**Looking for sweet home in captive waters: population genetic structure of the introduced *Arapaima* in Bolivia**

Fernando M. CARVAJAL-VALLEJOS1,2\*, Joachim CAROSFELD3, Paul VAN DAMME1, Selva MONTELLANO1, John NELSON4

1 FAUNAGUA, Cochabamba, Plurinational State of Bolivia.

2 Unidad de Limnología y Recursos Acuáticos (ULRA), Facultad de Ciencias y Tecnología (FCyT), Universidad Mayor de San Simón (UMSS), calle Sucre y Parque La Torre s/n, Zona Las Cuadras, Cochabamba, Plurinational State of Bolivia.

3 World Fisheries Trust (WFT), 831 Devonshire Road, #16 Victoria, BC V9A 4T5, Canada.

4 University of Victoria (UVic), Victoria, BC, Canada.

\*Author for correspondence: fmcvalle@yahoo.com

**Abstract**

Invasive aquatic species are increasing globally due to factors related to globalization and accelerated trade between regions. Introductions generally cause disturbance at the ecological level with subsequent local and/or regional socioeconomic changes. The paiche *Arapaima*, one of the largest fish in the Amazon, was introduced via Peru into Bolivia in the 1960s, and has generated significant changes in Amazonian fisheries and fish value chains. In recent years, it has been proposed that genus *Arapaima* is composed by different species distributed along the Amazon Basin. It is unknown if the individuals introduced in Bolivia corresponded to a single species, and if they belonged to a single population. The present study evaluated the genetic variability of the paiche in the Bolivian Amazon Basin (sub-basins of the Orthon, Madre de Dios and Beni rivers) to determine identity (NADH and CO1 - mtDNA) and population structure (microsatellites - nDNA). The mtDNA and nDNA results suggested that all analyzed individuals belong to the same entity, which is made up of three populations corresponding to the geographic sub-basins according to the microsatellite analysis. The genetic distance between populations was not significantly related to the genetic distance between collection sites. It is proposed that the set of individuals (founder population) of the species in Bolivia was composed of different populations, which crossed a bottleneck and dispersed in search of environmental conditions similar to those present in the habitats from which they were extracted in their natural area of distribution. Planning for the sustainable use of the species by fisheries should consider that there are different populations in the Bolivian sub-basins, and recruitment seems to depend on exchanges between nearby surrounding aquatic habitats rather than between sub-basins.

**Key words**: Invasive fish species, Osteoglossiformes, CO1, ATPase, microsatellites, Bolivian Amazon Basin, Madeira River basin.

**Resumen**

Las especies acuáticas invasoras están incrementando a nivel global debido a factores relacionados a la globalización, y el intercambio comercial acelerado entre regiones. Por lo general, las introducciones causan perturbaciones a nivel ecológico con subsecuentes cambios socioeconómicos locales y/o regionales. *Arapaima*, uno de los peces más grandes del Amazonas, fue introducido a través del Perú en Bolivia en los años sesenta, y ha generado cambios significativos en las pesquerías y cadenas de valor de pescado amazónico. En años recientes, se ha propuesto que *Arapaima* se compone por diferentes especies distribuidas en la Cuenca del Amazonas. Se desconoce si los individuos introducidos en Bolivia correspondían a una sola especie, y si ellos pertenecían a una sola población. El presente estudio evaluó la variabilidad genética de *Arapaima* en la Cuenca del Amazonas de Bolivia (subcuencas de los ríos Orthon, Madre de Dios y Beni) para conocer la identidad (NADH y CO1 - DNAmt) y la estructura poblacional (microsatélites - DNAn). Los resultados del DNAmt y DNAn sugieren que todos los individuos analizados pertenecen a una misma entidad, la cual se compone de tres poblaciones correspondientes a las subcuencas geográficas según el análisis de los microsatélites. La distancia genética entre poblaciones no se relacionó significativamente a la distancia genética entre sitios de colecta. Se propone que el conjunto de individuos (población fundadora) de la especie en Bolivia estuvo constituido por diferentes poblaciones, las cuales atravesaron un cuello de botella y se dispersaron en búsqueda de condiciones ambientales similares a las presentes en los hábitats de donde fueron extraídos en su área natural de distribución. La planificación del aprovechamiento sostenible de la especie por las pesquerías debe considerar que existen diferentes poblaciones en las subcuencas de Bolivia, y el reclutamiento parece depender de los intercambios entre hábitats cercanos circundantes más que entre subcuencas.

**Palabras clave**: Pez invasor, Osteoglossiformes, CO1, ATPase, microsatellites, Cuenca Amazónica Boliviana, Cuenca del río Madera.

**Introduction**

Human-mediated introductions of non-native invasive species experience population funding events, which following general theory establish only a fraction of the genetic variants occurred in the source population(s) and pass through generations of small effective population size (bottleneck) (Nei et al. 1975). The founder population of the invader could proceed from a simple stock (population) or from multiple stocks, and the mix of stocks would result in novel genetic combinations similar to observed in multiple introductions (Novak & Mack 2005). Association between novel combinations and genetic variation in invasions is poorly known yet, and this is a challenge for further studies in invasion biology.

Most freshwater ecosystems have been affected to a greater or lesser extent by biological invasions or introductions, whose are pervasive and may cause food web disruptions, biodiversity loss, and economic harm at different scales (Parker et al. 1999, Thomaz et al. 2015). Aquatic invasions rates are going up over time due to globalization and accelerated trade among naturally differentiated regions, and concerns about their impacts come mainly from observed disturbances at ecological level with subsequent socio-economic harm (Thomaz et al. 2015). According Salmenkova´s (2008) review and Havel et al. (2015), the term introduction belongs to the wider concept of biological invasion, which covers introduction of living organisms in ecosystems located beyond their initial (generally natural) range. These notions are closely related to the terms transplantation (transportation to a new location), acclimatization (transportation of individuals to a new biotop to form self-sustaining populations or new populations maintained by humans new populations), and naturalization (the formation by the alien species of stable self-sustained populations in the new area).

