Reproductive physiology of *Arapaima gigas* (Schinz, 1822) and development of tools for broodstock management

Lucas Simon Torati, BSc, MSc

July 2017

A Thesis Submitted for the Degree of Doctor of Philosophy

Institute of Aquaculture

University of Stirling

Scotland

Declaration

This thesis has been composed in its entirety by the candidate. Except where specifically acknowledged, the work described in this thesis has been conducted independently and has not been submitted for any other degree.

Candidate Name: Lucas Simon Torati

Signature: ……..

Date: ……..

Supervisor Name: Professor Hervé Migaud

Signature: ……..

Date: ……..

Abstract

Arapaima gigas is the largest scaled freshwater fish in the world reaching over 250 kg. With growth rates of 10 kg+ within 12 months, *A. gigas* is considered as a promising candidate species for aquaculture development in South America. However, the lack of reproductive control in captivity is hindering the industry expansion. The work carried out in this doctoral thesis therefore aimed to better understand the species' reproductive physiology, develop tools to identify gender and monitor gonad development, test hormonal therapies to induce ovulation and spawning and characterise the cephalic secretion for its potential roles in pheromone release and during parental care. Initially, a genomic study investigated the overall extent of polymorphism in *A. gigas*, which was found to be surprisingly low, with only 2.3 % of identified RAD-tags (135 bases long) containing SNPs. Then, a panel with 293 single nucleotide polymorphism (SNP) was used to characterise the genetic diversity and structure of a range of Amazon populations. Results revealed populations from the Amazon and Solimões appeared to be genetically different from the Araguaia population, while Tocantins population comprised individuals from both stocks. This data provided a tool for broodstock identification and future management. The PhD then aimed to evaluate the effects of slow-release mGnRH implants and different broodstock size pairings on maturation and spawning. Results showed that the implants stimulated the brain-pituitary-gonad axis resulting in increased plasma levels of testosterone (females) and 11-ketotestosterone in males, respectively regardless of pairing sizes. However, no spawning was observed. Results also showed the release of sex steroids with potential pheromonal action through the cephalic secretion, a biological fluid released from the adult head along the reproductive period. Thereafter, a non-surgical field endoscopy method was developed and validated for ovarian assessment and gender identification. The method was then used to describe the

female gonopore and obtain biopsy of the ovary through cannulation which allowed the description of oogenesis in *A. gigas.* Importantly, oocytes obtained by cannulation confirmed that adult females under investigation were maturing with oocytes in final maturation stage but failed to ovulate/spawn. Another hormonal induction trial was therefore performed in which a combination of GnRHa (mGnRHa/sGnRHa) was used by injection to induce ovulation and spawning in selected maturing females with effects on oocyte maturation monitored post-induction through biopsy. However, this trial appeared to not be successful at inducing ovulation or spawning. Finally, the peptidome and proteome of the cephalic secretion was further characterised through the comparison between parental and non-parental fish. Results highlighted the complex role of this biological fluid including potential roles on the developing offspring during the parental care period. Overall, this doctoral thesis provided new basic and applied data on *A. gigas* reproduction and tools that can be used in future studies to better understand the environmental and hormonal control of oogenesis and spawning.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my main supervisor Dr. Hervé Migaud, for his enthusiastic guidance, friendship, support and encouragement along this study both in Scotland and Brazil. I thank also Dr. John Taylor for his co-supervision, all support and corrections provided. I am also very grateful to Dr. Amaya Albalat for being so positive with her guidance, and advising me with the proteomic and peptidomic analyses and chapter corrections. Likewise, I'm grateful to Dr. John B. Taggart for his assistance with ddRAD library construction, MiSeq sequencing, genomic data processing, for providing a comparative sequencing dataset on other teleosts, advice and corrections in the chapter on population genomics. I also thank colleagues and friends who helped me with histology (Dr. Debbie Faichney and Dr. Luciana Ganeco), microscopy (Dr. James Bron), molecular biology, lab techniques and data analyses (Dr. Andrew Davie, Dr. Sarah-Louise Counter, Dr. Robyn Harris, Dr. Stefanie Wehner, Dr. Kate Howie, Dr. Rogelio Flores, Dr. Nicole Rhody, Dr. Rosiana Rodrigues). I wish to thank Dr. Ben Clokie also for precious help with R programming and for all his aid in Stirling during my studies.

