# Population genetic analysis of *Arapaima gigas*, one of the largest freshwater fishes of the Amazon basin: implications for its conservation

Tomas Hrbek[1,2†], Izeni P. Farias[1,3], Marcelo Crossa[4], Iracilda Sampaio[5], Jorge I. R. Porto[6] and Axel Meyer[1]

[1] Department of Biology, University of Konstanz, 78457 Konstanz, Germany
[2] Department of Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, MO, 63110 USA
[3] Universidade Federal do Amazonas, ICB, Manaus, Amazonas, Brazil
[4] Instituto de Pesquisa Ambiental da Amazônia (IPAM), Santarém, Pará, Brazil
[5] Universidade Federal do Pará, Campus de Bragança, Pará, Brazil
[6] Instituto Nacional de Pesquisas da Amazônia (INPA), CPBA, CP 478, Manaus, Amazonas, Brazil

**Abstract**

The present study reports the first population genetic analysis of *Arapaima gigas*, an important but critically over-exploited fish species of the Amazon basin. We sequenced two discontinuous mitochondrial DNA regions of 1204 base-pairs (bp) (NADH1 segment) and 1143 bp (ATPase segment) from 139 individuals of *A. gigas* representing eight localities spanning the Amazon basin from Iquitos, Peru to Macapá, Brazil. We discovered 34 haplotypes separated by 44 segregating sites. The two most common haplotypes are shared among all populations and isolation-by-distance appears to be the most important population dynamic, although there is no significant association between geographical distance and genetic differentiation. Observed patterns of differentiation appear to be attributable largely to anthropogenic factors associated with over-exploitation. Greatest haplotypic diversity exists upstream of Manaus and in areas away from large centres of population. The female variance and inbreeding effective population sizes are approximately 150 000 individuals and localities in the Amazon basin are connected by gene flow. Naturally low levels of population differentiation and relatively high between-population connectivity is encouraging for the conservation and management of *A. gigas*. If strategically placed biological reserves were created throughout the Amazon basin to act as sources of emigrants within a source–sink metapopulation model, we believe locally depleted populations can be re-populated and maintained by individuals immigrating from these reserves.

## INTRODUCTION

*Arapaima gigas*, locally known as pirarucu in Brazil or paiche in Peru and Bolivia, is one of the largest freshwater fishes of South America; it can grow to up to 3 m in length and can weigh over 200 kg (Saint-Paul, 1986; Nelson, 1994). It is a long-lived, solitary, paternal mouth-brooder, breeding once a year from the age of 4–5 years. It is also a representative of one of the oldest fish lineages, being one of only seven species in the family Osteoglossidae and belonging to the monotypic genus *Arapaima* (Nelson, 1994). *Arapaima* is distributed predominantly in the Amazon basin floodplain, although it has been also recorded from the upper Essequibo basins of Guyana (Lüling, 1964). Fossil evidence indicates that it was formerly found in Colombia's Río Magdalena basin as well (Lundberg & Chernoff, 1992). In spite of its commercial importance, more precise information about the distribution of the *Arapaima* within the Amazon basin

and especially information regarding its occurrence in the lower part of the basin is lacking (Crossa & Petrere Júnior, 1999). It is clear, however, that *Arapaima* is closely associated with the floodplain ecosystem (Goulding, Smith & Mahar, 2000; Goulding, Barthem & Ferreira, 2003).

Traditionally *Arapaima* formed a significant portion of the diet of Amerindians living near larger bodies of water. Starting as early as the late 18th century, a commercial fishery was developed for *Arapaima* to satisfy demand for bacalhau. Bacalhau (Portuguese) or bacalao (Spanish) is salted and dried codfish whose origins trace back to the dark ages when Basque fishermen created a market for it in Catholic Europe as a high quality and durable source of animal protein that could be consumed on religious holidays (Kurlansky, 1998). Its prominence increased during the colonial era as a cheap protein source fed to the slave population working on Caribbean and South American plantations (Kurlansky, 1998). Within the Amazon basin, *Arapaima* was discovered to be an excellent local substitute for cod due to its firm, white, flaky, practically boneless and nearly fat-free flesh.

---

†All correspondence to: Tomas Hrbek. Tel: +1 314 362 4198; Fax: +1 314 362 3595; E-mail: hrbek@pcg.wustl.edu

*Arapaima* was salted and dried and then exported out of Brazilian processing centres in Manaus, Santarém and Belém.

The commercial importance of *Arapaima* can be appreciated in the descriptions of Veríssimo (1895) who states that between the years 1885 and 1893, 11 540 tons of the dry and salted form of *Arapaima* were sold in Belém (1282 tons/year). Menezes (1951) cites that in the 7 years between 1918 and 1924, 6775 tons of pirarucu (968 tons/year) were landed in the State of Amazonas and between 1919 and 1921, 5729 tons (1920 tons/year) were landed in Belém alone. In spite of these levels of exploitation, *Arapaima* appears to have been relatively abundant near main Amazonian processing and export cities, such as Manaus in the State of Amazonas and Santarém and Belém in the State of Pará until the early 1960s (Crossa & Petrere Júnior, 1999).

*Arapaima* became increasingly scarce in the 1970s, until in the 1980s it became commercially extinct near larger Amazonian cities and, in some areas, appears to have disappeared completely (Goulding, 1980). As early as 1975 the world community recognised the plight of *Arapaima* and listed it as a CITES Appendix II fish. Formal protection of *Arapaima* spurred aquiculture efforts in the 1990s, but this did not prevent illegal fishing. In early 2001 the Brazilian government banned all fishing for *Arapaima*, except in the Mamirauá sustainable developmental reserve.

The ecology and recent demographic history of *Arapaima* allows us to test several specific hypotheses with genetic data. The first hypothesis concerns the population structure of *Arapaima*. Given that the Amazonian flood plain forms a largely continuous ecosystem, one might expect little geographical structuring between localities. Isolation by distance might, however, be a major population structuring factor. Demographic and census data indicate that *Arapaima* went though a dramatic population crash in the last 20 years and most probably was steadily decreasing over the last two centuries. Using this observation, we test a second hypothesis that the distribution of genetic diversity is in disequilibrium as a result of the recent demographic decline. A third hypothesis combines the two previous hypotheses and suggests that the observed population structuring, if any, could be an artefact of the demographic and genetic bottleneck. If this hypothesis is correct, one should observe a negative correlation between intra-population diversity and inter-population differentiation and a pattern of pair-wise population differentiation that is not correlated with geography.

In this study we also hope to make a contribution to the conservation and management of natural populations of *A. gigas*. Due to its commercial importance, it is necessary to formulate and implement a realistic conservation policy that takes local socioeconomic needs into account. Formulation of management policies in the absence of good data can have disastrous consequences for the species and those depending on it. Population genetic data are arguably the most important baseline component of any conservation and management plan (Frankham, Ballou & Briscoe, 2002), but such data are non-existent for *Arapaima*, as well as for nearly all other Amazonian species, regardless of their economic importance.

## METHODS

### Sampling

Geographical sampling was designed to include individuals from the extremes as well as the central areas of *Arapaima*'s distribution (Fig. 1). Sampled individuals come from Iquitos ($n = 16$), Letícia ($n = 22$), Mamirauá ($n = 33$), Manaus (Careiro Island) ($n = 13$), Madre de Dios ($n = 13$), Santarém ($n = 15$), Macapá ($n = 14$) and Marabá ($n = 13$), see Table 1. All samples were collected by researchers or obtained by researchers directly from artisanal fishermen. Tissue samples were preserved in 95% ethanol and stored in a cool place until delivered to the laboratory.