In the 1960s, *Arapaima* *gigas* (paiche), one of the biggest fish species in the Amazon, was transplanted and acclimatized for incipient aquaculture in Puerto Maldonado (Madre de Dios River, Upper Madeira Basin, Peru) (Carvajal-Vallejos et al. 2011). The precarious confinement and seasonal flooding allowed this species to escape and spread downstream towards Bolivian rivers and oxbow lakes where this species is non-native (Carvajal-Vallejos et al. 2011). The species successfully established and currently is one of the most important fisheries resources (Carvajal-Vallejos et al. 2011, Miranda-Chumacero et al. 2012, Van Damme et al. 2015). *Arapaima* was severely depleted by fisheries in its natural area of distribution since several decades ago, and some genetic studies (mt- and nDNA) were conducted in the framework of conservation programs for the species in Brazil and Peru (Hrbek et al. 2005, 2007; Read 2009; Araripe et al. 2013; Vitorino et al. 2015). The general conclusions were that *A*. *gigas* is constituted by a single taxonomic entity, with a moderate spatial population structure concordant with the geographical distances (Western to Eastern). However, recent studies propose that different species co-exist in the Amazon basin and new species should be described (Stewart 2013 a, b).

Our hypothesis is that after the invasion an unknown number of individuals with a low genetic variability established in Bolivia, resembling the event of a bottleneck. To test this hypothesis, we determined the genetic identity based on mt- and nDNA , and evaluated whether the genetic variability and structure of the species is concordant with a single event of introduction and source stock. Based on the interpretation of the local and regional differentiation caused by the genetic exchange that currently exists at different geographic scales, we provide recommendations for the sustainable use of the species in Bolivia.

**Methods**

Study area

The study area is situated in the Upper Madera (Madeira in Brazil) River basin in Bolivia, comprising the Abuná, Orthon, Madre de Dios, Beni, Mamoré and Iténez (or Guaporé) River basins. It has a tropical climate, with the rainy season between October to March. The rivers, characterized by seasonal flood pulses (Molina Carpio 2011), and their floodplains are species-rich (i.e. Maldonado & Carvajal 2005). Most of the rivers are meander-like, characterized by unstable bed, structural heterogeneity and high geomorphological complexity (Crespo & Van Damme 2011).

Samples were obtained from the Orthon, Madre de Dios and Beni rivers sub-basins. The Orthon River is formed by the confluence of a white (high water season) -black (low water season) mixed water river (Manuripi) (Navarro & Maldonado 2002) and a white-water river (Tahuamanu), which originate in low lands of Peru and enter Bolivia in the extreme North-West. The Manuripi River has alluvial plains with seasonally flooded surfaces of fluvial terraces and internal deltas, associated with the river bed (Navarro & Maldonado 2002). The Orthon River has a denudation rate of 0.033 ± 0.004 mm/yr (Wittmann et al. 2011), and discharges its waters to the Beni River after the latter's confluence with the Madre de Dios River, approximately 33 km downstream from the city of Riberalta. The Madre de Dios River is a river-floodplain system characterized by its white-waters originating in the Peruvian Andes (> 6 000 m of elevation), and has a denudation rate of 0.28 ± 0.13 mm/yr (Wittmann et al. 2011). The Beni River flows from the Andes in Bolivia and runs through territories of La Paz and Cochabamba departments in its upper portion. Since its departure from the Andes, at the height of the city of Rurrenabaque, the fluvial landscape has the characteristics of a river-floodplain system in its middle portion of the basin, and is characterized by being turbulent and carrying a large amount of suspended sediments (white-water) (Navarro & Maldonado 2002). Its denudation rate is 0.45 ± 0.07 mm/yr (Wittmann et al. 2011).



Figure 1. Representation of paiche *Arapaima* samples collection sites in northern Bolivia between the years 2011-2013. Number in parentheses represent the sample size in the site.

Sample collection

The collection points of paiche were the vicinity of the Barraca Hiroshima (latitude -11.40083, longitude -67.8525, 175 m), in the Manuripi-Orthon river basin (16 individuals - ind); El Mentiroso lagoon (latitude -11.096981, longitude -66.572905,130 m), in the Madre de Dios River basin (55 ind); Las Pozas (latitude -11.38857, longitude -66.36494, 130 m) (37 ind) and the Portachuelo lagoon (latitude -11.264171, longitude -66.288122, 125 m) (38 ind), in the Beni River basin. A total of 119 individuals from the three mentioned basins were analyzed (Figure 1). The samples were obtained during the accompaniment of commercial fishermen residing in communities with the right of access to the resources and bodies of water described. In this way, we worked with fishermen from Puerto Rico - urban-peasant community that operates in the Manuripi River - Manuripi Amazonian Wildlife National Reserve; Trinidacito - Tacana indigenous community of the Territorio Indígena Originario Campesino (TIOC) of the Territorio Indígena Multiétnico II (TIM II) that operates in El Mentiroso lagoon; Portachuelo Bajo - the Esse-Ejja indigenous community of the TIOC TIM II that operates in the Portachuelo lagoon, and Flor de Octubre - Tacana indigenous community of the TIOC TIM II that operates in Las Pozas, among other bodies of water. A photographic record of the body was made from each captured specimen and a sample of muscle tissue was taken from the dorsum that was preserved in a 15 ml tube with 95% ethyl alcohol. All samples were deposited in the fish collection of the Unidad de Limnología y Recursos Acuáticos (ULRA), of the University Mayor de San Simón (UMSS), Cochabamba, Bolivia.