I would also like to thank EMBRAPA for the funding and education opportunity, herein represented by Dr. Carlos Magno C. Rocha, Dr. Ariovaldo Luchiari, Dr. Eric Routledge and Dr. Alexandre A. Freitas. Very special thanks to Dr. Adriana Lima, Dr. Eduardo Varela and Dra. Luciana N. Ganeco for supporting the project in Brazil, and for all productive discussions we've had along this PhD. I thank Dr. Marta E. Ummus for elaborating maps in ArcGIS used to illustrate this thesis. Whereas in EMBRAPA Tropical Agroindustry (Fortaleza), I thank logistical and laboratorial support and advice received from Dr. Guilherme Zucolo, Dr. Edy S. Brito, Dr. Kirley Canuto, Dr. Lorena Maia, Dr. Tigressa Rodrigues and Dr. Patricia Bordallo. Special thanks also to Dr. Celli Muniz and

Dr. Luciana Ganeco for aid with SEM analyses of eggs. Still in Fortaleza, I was aided by Dr. Arlindo Moura and Dra. Paula Rodriguez who helped with preliminary proteomic analyses in Federal University of Ceará. Broodstock and infrastructure used along this thesis were provided by DNOCS, particularly Dr. Pedro Eymard C. Mesquita; and Mirador farm, represented by Marcelo Borba and Juarez Santos, to whom I owe my gratitude.

I also thank Dr. Juliana Araripe and Dr. Eduardo S. Varela for providing samples which enabled the genomic analyses of Arapaima wild populations (Chapter 2). Also, slow-release implants used in Chapter 3 were made available thanks to Dr. Yonathan Zohar and Dr. John Stubblefield. I wish to thank also H. Strattner, Karl Storz and Ana Paula Varges for providing the endoscopy system used in Chapter 4, and Dr. Adriana Lima who strongly supported the experiments performed in Chapter 4 – Parts B and C with also important ideas and discussions on Arapaima cannulation. Also, thanks to Dr. Willian Mullen (University of Glasgow), Dr. Mary K. Doherty (University of the Highlands and Islands) and Justyna Siwy (Mosaiques Diagnostics GmbH) who provided their analytical facilities and support with CE-MS and LC-MS/MS analyses. I thank also important help received from very skilled colleagues along fish netting and handling, namely Rogério Miranda Araújo, Francisco Valdécio M. de Sousa, Valberto M. de Sousa, Luiz Oliveira, Agenor Galvão, Joacir Xavier da Silva, Tiago Vieira da Costa, Emilio Pinho and staff from Mirador farm.

I acknowledge and thank financial support received from EMBRAPA, MAPA-CNPq (Grants. N. 457465/2012-3 and 434400/2016-5) and SEBRAE (FAPTO Grant N. 2538/2012).

Finally, I wish to thank all my family for supporting me along this project. I'm deeply grateful to my beloved wife Melina Cambi for her constant and unconditional

support throughout this journey, accepting to move from a city to another (and from a job to another), for aiding me in moments of stress, and for facing the new challenges together in her joyful manner.

Table of Contents

LIST OF ABBREVIATIONS

11-KT — 11-ketotestosterone

17α-OHP — 17α-hydroxyprogesterone

20β-S — 17α,20β,21-trihydroxy-4-pregnen-3-one

A1AT — α-1-antitrypsin homolog

ACN — acetonitrile

amu — atomic mass unit

ANOVA — analysis of variance

apoA-I — apolipoprotein A-I

BCA — bicinchoninic acid assay

BH — Benjamini and Hochberg

BIC — Bayesian Inference Criterion

bp — base pair

BPG — brain-pituitary-gonad

BW — body weight

C3 — complement C3

C8B — complement component C8 beta chain

CE-MS — capillary electrophoresis coupled to mass spectrometry

CID — collision induced dissociation

CITES — Convention on International Trade of Endangered Species of Wild Fauna and Flora