### DNA extraction and amplification

Total genomic DNA was extracted from muscle or liver tissue preserved in 95% ethanol. The tissue was dissolved and digested using a Proteinase K/SDS solution, followed by phenol and chloroform extraction, with the addition of 5M NaCl followed by 70% ethanol precipitation of the DNA product.

Polymerase chain reaction (PCR) amplifications were performed on total genomic DNA. Negative controls were performed for all reactions. The temperature profile for the 30-cycle amplification reaction consisted of denaturation at 94 °C for 35 s, annealing at 50 °C for 35 s and extension at 72 °C for 90 s per cycle. The resulting products were evaluated on a 1% agarose gel and then purified with Qiagen spin-columns (Hilden, Germany). Amplified mtDNA segments were sequenced from both the 5′ and 3′ ends. Sequencing reactions followed the standard Perkin Elmer Big Dye sequencing protocol for double-stranded cycle sequencing reactions. Sequences were determined on a Perkin Elmer ABI 3100 automatic DNA sequencer. The mtDNA region analysed included the 3′ end of the 16S rRNA, tRNA leucine and the complete NADH dehydrogenase subunit 1 genes (the NADH1 segment). The NADH1 segment was amplified using L3002 5′-TACGA-CCTCGATGTTGGATCAGG-3′ and H4419 5′-AAGCT-TTYGGGCCCATACC-3′ (Hrbek & Meyer, 2003) and sequenced using L3079 5′-ACGTGATCTGAGTTCA-GACCG-3′ (Hrbek & Meyer, 2003) and H4364 5′-GGAAGCACTARGAGTTTTGA-3′, designed for this study. The ATPase segment included the 3′ end of the cytochrome oxidase subunit 2, tRNA lysine, ATPase subunit 6, ATPase subunit 8 and the 5′ end of the cytochrome oxidase subunit 3. The ATPase segment was amplified with L8106 5′-TGGGTGTTAAAATAGATGC-3′ and H9264 5′-GAGGAGAGCRGCRGATGCCCC-3′ and sequenced with L8106, L8537 5′-TGAAACTGACCAT-GACACTAAG-3′ and H8516 5′-CTTAGTGTCATG-GTCAGTTTCA-3′. All ATPase primers were designed specifically for this study based on complete mitochondrial genomes of *Pantodon buchholzi* (NC_003096;
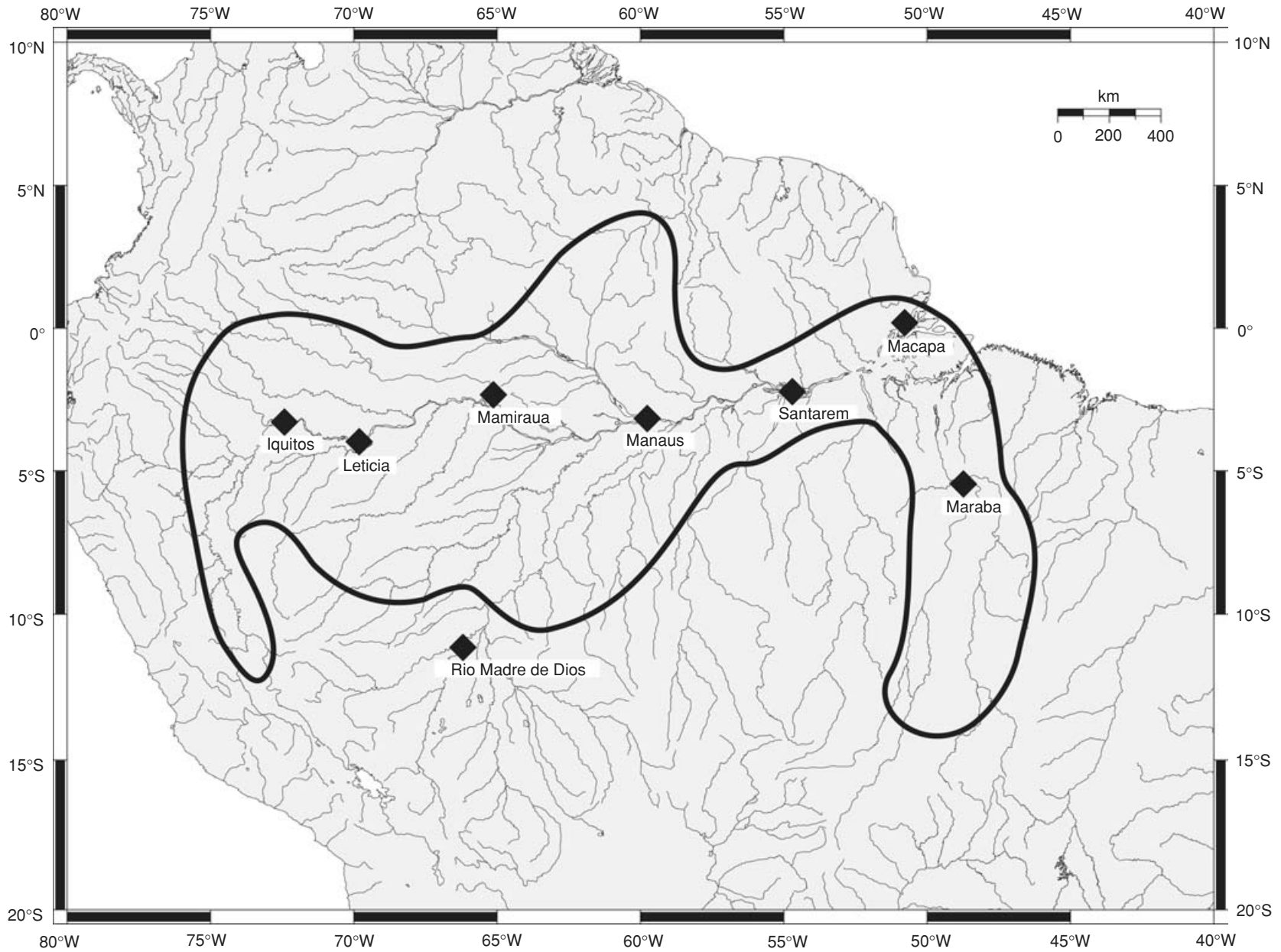
**Fig. 1.** Map of the approximate natural distribution of *Arapaima gigas* in the Amazon River basin showing the localities sampled for this study.

**Table 1.** Haplotypes and the number of individuals with these haplotypes present at each locality included in this study