DNA extraction

DNA in muscle tissue was extracted by incubating a small piece of sample in 100 μl/10% Chelex-100 (Bio-Rad), 0.2% SDS and 0.4 mg Proteinase K/ml for two hours at 55° C, then at 95° C for 15 min.

Mitochondrial DNA (mtDNA) amplification

Two mtDNA loci were amplified. The fragment of the genetic bar code (barcoding fragment) that corresponds to the cytochrome oxidase subunit 1 (CO1) to identify the taxonomic genetic identity of the paiche in relation to sequences available in the literature and databases (GenBank), and the fragment encoding ATPase for population analysis and comparisons with sequences previously obtained by Hrbek et al. (2005) for specimens caught in Brazil and Peru. The CO1 was amplified and sequenced with the FishF1 and FishR1 primers (Ward et al. 2005). ATPase was amplified with the L8106 and H9264 primers, and sequenced with L8537 and H8516 (Hrbek et al. 2005). The amplified ATPase fragment included the 3' cytochrome oxidase subunit 2, lysine tRNA, ATPase subunit 6, ATPase subunit 8, and the 5' of cytochrome oxidase subunit 3. Only five individuals from each locality collected were sequenced for CO1, while all those collected for the study of population genetics using the variation of ATPase. The amplification of the fragments by polymerase chain reaction (PCR) was carried out using the whole extracted genome and negative controls were included in each reaction.

The cocktail for the amplification of CO1 totaled 25 μl composed of 18.75 μl of ultrapure water, 2.25 μl of 10X buffer for PCR, 1.25 μl of MgCl2 (50 mM), 0.25 μl of each primer (0.01 mM), 0.125 of each dNTP (0.05 mM), 0.625 U of Tap polymerase, and 1 μl of DNA extraction. The temperature profile of the reaction consisted of an initial step at 95° C for 2 min, followed by 35 cycles composed of 0.5 min at 94° C, 0.5 min at 54° C, one minute at 72° C, and a final cycle at 72° C for 10 min before being kept at 4° C. The cocktail for the amplification of ATPase had a total volume of 25 µl composed of the same proportions of reagents used in CO1. The temperature profile for the reaction began with denaturation at 95° C for 2 min, followed by 30 cycles consisting of the sequence 94° C for 0.6 min, 50° C for 0.6 min, 90° C for 1.5 min, and one final cycle 72° C for 10 min before staying at 4° C.

Amplifications were carried out in a Mastercycler Eppendorf thermal cycler (Brinkmann Instrumments, Inc.). The PCR products were visualized on 1% agarose gels and the most intense products were selected for sequencing. Products were labeled using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Inc.) and bi-directionally sequenced using an ABI 3730 capillary sequencer following the manufacturer's instructions.

Nuclear DNA (nDNA) amplification

A total of 124 individuals were studied in the polymorphic loci (variables) of microsatellites AgCTm7, AgCAm2, AgCAm15 identified by Farias et al. (2003). The cocktail reaction was carried out in 10 μl volume containing 5.5 μl of ultrapure water, 1.0 μl of 10X buffer (100 mM Tris-HCl, 500 mM KCl, 15 nM MgCl2), 2.0 μl of each primer (0.2 μM each), 0.8 μl of a mixture of dNTPs (200 μM each dNTP), 0.2 U of Tap polymerase (AccuPrime, Invitrogen), and 1.0 μl of total DNA extraction. The temperature profile for the reaction began with a denaturation at 92° C for two minutes, followed by 35 consistent cycles of 94° C for 10 s, a specific temperature for each chosen primer (see Farias et al. 2003) by 10 s, 72° C for 30 s, and a final extension at 72° C for 30 min before standing at 4° C. The reactions were carried out in the same thermal cycler used for the mitochondrial fragments. The PCR products were visualized in 1% agarose gels and those more intense were selected to be subjected to 7% polyacrylamide gel electrophoresis of 1 mm thickness. The non-denaturing polyacrylamide was composed of acrylamide and bisacrylamide in a 19:1 ratio. The gels were prepared with 2X TAE buffer (Sambrook et al. 1989), the same one that was used to immerse the chamber with the gels and close the electric field. The products were mixed with 3 μl of 10X running dye (50 mM EDTA [pH 8.0], 30% glycerol, 0.25% bromophenol blue) to conglomerate the products, and 10 μl of this solution was loaded into the polyacrylamide cells. After the fragments were subjected to the electric field and migrated overnight (12-14 h), they were stained in an aqueous solution of Ethidium Bromide (0.5 µg / ml) and the bands were visualized on an ultraviolet ray transluminator. The digital images of the gels were obtained with an Eagle-Eye system (Stratagene Corp., San Diego, California), and the gels were manually read and evaluated using the GelAnalyzer program (2010, www.gelanalyzer.com). Each electrophoretic run considered one column for the blank reactions and two for the molecular weight marker (at the ends) that varied every 100 base pairs (Promega) in the reading zone.

CO1 analysis

The CO1 sequences were aligned and edited with the BioEdit program (Hall 1999) with the implemented tool Clustal W. The search and comparison with other sequences available in literature was carried out with the BLAST (Basic Local Alignment Search Tool) program in the base GenBank online data set (National Center for Biotechnology Information - http://www.ncbi.nlm.nih.gov/BLAST/Blast.cgi?CMD=Web&PAGETYPE=BLASTHo-me). The publications available for comparisons were Hrbek et al. (2007) and Ardura et al. (2010). The haplotype and nucleotide diversity was calculated with the DnaSP version 5 program (Librado & Roszas2009).