CN — chromatin-nucleolus

CPE — carp pituitary extract

CSF — cerebrospinal fluids

DAPC — discriminant analysis of principal components

ddRAD — double digest restriction-site-associated DNA

DHP — 17α,20β,dihydroxy-4-pregnen-3-one

DNA — deoxyribonucleic acid

dpi — days post implantation

dps — days post spawning

DTT — dithiothreitol

E_2 — 17β-oestradiol

EIA — enzyme immune assay

ELISA — enzyme-linked immunosorbent assay

emPAI — exponentially modified Protein Abundance Index

ESI — electrospray ionization

EVAc — ethylene-vinyl acetate polymer

FDR — false discovery rate

FGF-3 — fibroblast growth factor 3

F_{IS} — coefficient of inbreeding

FOM—final oocyte maturation

FSH — follicle stimulating hormone

FSH-R — follicle stimulating hormone receptor

F_{ST} — fixation index

GABA — gamma-amino-butyric acid

GBS — genotyping by sequencing

GDF6 — growth/differentiation factor 6-A

GeLC-MS/MS — gel electrophoresis liquid chromatography-mass spectrometry/MS

GH — growth hormone

GnIH — Gonadotropin-inhibitory hormone

GnRH — gonadotrophin releasing hormone

GnRHa — gonadotropin releasing hormone analogue

GO — gene ontology

HCDMS2 — high energy collision dissociation in mass spectrometer 2

HCG — human chorionic gonadotrophin

H_E — expected heterozygosity

HE — haematoxylin – eosin

H_O — observed heterozygosity

HPLC — high performance liquid chromatography

HWE — Hardy-Weinberg Equilibrium

I — Shannon's Information Index

IGF — insulin-like growth factor

IP — intraperitoneal

IPP — inositol pentaphosphate

ISSR — inter simple sequence repeats

K — Fulton’s condition factor

KiSS — kisseptin

LC — leading cohort

LD — linkage disequilibrium

LH — luteinizing hormone

LH-R — luteinizing hormone receptor

MAC — membrane attack complex

MCMC — Markov chain Monte Carlo

micro-TOF — micro-Time of Flight

MIS — maturation inducing steroids

MS — mass spectrometry

MT — metric tons

mtDNA — mitochondrial DNA

MW — molecular weight

NGS — next generation sequencing

NJ — Neighbor-Joining

OD — oocyte diameter

OM — oocyte maturation

OV — ovulation

PC — principal components

PCR — polymerase chain reaction

PG — primary growth

PIT — passive integrated transponder

PRL — prolactin

qPCR — quantitative PCR

QTL — quantitative trait loci

RAD — restriction-site-associated DNA

RIA — radioimmunoassay

SD — standard deviation

SDS — sodium dodecyl sulphate

SE — standard error of the mean

SEM — scanning electron microscopy

SG — secondary growth

SNP — single nucleotide polymorphism

SNR — signal to noise

STC — stanniocalcin

T — testosterone

TF — serotransferrin

TL — total length

UP — urogenital papilla

UPLC — Ultra performance liquid chromatography

VLDL — very low-density lipoproteins

VTG — vitellogenin

LIST OF SPECIES

Anguilla japonica (**Japanese eel**)

Anhinga anhinga (**Snake bird**)

Arapaima agassizii (**Arapaima**)

Arapaima arapaima (**Arapaima**)

Arapaima gigas (**Arapaima**)

Arapaima leptosoma (**Arapaima**)

Arapaima mapae (**Arapaima**)

Argyrosomus regius (**Meagre**)

Brycon orbignyanus (**Piracanjuba**)

Centropomus undecimalis (**Common snook**)

Cichlasoma citrinellum (**Midas cichlid**)

Clarias anguillaris (**Mudfish**)

Clarias gariepinus (**African catfish**)

Ctenolabrus rupestris (**Goldsinny**)