| | Iquitos (16/7) | Letícia (22/7) | Mamirauá (33/15) | Manaus (13/3) | Río Madre de Dios (13/1) | Santarém (15/5) | Macapá (14/2) | Marabá (13/8) | Total (139/34) |
|---|---|---|---|---|---|---|---|---|---|
| Hp01 | 7 | 13 | 7 | 1 | – | 4 | – | – | 32 |
| Hp02 | 1 | 2 | 9 | 5 | – | 7 | – | 4 | 28 |
| Hp03 | – | 1 | – | – | – | – | – | – | 1 |
| Hp04 | – | – | 2 | – | – | – | – | 1 | 3 |
| Hp05 | – | – | – | – | – | – | – | 1 | 1 |
| Hp06 | – | – | – | – | – | – | 1 | – | 1 |
| Hp07 | – | – | – | – | – | – | 13 | – | 13 |
| Hp08 | 1 | – | – | – | – | – | – | – | 1 |
| Hp09 | – | – | 1 | – | – | – | – | – | 1 |
| Hp10 | 2 | – | – | – | 13 | – | – | – | 15 |
| Hp11 | 1 | – | – | – | – | – | – | – | 1 |
| Hp12 | 3 | 2 | – | – | – | – | – | – | 5 |
| Hp13 | – | – | – | – | – | – | – | 1 | 1 |
| Hp14 | – | – | – | – | – | 1 | – | – | 1 |
| Hp15 | – | – | – | – | – | – | – | 1 | 1 |
| Hp16 | – | – | 1 | – | – | – | – | 1 | 2 |
| Hp17 | – | – | 1 | – | – | – | – | 3 | 4 |
| Hp18 | – | – | 1 | – | – | – | – | – | 1 |
| Hp19 | – | – | 1 | – | – | – | – | – | 1 |
| Hp20 | – | 1 | – | – | – | – | – | – | 1 |
| Hp21 | – | – | – | – | – | 2 | – | – | 2 |
| Hp22 | – | 2 | – | – | – | – | – | – | 2 |
| Hp23 | – | 1 | – | – | – | – | – | – | 1 |
| Hp24 | – | – | – | – | – | – | – | 1 | 1 |
| Hp25 | – | – | – | – | – | 1 | – | – | 1 |
| Hp26 | – | – | 1 | – | – | – | – | – | 1 |
| Hp27 | – | – | 1 | – | – | – | – | – | 1 |
| Hp28 | – | – | 2 | – | – | – | – | – | 2 |
| Hp29 | – | – | 1 | – | – | – | – | – | 1 |
| Hp30 | – | – | 2 | – | – | – | – | – | 2 |
| Hp31 | 1 | – | – | – | – | – | – | – | 1 |
| Hp32 | – | – | – | 7 | – | – | – | – | 7 |
| Hp33 | – | – | 1 | – | – | – | – | – | 1 |
| Hp34 | – | – | 2 | – | – | – | – | – | 2 |

Locality names correspond to those in Fig. 1 and are followed in parentheses by number of individuals sampled/haplotypes observed. The numbers in the table body represent the number of individuals carrying a specific haplotype sampled at a specific locality. GenBank accession numbers range from AY081867 to AY081890 for NADH1 segment haplotypes and AY081891 to AY081903 for ATPase segment haplotypes.

AB043068) and *Osteoglossum bicirrhosum* (NC_003095; AB043025) available in GenBank. The primer designations correspond to their 3′ position in the human mitochondrial genome (Anderson *et al.*, 1981) by convention. H and L designates the heavy and light strand, respectively. The mitochondrial control region was not chosen because its base-pair composition in the Osteoglossoidei prevents efficient amplification and sequencing (Inoue *et al.*, 2001).

## Data analysis

Homologous protein-coding regions were aligned manually and confirmed by translating DNA data into putative amino acid sequences using the program BioEdit (Hall, 1999). Alignment of transfer RNAs was constructed manually based on secondary structural models

(Kumazawa & Nishida, 1993). All sequence data are deposited in GenBank under accession numbers AY081867 to AY081890 and AY081891 to AY081903.

Relationships between haplotypes were inferred from minimum-spanning networks. The haplotype network for the combined gene regions was reconstructed using the program TCS, version 1.13 (Clement, Posada & Crandall, 2000), which implements the algorithm described by Templeton, Crandall & Sing (1992). Alternative branching orders in the TCS-generated network were resolved by choosing the most likely relationship between haplotypes as favoured by 25 maximum likelihood searches in PAUP* (Swofford, 2001) under the HKY85 model of molecular evolution (Hasegawa, Kishino & Yano, 1985). Clades were nested into higher level categories according to the nested-clade design of Templeton & Sing (1993). We then

used the program GeoDis version 2.0 (Posada, Crandall & Templeton, 2000) to test for geographical association of clades and nested clades using a permutation contingency analysis (Templeton & Sing, 1993).

Results of the Nested Clade Analysis (NCA) were used to infer the population genetic structure of *Arapaima* and to infer which clades are connected by gene flow. Differentiation among populations was further tested using Wright's *F* statistics (Wright, 1951, 1969) and by the Mantel test (Mantel, 1967) of genetic and geographical correlation. Among-locality gene flow was estimated from the parameter *Nm* using the method of Beerli & Felsenstein (1999) implemented in the program MIGRATE. MIGRATE default search strategy values were used, except that we ran 20 short chains and five long chains and we used a geographical distance file.

Total population analyses included a test of genetic equilibrium of mtDNA alleles (Watterson, 1978; Tajima, 1989; Rogers & Harpending, 1992; Rogers, 1995; Fu, 1997). Although most of these tests have been formally designed to test for selection, a significant deviation from genetic equilibrium can best be explained by demographic events in situations where no selective advantage between haplotypes exists (Hartl & Clark, 1997). The data were also analysed using generalised skyline plots (Strimmer & Pybus, 2001; Pybus & Rambaut, 2002) to differentiate between alternate demographic models, including population bottlenecks, and verified by the phylogenetic method of Kuhner, Yamato & Felsenstein (1995). Variance effective population size was estimated from the population genetic parameter theta ($\Theta$) (Watterson, 1975) and inbreeding effective population size was estimated from a generalised skyline plot (Strimmer & Pybus, 2001; Pybus & Rambaut, 2002).

Indexes of genetic diversity were also estimated for individual localities. These included Watterson's $\Theta$, gene diversity ($\hat{H}$) calculated as the probability that two randomly chosen samples are different (Nei, 1987) and haplotype diversity calculated as the number of haplotypes divided by number of individuals sampled.

Since *Arapaima* occurs in two distinct forms, orange-fleshed and white-fleshed individuals, which according to some authors (e.g. Nunes-Pereira, 1934) are thought to represent two distinct species, we tested for a significant association of haplotypes with colour of individuals.

Statistical analyses of the data were performed using the programs GeoDis version 2.0 (Posada, Crandall & Templeton, 2000), GENIE version 3.1 (Pybus & Rambaut, 2002), FLUCTUATE version 1.4 (Kuhner, Yamato & Felsenstein, 1998), MIGRATE version 1.6.9 (Beerli & Felsenstein, 2001) and Arlequin version 2000 (Schneider, Roessli & Excoffier, 2000). Alternatively, equations were solved using the program MathCad 2001 (MathSoft Inc., 2002).

## RESULTS

Based on the analysis of 139 individuals, we discovered a total of 25 haplotypes separated by 27 segregating sites in the NADH1 region and 13 haplotypes separated by 17 segregating sites in the ATPase region. The NADH1 and ATPase do not show any conflict (Partition Homogenetity Test: $P = 1.00$); thus, the two mitochondrial regions were combined in all further analyses. The combined data resulted in 34 haplotypes separated by 44 segregating sites. In the NADH1 segment, no mutations resulted in amino acid replacements; however, two mutations separating the two major haplotypes strictly covary. In the ATPase segment, five out of the 17 mutations result in amino acid replacements; in addition five out of the 17 mutations strictly covary and separate the two main haplotypes and three of these mutations occur in tandem. For these reasons, these covarying mutations were reduced to one effective mutation in all analyses, although treating these mutations as independent does not qualitatively affect the results or conclusions drawn from the results. Only one homoplastic mutation occurred in the whole network, thus the assumptions of infinite sites and infinite alleles models upon which sequence statistics are based appear not to be violated severely.