ATPase analysis

The ATPase sequences were edited and aligned following the same procedure as CO1. The haplotypes identified were compared with the haplotypes identified by Hrbek et al. (2005), who analyzed 110 samples from Brazil (different points along the main axis of the Amazon), 16 from Peru (Iquitos), and 13 from Bolivia (Madre de Dios River basin without specific localities)

Microsatellites (nDNA) analysis

The genetic variation was described by the classical parameters of population genetics such as allelic frequency per locus, level of heterozygosity observed (Ho) and expected without bias (Hn) (Nei 1978). The inbreeding index (F*IS*) (Wright 1978), estimated by ƒ of Weir & Cokerham (1984), was calculated to assess deviation to panmyxia. The significance of the deviations was evaluated with 1 000 matrices generated at random by permutations between alleles. The general differentiation between geographic samples was evaluated considering all the loci together, using comparisons between the values of the fixation index F*ST*, estimated by the value of theta - *θ* (Weir & Cokerham 1984). A genetic population is understood as a panmictic unit (the value of F*IS* not significantly different from zero), differentiated from the others by significant values of F*ST*. All the above mentioned procedure and indices were generated with the Genetix 4.05 program (Belkhir et al. 2004).

To find out if the genetic differentiation between the geographic samples could be related to their spatial position, the relationship between the genetic distance obtained (F*ST* / 1-F*ST*, Rousset 1997) *vs*. the linear distance (km) between the studied localities and the distance along the (sinuous) course of the rivers. The degree of relationship was measured with the Pearson correlation coefficient (*r*) implemented in Excel (Microsoft Office Professional Plus 2016), and the geographical distances were obtained using Google Earth.

**Results**

CO1 and ATPase

The obtained CO1 fragment consisted of 651 base pairs (bp). Haplotypes (different sequences), differentiated by a polymorphic site, were identified in the samples from the three studied locations. The haplotype diversity of the 12 samples (Hd) was 0.303 ± 0.147, and the nucleotide diversity (Pi) was 0.0005 ± 0.0002. These sequences varied at three sites when compared to a single sequence identified in two individuals by Ardura et al. (2010) and two individuals by Hrbek et al. (2008) from the Brazilian Amazon Basin.

A single 1 119 bp length haplotype was identified in 90 individuals from Bolivia screened for the ATPase locus. The sequence identified in Bolivia was different from the 13 haplotypes identified for 126 individuals from the main axis of the Amazon from Brazil to Peru, and Bolivia (Hrbek et al. 2005).

Microsatellites

The highest allelic variation observed was at the AgCAm2 locus (4) and the lowest at AgCTm7 (2). The lowest allelic diversity was observed in the Orthon River basin, which showed the presence of six alleles; none of them exclusive to the sampling location. The other two locations were more diverse (8 alleles), but presented exclusive alleles. In the Madre de Dios River basin an exclusive allele was identified at the AgCAm2 locus (310), and in the Beni River basin at the Ag-CAm15 locus (238) (Table 1, Figure 2).

Table 1. Genetic variability of the paiche *Arapaima* in three microsatellite loci for three localities of the Bolivian Amazon Basin. MD: Madre de Dios River basin; BE: Beni River basin; OR: Orthon River basin. N: number of individuals. Ho: observed heterozygosity; He: expected heterozygosity; SD: Standard deviation. FIS: Inbreeding index; ns: not significant (*p* > 0.05).

|  |  |  |  |
| --- | --- | --- | --- |
| **Alelos** | **MD** | **BE** | **OR** |
| N | 55 | 38 | 16 |
| **Locus AgCAm2** |
| 300 | 0.164 | 0.013 | 0.000 |
| 306 | 0.073 | 0.105 | 0.125 |
| 310 | 0.027 | 0.000 | 0.000 |
| 316 | 0.736 | 0.882 | 0.875 |
| Ho | 0.455 | 0.237 | 0.250 |
| Hn | 0.430 | 0.214 | 0.226 |
| FIS AgCAm2 | -0.061 ns | -0.106 ns | -0.111 ns |
| **Locus AgCTm7** |
| 290 | 0.473 | 0.176 | 0.344 |
| 300 | 0.527 | 0.824 | 0.656 |
| Ho | 0.582 | 0.297 | 0.313 |
| Hn | 0.503 | 0.294 | 0.466 |
| FIS AgCTm7 | -0.158 ns | -0.013 ns | 0.336 ns |
| **Locus AgCAm15** |
| 228 | 0.827 | 0.763 | 0.531 |
| 238 | 0.000 | 0.197 | 0.000 |
| 244 | 0.173 | 0.040 | 0.469 |
| Ho | 0.346 | 0.395 | 0.313 |
| Hn | 0.288 | 0.382 | 0.514 |
| FIS AgCAm15 | -0.200 ns | -0.034 ns | 0.400 ns |
| **Global** |
| FIS | -0.134 ns | -0.044 ns | 0.281 ns |
| Ho | 0.461 | 0.310 | 0.292 |
| DE (Ho) | 0.118 | 0.080 | 0.036 |
| Hn | 0.407 | 0.298 | 0.402 |
| DE (Hn) | 0.109 | 0.084 | 0.154 |

The panmixia deviation tests for the three loci together (F*IS* global = 0.025) and each of them (F*IS* locus AgCAm2: -0.046, AgCTm7: 0.030, AgCAm15: 0.079), considered as a single set (without taking into account the localities), showed values that do not deviate significantly from zero (*p* > 0.05) (Table 1). Therefore, the possibility that the three samples together are part of a single large panmictic unit is not rejected.

The inbreeding index (F*IS*) values for each geographic sample showed that panmixia cannot be rejected in each one of them at the multilocus and by locus level (Table 1).

The differentiation test between geographic samples (*θ*, Weir & Cockerham 1984) showed that there is a significant differentiation (*p* < 0.01) between the localities, denoting that there is an important variation between each one of them and they can be considered as different genetic populations (Table 2). The greatest proximity (least genetic distance, F*ST* / 1-F*ST* - Rousset 1997) was observed between the population of the Madre de Dios River and the population of the Orthon River, and the lowest between the latter and the population of the Beni River (Table 2)



Figure 2. Genetic variability of the paiche *Arapaima* in three microsatellite loci (AgCAm2, AgCTm7, AgCAm15) for three localities of the Bolivian Amazon Basin. MD: basin of the Madre de Dios River; BE: Beni River basin; OR: Orthon River basin.