Cyprinus carpio (**Common Carp**)

Danio rerio (**Zebrafish**)

Gadus morhua (**Atlantic Cod**)

Gymnarchus niloticus (**African electric fish**)

Heterotis niloticus (**African arowana**)

Hippoglossus hippoglossus (**Halibut**)

Melanogrammus aeglefinus (**Haddock**)

Melanosuchus niger Amazon (**Alligator**)

Morone saxatilis (**Striped Bass**)

Oncorhynchus mykiss (**Rainbow Trout**)

Oreochromis mossambicus (**Mozambique tilapia**)

Oreochromis niloticus (**Nile Tilapia**)

Panthera onca (**Jaguar**)

Phalacrocorax brasilianus (**Biguá**)

Pleuronectes ferrugineus (**Yellowtail Flounder**)

Poecilia reticulata (**Guppy**)

Pseudoplatystoma corruscans (**Amazon catfish**)

Salmo salar (**Atlantic salmon**)

Dicentrarchus labrax (**European Seabass**)

Erythrinus erythrinus (**Traíra**)

Esox lucius (**Pike**)

Sprattus sprattus (**Sprat**)

Symphysodon aequifasciata (**Discus fish**)

Thunnus thynnus (**Atlantic Bluefin Tuna**)

Scophthalmus maximus (**Turbot**)

Solea senegalensis (**Senegalese sole**)

Sparus aurata (**Gilthead Seabream**)

Umbra krameri (**Mudminnow**)

Xyrauchen texanus (**Razorback sucker**)

LIST OF FIGURES AND TABLES

Chapter 1: Figures

from pituitary, controlling vitellogenesis and spermatogenesis through synthesis of sexual steroids. Synthesis of melatonin in the pineal can interact with the BPG axis through dopaminergic inhibition of FSH and LH production. Energy homeostasis messengers such as insulin-like growth factor (IGF), growth hormone (GH), ghrelin and leptin can interact with reproductive cycle through unclear pathways. Figure extracted from Migaud et al. (2010).

Chapter 2: Figures

Figure 2.1. Map of Amazon region depicting approximate natural distribution of *Arapaima gigas* after Castello & Stewart (2010) (*by* Marta E. Ummus using software ArcGIS). Sampling points include two wild populations from Amazon river Basin (Amazon - 1 and Solimões -2), two wild populations from Tocantins-Araguaia Basin (Tocantins – 3 and Araguaia - 4), and one captive population (Captive - 5) from Tocantins state, Brazil.

Figure 2.2. Summary of ddRAD sequencing for *Arapaima gigas*, scheme modified from Brown et al. (2016). Workflow of data processing from the obtained raw reads (upper disk) down to the markers used to investigate genomic diversity and structure in populations.

Figure 2.3. Loci analysis for 392 SNPs across 60 individuals analysed across the five populations of *Arapaima gigas.* A. Frequency distribution of F_{ST} values. B. Detection of loci under selection (outlier) using hierarchical structure model implemented in Arlequin v. 3.5.2.2. Outlier loci are indicated in red ($P<0.01$) and blue ($P<0.05$) dots.

Figure 2.4. Within population pairwise mean relatedness (r) for *Arapaima gigas* (n=12 individuals each population). Calculations followed method of Lynch & Ritland (1999) with confidence intervals of 95% denoted by the U (upper) and L (lower) marks,

calculated after 1,000 bootstrap resamplings and 1000 permutations in GenAlEx version 6.5.

Figure 2.5. Bayesian clustering representation for populations of *Arapaima gigas* using 392 SNP markers in STRUCTURE v. 2.3.4 (Pritchard et al., 2000). A. Analysis of five populations (n=60 individuals; optimal Evanno's K = 2). B. Graphical representation of optimal number of clusters (K) across the five studied populations determined by Evanno's method (Evanno et al., 2005), indicated by delta K peaking at K=2. C. Geographical representation of structure results for the global population. D. Structure analysis for Amazon river basin (Amazon and Solimões, optimal Evanno's K = 2). E. Structure analysis for Araguaia-Tocantins basin (Tocantins, Araguaia and Captive, optimal Evanno's K = 2).

