Genetic diversity of *Arapaima gigas* (Schinz, 1822) (Osteoglossiformes: Arapaimidae) in the Araguaia-Tocantins basin estimated by ISSR marker

Carla A. Vitorino[1], Renata C. C. Oliveira[1], Vladimir P. Margarido[2] and Paulo C. Venere[1]

The genetic diversity of the specimens of four natural populations of *Arapaima* from Araguaia-Tocantins basin was assessed within and among these stocks, using five primers for ISSR. COI (*cytochrome c oxidase subunit I*) partial sequences confirmed that the specimens belongs to *Arapaima gigas*. The ISSR provided 168 loci, of which 165 were polymorphic. However, the number of loci for each population and expected heterozygosity values were low. AMOVA showed 52.63% intra-population variation and 47.37% inter-population variation. The F_{ST} was high among all populations ($F_{ST} \geq 0.25$), however, the cluster analysis (PCoA) and Bayesian inference showed three major groups: Araguaiana-MT + São Félix do Araguaia-MT, Novo Santo Antônio-MT and Itupiranga-PA. The genetic distance was not correlated with geographical distance. The ISSR marker revealed that the populations of the Araguaia-Tocantins are structured and have a low genetic diversity. These are the first data from a population analysis using molecular markers for *A. gigas* of Araguaia-Tocantins basins and may be used to define the best management strategies and conservation projects for this species.

A diversidade genética dos espécimes de quatro populações naturais de *Arapaima* coletados na bacia do Araguaia-Tocantins foi avaliada com base em cinco *primers* para marcadores moleculares ISSR. A sequência parcial do COI (*cytochrome c oxidase subunit I*) confirmou que os espécimes pertencem à espécie *Arapaima gigas*. Os ISSR forneceram 168 *loci*, dos quais 165 polimórficos. No entanto, para cada população, os valores de heterozigosidade esperada foram baixos. A AMOVA mostrou 52,63% de variação intrapopulacional e 47,37% interpopulacional. O F_{ST} foi alto entre todas as populações ($F_{ST} \geq 0,25$); entretanto, a análise de agrupamento e a inferência Bayesiana mostraram três grandes grupos: Araguaiana-MT + São Félix do Araguaia-MT, Novo Santo Antônio-MT e Itupiranga-PA. A distância genética não teve correlação com a distância geográfica. Os ISSRs se mostraram eficientes para determinar a diversidade genética para a *A. gigas*, revelando que as populações da bacia Araguaia-Tocantins estão estruturadas e com baixa diversidade genética. Estes são os primeiros dados de análise populacional utilizando ISSR para *A. gigas* da bacia Araguaia-Tocantins e poderão ser utilizados para definir as melhores estratégias de manejo e projetos de conservação dessa espécie.

Keywords: Bottleneck, Endogamy, F_{ST}, Pirarucu, Population genetics.

Introduction

The genetic diversity of a population is considered raw material for evolution by natural selection (Fisher, 1930; Hughes *et al*., 2008). However, the gene pool of a population may change over time because the population size may vary due to birth rates and mortality, migration and contact with other populations (Klug *et al*., 2012). These changes may compromise their evolutionary success, since the existence of a species is closely related to genetic variability (Woodruff, 2001). Small populations tend to reduce genetic diversity due to high rates of inbreeding, inducing a high rate of homozygosity (Woodruff, 2001; Frankham *et al*., 2008). In this context, the loss of habitat has direct consequences in reducing population size, and this is more evident, especially for species of fish from inland waters (Barletta *et al*., 2010), since they tend to be isolated in drains, resulting in distinct populations (Allan & Flecker, 1993).

[1]Laboratório de Citogenética e Genética Animal, Instituto de Biociências, Universidade Federal de Mato Grosso, Av. Fernando Correa da Costa s/n, Coxipó, 78060-900 Cuiabá, MT, Brazil. (CAV) carlanvit@gmail.com, (RCCO) renataclaudino.oliveira@gmail.com, (PCV) pvenere@uol.com.br (corresponding author)
[2]Centro de Ciências Biológicas e da Saúde, Universidade Estadual do Oeste do Paraná, Rua Universitária, 2.069, Jardim Universitário, 85819-110 Cascavel, Paraná, Brazil. vladimir.margarido@unioeste.br

In addition to habitat loss, commercial overexploitation, especially of large species, has rapidly degraded native fish stocks, threatening the diversity of fish (Allan *et al*., 2005; Castello *et al*., 2011), and driving some species to population bottleneck situations that are almost irreversible. Overexploitation has been a major threat to the "pirarucu", *Arapaima gigas* (Schinz, 1822) (Castello & Stewart, 2010; Castello *et al*., 2014), one of the largest species of freshwater fish in the world, which reaches up to 2½ meters in length and 250 to 300kg (Nelson, 2006). *Arapaima gigas* was considered the only species of the genus *Arapaima*, however, recent studies have described a new species, *Arapaima leptosoma* Stewart, 2013 and redescribed *Arapaima agassizii* (Valenciennes, 1847), based on morphological characteristics (Stewart, 2013a, 2013b). *Arapaima gigas* can be found mainly in lentic environments of the Amazon, Araguaia-Tocantins and Essequibo basins (Queiroz, 2000; Castello & Stewart, 2010). It is one of the few species of fish listed in the appendices II of Convention of International Trade in Endangered Species of Wild Fauna and Flora (CITES).

Another threat to the genetic integrity of *A. gigas* refers to transposition of specimens (transporting specimens from one region to another) that is routinely done by aquaculture enterprises, often over hundreds of kilometers, which can homogenize the gene pool or even eliminate locally adapted races (Castello & Stewart, 2010). Several studies about genetic diversity in *A. gigas* have been carried in the Amazon region, with the purpose of determining the risk of extinction of this species and helping to define appropriate strategies for their management (Farias *et al*., 2003; Hrbek *et al*., 2005; Hrbek & Farias, 2008; Hamoy *et al*., 2008; Araripe *et al*., 2013). Thus, while there are already several studies concerning the stocks of the Amazonian plain, there is only one study based on chromosomal/genetic markers in the plain of the rio Araguaia, and this study was conducted at a single location in the middle rio Araguaia (Marques *et al*., 2006). This study indicated that the samples collected were part of a single population with high intra-population genetic diversity.

The possibility of estimating the genetic variability by polymorphisms in the DNA of an organism, as well as of the popularization of molecular techniques, has encouraged progress in studies of population genetics (Antunes *et al*., 2010; Moresco *et al*., 2013). These molecular markers are important tools for estimating parameters, such as levels of genetic diversity within populations and magnitudes of gene flow between populations (Avise, 1996; Woodruff, 2001).

A variety of techniques can be used to evaluate the genetic diversity of a population or species, using both markers of mitochondrial DNA and nuclear DNA. The mitochondrial markers are widely employed in studies of genetic diversity, with the "DNA barcode" most recently being used. The sequence of the mitochondrial gene cytochrome c oxidase subunit I (COI) the most widely used marker for "DNA barcode", which is considered an efficient, fast, accurate and globally accessible tool for the identification of species (Hebert & Gregory, 2005; Hajibabaei *et al*., 2007). Among the molecular markers, microsatellites (or *simple sequence repeats* – SSRs) have become more commonly used in population studies (Frankham *et al*., 2008).