The haplotype network shows a relatively tight clustering of individual haplotypes; two common haplotypes predominate, while the majority of the other haplotypes are composed of rare haplotypes or singletons (Table 1, Fig. 2). NCA of *A. gigas* indicated little population structuring. Following the key of Templeton (1998), significant events in the history of *Arapaima* are restricted gene flow with isolation by distance between Manaus and other portions of the Amazon basin, fragmentation of Macapá from rest of the Amazon and range expansion into Río Madre de Dios in Bolivia (Table 2). Among population differentiation was further tested by pair-wise $F_{ST}$ analysis, which revealed significant differentiation between some localities (Table 3), however, the isolation-by-distance model as a significant structuring factor is not supported (Mantel test of association of geographical distances and genetic differentiation, $r = -0.047$, $P = 0.585$). Analysis of gene flow using a coalescent algorithm implemented in the program MIGRATE (Beerli & Felsenstein, 2001) revealed a mixed pattern of gene flow among populations. For example, very high gene flow exists between Mamirauá and Marabá, both regions of high genetic diversity. Mamirauá and Marabá are at the geographical extremes where migrants must pass through intervening regions with which they experience near zero gene flow levels (Table 4).

Historical data indicate that *A. gigas* underwent a severe demographic bottleneck in the latter part of the last century. Genetic data support this hypothesis. The Ewens–Watterson test (Watterson, 1978), as well as Slatkin's (1994, 1996) version of this test show that the haplotype frequency distribution deviates significantly from the expectation of a genetic steady-state (Watterson's $P = 0.012$; Slatkin's $P = 0.001$). Common alleles occur more frequently than expected by chance alone and there are also significantly more singletons than expected by chance alone resulting in a significantly large *F* value (Watterson, 1978). Both Tajima's *D* (Tajima, 1989) and Fu's *Fs* (Fu, 1997) statistics are significantly negative
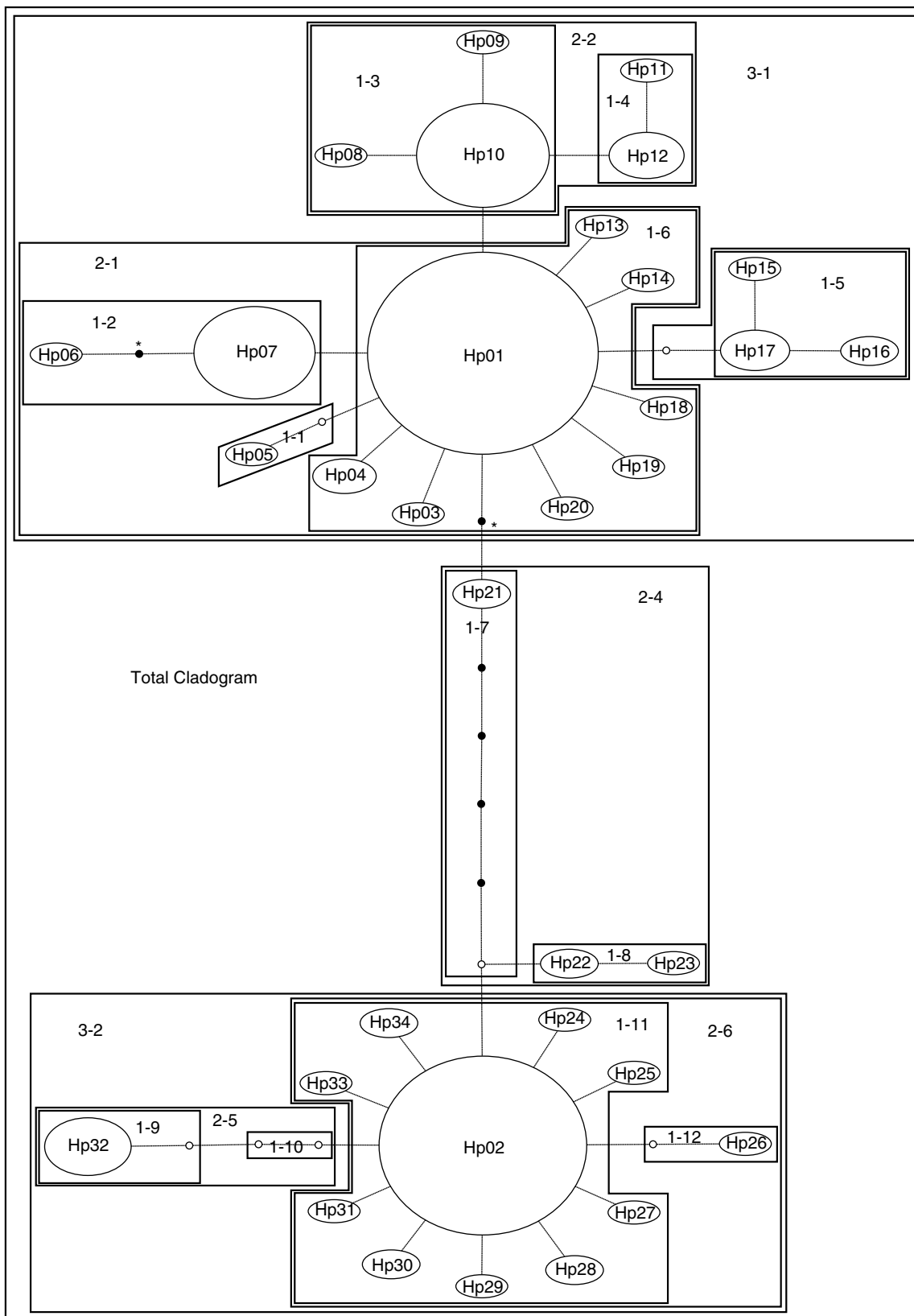
**Fig. 2.** Minimum-spanning haplotype network of haplotypes from the NADH1 and the ATPase segments. Circles with numbers represent observed haplotypes, while unlabelled circles represent missing haplotypes. Filled circles indicate correlated mutational changes. Asterisks represent a homoplaseous mutation in the NADH1 segment. Haplotypes are connected by single mutational steps represented by a line. Haplotype nesting was carried out as described by Templeton & Sing (1993).

**Table 2.** Nested contingency analysis of geographical association (Templeton *et al.*,1995) of *Arapaima gigas* haplotypes of the combined NADH1 and ATPase segments

| Total clade | Chi-square statistic | Probability | Interpretation |
|---|---|---|---|
| 1–3 | 21.9111 | 0.0228* | Continuous range expansion (Madre de Dios) |
| 1–4 | 0.6000 | 1.0000 | H₀ not rejected |
| 1–5 | 0.8750 | 1.0000 | H₀ not rejected |
| 1–6 | 46.9170 | 0.2059 | H₀ not rejected |
| 1–11 | 43.4643 | 0.5563 | H₀ not rejected |
| 2–1 | 73.7561 | 0.0001* | Past fragmentation (Macapá) |
| 2–2 | 14.1092 | 0.0003* | Insufficient genetic resolution to discriminate between range expansion/ colonisation and restricted dispersal/gene flow (Madre de Dios) |
| 2–4 | 5.0000 | 0.1000 | H₀ not rejected |
| 2–6 | 1.1868 | 1.0000 | H₀ not rejected |
| 3–1 | 90.1611 | 0.0000* | Continuous range expansion (Madre de Dios) |
| 3–2 | 24.5854 | 0.0004* | Restricted gene flow with isolation by distance (Manaus) |
| Entire cladogram | 67.7625 | 0.0000* | Restricted gene flow with isolation by distance |

Clade refers to the nesting clades shown in Fig. 2. Permutational chi-square probability is assessed by randomly permuting the lower level clade categories within the nesting clade *versus* geographical locality 10 000 times; significant results are indicated by *. Interpretation is based on Templeton *et al.* (1995) and Templeton (1998). H₀ = no geographical association of haplotypes.