Figure 2.6. A. Discriminant analysis of principal components (DAPC) using 392 SNP markers in *Adegenet* v. 2.0.1 (Jombart, 2008) for five population samples of *Arapaima gigas* (n = 60 individuals). B. Selection of number of cluster was based on value of Bayesian Inference Criterion (BIC), which indicates 3 clusters for data summarization (elbow drop). C. Membership probabilities (red = 1, white = 0) for individuals into clusters, blue crosses indicate the prior cluster provided into DAPC.

Figure 2.7. Neighbor-Joining phylogram showing relationship among 60 individuals of *Arapaima gigas* built in MEGA version 7.0.21. The branch length of the optimal tree is 3,574.9. The tree is drawn to scale, with branch lengths (next to the branches) based on a matrix of genetic distance (GD) among individuals produced in GenAlEx v 6.5.

Chapter 2: Tables

Table 2.1. Studied populations of *Arapaima gigas*. Information on site locations and geographical coordinates to four wild (Amazon, Solimões, Tocantins and Araguaia) and

into ponds) to day 181 (119 days post GnRHa implantation) in the four experimental groups: A. Control couples (n = 5), B. Large implanted couples (n = 5), C. Small implanted couples (n = 3) and D. Mixed size implanted couples (n = 5). Values are presented as mean ± SE. Lowercase and uppercase superscripts denote time effects in blood plasma and cephalic secretion levels, respectively ($P<0.05$). Arrows indicate GnRHa implantation time.

Figure 3.7. Sex steroid correlations in male *Arapaima gigas* broodstock. A. Correlation between blood plasma levels ($ng.ml^{-1}$) of Testosterone (T) and 11-ketotestosterone (11-KT). B. Correlation between 11-KT levels in blood plasma ($ng.ml^{-1}$) and cephalic secretion ($pg.ml^{-1}$). Pearson Product Correlation Coefficients were calculated on log-transformed data.

Figure 3.8. Levels of testosterone (T) in the plasma ($ng.ml^{-1}$) and cephalic secretion ($pg.ml^{-1}$) in females of *Arapaima gigas* from day 0 (couple pairing and stocking into ponds) to day 181 (119 days post GnRHa implantation) in the four experimental groups: A. Control couples (n = 5), B. Large implanted couples (n = 5), C. Small implanted couples (n = 3) and D. Mixed size implanted couples (n = 5). Values are presented as mean ± SE. Lowercase and uppercase superscripts denote time effects in blood plasma and cephalic secretion levels, respectively ($P<0.05$). Arrows indicate GnRHa implantation time.

Figure 3.9. Levels of 17β-oestradiol (E_2) in the plasma ($ng.ml^{-1}$) and cephalic secretion ($pg.ml^{-1}$) in females of *Arapaima gigas* from day 0 (couple pairing and stocking into ponds) to day 181 (119 days post GnRHa implantation) in the four experimental groups: A. Control couples (n = 5), B. Large implanted couples (n = 5), C. Small implanted couples (n = 3) and D. Mixed size implanted couples (n = 5). Values are presented as mean ± SE. Arrows indicate GnRHa implantation time.

Chapter 4: Part B - Tables

Table 4.B.2. Endoscopy procedure in adult females of *Arapaima gigas*. Females are numbered as in presented in Figure 7, and differently from Table I which correspond to broodstock animals examined under field operational conditions. Female weight (kg), total length (TL; cm), time elapsed to observe ovary (minutes), ovary predominant colour, diameter and stage of leading cohort (LC, µm) and ovary stage *according to Núñez & Duponchelle (2009). OMgvm = Oocyte maturation germinal vesicle migration, SGfg = Secondary growth full-grown, OV = Ovulation.

Chapter 4: Part C - Figures

Figure 4.C.1. Intraperitoneal (IP) injection of mGnRHa and sGnRHa 50:50 (v:v) at the base of the right pelvic fin in *Arapaima gigas* broodstock.