The segment of DNA amplified by the inter-SSR markers (ISSR) includes the nucleotide sequence located between two blocks of microsatellites, oriented in opposite directions, producing a multilocus maker, which is highly polymorphic, and is useful for the analysis of genetic diversity, making it a good choice for DNA fingerprinting (Bornet & Branchard, 2001; Reddy *et al*., 2002; Maltagliati *et al*., 2006). Due to its ease of use, low cost and high sensitivity, ISSR markers have been employed in studies on genetic diversity of several species of plants and animals (Luque *et al*., 2002; Bornet & Branchard, 2004; Askari *et al*., 2011; Moresco *et al*., 2013), as well as studies of sexual differentiation in plants (Ehsanpour *et al*., 2008) and identification of hybrid fish (Bignotto *et al*., 2009; Almeida-Ferreira *et al*., 2011). This marker has allowed the analysis of the main parameters used for determining the genetic diversity and the level of differentiation between species and natural populations of Neotropical fish (amount of polymorphic loci, expected or average heterozygosity and the number of migrants per generation) (Paiva *et al*., 2006; Sofia *et al*., 2006; Lopes *et al*., 2008; Antunes *et al*., 2010; Almeida-Ferreira *et al*., 2011; Domingos *et al*., 2014).

Given the possibility of other *Arapaima* species, and in order to make more robust work, we used the DNA barcode (COI gene sequence) to check whether natural populations of specimens of *A. gigas*, distributed along an environmental gradient, from upstream lakes in the middle rio Araguaia region to the lower rio Tocantins portions, belonged to a single species – *Arapaima gigas*. The work concentrated on estimating the genetic variability within and between each sampled stock, in order to verify the level of genetic diversity within and among them. Although there are specific primers for *A. gigas* microsatellites, the use of ISSRs was chosen in this study due to its easy application and low cost.

Material and Methods

Samples of muscle and fin tissue from one hundred and forty nine *Arapaima gigas* individuals were collected from four different sites distributed along the Araguaia and Tocantins rivers (Fig. 1): 45 samples from the lake Mimoso, municipality of Araguaiana-MT (15°23'58,81"S/51°43'00,08"W), 45 samples from lakes located in the municipality of Novo Santo Antônio-MT (12°18'56.11"S/50°58'24.78"W), 42 samples from lakes located in the municipality of São Félix do Araguaia-MT (11°38'26.36"S/50°40'43.44"W) and 14 samples from lakes near the city of Itupiranga-PA (05°10'46.22"S/49°21'44.45"W).
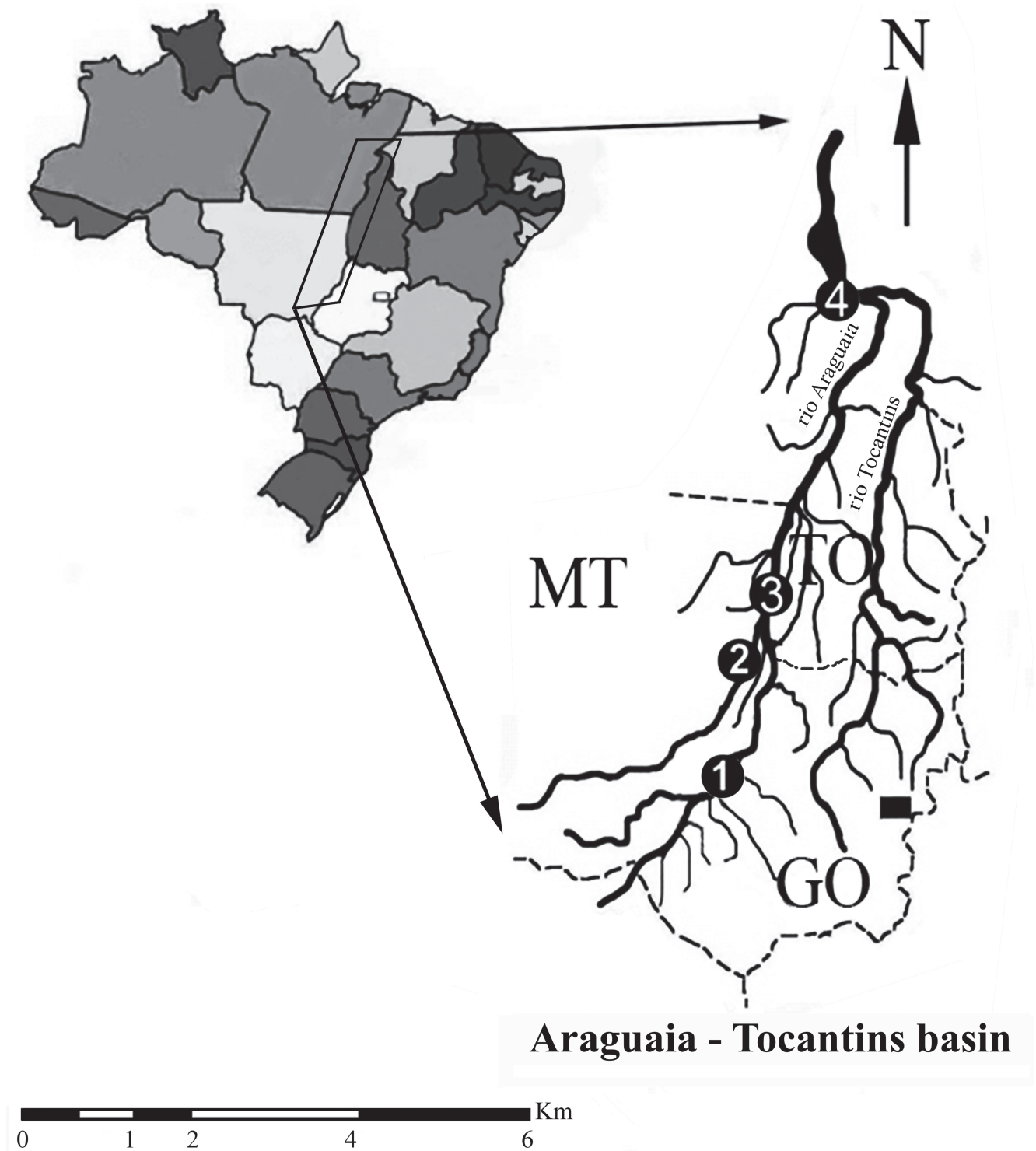
Fig. 1. Map of Brazil, with emphasis on the Araguaia-Tocantins basin. Collection sites of *Arapaima gigas*: (1) Araguaiana-MT, (2) Novo Santo Antônio-MT, (3) São Félix do Araguaia-MT, and (4) Itupiranga-PA.

Samples of specimens from Araguaiana-MT were obtained during a rescue operation, where the fish were taken from a seasonal lake and placed in a perennial lake. During this operation, small pieces of fin were removed. All other tissue samples were obtained from specimens killed for sale, along with fishermen, so no specimens of *A. gigas* on each location was sacrificed specifically for this study. All tissue sampling (muscles or fin) were preserved in 100% alcohol and deposited in the laboratory of the Grupo de Estudos em Peixes do Médio Araguaia (GEPEMA/CNPq/UFMT) for DNA extraction.