**Table 3.** Estimate of pair-wise $F_{ST}$ values

Population pairwise $F_{ST}$
Distance method: Pairwise difference

|  | 1: Iquitos | 2: Letícia | 3: Mamirauá | 4: Manaus | 5: Madre | 6: Santarém | 7: Macapá | 8: Marabá |
|---|---|---|---|---|---|---|---|---|
| 1: Iquitos | 0.00000 | | | | | | | |
| 2: Letícia | 0.05232 | 0.00000 | | | | | | |
| 3: Mamirauá | 0.22983* | 0.15932 | 0.00000 | | | | | |
| 4: Manaus | 0.58293* | 0.56792* | 0.33735* | 0.00000 | | | | |
| 5: Madre | 0.24473* | 0.44400* | 0.47536* | 0.77922* | 0.00000 | | | |
| 6: Santarém | 0.28969 | 0.19468 | − 0.02781 | 0.38775* | 0.66707* | 0.00000 | | |
| 7: Macapá | 0.51817* | 0.47904* | 0.47444* | 0.77154* | 0.96418* | 0.64906* | 0.00000 | |
| 8: Marabá | 0.18756 | 0.13353 | 0.66707* | 0.40709* | 0.47778* | 0.09950 | 0.40709* | 0.00000 |

Asterisk (*) indicates a significant $F_{ST}$ value at the 0.05 level after Bonferroni correction for multiple comparisons.

**Table 4.** MIGRATE analysis showing pair-wise estimates of gene flow

MCMC estimates Ln(L) of estimates = 25.187

| Population | Theta [4Ne mu] | 4Nm [x = receiving [population] | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 1,x | 2,x | 3,x | 4,x | 5,x | 6,x | 7,x | 8,x |
| 1: Iquitos | 0.00059 | – | 1146.2850 | 3.65e-13 | 1.08e-13 | 4.06e-13 | 2.20e-14 | 1.30e-11 | 0.0020 |
| 2: Letícia | 0.00027 | 290.5848 | – | 1129.6998 | 8.57e-15 | 2.66e-17 | 7.57e-15 | 2.71e-17 | 1319.5606 |
| 3: Mamirauá | 0.00321 | 3032.9243 | 1.33e-07 | – | 486.1684 | 2.72e-06 | 1.80e-13 | 4.83e-14 | 49278.7280 |
| 4: Manaus | 0.00035 | 9.22e-15 | 1186.6377 | 3.71e-13 | – | 7.39e-08 | 2.34e-11 | 5.60e-14 | 8.90e-13 |
| 5: Madre | 5.43e-06 | 126.5687 | 2.10e-07 | 3.54e-06 | 7.82e-10 | – | 3.15e-04 | 3.66e-04 | 3.86e-04 |
| 6: Santarém | 0.00011 | 0.0065 | 2.91e-10 | 535.8355 | 0.0013 | 0.0044 | – | 288.528 | 6.25e-13 |
| 7: Macapá | 4.40e-06 | 7.48e-09 | 7.98e-10 | 130.1448 | 9.41e-11 | 6.54e-12 | 4.48e-19 | – | 1.40e-04 |
| 8: Marabá | 0.00170 | 1.42e-05 | 4.46e-06 | 486.4576 | 4.28e-10 | 4.25e-07 | 5.78e-15 | 1.11e-08 | – |

Populations can have different numbers of immigrating and emigrating individuals.

($D = -1.648$; $P = 0.020$ and $Fs = -21.366$; $P < 0.001$); both statistics support a hypothesis of a significant deviation from an equilibrium condition. A generalised skyline plot analysis of population size (Strimmer & Pybus, 2001; Pybus & Rambaut, 2002) suggests a rapid step-wise change in population size to 180 000 female

**Table 5.** Characteristics and important summary statistics of the NADH1 segment and the ATPase segment

Population structure

*Mantel test*
r = −0.047, P = 0.585

Population history

*Ewens-Watterson tests of selective neutrality*
F = 0.2121, Watterson's P = 0.019*, Slatkin's P = 0.001*
*Mismatch distribution*
Raggedness index = 0.025, P = 0.631
*Tajima's test of selective neutrality*
D = −1.648, P = 0.020*
*Fu's test of selective neutrality*
Fs = −21.366, P = 0.000*
*Migration estimate (Θ = 7.08, F = 0.39, μ = 1 × 10$^{-8}$)*
Nm (island model) ∼ 1.45
Nm (stepping-stone model) ∼ 6.34

Phenotype differentiation

*Color phenotype association*
$F_{ST}$ = 0.085, P = 0.179

Significant results are indicated by * and are based on the P = 0.05 level.

**Table 6.** Characterisation of intra-populational genetic diversity

|  | Θ | Ĥ | Haplotypes/individual |
|---|---|---|---|
| Iquitos | 2.411 | 0.792 | 0.438 |
| Letícia | 2.469 | 0.390 | 0.318 |
| Mamirauá | 4.435 | 0.884 | 0.455 |
| Manaus | 2.256 | 0.603 | 0.231 |
| Río Madre de Dios | 0.000 | 0.000 | 0.077 |
| Santarém | 1.538 | 0.733 | 0.333 |
| Macapá | 0.314 | 0.143 | 0.143 |
| Marabá | 3.545 | 0.885 | 0.615 |

individuals. Variance effective population size, estimated from the population genetic parameter Θ (Watterson, 1978) and identity-by-descent F (Watterson, 1978), suggests approximately 150 000 female individuals. The program FLUCTUATE (Kuhner *et al.*, 1998) also supports the hypothesis of a sudden step-wise change in population size as the most likely demographic scenario for *Arapaima*. Only the mismatch distribution pattern (Rogers & Harpending, 1992) of the *Arapaima* data is not significantly different from the equilibrium expectation (P = 0.631).

In spite of suggestions that orange-fleshed and white-fleshed individuals might represent two different species, there is no association of haplotypes or haplotype clusters with flesh colour phenotypes ($F_{ST}$ = 0.085, P = 0.179). Population statistics are summarised in Table 5 and indexes of genetic diversity for individual sampling localities are listed in Table 6.

## DISCUSSION

### Population structuring

Inference based on the results of the NCA (Templeton, Routman & Phillips, 1995; Templeton, 1998) supports the hypothesis of restricted gene flow with isolation by distance. This inference pertains to the highest nesting level and applies to Amazonia as a whole; it is also observed at the second-highest level, which contrasts a subset of Manaus haplotypes with haplotypes from the rest of Amazonia. Continuous range expansion and long-distance colonisation is inferred for contrasts with Río Madre de Dios samples and the differentiation of Macapá from other Amazonian localities is inferred to be the result of a past fragmentation event.

Within Amazonia no simple pattern of differentiation is evident. Pair-wise $F_{ST}$ analysis shows a number of populations to be significantly differentiated from each other (Table 3), however, isolation-by-distance (Mantel test of association of geographical distances and genetic differentiation, r = −0.047, P = 0.585) does not appear to be a factor involved in genetic structuring. For example, Iquitos and Marabá are not significantly differentiated although they represent our geographically most extreme sampling points (Table 3). However, fishes from these areas are significantly differentiated from geographically intermediate localities. A similar pattern is observed in the gene flow analysis (Table 4) using a coalescent algorithm implemented in the program MIGRATE (Beerli & Felsenstein, 2001), although the model implemented in MIGRATE (Beerli & Felsenstein, 2001) is largely free of the assumptions underlying the $F_{ST}$-based methods for deriving Nm values. Indeed there is no significant association of $F_{ST}$ and emigrant Nm values (paired *t*-test, P = 0.085) or $F_{ST}$ and immigrant Nm values (paired *t*-test, P = 0.285). The observed differentiation and emigration/immigration patterns have little biological plausibility and may be the result of past anthropogenic activities.