Figure 4.C.2. Changes in mean oocyte diameter of leading oocyte cohort (LC) between baseline (day 0) and 12 days post GnRHa injections (day 15) from injected and non-injected females of *A. gigas* treated with a combination of mGnRHa and sGnRHa 50:50 (v:v). Different letters denote significant change in female LC between sampling dates (Paired t-tests; $P<0.05$).

Figure 4.C.3. Change in plasma sex steroid levels ($ng.ml^{-1}$) after 12 days post GnRHa injections in *A. gigas*. No significant difference (2-sample t-test; $P>0.05$) was seen comparing non injected and injected groups treated with a combination of mGnRHa and sGnRHa 50:50 (v:v). A. 17β-Oestradiol (E_2) in females. B. 11-ketotestosterone (11-KT) in males.

Chapter 5: Figures

Figure 5.1. Compiled CE-MS peptide fingerprints from the cephalic secretion of *Arapaima gigas.* A. Non-parental males (M; n = 9). Parental males (PM; n = 5). C. Non-

parental females (F; n = 11). D. Parental females (PF; n = 5). Migration time of capillary electrophoresis (CE) is shown in X axis whereas Y is the logarithmic scale of molecular mass (kDa) and Z is the mean signal intensity.

Figure 5.2. Gene ontology (GO) comparison of peptidome (7009 peptides) and proteome (422 proteins) identified in the cephalic secretion of *Arapaima gigas*. A, B and C. Peptidome GO for biological process, molecular function ad cellular component, respectively. D, E and F. Proteome GO for biological process, molecular function ad cellular component, respectively. Analyses were conducted in STRAP *v. 1.5* (Bhatia et al., 2009).

Figure 5.3. Number of unique and shared proteins in the cephalic fluid of *Arapaima gigas* comparing parental (PM) and non-parental males (M), and parental (PF) and non-parental females (F). A. Total of 422 proteins catalogued after GeLC-MS/MS and the 28 secreted (extracellular) proteins (B) revealed after GO analysis. Venn diagrams were produced in Venny (Oliveros, 2007).

Chapter 5: Tables

Table 5.1. List of 28 secreted proteins (extracellular) present in the proteome of *Arapaima gigas.* Putative functions, detected groups and relative measure of concentration (emPAI) are given for pools of males (M), females (F), parental males (PM) and parental females (PF).

Table S5.1. List of peptides. Peptides identified with CE-MS and LC-MS/MS and comparisons among studied groups.

Table S5.2. List of proteins. Information on the 422 proteins identified through GeLC-MS/MS in all studied groups.

CHAPTER 1

1. Chapter 1

GENERAL INTRODUCTION

1.1. Biology of *Arapaima gigas* and prospects for aquaculture

1.1.1. Natural history

The order Osteoglossiformes is a primitive group in the evolution of teleosts, encompassing species unique for having a bony plaque "tongue" forming a biting apparatus with the parasphenoid bone in the mouth, and displaying elongated dorsal and anal fins (Nelson et al., 2016). Some of these species are emblematic for their unusual biology, such as the insectivorous Arowanas (*Osteoglossum* spp.), the African elephant nose (*Mormyrus* spp.), the African butterfly fish (*Pantodon* spp.) and the African electric fish *Gymnarchus niloticus* (Nelson et al., 2016). Within the Osteoglossiformes, the Osteoglossidae family includes four genera and sixteen freshwater species with circumtropical distribution, characterised by the presence of a suprabranchial organ that enable these aquatic animals to breath atmospheric air (Nelson et al., 2016). Among the osteoglossids, the giant Amazon *Arapaima gigas* (Schinz, 1822) (Fig. 1.1) is considered as the largest scaled fish in the world with adults reaching up to 250 kg and 2.5 meters in total length (TL) (Nelson et al., 2016). In Brazil, *A. gigas* is popularly known as the *Pirarucu* ("red fish"), in Peru as *Paiche* and in Guyana as *Warapaima* (Queiroz, 2000). Previous phylogenetic studies found that the closest living species to *A. gigas* is the African *Heterotis niloticus*, and palaeontological records indicated their separation occurred before the Afro-South American drift, with the *Arapaima* lineage dating from the Jurassic period (Lundberg & Chernoff, 1992; Zhang, 2006; Lavoue, 2016). Several investigations have shown primitive and unique characteristics in *A. gigas*, such as the presence of inositol pentaphosphate (IPP) in the blood, a compound widespread in reptiles, birds and few elasmobranchs, but not detected in any other teleosts studied so far; or also unique profiles of bile alcohols and acids not found before in any other vertebrate species (Haslewood & Tökés, 1972; Val, 2000; Sato et al., 2016).