DNA extraction from muscular tissue or fin followed a salt extraction protocol of Aljanabi & Martinez (1997), with the following modifications. In microtubes containing tissue fragments, 440µL of lysis buffer (10mM Tris-HCl, 2mM EDTA, 400mM NaCl, 2% SDS) and 10 µL of proteinase K (10mg/mL) were added, and then incubated in a water bath at 55°C for approximately 1:30h. DNA was precipitated using 300µL of NaCl (5M) and centrifuged for 10min at 10.000rpm. Supernatant containing the DNA was transferred to micro tubes and precipitated with 500µl of 100% isopropanol. DNA was centrifuged for 10min at 10.000rpm, washed with 700µL of 70% ethanol, dried and resuspended in 50 µL sterile dH_2O. After that, 5µL of RNAse (10mg/mL) was added to each sample, which were incubated at 37°C for 30 min and stored at -20°C. DNA quantification and quality analysis were conducted using Eppendorf Biophotometer Plus (Eppendorf Hamburg, Hamburg, Germany). Subsequently, DNA samples were diluted to a final concentration of 50 ng/µL.

Partial sequences of the cytochrome c oxidase I gene (COI) were amplified with COI FishF1 and COI FishR1 primers (Ward *et al.*, 2005). The PCR reaction contained the final volume of 12.5 µl, including 1.25µl 10x PCR Buffer, 0.37µl MgCl2 (50mM), 0.25 µl of each primer (10µM), 0.2µl dNTPs mix (2.5mM), 0.1 µl Taq polymerase (Invitrogen), 1.0 µl template DNA (50ng/µl) and 15.6 µl ultrapure water. Amplifications were made using the cycle that corresponds to an initial step of 5 min at 94ºC, followed by 35 cycles of 30 sec at 94ºC, 45 sec at 52ºC and 1 min at 72ºC, followed by 7 min at 72ºC. PCR products were visualized on 1% agarose gel and sequenced following the procedure of Sanger *et al.* (1977), using automated sequencer ABI 3500 Life Technologies.

The COI sequences were analyzed, aligned and edited with BioEdit software (Hall, 1999) with the Clustal W tool, researched and aligned with sequences available in GenBank (National Center for Biotechnology Information), http://www.ncbi.nlm.nih.gov/BLAST/ Blast.cgi?CMD=Web&PAGETYPE=BLASTHome) using BLAST (Basic Local Alignment Search Tool), and used BOLD Systems bioinformatics platform (http://www.boldsystems.org/index.php/IDS_ IdentificationRequest#) for comparison of the sequences in order to confirm the individuals identity used in the study. Genetic distance was calculated using the model Kimura-2-Parameter (K2P) (Kimura, 1980). Neighbor-joining dendogram (NJ) was done with K2P model, with support 1.000 bootstrap. Analysis of genetic distance and dendogram were made with MEGA 6.0 software (Tamura *et al.*, 2013).

For ISSR amplification, the following primers were previously selected and used $(GGAC)_4$, $(GGAC)_3A$, $(GGAC)_3T$, $(GGAC)_3C$, and $(AACC)_4$. Polymerase chain reaction (PCR) markers were developed according to Fernandes-Matioli *et al.* (2000). Each amplification reaction contained 50ng of DNA, 0.5 µM primer, 0.2 mM dNTP, 1X buffer 200 mM Tris-HCl (pH 8.4), 500 mM KCl), 1.5 mM MgCl2, 0.6 Recombinant Taq DNA polymerase (Invitrogen) and enough water to make up a volume of 13µL. Negative controls without DNA were included in each set of amplifications. The amplification reactions were performed in Eppendorf Mastercycler Gradient thermocycler scheduled for 5 cycles of 45s at 94°C, 1 min at 51°C and 1 min at 72°C, followed by 30 cycles of 45s at 94°C, 1 min at 48°C and 1 min at 72°C. After the last cycle of amplification, the reaction mixture was cooled and maintained at 4°C. PCR amplification reproducibility was tested in at least five independent reactions. After amplification, samples consisting of 3 µL of PCR reaction mixture were subjected to electrophoresis on 10% polyacrylamide gel stained with silver nitrate. Gels were photodocumented using a Mini BIS image analysis system (DNR Bio-Imaging Systems Ltd., Kiryat Anavim, Israel) for posterior analysis.

Each amplified ISSR fragment was considered an independent allele and judged as binary characters: present or absent for each specimen. Thus, a matrix based on the presence (1) or absence (0) of bands on gels was generated and used to calculate genetic distance, and intra- and interpopulation variation. Program POPGENE 1.32 (Yeh *et al.*, 1999) was used to calculate the percentage of polymorphic loci. Pairwise genetic distance matrix between individuals was obtained by the Jaccard similarity index, and used to construct the Neighbor-joining dendrogram with the program FreeTree (Hampl *et al.*, 2001) and MEGA 6.0 (Tamura *et al.*, 2013). Scatter plot of principal coordinates was constructed using the programs DistPCOA (Legendre & Anderson, 1998) and Statistica 7.1 (StatSoft, 2005). Genetic differentiation was examined by applying the Mantel test, with 10.000 permutations for the Jaccard similarity matrix using the Mantel-Struct 1.0 program (Miller, 1999). Analysis of molecular variance (AMOVA), expected heterozygosity, the value of genetic differentiation (F_{ST}) and the estimated number of migrants per generation were obtained using the program Arlequin 3.5.1.2 (Excoffier & Lischer, 2010). Mantel test was applied to determine the relationship between genetic differentiation (F_{ST}) and geographic distance. Mantel test was applied to determine the relationship between genetic differentiation (F_{ST}) and geographic distance. Geographic distance was estimated following the main river channel.

The probability of a given number of stocks based on a Bayesian approach was performed using the program STRUCTURE version 2.3.3 (Pritchard *et al.*, 2000). The number of presumed populations (K) was set from 1 to 5. Analyses had a burn-in and Monte Carlo Markov Chain (MCMC) set to 50,000 and 100,000 respectively, and a model without admixture and allele frequencies was used. Number of populations was defined based on the value of delta k, using the program STRUCTURE HARVESTER (Earl & vonHoldt, 2012).

Results

The results obtained for the barcode DNA were based on partial sequences of 670 bp of mitochondrial gene COI of 21 samples, five specimens of Araguaiana-MT, five specimens of Novo Santo Antônio-MT, four specimens of São Félix do Araguaia-MT and three specimens of Itupiranga-PA. The dendrogram of similarity showed only a group and was not a verified difference greater than 2% between the samples of the different locations, nor with the sequences deposited in GenBank, confirming that the fish studied all belonged to the species *A. gigas*. A comparison of the data in BOLD Systems platform showed a 99.7% similarity percentage with *A. gigas*.

After PCR amplification of the five ISSR primers, 168 loci (different bands sizes) were obtained for specimens collected in Araguaia-Tocantins basin. In analyzing the 146 specimens as a single population, 165 (98.21%) of the 168 loci were polymorphic. The total number of loci amplified by each primer was: 31, 28, 46, 35 and 28 for primers $(GGAC)_4$, $(GGAC)_3A$, $(GGAC)_3T$, $(GGAC)_3C$ e $(AACC)_4$, respectively.