Historical data indicate that *A. gigas* was over-exploited in the latter part of the last century. Level of over-exploitation should be negatively correlated with intra-locality genetic diversity. Because of elevated intra-locality genetic drift in these localities, the level of over-exploitation should be positively correlated with increased inter-locality genetic differentiation. This hypothesis is significantly supported when inter-locality genetic differentiation is measured as $F_{ST}$ and intra-locality genetic diversity is measured as Ĥ (r = −0.7961, P = 0.018), when intra-locality haplotype diversity is measured as Watterson's Θ (r = −0.8092, P = 0.015) or when intra-locality haplotype diversity is measured as the number-of-alleles divided by the number-of-individuals (r = −0.8528, P = 0.003).

An alternative explanation for the observed NCA inference of past fragmentation of the Macapá population may be Pleistocene sea level changes. Macapá lies near the eastern limit of the distribution of *Arapaima*, however, there is no evidence that lowering of sea levels would have isolated this area. Rather, during the Pleistocene the Amazon River extended further east and with it the

distribution of *Arapaima* should have extended further east as well.

These results would imply that the pattern of differentiation observed between sampling localities of *Arapaima* is most probably an artefact of over-exploitation. The pattern of restricted gene flow with some isolation by distance – although this pattern may be partially confounded by haplotype extinctions (Masta, Laurent & Routman, 2003) – inferred from NCA is likely to be the only relevant population structuring factor. *Arapaima gigas* most probably behaves as a panmictic population, but due to the vastness of its geographical range, some restricted gene flow and/or dispersal within the Amazon basin probably exists. Indeed estimates of gene flow based on the probability-of-identity $F$ values suggest relatively high levels of among-locality gene flow (Island model $Nm \sim 1.5$; Stepping-stone model $Nm \sim 6.3$).

## Population history

There has been a long and well documented history of exploitation of *A. gigas* in the Amazon basin, although severe effects of overexploitation did not become evident until the 1960s. Stocks near major Amazonian cities became depleted and commercial fishermen had to move further and further from traditional grounds to catch sufficient quantities of *Arapaima*; in some areas *Arapaima* went extinct (Goulding, 1979, 1980; Crossa & Petrere Júnior, 1999; Goulding *et al.*, 2003). The combined effect of these events impacted the current pattern of distribution of genetic diversity in *Arapaima*.

The *Arapaima* population is not in a genetic equilibrium predicted under the mutation–migration–drift equilibrium models. This observation is concordant with our hypothesis that *Arapaima* suffered a severe population bottleneck. However, the direction of the deviation from genetic equilibrium is in the direction opposite to expectation. In our data set we observe significantly more common and rare alleles in relation to medium-frequency alleles than expected by chance alone. Traditionally, the excess of rare alleles resulting in significantly negative values of $D$ and $Fs$ statistics and a significantly large probability-of-identity coefficient $F$ is interpreted as a genetic signature of diversifying selection or a recent population expansion at a broad geographical scale (Rand, 1996). While these statistical tests are able to recover the signal of a mutation–migration–drift disequilibrium in *A. gigas*, we suggest that the direction of the signal is the result of an unequal severity of the demographic bottleneck throughout the range of *Arapaima* and thus warrants an alternate inference.

The directionality of the statistical tests is consistent with the hypothesis that a demographic bottleneck occurred near *Arapaima*'s geographical centre. In situations where some isolation-by-distance and restricted gene flow exist, a species' geographical centre is statistically the oldest and it harbours the oldest haplotypes as well as the greatest diversity of haplotypes (Castelloe & Templeton, 1994). In addition to the most common haplotypes that are likely to be found everywhere, the centre also contains the majority of the medium-frequency haplotypes, as well as the rare, recently arisen haplotypes (Castelloe & Templeton, 1994). A severe bottleneck in the species' geographical centre would therefore eliminate all but the most common haplotype(s) from the centre, since the probability of allelic elimination is inversely proportional to its frequency. Indeed, genetic diversity (Watterson's $\Theta$) in the lower central Amazon, which has experienced heavy fishing pressure – see Introduction, above – is significantly reduced compared to other areas of the Amazon basin (paired $t$-test, $P = 0.012$). Even if all individuals were eliminated, the subsequently immigrant individuals will again most probably bear the most common haplotype(s), since the probability of a particular haplotype colonising an area is also inversely proportional to its frequency. In the context of the whole population, we would observe a relative deficit of medium-frequency haplotypes or, viewed in another way, an overall excess of rare and common haplotypes. The lack of the medium-frequency haplotypes would increase the value of $F$ (Watterson, 1986) and, simultaneously, change the relationship of the number of segregating sites to the average pair-wise sequence (Tajima, 1989; Fu, 1997) in the same way that is usually interpreted as the genetic signature of an expanding population.

The presence of a demographic bottleneck in the centre of *Arapaima*'s distribution is supported by fisheries data. Most of the fishing for *Arapaima* occurred in the main channel of the Amazon River and its major tributaries and especially near major urban centres (Veríssimo, 1895; Crossa & Petrere Júnior, 1999). In many of these regions *Arapaima* became commercially extinct as early as the 1960s or even became completely eliminated (Goulding, 1979, 1980). These regions are precisely those areas that theoretically should have harboured the greatest and the oldest genetic diversity (Castelloe & Templeton, 1994). Populations analysed from around major population centres and those from the lower reaches of the Amazon River have decreased levels of genetic diversity (Table 6). The relatively high frequency of unique haplotypes in Macapá and Manaus may represent what were historically medium-frequency haplotypes that now became dominant in these populations due to genetic drift. Populations from less densely settled areas (Iquitos, Letícia and Mamirauá) and from areas away from the main stream (Marabá), have a relatively higher diversity of genotypes including medium-frequency haplotypes compared to Santarém, Macapá and Manaus (paired $t$-test, $P = 0.012$).

Population size changes are also inferred from the generalised skyline plot analysis (Strimmer & Pybus, 2001; Pybus & Rambaut, 2002) which, based on distribution of allelic frequencies, suggests a rapid step-wise change in population size to 180 000 female individuals. The program FLUCTUATE (Kuhner *et al.*, 1998) also supports the hypothesis of a sudden step-wise change in population size. Both analyses confirm demographic data.

Variance effective population size derived from Watterson's $\Theta$ (Watterson, 1978) and identity-by-descent $F$ (Watterson, 1978) suggests a current population of approximately 150 000 female individuals. Although

150 000 female effective individuals may not seem small, it should be noted that records indicate that yearly commercial harvest of *Arapaima* in the lower Amazon during the last quarter of the 19th and the first quarter of the 20th centuries easily exceeded 3000 tons of salted and dried *Arapaima* flesh. Assuming each *Arapaima* produces 20 kg of salted dry flesh, a 3000 tons/year harvest represents the killing of 150 000 individuals/year. The population size from which these numbers were taken must have been huge compared to the current population size even if current census numbers exceed effective population numbers by an order of magnitude.