Table 2. Genetic differentiation of paiche *Arapaima* samples from three localities in the Bolivian Amazon Basin. MD: Madre de Dios River basin; BE: Beni River basin; OR: Orthon River basin. F*ST*: Fixation index (*θ*, Weir & Cockerham 1984); D: Genetic distance defined as F*ST* / 1-F*ST* (Rousset 1997). \*\*: *p* <0.01 (highly significant).

|  |  |  |  |
| --- | --- | --- | --- |
| D\FST | **MD** | **BE** | **OR** |
| **MD** |  | 0.107\*\* | 0.080\*\* |
| **BE** | 0.120 |  | 0.127\*\* |
| **OR** | 0.087 | 0.146 |  |

Genetic distance between populations when compared to the geographical distance following the course of the rivers (r = 0.14) and linear (r = 0.16) between localities, not showed a significant relationship. This lack of relationship denotes that genetic differentiation does not depend on spatial distance but on other factors involved (e.g. ecological).

**Discussion**

CO1 information

The variation and divergence < 2% (3/651) of the CO1 identified for the haplotypes of the present study and one haplotype identified for the Central Amazon in Brazil (Hrbek et al. 2008; Burning et al. 2010) suggests that the genetic identity present in Bolivia corresponds to the same identity identified for Brazil by the aforementioned authors. The limit to differentiate species using the variation of CO1, which resembles a genetic barcode (barcoding), is approximately 2% (Pereira et al. 2013).

However, the precision of barcode locus is not absolute and some species recognized at the morphological level cannot be differentiated at this level. Although some examples are known where all the species considered presented a divergence and satisfactory assignment (e.g. Ward et al. 2005; Nwani et al. 2011), or new taxa were discovered or redefined (e.g. Cerutti-Pereyra et al. 2012; Amaral et al. 2013; Pereira et al. 2013; Castro et al. 2014), a general balance showed that 93-98% of the known fish species can be differentiated through this methodology (Ward et al. 2009), and in some cases the accuracy may be lower. In the Amazon basin there are some cases in fish in which the CO1 information is not enough to differentiate species (e.g. Toffoli et al. 2008; Garcia-Davila et al. 2013; Farias et al. 2018 – *Serraslamus rhombeus* group).

Similar diagnosis on a single entity in Bolivia can be inferred from the two additional markers that were used (ATPase and microsatellites), and the content of their information is commented below.

ATPase information

The uniformity found in all the sequences obtained for the three geographical samples from Bolivia in the present study suggests that a bottleneck occurred when the paiche reached the natural environment in the Madre de Dios River basin in Peru. Despite the fact that the ATPase sequences correspond to a coding locus, and therefore their variability is quite low (conserved locus), none of them were identical to the haplotypes identified by Hrbek et al. (2005), who found seven haplotypes in 17 individuals from the Peruvian portion (Iquitos) of the Amazon Basin. Surprisingly, the sequences were differentiated by a mutation that was not present in any of the sequences obtained by Hrbek et al. (2005). The total number of polymorphic sites in Hrbek's work et al. (2005) was 18, and they add up to 19 when the unique and different sequence identified for the samples of the present study is added. The average number of nucleotide differences between the sequences of Hrbek et al. (2005) and the present study was 4,692, which denotes a low variability in this locus as expected. The difference in the Bolivian sequences is due to a mutation of Thymine by Cytosine, at the level of the ..AAGCTTCTTTGATCAA .. (Bolivia) motif, which is ..AAGCTTTTTTGATCAA .. in all the sequences from Brazil, Peru and Bolivia found by Hrbek et al. (2005). The origin of this difference may be due to a rapid mutation in the locus, but it is unlikely due to the mutation rate of 1.3% per million years that this fragment has in fish (Bermingham et al. 1997). One possibility for the presence of a mutation only in the Bolivian samples from the three locations is that the sequencers used have made an erroneous reading. If this were the case for the sequencer used in the present study, the only sequence identified would be the same as the haplotype gb | AY081891.1 (Gen Bank) found by Hrbek et al. (2005), the only haplotype identified in Iquitos (two of 16 individuals) and Madre de Dios River (all 13 individual) following supplementary information at https://github.com/legalLab/publications/commit/53529136027435a0bcf4e2a9ea7831b2a594b682. Under this last possibility, haplotype distribution is coincident with a founder event of the invader in Madre de Rios River with a stock that included individuals from around Iquitos, which is concordant with the history of the introduction in Bolivia as reconstructed by Carvajal-Vallejos et al. (2012). In the light of this uncertainty whether the C or T base is correct, it is recommended to obtain sequences for a group of individuals already studied in another sequencer.