The number of polymorphic loci, the amount of exclusive loci, the values of expected heterozygosity and the molecular diversity index of the intrapopulation haplotype, based on average gene diversity over all haplotype loci for the four samples, are reported in Table 1. Among the four populations studied, the Itupiranga-PA has the smallest sample (14 specimens), but maintains the highest level of variability (polymorphism = 56.5%, expected heterozygosity = 0.190, average gene diversity of all loci = 0.197). On the other hand, populations of Araguaiana-MT and Novo Santo Antônio-MT are the most homogeneous (Table 1). The intrapopulational dissimilarity generated by the Jaccard index was almost the same for the populations of Araguaiana-MT (0.255) and Novo Santo Antônio-MT (0.261), and higher for the populations of São Félix do Araguaia-MT (0.338) and Itupiranga-PA (0.351) (Table 1).

The results of the molecular analysis of variance (AMOVA) showed 52.63% interpopulation variation and 47.37% intrapopulation variation. The interpopulation genetic differentiation index (F_{ST}) was 0.52634 (p= 0.00000), which is considered very high for natural populations ($F_{ST} \geq 0.25$). The smallest genetic distance observed between populations, based on the values of F_{ST}, was between Araguaiana-MT and São Félix do Araguaia-MT, at 0.28639 (Table 2). The estimated number of migrants (Nm) between populations ranged from 0.18 to 1.2, with the lowest value between populations of Novo Santo Antônio and Itupiranga, and the higher of Araguaiana *versus* São Félix do Araguaia (Table 2).

The scatter plot constructed with the two largest eigenvectors (0.186 and 0.127 of variation, respectively), also obtained with the dissimilarity index of Jaccard, separated the populations into three main groups (Fig. 2).

Table 1. Comparison of key parameter indicators genetic diversity within each population *Arapaima gigas*. (Np) Number of polymorphic loci; Number of exclusive loci, (He) expected heterozygosity; average gene diversity of all haplotypic loci and intra-population genetic dissimilarity, using the Jaccard coefficient. (ARA) Araguaiana-MT, (NSA) Novo Santo Antônio-MT, (SFA) São Félix do Araguaia-MT and (ITU) Itupiranga-PA.

Populations	Np	Exclusive	He	Diversity	Dissimilarity
ARA	76 (45.2%)	0	0.08028 (±0.133)	0.0766 (±0.039)	0.255112
NSA	67 (39.9%)	8	0.08660 (±0.152)	0.0853 (±0.043)	0.260790
SFA	86 (51.2%)	5	0.10836 (±0.145)	0.106 (±0.053)	0.338271
ITU	95 (56.5%)	37	0.19052 (±0.199)	0.197 (±0.103)	0.350624

Table 2. Pairwise comparison of *Arapaima gigas* populations. (Dist) Approximate distance in km; (F_{ST}) inter-population genetic differentiation index; (p) significance level; (r) correlation of two matrices; (Z) robustness of the relationship with the original data and (Z10,000) after 10,000 permutations; genetic dissimilarity using the Jaccard coefficient; (Nm) estimated number of migrants per generation. (ARA) Araguaiana-MT, (NSA) Novo Santo Antônio-MT, (SFA) São Félix do Araguaia-MT and (ITU) Itupiranga-PA.

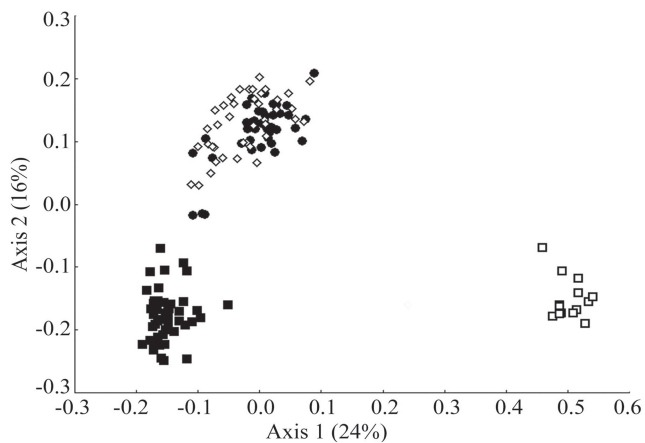| Pairwise comparisons	Dist	F_{ST}	p	r	Z	$Z_{10,000}$	Dissimilarity	N_m
ARA x NSA	676.8	0.483	0.001	0.774	881.03	703.64	0.435	0.535
ARA x SFA	787.8	0.286	0.001	0.424	724.81	640.93	0.383	1.246
ARA x ITU	173.5	0.702	0.001	0.922	418.53	259.30	0.664	0.212
NSA x SFA	111.0	0.424	0.001	0.708	882.82	723.65	0.467	0.679
NSA x ITU	105.7	0.727	0.001	0.960	456.88	274.73	0.725	0.188
SFA x ITU	946.7	0.664	0.001	0.888	412.06	280.64	0.701	0.252

Fig. 2. Scatter plot, ordered through the main coordinates (PCoA) based on the Jaccard dissimilarity index. Specimens of *Arapaima gigas* belonging to (black circle) Araguaiana, (black square) Novo Santo Antônio, (white lozenge) São Félix do Araguaia and (white square) Itupiranga.

A similar result was obtained with a Bayesian inference, which indicated a value of *k* equal to three populations. These populations are divided according to the sampling sites. The specimens of Novo Santo Antônio and Itupiranga are two separate populations and the specimens of São Félix do Araguaia and Araguaiana form a single population (Fig. 3).
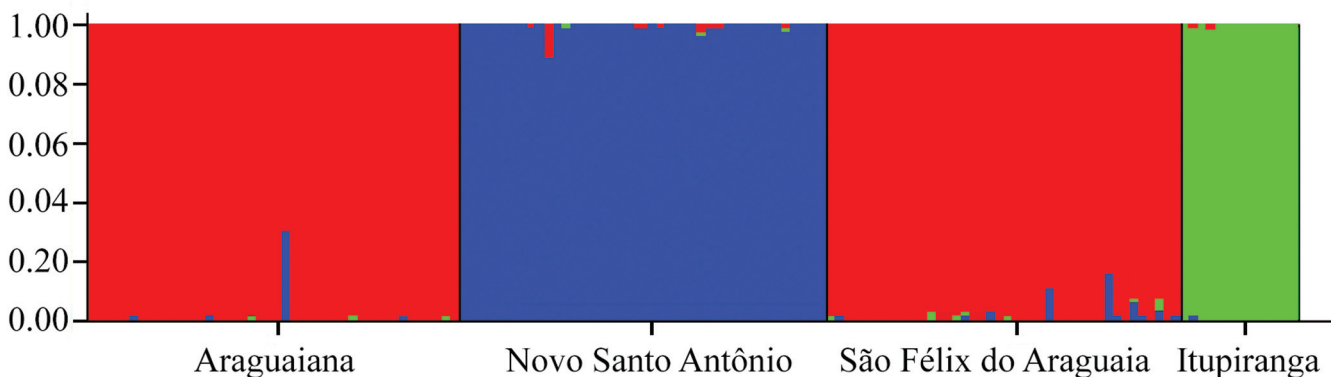
The dendrogram of Neighbor-joining, also built with the Jaccard similarity index, showed that only three specimens from the Araguaiana population clustered with specimens of São Félix do Araguaia (Fig. 4). The Mantel test did not show a significant correlation between geographic distance and the value of F_{ST} between populations (p = 0.15500, r = 0.646095), even when the analysis was performed excluding one stock (population) at a time. This was done to see if any of the locations led to a deviation from this correlation between genetic and geographic distances.