### Origin of Río Madre de Dios *Arapaima*

The Río Madre de Dios is one of three major tributaries of the upper Rio Madeira within the Bolivian basin. While descending from the Bolivian basin between Puerto Paez, Bolivia and Porto Velho, Brazil, the Madeira River encounters a series of 12 major and minor cataracts that form a major aquatic barrier to migration (Goulding, 1979). The cataracts also form faunal boundaries for a number of aquatic species (e.g. Hamilton *et al.*, 2001). Until recently no information existed on the presence of *A. gigas* in the Río Madre de Dios drainage, or any drainage above these cataracts. A consultation with Michael Goulding (20/5/2002, pers. comm.) confirmed that *A. gigas* was introduced above the Teotonio cataracts in the late 1970s and early 1980s, probably from Iquitos or from Rio Madeira just below Porto Velho. The samples we obtained from the Río Madre de Dios locality (Fig. 1) represent genetically the most depauperate population. All individuals share only one haplotype (Hap10) linked by one mutational step to one (Hap01) of the two most common haplotypes. NCA (Templeton *et al.*, 1995; Templeton, 1998) suggests a range expansion into the Río Madre de Dios region from the Iquitos area, supporting the hypothesis of a recent introduction from the Iquitos area. It is interesting to note that the introduced *Arapaima* appear to have adapted well and now have a large enough population to be occasionally exploited by commercial fishermen.

### Conservation genetics

Despite severe overexploitation and depletion of genetic diversity in the centre of its distribution, *Arapaima* possesses a number of factors favouring its recovery. NCA indicates that *Arapaima* in the Amazon basin forms one large panmictic population, with some isolation-by-distance and fragmentation (Table 2). Population fragmentation indicated by NCA does not mean there is no gene flow, or no potential gene flow; in this case it merely reflects the appearance of differentiation caused by a bottleneck and by genetic drift. This phenomenon is also observed in the result of the MIGRATE analysis, where geographically adjacent populations often experience no gene flow, but some geographically extreme, genetically diverse populations experience gene flow levels equivalent to thousands of effective individuals that must first pass through their neighbours with whom no apparent genetic exchange exists (Table 4). Due to artificial differentiation of localities from one another driven by anthropogenic population fragmentation, overall gene flow levels are likely to be much higher than estimated from current data. A gene flow estimate for the species as a whole based on the probability-of-identity $F$ also suggests relatively high levels of among-locality gene flow (Island model $Nm \sim 1.5$; Stepping-stone model $Nm \sim 6.3$). Because of the violation of the genetic equilibrium assumption, these estimates underestimate the actual gene flow levels and overestimate the effects of isolation-by-distance. High gene flow is indirectly supported by *Arapaima*'s high vagility; radio-telemetry tracking experiments indicate that individuals migrate up to 35 km per month (Crossa & Petrere Júnior, 1999). The observed high vagility of individuals and high gene flow levels would seem to create opportunities for the colonisation of areas from which *Arapaima* has been extirpated previously. For example, Goulding (1979) reported that *Arapaima* became nearly extinct in the Madeira River system; however, it appears that once again *Arapaima* are being hunted by local fishermen.

Lack of strong population structuring inferred from mtDNA also suggests a lack of local adaptation and is indirectly confirmed by the Bolivian introduction case. Although selection could potentially operate on nuclear genes, selection would have to be very strong to overcome the high levels of observed mtDNA gene flow. Lack of population structuring and local adaptation is highly conducive to the re-colonisation of locally depleted populations by immigrants from other areas and also to managed re-introductions from any area with high genetic diversity, such as the Mamirauá sustainable biological reserve. However, adaptive genetic variation is not the same as neutral genetic variation, thus despite the fact that there is genetic interconnection between areas, re-introductions should not be done without additional research and careful monitoring.

### *In situ* conservation

In spite of Brazilian federal laws that prohibit all commercial fishing, *Arapaima* fishing continues throughout the year, potentially threatening the survival of this species. *Arapaima* populations around major Amazonian cities and in the centre of its distribution are genetically depleted and the whole population is in genetic disequilibrium. According to Crossa & Petrere Júnior (1999) there is a trend of decreased size at capture. In addition, some individuals less than 120 cm total length (< 17 kg of total weight) were already reproductive or were with juveniles, suggesting an adaptive life-history response to this heavy fishing pressure. *Arapaima* usually do not breed until reaching 150 cm or more of total length (Crossa & Petrere Júnior, 1999).

However, the future of *Arapaima* may not be as bleak as may appear at first. Our study also identified areas of relatively high genetic diversity that can act as

centres for recovery of genetic diversity. One such area is the Mamirauá sustainable development reserve located northwest of the town of Tefé. This area consists of a flooded forest with a series of lakes, the várzea type habitat preferred by *Arapaima*. *Arapaima* in Mamirauá is fished on a rotational basis, thus preventing overfishing in any one area (Queiroz, 1999; Queiroz & Crampton, 1999). This sound management combined with a genetically diverse founding population is positively reflected in our study, since Mamirauá has the highest genetic diversity of all the areas investigated. There are other areas of the Amazon basin that show high levels of genetic diversity such as the Tocantins River near its confluence with the Araguaia River (Marabá) and the upper Amazon of Peru (Iquitos and Letícia). Additional areas of potentially high genetic diversity not investigated by us, such as the recently created Piagaçu-Purús sustainable development reserve or the Meio Juruá extractive reserve, are likely to be present throughout the Amazon basin. However, with these exceptions, most areas are either unprotected or are in the form of national parks. In both cases there is little incentive for local people to exploit these areas in a sustainable manner and industrial fishing and poaching tend to be a serious problem. Therefore, we recommend additional studies to identify várzea forest regions that contain high genetic diversity and the subsequent creation of sustainable development reserves (Reservas de Desenvolvimento Sustentável – RDS) for the sustainable management and exploitation of *Arapaima* and other fauna and flora of the flooded forest.

### Captive-breeding programs

Captive breeding has been initiated as a source of farm-raised *Arapaima* for the Amazonian market. *Arapaima* is omnivorous, tolerates hypoxia well, grows rapidly and reproduces well under semi-natural conditions, thus lending itself to captive breeding efforts (Palmeira, 2001). Based on our data, these efforts also can draw from a large pool of founding individuals irrespective of their geographical origin, thus minimising the potential loss of genetic diversity in these captive populations *via* the founder effect. In addition, captive *Arapaima* can be managed to maintain maximum genetic diversity by periodically supplementing these captive populations with wild individuals from selected areas. The captive-bred populations themselves can also be used as a source of potential animals in a release or a supplementation program if this need ever arises. However, we would like to stress once again that local adaptation may exist and that genetic adaptation to captivity can be a serious problem in reintroduction programs. Therefore, although managing a captive population for maximum genetic diversity is commendable, we do not recommend the release of captive-bred individuals into the wild without additional research and then only if there is a serious need to do so. We recommend rather the creation of multiple additional sustainable biological reserves or extractive reserves throughout the Amazon basin that will maintain healthy, genetically diverse populations of *Arapaima* and

act as sources of immigrants to recolonise locally depleted areas.

### REFERENCES

Anderson, S. A., Bankier, T., Barrell, B. G., de Bruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R. & Young, I. G. (1981). Sequence and organization of the human mitochondrial genome. *Nature* **290**: 457–465.

Beerli, P. & Felsenstein, J. (1999). Maximum likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics* **152**: 763–773.

Beerli, P. & Felsenstein, J. (2001). Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proc. Natl. Acad. Sci. USA* **98**: 4563–4568.

Castelloe, J. & Templeton, A. R. (1994). Root probabilities for intraspecific gene trees under neutral coalescent theory. *Mol. Phylogenet. Evol.* **3**: 102–113.

Clement, M., Posada, D. & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* **9**: 1657–1659.

Crossa, M. & Petrere Júnior, M. (1999). Morphometric relationships and indirect determination of the length frequency structure of the pirarucu *Arapaima gigas* Cuvier. *Fish. Manage. Ecol*. **6**: 173–186.