Microsatellite information

The analysis of the variation of the microsatellites showed that the analyzed samples are part of a large population existing in three of the major systems of the Bolivian Amazon invaded by the paiche. This result is consistent with the unique escape of a small group of specimens committed to a founder effect, which is reflected in the observed variation of CO1 and ATPase, which were successful in their expansion and reproduction since their introduction to the natural environment and advance towards water bodies in Bolivia. The fact that a global deviation to panmixia is not seen may be due to a rapid increase in abundance and a short period insufficient to accumulate notable differentiation at the global level. Conservation genetics readings suggest that reduced genetic variation due to gene drift or founder effects (bottlenecks) limits the adaptive capacity of populations, and reduced population sizes increase the risk of extinction due to inbreeding depression and reduced fitness (Frankham & Ralls 1998; Allendorf & Lundquist 2003). However, each geographic sample analyzed turned out to be a differentiated genetic unit (genetic population), with differences at the level of allele frequencies and allelic composition (exclusive alleles) .The presence of exclusive alleles could be the result of recent mutations in microsatellites by its hypermutable nature, and therefore hypervariable, but this is unlikely given the mutation rate that is known for this marker (1/100 to 1/1 000 000 per locus and per generation - Ellegren 2000; Sia et al. 2000 ), the recent presence of the paiche in Bolivia (around 50 years), and a four-year-old generation of the species. This result denotes the existence of a genetic microstructure where the resulting populations are not distributed randomly, they have their own characteristics (variability) and there are factors that are conditioning their structure. Rapid evolution of adaptive traits is known to sometimes occur in populations exposed to ecologically divergent environments (Reznik et al. 1997; Hendry & Kinnison 1999; Dlugosch & Parker 2008), and there is not necessarily a loss of genetic variability (Roman & Darling 2007). The fact that the variation is not uniform and random is coinciding with the sedentary behavior that has been described for the species in its natural environment (Castello 2007; Arantes et al. 2013; Nunez-Rodriguez et al. 2015), and at the same time that there is a preference for some habitats or the degree of exchange depends on the spatial distance. Regarding the latter, it has been seen that there is no relationship with geographical distances (linear and by the course of rivers), so other factors are influencing the differentiation of populations. It is interesting to note that there is a correspondence between the observed population structure with the type of system where the populations are found. There are examples of successful invaders that may have developed traits associated with environmental fluctuations and bottlenecks, which allowed them to colonize new areas with relatively low levels of genetic diversity (Gelembiuk et al. 2006). Each system from which the populations come has characteristics that differentiate them, for example, although the Madre de Dios and Beni rivers are white water, the latter is characterized by transporting a large amount of suspended sediment and dissolved solids, in the order of almost four times higher in relation to the first (Guyot et al. 1996). In turn, the Manuripi-Orthon river system is known for having intermediate characteristics between the white water systems (dry season) and clear waters (rainy season), which is why it has been called mixed waters (Navarro & Maldonado 2002). This correspondence with the revealed population structure suggests that the species is responding to certain dominant factors of the systems and possibly the set of individuals that invaded Bolivia was made up of specimens from different stocks in Peru with preferences for certain types of habitats. This means that the fish confined in Peru and transported to the Valencia and Sandoval lagoons (Carvajal-Vallejos et al. 2011) might have been gathered from different localities and / or tributaries, and therefore from different populations, of the Peruvian portion of the main axis of the Amazon and which sought the most optimal conditions and similar to those of their natural environment from which they were mined to settle down. It is known that fish farming stocks in Peru were made up for several decades by specimens from systems of white waters (e.g. Ucayalí, Marañón) and clear waters (e.g. Nanay), and no differentiation was made based on their origin to keep them together or separate (García-Dávila et al. 2011; personal observation in fish farming in Iquitos and Imiría lagoon 2006-2010). It is known that founding groups made up of individuals from different genetic populations have greater potential to expand and multiply, which makes them successful in a process of introduction and invasion (Kolbe et al. 2004). Another potential explanation for the identified population structure is that the species followed a rapid expansion process where individuals that colonized new environments began to protect their territory, and to form groups or families with related genotypes and phenotypes. The most aggressive genotypes were at the forefront of the invasion line and the most sedentary or shy genotypes were occupying the different habitats throughout the expansive wave, with a subsequent rapid change in allelic frequencies in their population constitution due to sedentary lifestyle and territorialism (at least during reproduction season – Saavedra et al. 2005) that they possess and the bottlenecks that occurred progressively. However, the presence of exclusive alleles in two of the populations studied shows that there is a relationship between the genotypic constitution and the habitats that the populations occupy, and it is most likely that there is a combination between the walking or sedentary behavior of some genotype, and the preference for establishing and reaching specific habitats.

Recommendations for the management

The results of the present work show that the paiche present in the Madre De Dios, Orthon and Beni rivers originate from a stock in Perú, or from a mixture of stocks, that reached the natural environment as described in the literature (see Carvajal-Vallejos et al. 2011) and gave rise to several populations resulting from multiple bottlenecks throughout the invasion wave in the upper Madeira basin of Bolivia. This structuring is related to an important ecological feature of the species, which is the sedentary lifestyle and the short migratory movements carried out throughout its life, and which is described and commented on by fishermen in Bolivia. The paiche does not perform large movements, and after water retreat at the end of the high water season it returns to its place of origin to feed and reproduce during the low water period. Therefore, the management and exploitation of the species should not consider all stocks of different geographic origins or systems as part of a single large population with the same genetic traits. Some features of the systems are related to the genetic variation at the geographic level. In this way, we suggest that future regional or local management plans should consider the recruitment in lagoons and the dynamics of the populations as sustained by stocks in the same lagoons subject to fishing and surrounding nearby aquatic habitats. The specimens in these surrounding smaller environments and less accessible for fishing, most likely, have a close genetic and exchange relationship with the specimens present in permanent and temporal lagoons, or streams more accessible for the fisheries.

**Acknowledgements**

A special thanks to the fishermen of the indigenous communities of Trinidacito Portachuelo Bajo (Pando Department), Flor de Octubre (Beni Department) (TIOC TIM II), and fishermen of the community of Puerto Rico (Pando) for their collaboration in collecting samples. A special knowledge of the work of Adalid Argote during the collection of samples. To Alison Magnaugthon, Tiffanie Rainville (WFT), Ben Koop (UVic), for their interest in carrying out the work and logistical support in the field and laboratory phases. To Pablo Prado (ULRA-UMSS) for the cartographic design of the map. This work was carried out with the help of a grant from the International Development Research Center (IDRC), Ottawa, Canada, and with financial support from the Government of Canada, provided through the Department of Public Affairs. External, Trade and Development (Department of Foreign Affairs, Trade and Development - DFATD). The work was co-financed in its final stage by the Instituto Nacional de Innovación Agraria y Forestal (INIAF) through financial support of the Plurinational State of Bolivia and the World Bank, within the framework of the subproject “Bases Técnicas para el Aprovechamiento Integral del Paiche (*Arapaima gigas*) en el Norte del Departamento del Beni".