Discussion

The comparison of the mitochondrial DNA region (COI) showed that the analyzed samples of *Arapaima* belonged to a single species (*A. gigas*), since the genetic divergence values found for the COI region were much smaller than those accepted for the separation of species (approximately 2%) (Hebert *et al*., 2003). Comparing the sequence obtained in this study with those of Systems Bold data, there was 99.7% similarity to the sequence of the gene COI of *A. gigas*.

Because of the microsatellite regions having a higher mutation rate than did the COI gene, the molecular marker ISSR was efficient in detecting genetic variation among populations of *A. gigas*.

Fig. 3. Genetic contribution profile of each stock considering *k* = 3 for each subject analysis. Each individual is represented by a thin vertical line, which is partitioned into K segments that represent its estimated population group membership fractions. Black lines separate individuals from geographical site locations. The red, blue and green colors represent each stock (gene pool).

Although the percentage of polymorphic loci found in the present work for *A. gigas* was high (98.21%), the genetic variability values determined by the percentage of polymorphic loci for each population indicated a low intrapopulational genetic diversity, which can also be checked with the expected heterozygosity values (range 0.08-0.19), by the molecular diversity index of intrapopulation haplotype (range 0.076-0.197) and by intrapopulational dissimilarity values (between 0.255 and 0.350).

Generally, low expected heterozygosity values were found for populations of threatened species that underwent population bottlenecks. This indicates loss of genetic diversity, which is a consequence of inbreeding and reduced reproductive success, also resulting in the loss of the evolutionary potential of the species (Frankham *et al*., 2008). Other studies using ISSR markers in Neotropical fish stocks, such as species of the genus *Cichla* in the Amazon basin (Almeida-Ferreira *et al*., 2011) and *Hypostomus ancistroides* in the urban

streams of Londrina, Paraná, Brazil (Sofia *et al*., 2008), also detected low genetic diversity. In the case of *H. ancistroides*, this may be related to the sedentary habits of the species, which reduces possibilities of gene flow between populations of different locations. This fact may also lead to the low genetic variation found for *A. gigas*, reflecting an inbreeding process between specimens of each locality, since this species also has a sedentary habit and a preference for lentic environment (Castello & Stewart, 2010). In addition to these biological factors, human actions, such as overexploitation and environmental degradation, can lead to the formation of refuges, where small populations of species persist without exchanging genes due to the fact that the impacted areas around them act as barriers to migration (Solé-Cava, 2001).

The localities where the *A. gigas* samples were collected for this study were strongly impacted by fishing and the use of land for agriculture and livestock, even though these areas are considered as some of the last refuges for the Cerrado fauna (Latrubesse & Stevaux, 2006).

As gene flow reduces genetic differences between populations and increases the variation within populations (Allendorf & Luikart, 2007), the data presented corroborate the idea that the populations of *A. gigas* studied have restricted gene flow and are highly endogamic. Low gene flow has also been proposed for *A. gigas* populations of the Amazon region, and in this case, genetic differentiation found between "pirarucus" from different locations is considered a consequence of overexploitation (Hrbek *et al*., 2005).

In addition to overfishing, another factor that may contribute to the high inter-population genetic differentiation found in this study is the hydrological variability of the Araguaia-Tocantins region, which is characterized by two distinct seasons (wet and dry) that control the discharge of variations in the river. During periods of drought, the lakes belonging to the floodplain lose the connection to the river (Morais *et al*., 2005; Latrubesse & Stevaux, 2006), retaining various fish species. In the Amazon region where these fish are best studied, the cycles of floods and outflow are longer. The floodplain inundation areas remain at very high rates for an extended period, allowing for the movement of "pirarucus" for a longer period (called the process lateral migration) (Castello, 2008a), since the lakes remain connected for a relatively long time. Thus, Hrbek *et al.* (2005, 2007) and Araripe *et al.* (2013) results demonstrated that the levels of genetic diversity among stocks from different localities of the Amazon basin, although reflecting a consistent population structure, show that there is an effective gene flow. On the other hand, the results for the middle rio Araguaia region are interesting as they show little genetic variability and low gene flow, indicating that some of the stocks studied apparently represent consistent family groups.
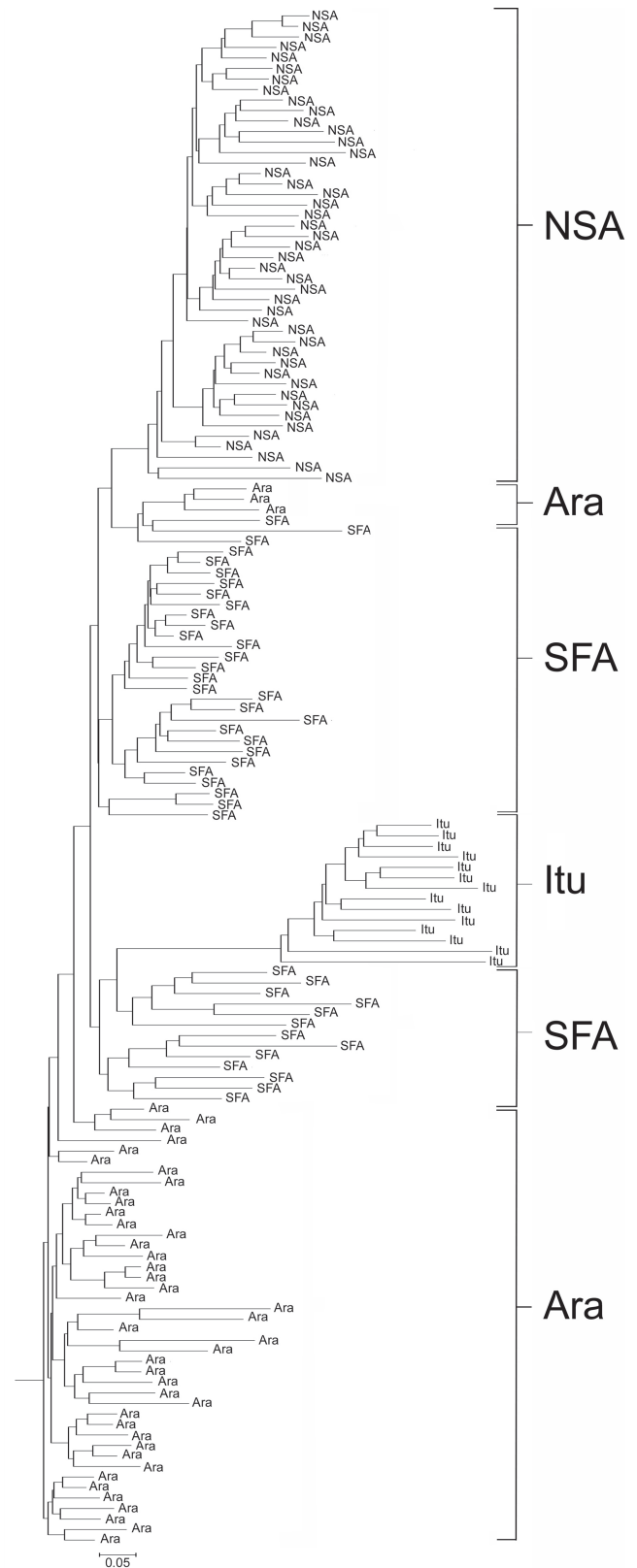