Figure 1.1. Lateral view of an adult male weighting 55 kg illustrating the reddish colour pattern of *Arapaima gigas*.

The natural geographic distribution of *A. gigas* spreads from the tributaries of the Amazon, Essequibo and Araguaia-Tocantins rivers in Brazil to Colombia, Ecuador, Guyana and Peru (Fig. 1.2) (Ortega et al., 2006; Castello & Stewart, 2010). Commercialised for its flesh in aquaculture and as an ornamental species over the past few decades, *A. gigas* was introduced in the Brazilian states of Bahia, Ceará, Espírito Santo, Minas Gerais, Rio Grande do Norte, Pernambuco and São Paulo (Fontenele, 1948; Oliveira et al., 2012; Carvalho et al., 2015). Accidentally, escapes and establishment of *A. gigas* into non-indigenous areas occurred in the upper Paraná river basin in southeast Brazil, southeast areas of Peru and northern Bolivia (Miranda-Chumacero et al., 2012; Carvalho et al., 2015; Damme et al., 2015). However, information regarding population sizes and potential impacts of these introductions are lacking. Worldwide introductions of *A. gigas* for ornamental, recreational and aquaculture projects have been reported in China, Cuba, Mexico, Philippines, Singapore, Thailand and USA (Lawson et al., 2015; FAO, 2016).