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

Frankham, R., Ballou, J. R. & Briscoe, D. A. (2002). *Introduction to conservation genetics*. Cambridge: Cambridge University Press.

Fu, Y.-X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**: 915–925.

Goulding, M. (1979). *Ecologia de Pesca do Rio Madeira*. Manaus, Brazil: INPA.

Goulding, M. (1980). *Fishes and the forest*. Los Angeles, CA: University of California Press.

Goulding, M., Barthem, R. & Ferreira, E. J. G. (2003). *The Smithsonian atlas of the Amazon*. Washington, DC: Smithsonian Institution Press.

Goulding, M., Smith, N. J. H. & Mahar, D. J. (2000). *Floods of fortune*. New York, NY: Columbia University Press.

Hall, T. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp. Ser.* **41**: 95–98.

Hamilton, H., Caballero, S., Collins, A. G. & Brownell Jr, R. L. (2001). Evolution of river dolphins. *Proc. Roy. Soc. Lond. Ser. B.* **268**: 549–556.

Hartl, D. L. & Clark, A. G. (1997). *Principles of population genetics,* third edition. Sunderland, MA: Sinauer Associates.

Hasegawa, M., Kishino, H. & Yano, T. A. (1985). Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**: 160–174.

Hrbek, T. & Meyer, A. (2003). Closing of the Tethys Sea and the phylogeny of Eurasian killifishes (Cyprinodontiformes: Cyprinodontidae). *J. Evol. Biol.* **16**: 7–26.

Inoue, J. G., Miya, M., Tsukamoto, K. & Nishida, M. (2001). A mitogenomic prespective on the basal teleostean phylogeny: resolving higher-level relationships with longer DNA sequences. *Mol. Phylogenet. Evol.* **20**: 275–285.

Kuhner, M. K., Yamato, J. & Felsenstein, J. (1995). Estimating effective population size and mutation rate from sequence data using Metropolis–Hastings sampling. *Genetics* **140**: 1421–1430.

Kuhner, M. K., Yamato, J. & Felsenstein, J. (1998). Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics* **149**: 429–434.

Kumazawa, Y. & Nishida, M. (1993). Sequence evolution of mitochondrial tRNA genes and deep-branch animal phylogenetics. *J. Mol. Evol.* **37**: 380–398.

Kurlansky, M. (1998). *Cod: a biography of the fish that changed the world*. London: Jonathan Cape.

Lüling, V. (1964). Zur Biologie und Ökologie von *Arapaima gigas* (Pisces, Osteoglossidae). *Z. Morph. Öko. Tiere* **54**: 436–530.

Lundberg, J. & Chernoff, B. (1992). A miocene fossil of the amazonian fish *Arapaima* (Teleostei: Arapaimidae) from the Magdalena river region of Colombia: biogeographic and evolutionary implications. *Biotropica* **24**: 2–14.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**: 209–220.

Masta, S. E., Laurent, N. M. & Routman, E. J. (2003). Population genetic structure of the toad *Bufo woodhousii*: an empirical assessment of the effects of haplotype extinction on nested cladistic analysis. *Mol. Ecol.* **12**: 1541–1554.

MathSoft Inc. (2002). *MathCad*. Cambridge, MA: MathSoft Engineering & Education, Inc.

Menezes, R. S. (1951). Notas biológicas e econômicas sobre o pirarucu *Arapaima gigas* (Cuvier) (Actinopterygii, Arapaimidae). *Serv. Inform. Agrí./Minist. Agric. Ser. est. téc.* **3**: 9–39.

Nei, M. (1987). *Molecular evolutionary genetics*. New York, NY: Columbia University Press.

Nelson, J. S. (1994). *Fishes of the World, 3rd edition*. New York, NY: John Wiley and Sons, Inc.

Nunes-Pereira, A. (1934). A pesca no rio Purus. A voz do mar. *Bol. Minist. Agricultura* 178–183.

Palmeira, I. E. (2001). Production potential of pirarucu, *Arapaima gigas*, in captivity. *Acta Amazon.* **31**: 299–316.

Posada, D., Crandall, K. A. & Templeton, A. R. (2000). GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Mol. Ecol.* **9**: 487–488.

Pybus, O. G. & Rambaut, A. (2002). GENIE: estimating demographic history from molecular phylogenies. *Bioinformatics* **18**: 1404–1405.

Queiroz, H. L. (1999). Artisanal fisheries of Pirarucu at the Mamirauá Ecological Station. In: *Várzea: diversity, development, and conservation of Amazonia's whitewater floodplains. Section 1: Fish and fisheries*: 83–99. Padoch, C., Ayres, J. M., Pinedo-Vasquez, M. & Henderson, A. (eds). Bronx, NY: The New York Botanical Garden Press (NYBG).

Queiroz, H. L. & Crampton, W. G. R. (eds) (1999). *Estratégias para Manejo de Recursos Pesqueiros em Mamirauá*. Brasilia, Brazil: Sociedade Civil Mamirauá, MCT – CNPq.

Rand, D. M. (1996). Neutrality tests of molecular markers and the connections between DNA polymorphism, demography, and conservation biology. *Conserv. Biol.* **10**: 665–671.

Rogers, A. R. (1995). Genetic evidence for a Pleistocene population explosion. *Evolution* **49**: 608–615.

Rogers, A. R. & Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* **99**: 552–569.

Saint-Paul, V. (1986). Potential for aquaculture of South American freshwater fishes: a review. *Aquaculture* **5**: 205–240.

Schneider, S., Roessli, D. & Excoffier, L. (2000). *Arlequin ver. 2000: A software for population genetic data analysis*. Geneva, Switzerland: Genetics and Biometry Laboratory, University of Geneva.

Slatkin, M. (1994). An exact test for neutrality based on the Ewens sampling distribution. *Genet. Res.* **64**: 71–74.

Slatkin, M. (1996). A correction to the exact test based on the Ewens sampling distribution. *Genet. Res.* **68**: 259–260.

Strimmer, K. & Pybus, O. G. (2001). Exploring the demographic history of DNA sequences using the generalized skyline plot. *Mol. Biol. Evol.* **18**: 2298–2305.

Swofford, D. L. (2001). *PAUP\*. Phylogenetic Analysis Using Parsimony (\* and Other Methods)*, Beta Version v4.10b. Sunderland, MA: Sinauer Associates.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.

Templeton, A. R. (1998). Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Mol. Ecol.* **7**: 381–397.

Templeton, A. R., Crandall, K. A. & Sing, C. F. (1992). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data: III. Cladogram estimation. *Genetics* **132**: 619–633.

Templeton, A. R., Routman, E. & Phillips, C. A. (1995). Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics* **140**: 767–782.

Templeton, A. R. & Sing, C. F. (1993). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping: IV. Nested analyses with cladogram uncertainty and recombination. *Genetics* **134**: 659–669.

Veríssimo, J. (1895). *A Pesca na Amazônia*. Rio de Janeiro, Brazil: Livraria Clássica, de Alves & C.

Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theoret. Pop. Biol.* **7**: 256–276.

Watterson, G. A. (1978). The homozygosity test of neutrality. *Genetics* **88**: 405–417.

Watterson, G. A. (1986). The homozygosity test after a change in population size. *Genetics* **112**: 899–907.

Wright, S. (1951). The genetical structure of populations. *Ann. Eugen.* **15**: 323–354.

Wright, S. (1969). *Evolution and the Genetics of Populations, vol. 2. The Theory of Gene Frequencies*. Chicago, IL: University of Chicago Press.