Fig. 4. Neighbor-joining dendrogram of *Arapaima gigas* generated with data from ISSR based on Jaccard similarity index with 10,000 bootstrap. (Ara) Araguaiana-MT, (NSA) Novo Santo Antônio-MT, (SFA) São Félix do Araguaia-MT and (Itu) Itupiranga-PA.

The values found for the molecular fixation index (F_{ST}) also support this hypothesis, since all *A. gigas* populations showed $F_{ST} > 0.25$. The fixation index is genetic differentiation between populations, ranging from 0-1. However Wright (1978) proposed that, for natural populations, values for $F_{ST} > 0.25$ are indicative of very high genetic differentiation, values for F_{ST} between 0.15 and 0.25 indicate large differentiation, between 0.05 and 0.15 indicate moderate differentiation and $F_{ST} < 0.05$ is indicative of little genetic differentiation. In addition, the Mantel test (r = 0.646; p > 0.05) showed that it is not only the geographical distance that is leading these populations to high genetic differentiation. Data suggest that such occurrences may be associated with bottlenecks events, leading to a marked reduction in the size of the population groups. There was also no significant association between geographic distance and genetic differentiation (determined with mitochondrial markers) among "pirarucus" in the Amazon basin (Hrbek *et al.*, 2005); however, in these cases, the analysis suggested the presence of an intense gene flow between the populations sampled. Moreover, the effects of the geographic distance that structured the "pirarucus" populations can be observed between location distances to at least 2,500 km (Hrbek *et al.*, 2007).

When considering the effect of distance on the genetic structure of a population, inferences are made about the limitations of gene flow between them, which may occur by migration or dispersal of juveniles. In this work, only a couple of populations showed a number of migrants value greater than one (Araguaiana and São Félix do Araguaia, Nm = 1.24), while all other pairs had values between 0.18 and 0.67. These low values indicate a very limited gene flow between populations, since the estimated number of potential migrants is an indirect way to estimate gene flow between populations (Neigel, 1997). When the number of migrants is less than or equal to one, the differentiation between populations can be explained by genetic drift in populations; however, if more than one individual migrates between populations, gene flow does not allow any allele to lock in the population as a result of genetic drift (Wright & Contents, 1931; Slatkin, 1985).

In the case of fragmented natural populations, smaller fragments suffer more severe genetic drift than do larger fragments. These populations lose genetic diversity and tend to achieve high levels of homozygosity faster than do large populations (Allendorf & Luikart, 2007; Frankham *et al.*, 2008), which may explain the low expected heterozygosity values found for the populations of the Araguaia-Tocantins basin.

Populations of Araguaiana and São Félix do Araguaia showed the lowest values for F_{ST}, dissimilarity genetics and the largest estimated number of migrants; furthermore, these populations appear overlapped on the scatter plot and in the histogram of the Bayesian analysis, suggesting that they share a common gene pool (Figs. 2-3). Despite the F_{ST} value being the lowest in the present study ($F_{ST} = 0.286$),

this value indicates high levels of genetic differentiation between populations. This can be seen in the dendrogram of Neighbor-joining (Fig. 4), where only three specimens of Araguaiana are grouped with specimens of São Félix do Araguaia. These two locations are separated by 787.8 km.

For *A. gigas* populations from the Amazon, high levels of genetic differentiation are found in locations that are separated by distances greater than 1,300 km. For populations that are separated on a mesoscale (100 km), there is low to moderate genetic differentiation (Araripe *et al.*, 2013). This relationship was not found in this study, because the localities of Novo Santo Antônio and São Félix do Araguaia are separated by 111 km and have a much higher genetic differentiation ($F_{ST} = 0.424$).

One possible explanation for this is the very location of the sampling points in the region of Novo Santo Antônio, in the rio das Mortes (Mato Grosso State), the main tributary of the left bank of the rio Araguaia. This locality lies in a region that represents the beginning of the Araguaia Pantanal, an indoor floodplain where the species has large areas to disperse and not suffer as much from severely dry periods that kill many families of *Arapaima* in marginal lakes in the main axis of the rio Araguaia. For this reason, it is expected that few specimens of *A. gigas* leave their reproduction and growth homes for this tributary, since apparently no severe natural environmental stresses, which induce such displacement, occurs with fish that are retained in the marginal lakes of the upper stretches of the rio Araguaia. Thus, new stocks originating from Novo Santo Antônio are residents, while the populations bordering the main river are more vulnerable to adversity produced by the regime of flood and severe droughts and, they risk larger downward migrations along the main channels of the rio Araguaia, as has been empirically shown from fishermen in the middle rio Araguaia region.

In addition to the above hypothesis, the genetic similarity between the Araguaiana and São Félix do Araguaia populations may also be associated with the transposition of specimens between these locations, a common practice in the region, which was confirmed by a single fish farmer who bought a lot of juvenile *Arapaima*, coming from Rondônia, Brazil. In this way, this hypothesis was confirmed by researchers working with the species in the rio Araguaia and participated in locus of transposition of six specimens among lakes of the middle rio Araguaia (P. C. Venere, *pers. obs.*). Since this species usually migrate, on average, no more than 10km, and has no fidelity to nesting place (Queiroz, 2000; Castello, 2008b). Fishermen and state agencies routinely perform the transposition of specimens of a lake to another during rescue operations, when specimens are removed from temporary lakes and transferred to perennial lakes. Marques (2003) indicated that the transposition of specimens was one of the factors contributing to the high similarity found between the two samples of *A. gigas* in the Araguaia-Tocantins basin during different years.

The results led to the conclusion that the populations of *A. gigas* in the Araguaia-Tocantins basin have low genetic diversity and are endogamic, having restricted gene flow, possibly leading them to inbreeding depression, a major threat to endangered species. The first evaluation of the genetic variability of *A. gigas* in the Araguaia-Tocantins basin presented in this paper indicates low levels of genetic variability, suggesting periodic monitoring to check eventual reducing levels of population genetic variability and the establishment of management plan. The use of ISSR markers can be a good strategy for the preliminary analysis of genetic diversity in *A. gigas* conservation programs since it is a low cost technique that has proved very effective for evaluating the genetic diversity of this species. It is emphasized that the data also suggest that there is an urgent need for studies that explain the population dynamics of this species in the Araguaia-Tocantins basin.

