecular Ecology Notes 6: 971–973.

Palumbi, S. R., 1996. Nucleic acids II: the polymerase chain reaction. In Hillis, D., C. Moritz & B. Mable (eds.), Molecular Systematics. Sinauer Associates Inc., Sunderland: 205–247.

Porto-Foresti, F., D. T. Hashimoto, J. A. Senhorini & F. Foresti, 2010. Hibridação em piscicultura: monitoramento e

perspectivas. In Baldisserotto, B. & L. C. Gomes (eds.), Espécies Nativas Para Piscicultura no Brasil. UFMS, Mato Grosso do Sul: 589–606.

Prado, F. D., B. G. Pardo, J. Guerra-Varela, J. A. Senhorini, P. Martínez, F. Foresti & F. Porto-Foresti, 2014. Development and characterization of 16 microsatellites for the Neotropical catfish *Pseudoplatystoma reticulatum* and cross species analysis. Conservation Genetics Resources 6: 679–681.

Prado, F. D., D. T. Hashimoto, F. F. Mendonça, J. A. Senhorini, F. Foresti & F. Porto-Foresti, 2011. Molecular identification of hybrids between Neotropical catfish species *Pseudoplatystoma corruscans* and *Pseudoplatystoma reticulatum*. Aquaculture Research 42: 1890–1894.

Prado, F. D., D. T. Hashimoto, J. A. Senhorini, F. Foresti & F. Porto-Foresti, 2012a. Detection of hybrids and genetic introgression in wild stocks of two catfish species (Siluriformes: pimelodidae): the impact of hatcheries in Brazil. Fisheries Research 125–126: 300–305.

Prado, F. D., T. L. Nunes, J. A. Senhorini, J. Bortolozzi, F. Foresti & F. Porto-Foresti, 2012b. Cytogenetic characterization of F1, F2 and backcross hybrids of Neotropical catfish species *Pseudoplatystoma corruscans* and *P. reticulatum* (Pimelodidae, Siluriformes). Genetics and Molecular Biology 35: 57–64.

Pritchard, J. K., M. Stephens & P. Donnelly, 2000. Inference of population Structure using multilocus genotype data. Genetics 155: 945–959.

R Development Core Team, 2014. The R project for statistical computing, Version 3.2.3. http://www.r-project.org.

Raymond, M. & F. Rousset, 1995. GENEPOP (ver. 3.3): a population genetics software for exact test and ecumenicism. Journal of Heredity 86: 248–249.

Resende, E.K., A.C. Catella, F.L. Nascimento, S.S. Palmeiras, R.A.C. Pereira, M.S. Lima & V.L.L. Almeida, 1996. Biologia do curimbatá (*Prochilodus lineatus*), pintado (*Pseudoplatystoma coruscans*) e cachara (*Pseudoplatystoma fasciatum*) na bacia hidrográfica do rio Miranda, Pantanal do Mato Grosso do Sul. EMBRAPA – CPAP, Boletim de Pesquisa 02, Mato Grosso do Sul, Brasil.

Revaldaves, E., L. H. G. Pereira, F. Foresti & C. Oliveira, 2005. Isolation and characterization of microsatellite loci in *Pseudoplatystoma corruscans* (Siluriformes: Pimelodidae) and cross-species amplification. Molecular Ecology Notes 5: 463–465.

Rhymer, J. M. & D. Simberloff, 1996. Extinction by hybridisation and introgression. Annual Review of Ecology, Evolution, and Systematics 27: 83–109.

Rice, W. R., 1989. Analyzing tables of statistical tests. Evolution 43: 223–225.

Roques, S., P. Duchesne & L. Bernatchez, 1999. Potential of microsatellites for individual assignment: the North Atlantic redfish (genus *Sebastes*) species complex as a case study. Molecular Ecology 8: 1703–1717.

Rousset, F., 1996. Equilibrium values of measures of population subdivision for stepwise mutation processes. Genetics 142: 1357–1362.

Sanz, N., R. M. Araguas, R. Fernández, M. Vera & J. L. García-Marín, 2009. Efficiency of markers and methods for detecting hybrids and introgression in stocked populations. Conservation Genetics 10: 225–236.

Toledo-Filho, S.A., L.F. Almeida-Toledo, F. Foresti, G. Bernardino & D. Calcagnotto, 1994. *Monitoramento e conservação genética em projeto de hibridação entre pacu e tambaqui.* CCS/USP, Cadernos de Ictiogenética 2, São Paulo, Brasil.

Toledo-Filho, S.A., L.F. Almeida-Toledo, F. Foresti, D. Calcagnotto, S.B.A.F. Santos & G. Bernardino, 1998. *Programas Genéticos de Seleção, Hibridação e Endocruzamento Aplicados à Piscicultura.* CCS/USP, Cadernos de Ictiogenética 4, São Paulo, Brasil.

Tomoko, O., 1983. On the evolution of multigene families. Theoretical Population Biology 23(2): 216–240.

Torrico, J. P., N. Hubert, E. Desmarais, F. Duponchelle, J. Nuñez-Rodriguez, C. G. Montoya-Burgos, F. M. Davila, A. A. Carvajal-Vallejos, F. Bonhommed Grajales & J. F. Renno, 2009. Molecular phylogeny of the genus *Pseudoplatystoma* (Bleeker, 1862): biogeographic and evolutionary implications. Molecular Phylogenetics and Evolution 51: 588–594.

Trigo, T. C., F. P. Tirelli, T. R. O. Freitas & E. Eizirik, 2014. Comparative assessment of genetic and morphological variation at an extensive hybrid zone between two wild cats in southern Brazil. PloS One 9: e108469.

Vähä, J. P. & C. Primmer, 2006. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. Molecular Ecology 15: 63–72.

Vaini, J. O., A. B. Grisolia, F. D. Prado & F. Porto-Foresti, 2014. Genetic identification of interspecific hybrid of Neotropical catfish species (*Pseudoplatystoma corruscans* vs. *Pseudoplatystoma reticulatum*) in rivers of Mato Grosso do Sul State. Brazil. Neotropical Ichthyology 12: 635–641.

Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills & P. Shipley, 2004. Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4: 535–538.

Vilas, R., C. Bouza, J. Castro, A. López & P. Martínez, 2010. Management units of brown trout from Galicia (NW: Spain) based on spatial genetic structure analysis. Conservation Genetics 11: 897–906.

Walter, R. P., E. S. Gnyra, L. I. Söderberg & D. D. Heath, 2014. Rapid genetic identification of brown bullhead (*Ameiurus nebulosus*), black bullhead (*Ameiurus melas*) and their hybrids. Conservation Genetics Resources 6: 507–509.

Weir, B. S. & C. C. Cockerham, 1984. Estimating F-statistics for the analysis of population structure. Evolution 38: 1358–1370.

White, T. J., T. Bruns, S. Lee & J. W. Taylor, 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In Innis, M. A., D. H. Gelfand, J. J. Sninsky & T. J. White (eds.), PCR Protocols: A Guide to Methods and Applications. Academic Press, New York: 315–322.

Willis, S. C., J. Macrander, I. P. Farias & G. Ortí, 2012. Simultaneous delimitation of species and quantification of interspecific hybridization in Amazonian peacock cichlids (genus *Cichla*) using multi-*locus* data. BMC Evolutionary Biology 12: 96.

Zhang, Z., K. A. Glove, V. Wennevik, T. Svasand, A. G. E. Sorvik, P. Fiske & S. Karlsson, 2013. Genetic analysis of Atlantic salmon captured in a netting station reveals multiple escapement events from commercial fish farms. Fisheries Management and Ecology 20: 42–45.