**References**

Allendorf F.W., Lundquist L.L. 2003. Introduction: population biology, evolution, and control of invasive species. Conservation Biology, 17: 24–30.

Amaral C.R.L., Brito P.M., Silva D.A., Carvalho E.F. 2013. A new cryptic species of South American freshwater Pufferfish of the genus *Colomesus* (Tetraodontidae), based on both morphology and DNA data. PLoS ONE, 8(9): e74397.

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

Arantes C.C., Castello L., Cetra M., Schilling A. 2013. Environmental influences on the distribution of arapaima in Amazon floodplains. Environmental Biology of Fishes, 96: 1257-1267.

Ardura A., Linde A.R., Moreira J.C., Garcia-Vazquez E. 2010. DNA barcoding for conservation and management of Amazonian commercial fish. Biological Conservation, 143: 1438-1443.

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

Bermingham E.S., McCafferty A., Martin A. 1997. Fish biogeography and molecular clocks: perspectives from the Panamian Isthmus. pp. 112–126. In: Kocher T., Stepien C. (Eds.). Molecular systematics of fishes. Academic Press, New York, USA.

Carvajal-Vallejos F.M., Van Damme P.A., Córdova L., Coca C. 2011. La introducción de *Arapaima gigas* (paiche) en la Amazonía boliviana. pp. 367-395. In: Van Damme P.A., Carvajal-Vallejos F.M., Molina Carpio J. (Eds.). Los peces y delfines de la Amazonía Boliviana: hábitats, potencialidades y amenazas. Editorial INIA, Cochabamba, Bolivia.

Castello L. 2007. Lateral migration of *Arapaima gigas* in floodplains of the Amazon. Ecology of Freshwater Fish, 17: 38-46.

Castro F.P., Batista J., Porto J.I. 2014. DNA barcodes of Rosy Tetras and allied species (Characiformes: Characidae: *Hyphessobrycon*) from the Brazilian Amazon Basin. PLoS ONE, 9: e98603.

Cerutti-Pereyra F., Meekan M.G., Wei N.W.V., O’Shea O., Bradshaw C.J.A., Austin C.M. 2012. Identification of rays through DNA barcoding: An application for ecologists, PLoS ONE, 7: e36479.

Crespo A., Van Damme P.A. 2011. Patrones espaciales de inundación en la cuenca amazónica de Bolivia. pp. 15-27. In: Van Damme P.A., Carvajal-Vallejos F.M., Molina Carpio J. (Eds.). Los peces y delfines de la Amazonía Boliviana: hábitats, potencialidades y amenazas. Editorial INIA, Cochabamba, Bolivia.

Dlugosch K.M., Parker M. 2008. Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. Molecular Ecology, 17: 431-449.

Ellegren H. 2000. Microsatellite mutations in the germline: implications for evolutionary inference. Trends in Genetics, 16: 551–558.

Farias I.P., Hrbek T., Brinkmann H., Sampaio I., Meyer A. 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.

Frankham R., Ralls K. 1998. Conservation biology: inbreeding leads to extinction. Nature, 392: 441–442.

García-Dávila C.R., Castro-Ruiz D., Chota-Macuyama W., Biffi C., Deza S., Bazan R., García J., Rebaza M., Rebaza C., Chavez C., Chu-Koo F., Duponchelle F., Nuñez J., Renno J.-F. 2011. Caracterización genética de ejemplares de paiche *Arapaima gigas* (Cuvier, 1829) utilizados en el repoblamiento del lago Imiría (cuenca del río Ucayalí). Folia Amazonica, 20: 67-75.

García-Dávila C.R., Duponchelle F., Castro-Ruiz D., Villacorta J., Quérouil S., Chota-Macuyama W., Núñez J., Römer U., Carvajal-Vallejos F.M., Renno J.-F. 2013. Molecular identification of a cryptic species in the Amazonian predatory catfish genus *Pseudoplatystoma* (Bleeker, 1962) from Peru. Genetica, 141: 347-358.

Gelembiuk G.W., May G.E., Lee C.E. 2006. Phylogeography and systematics of zebra mussels and related species. Molecular Ecology, 15: 1033–10.

Guyot J.L., Filizola N., Quintanilla J., Cortez J. 1996. Dissolved solids and suspended sediment yields in the Rio Madeira basin, from the Bolivian Andes to the Amazon. pp. 55-63. In: Erosion and Sediment yield: Global and Regional Perspectives, Exeter, July 1996. IAHS Publ. 236.

Havel J.E., Kovalenko K.E., Thomaz S.M., Amalfitano S., Kats L.B. 2015. Aquatic invasive species: challenged for the future. Hydrobiologia, 750(1): 147-170.

Hendry A.P., Kinnison M.T. 1999. The peace of modern life: Measuring rates of contemporary microevolution. Evolution, 53: 1637-1653.

Hrbek T., Farias I.P., Crossa M., Sampaio I., Porto J.I.R., Meyer A. 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.

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

Kolbe J.J., Glor R.E., Rodríguez L., Chamizo A., Larson A., Losos J.B. 2004. Genetic variation increases during biological invasion by a Cuban lizard.

Nature, 431: 177-181.

Leão A. 2009. Análise de la variabilidade genética das populações de pirarucu (*Arapaima gigas*, Schinz 1822) dos principais tributários do rio Amazonas através do uso de marcadores microssatélites. Tesis de doctorado. Instituto Nacional de Pesquisas da Amazônia – INPA. Universidade Federal do Amazonas – UFAM, Programas de Pós Graduação do Instituto Nacional de Pesquisas da Amazônia – PPGINPA, Programa de Pós Graduação em Genética, Conservação e Biologia Evolutiva – PPG-GCBEv. Manaus (AM), Brasil.