Acknowledgements

References

Aljanabi, S. M. & I. Martinez. 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acids Research, 25: 4692-4693.

Allan, J. D., R. Abell, Z. Hogan, C. Revenga, B. W. Taylor, R. L. Welcomme & K. Winemiller. 2005. Overfishing of Inland Waters. BioScience, 55: 1041-1051.

Allan, J. D. & A. S. Flecker. 1993. Biodiversity conservation in running waters: identifying the major factors that threaten the destruction of riverine species and ecosystems. BioScience, 43: 32-43.

Allendorf, F. W. & G. Luikart. 2007. Conservation an the genetics of populations. Malden, MA, Blackwell Publishing. 642p.

Almeida-Ferreira, G. C., A. V. Oliveira, A. J. Prioli & S. M. A. P. Prioli. 2011. Spar genetic analysis of two invasive species of *Cichla* (Tucunaré) (Perciformes: Cichlidae) in the Paraná river basin. Acta Scientiarum. Biological Sciences, 33: 79-85.

Antunes, R. S. P., V. N. Gomes, S. M. A. P. Prioli, R. A. Prioli, H. F. Júlio Jr., L. M. Prioli, C. S. Agostinho & A. J. Prioli. 2010. Molecular characterization and phylogenetic relationships among species of the genus *Brycon* (Characiformes: Characidae) from four hydrographic basins in Brazil. Genetics and Molecular Research, 9: 674-684.

Araripe, J., P. S. Rêgo, H. Queiroz, I. Sampaio & H. Schneider. 2013. Dispersal capacity and genetic structure of *Arapaima gigas* on different geographic scales using microsatellite markers. PloS ONE, 8: e54470 (p. 1-7).

Askari, N., M. M. Abadi & A. Baghizadeh. 2011. ISSR markers for assessing DNA polymorphism and genetic characterization of cattle, goat and sheep populations. Iranian Journal of Biotechnology, 9: 222-229.

Avise, J. C. 1996. Introduction: the scope of conservation genetics. Pp. 1-9. In: Avise, J. C. & J. L. Hamrick (Eds.). Conservation genetics: case histories from nature. New York, Chapman & Hall.

Barletta, M., A. J. Jaureguizar, C. Baigun, N. F. Fontoura, A. A. Agostinho, V. M. F. Almeida-Val, A. L. Val, R. A. Torres, L. F. Jimenes-Segura, T. Giarrizzo, N. N. Fabré, V. S. Batista, C. Lasso, D. C. Taphorn, M. F. Costa, P. T. Chaves, J. P. Vieira & M. F. M. Corrêa. 2010. Fish and aquatic habitat conservation in South America: a continental overview with emphasis on neotropical systems. Journal of Fish Biology, 76: 2118-2176.

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. Júlio Júnior & L. M. Prioli. 2009. Genetic divergence between *Pseudoplatystoma corruscans* and *Pseudoplatystoma reticulatum* (Siluriformes: Pimelodidae) in the Paraná River Basin. Brazilian Journal of Biology, 69(2, Suppl.): 681-689.

Bornet, B. & M. Branchard. 2001. Nonanchored Inter Simple Sequence Repeat (ISSR) markers: reproducible and specific tools for genome fingerprinting. Plant Molecular Biology Reporter, 19: 209-215.

Bornet, B. & M. Branchard. 2004. Use of ISSR fingerprints to detect microsatellites and genetic diversity in several related *Brassica taxa* and *Arabidopsis thaliana*. Hereditas, 140: 245-248.

Castello, L. 2008a. Lateral migration of *Arapaima gigas* in floodplains of the Amazon. Ecology of Freshwater Fish, 17: 38-46.

Castello, L. 2008b. Nesting habitat of *Arapaima gigas* (Schinz) in Amazonian floodplains. Journal of Fish Biology, 72: 1520-1528.

Castello, L., C. C. Arantes, D. G. McGrath, D. J. Stewart & F. S. Sousa. 2014. Understanding fishing-induced extinctions in the Amazon. Aquatic Conservation: Marine and Freshwater Ecosystems. DOI: 10.1002/aqc.2491.

Castello, L. & D. J. Stewart. 2010. Assessing CITES non-detriment findings procedures for Arapaima in Brazil. Journal of Applied Ichthyology, 26: 49-56.

Castello, L., D. J. Stewart & C. C. Arantes. 2011. Modeling population dynamics and conservation of arapaima in the Amazon. Reviews in Fish Biology and Fisheries, 21: 623-640.

Domingos, T. J., L. N. Moraes, R. M. Moresco, V. P. Margarido & P. C. Venere. 2014. Genetic and morphological diversity of *Moenkhausia oligolepis* (Characiformes: Characidae) populations in the tributaries of the Araguaia River, Brazil: implications for taxonomy and conservation. Genetics and Molecular Research, 13: 7979-7991.

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.

Ehsanpour, A. A., M. Tavassoli & L. Arab. 2008. Sex determination of *Pistacia vera* L. using ISSR markers. Malaysian Applied Biology, 37: 25-28.

Excoffier, L. & H. E. L. Lischer. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources, 10: 564-567.

Farias, I. P., T. Hrbek, H. Brinkmann, I. Sampaio & A. Meyer. 2003. Characterization and isolation of DNA microsatellite primers for *Arapaima gigas*, an economically important but severely over-exploited fish species of the Amazon basin. Molecular Ecology Notes, 1: 128-130.

Fernandes-Matioli, F. M. C., S. R. Matioli & L. F. Almeida-Toledo. 2000. Species diversity and geographic distribution of *Gymnotus* (Pisces: Gymnotiformes) by nuclear (GGAC)$_n$ microsatellite analysis. Genetics and Molecular Biology, 23: 803-807.

Fisher, R. A. 1930. The genetical theory of natural selection. Oxford, Clarendon Press. 272p.

Frankham, R., J. D. Ballou & D. A. Briscoe. 2008. Fundamentos de genética da conservação. Ribeirão Preto, Sociedade Brasileira de Genética. 262p.

Hajibabaei, M., G. A. C. Singer, E. L. Clare & P. D. N. Hebert. 2007. Design and applicability of DNA arrays and DNA barcodes in biodiversity monitoring. BMC Biology, 5: 24 (p. 1-7).

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

Hamoy, I. G., E. J. M. Santos & S. E. B. Santos. 2008. Rapid and inexpensive analysis of genetic variability in *Arapaima gigas* by PCR multiplex panel of eight microsatellites. Genetics and Molecular Research, 7: 29-32.

Hampl, V., A. Pavlícek & J. Flegr. 2001. Construction and bootstrap analysis of DNA fingerprinting-based phylogenetic trees with a freeware program FreeTree: application to trichomonad parasites. International Journal of Systematic and Evolutionary Microbiology, 51: 731-735.

Hebert, P. D. N., A. Cywinska, S. L. Ball & J. R. deWaard. 2003. Biological identifications through DNA barcodes. Proceedings of the Royal Society of London B, Biological Sciences, 270: 313-321.

Hebert, P. D. N. & T. R. Gregory. 2005. The promise of DNA barcoding for taxonomy. Systematic Biology, 54: 852-859.

Hrbek, T., M. Crossa & I. P. Farias. 2007. Conservation strategies for *Arapaima gigas* (Schinz, 1822) and the Amazonian várzea ecosystem. Brazilian Journal of Biology, 67(4, Suppl.): 909-917.

Hrbek, T. & I. P. Farias. 2008. The complete mitochondrial genome of the pirarucu. Genetics and Molecular Biology, 31(1, Suppl.): 293-302.

Hrbek, T., I. P. Farias, M. Crossa, I. Sampaio, J. I. R. Porto & A. Meyer. 2005. Population genetic analysis of *Arapaima gigas*, one of the largest freshwater fishes of the Amazon basin: implications for its conservation. Animal Conservation, 8: 297-308.

Hughes, A. R., B. D. Inouye, M. T. J. Johnson, N. Underwood & M.Vellend. 2008. Ecological consequences of genetic diversity. Ecology Letters, 11: 609-623.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution, 16: 111-120.

Klug, W. S., M. R.Cummings, C. A. Spencer & M. A. Palladino. 2012. Concepts of genetics. 10th ed. San Francisco, Pearson Education, 742p.

Latrubesse, E. M. & J. C. Stevaux. 2006. Características físico-bióticas e problemas ambientais associados à planície aluvial do Rio Araguaia, Brasil Central. Revista UnG - Geociências, 5: 65-73.

Legendre, P. & M. J. Anderson. 1998. Program DistPCOA. Montreal, Département de Sciences Biologiques, 10p. Available from: http://adn.biol.umontreal.ca/numericalecology/old/distpcoa-doc.html.

Lopes, T. S., R. P. Ribeiro, N. M. L. Barrero, R. N. Sirol, J. A. Povh, P. C. Gomes & L. Vargas. 2008. Caracterização genética de estoques de curimba (*Prochilodus lineatus*) utilizados em programas de repovoamento. Revista Brasileira de Saúde e Produção Animal, 9: 652-661.

Luque, C., L. Legal, H. Staudter, C. Gers & M. Wink. 2002. ISSR (Inter Simple Sequence Repeats) as genetic markers in *Noctuids* (Lepidoptera). Hereditas, 136: 251-253.

Maltagliati, F., T. Lai, M. Casu, S. Valdesalici & A. Castelli. 2006. Identification of endangered Mediterranean cyprinodontiform fish by means of DNA inter-simple sequence repeats (ISSRs). Biochemical Systematics and Ecology, 34: 626-634.

Marques, D. K. S. 2003. Caracterização genética do pirarucu *Arapaima gigas* (Cuvier) (Teleostei, Osteoglossidae) da bacia Tocantins-Araguaia, Estado do Mato Grosso. Unpublished Ph. D. Thesis, Universidade Federal de São Carlos, São Carlos, 66p.

Marques, D. K., P. C. Venere & P. M. Galetti Junior. 2006. Chromosomal characterization of the bonytongue *Arapaima gigas* (Osteoglossiformes: Arapaimidae). Neotropical Ichthyology, 4: 215-218.

Miller, M. P. 1999. MANTEL-STRUCT: a program for the detection of population structure via mantel tests. Journal of Heredity, 90: 258-259.

Morais, R. P., L. G. Oliveira, E. M. Latrubesse & R. C. D. Pinheiro. 2005. Morfometria de sistemas lacustres da planície aluvial do médio rio Araguaia. Acta Scientiarum, Biological Sciences, 27: 203-213.

Moresco, R. M., T. C. Maniglia, C. de Oliveira & V. P. Margarido. 2013. The pioneering use of ISSR (Inter Simple Sequence Repeat) in Neotropical anurans: preliminary assessment of genetic diversity in populations of *Physalaemus cuvieri* (Amphibia, Leiuperidae). Biological Research, 46: 53-57.

Neigel, J. E. 1997. A comparison of alternative strategies for estimating gene flow from genetic markers. Annual Review of Ecology, Evolution, and Systematics, 28: 105-128.

Nelson, J. S. 2006. Fishes of the world. 4th ed. Hooboken, NJ, J. Wiley & Sons. 601p.

Paiva, S. R., J. A. Dergam & F. Machado. 2006. Determining management units in southeastern Brazil: the case of *Astyanax bimaculatus* (Linnaeus, 1758) (Teleostei: Ostariophysi: Characidae). Hydrobiologia, 560: 393-404.

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

Queiroz, H. L. 2000. Natural history and conservation of pirarucu, *Arapaima gigas*, at the Amazonian Várzea: red giants in muddy waters. Unpublished Ph.D. Thesis University of St. Andrews, Scotland, 226p.

Reddy, M. P., N. Sarla & E. A. Siddiq. 2002. Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. Euphytica, 128: 9-17.

Sanger, F., S. Nicklen & A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences of United states of America, 74: 5463-5467.

Slatkin, M. 1985. Gene flow in natural populations. Annual Review of Ecology and Systematics, 16: 393-430.

Sofia, S. H., B. A. Galindo, F. M. Paula, L. M. K. Sodré & C. B. R. Martinez. 2008. Genetic diversity of *Hypostomus ancistroides* (Teleostei, Loricariidae) from an urban stream. Genetics and Molecular Biology, 31(1, Suppl.): 317-323.

Sofia, S. H., C. R. M. Silva, B. A Galindo, F. S. Almeida, L. M. K. Sodré & C. B. R. Martinez. 2006. Population genetic structure of *Astyanax scabripinnis* (Teleostei, Characidae) from an urban stream. Hydrobiologia, 553: 245-254.

Solé-Cava, A. M. 2001. Biodiversidade molecular e genética da conservação. Pp. 172-192. In: S. R. Matioli (Ed.). Biologia molecular e evolução. Ribeirão Preto, Holos Editora.

StatSoft, Inc. 2005. STATISTICA (data analysis software system), version 7.1. Available from http://www.statsoft.com (25 July 2014).

Stewart, D. J. 2013a. A New Species of *Arapaima* (Osteoglossomorpha: Osteoglossidae) from the Solimões River, Amazonas State, Brazil. Copeia, 3: 470-476.

Stewart, D. J. 2013b. Re-description of *Arapaima agassizii* (Valenciennes), a rare fish from Brazil (Osteoglossomorpha: Osteoglossidae). Copeia, 1: 38-51.

Tamura, K., G. Stecher, D. Peterson, A. Filipski & S. Kumar. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution, 30: 2725-2729.

Ward, R. D., T. S. Zemlak, B. H. Innes, P. R. Last & P. D. N. Hebert. 2005. DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 360: 1847-1857.

Woodruff, D. S. 2001. Declines of biomes and biotas and the future of evolution. PNAS, 98: 5471-5476.

Wright, S. 1978. Evolution and the genetics of populations. Chicago, University of Chicago Press, v. 4.Variability within and among natural populations. 590p.

Wright, S. 1931. Evolution in mendelian populations. Genetics, 16: 97-159.

Yeh, F., R. Yang & T. Boyle. 1999. Popgene version 1.32: Microsoft Windows-based freeware for population genetic analysis. Edmonton. University of Alberta. Available from https://www.ualberta.ca/~fyeh/popgene_download.html (12 August 2014).