Librado P., Rozas J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25: 1451-1452.

Machado V.N., Collins R.A., Ota R.P., Andrade M.C., Farias I.P., Hrbek T. 2018. One thousand DNA barcodes of pirhanas and pacus reveal geographic structure and unrecognised diversity in the Amazon. Scientific Reports, 8 (8387), https://doi.org/10.1038/s41598-018-26550-x

Maldonado M., Carvajal F. 2005. La ictiofauna lacustre de la llanura de inundación del río Ichilo (Bolivia). Revista Boliviana de Ecología y Conservación Acuática, 17: 15-32.

Miranda-Chumacero G., Wallace R., Calderón H., Calderón G., Willink P., Guerrero M., Siles T., Lara K., Chuqui D. 2012. Distribution of arapaima (*Arapaima gigas*) (Pisces: Arapaimatidae) in Bolivia: implications in the control and management of a non-native population. BioInvasions Records, 1(2): 129-138.

Navarro G., Maldonado M. 2002. Geografía Ecológica de Bolivia: vegetación y ambientes acuáticos. Editorial Centro de Ecología y Difusión Simón I. Patiño, Santa Cruz, Bolivia. 719 p.

Nei M., Maruyama T., Chakraborty R. 1975. The bottleneck effect and genetic variability in populations. Evolution, 29: 1-10.

Nei M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics, 89: 583–590.

Nuñez-Rodríguez J., Duponchelle F., Cotrina-Doria M., Renno J.-F., Chavez-Ventilla C., Rebaza C., Deza S., García-Dávila C., Chu-Koo F., Tello S. 2015. Movement patterns and home range of wild and re-stocked *Arapaima gigas* (Schinz, 1822) monitored by radio-telemetry in Lake Imiria, Peru. Journal of Applied Ichthyology, 31: 10-18.

Parker I.M., Simberloff D., Lonsdale W.M., Goodell K., Wonham M., Kareiva P.M., Williamson M.H., Von Holle B., Moyle P.B., Byers J.E., Goldwasser L. 1999. Impact: Toward a framework for understanding the ecological effects of invaders. Biological Invasions, 1:3-19.

Pereira L.H.G., Hanner R., Foresti F., Oliveira C. 2013. Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? BMC Genetics, 14:20.

Reznik D.N., Shaw F.R., Rodd F.H., Shaw R.G. 1997. Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). Science, 275: 1934-1937.

Roman J., Darling J.A. 2007. Paradox lost: genetic diversity and the success of aquatic invasions. Trends in Ecology and Evolution, 22: 454-464.

Rousset F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics, 145: 1219.

Saavedra E.A., Quintero L.G., Landines M.A. 2005. Aspectos reproductivos. pp. 31-40. In: Sanabria A.I., Beltrán I.C., Victoria P. (Eds.). Biología y cultivo del pirarucú *Arapaima gigas* (Schinz, 1822) (Pisces: Arapaimatidae). Bases para un aprovechamiento sostenible. Imprenta Nacional de Colombia, Bogota, Colombia.

Salmenkova E.A. 2008. Population genetic processes in introduction of fish. Russian Journal of Genetics, 44(7): 758-766.

Sambrook J., Fritsch E.F., Maniatis T. 1989. Molecular cloning: a laboratory manual. Coldspring Harbor Laboratory Press, 2nd edition, New York, USA.

Sia E.A., Butler C.A., Dominska M., Greenwell P., Fox T.D., Petes T.D. 2000. Analysis of microsatellite mutations in the mitochondrial DNA of *Saccharomyces cerevisiae*. Proceedings of the National Academy of Sciences, 97: 250-255.

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

Stewart D. 2013b. A new species of *Arapaima* (Osteoglossomorpha: Osteoglossidae) from the Solimões River, Amazonas State, Brazil. Copeia, 3: 470-476.

Thomaz S.M., Kovalenko K.E., Havel J.E., Kast L.B. 2014. Aquatic invasive species: general trends in the literature and introduction to the special issue. Hydrobiologia, 746: 1-12.

Toffoli D., Hrbek T., Araújo M.L., Almeida M., Charvet-Almeida P., Farias I.P. 2008. A test of the utility of DNA barcoding in the radiation of the freshwater stingray genus *Potamotrygon* (Potamotrygonidae, Myliobatiformes). Genetics and Molecular Biology, 31: 324-336.

Van Damme P.A., Coca C., Zapata M., Carvajal-Vallejos F.M., Carosfeld J., Olden J. 2015. The expansion of *Arapaima* cf. *gigas* (Osteoglossiformes: Arapaimidae) in the Bolivian Amazon as informed by citizen and formal science. Management of Biological Invasions, 6 (4): 375-383.

Vitorino C.A., Oliveira R.C.C., Margarido V.P., Venere P.C. 2015. Genetic diversity of *Arapaima gigas* (Schinz, 1822) (Osteoglossiformes: Arapaimidae) in the Araguaia-Tocantins basin estimated by ISSR marker. Neotropical Ichthyology, 13(3): 557-568.

Ward R.D., Hanner R., Hebert P.D.N. 2009. The campaign to DNA barcode all fishes, FISH-BOL. Journal of Fish Biology, 74: 329-356.

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

Wittmann H., von Blanckenburg F., Guyot J.L., Maurice L., Kubik P. 2011. Quantifying sediment discharge from the Bolivian Andes into the Beni foreland basin from cosmogenic 10Be-derived denudation rates. Revista Brasileira de Geociências, 41(4): 629-641.

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

Wright, S. 1978. Evolution and the Genetics of Populations. University of Chicago Press, Chicago, USA.