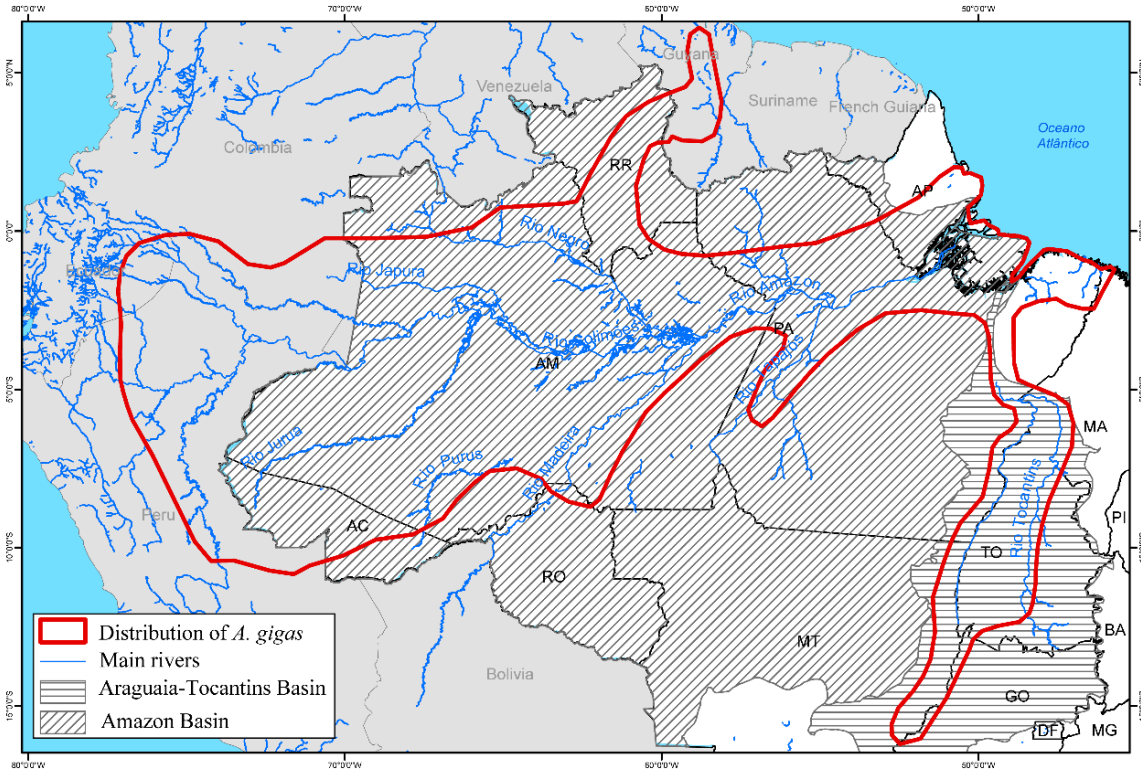

Figure 1.2. Natural distribution of *Arapaima gigas* in the Neotropical region demarked by red lines according to Castello & Stewart (2010), depicting Araguaia-Tocantins basin (horizontal lines background) and Amazon Basin (diagonal lines background). Map elaborated by Marta E. Ummus using ArcGIS.

Arapaima has been considered as a monotypic genus with *A. gigas* since 1868. However, a taxonomic revision based on morphological characters (teeth number, orbit diameter, anal fin ray number among other meristic characters) concluded *A. agassizii* should be considered as a valid distinct species, though known only from its holotype collected from the "lowlands of the Brazilian Amazon" (Stewart, 2013b). Likewise based on a comparative morphological analysis, the new species *A. leptosoma* was described with material collected from the confluence of Solimões and Purus rivers in Amazonas State, Brazil (Stewart, 2013a). In these studies, the validity of *A. mapae* and *A. arapaima* was also suggested but the conclusions were limited by the number of samples analysed

(Stewart, 2013a; Stewart, 2013b). On the other hand, such new taxonomic revisions are also not supported by previous phylogeographic investigations which used two discontinuous mitochondrial DNA regions of 1204 base-pairs (bp) (NADH1 segment) and 1143 bp (ATPase segment) (Hrbek et al., 2005), and 14 variable microsatellite markers on a larger number of specimens collected in the Amazon, Solimões and Tocantins rivers, whose results suggested a single panmictic population (Hrbek et al., 2005; Hrbek et al., 2007; Araripe et al., 2013). Further studies on population genetics analysing samples from rivers Essequibo and Araguaia-Tocantins using inter simple sequence repeats (ISSR) and microsatellite markers did not again support multiple species in *Arapaima* given limited sampling assemblage, though patterns of structuration were present in such populations (Vitorino et al., 2015; Watson et al., 2016). In these recent publications, it has been suggested that comparisons between populations from different river basins using other nuclear molecular markers could help clarify whether or not *Arapaima* comprises a panmictic population (Watson et al., 2016). In this thesis, the species name *A. gigas* will be used while a complete revision of *Arapaima* awaits publication to sort out existing doubts on species diversity.

1.1.2. General physiology and life-history

Air-breathing mechanisms evolved multiple times in separate fish lineages, possibly in response to aquatic hypoxia in freshwater habitats, and such mechanisms have fascinated researchers for more than a century (Lefevre et al., 2014). In *A. gigas*, the swimbladder has been modified as a highly vascularised lung with breathing requiring a buccopharyngeal pump (Greenwood & Liem, 1984). During larval development, the gill lamellae become progressively reduced in size, consequently decreasing their importance for oxygen uptake (Gonzalez et al., 2010; Ramos et al., 2013). It is estimated that the

respiratory surface of the swimbladder is 2.8 times greater than the gill, consequently 78 % of oxygen uptake in adult fish derives from aerial breathing, and this value can reach 100 % when animals are exposed to anoxic environments (Randall et al., 1978; Stevens & Holeton, 1978; Brauner & Val, 1996; Fernandes et al., 2012). For this reason, adult *A. gigas* need to gulp air at intervals ranging from 5 to 15 minutes and the gill branches are still responsible for removing 85 % of carbon dioxide into the water (Brauner & Val, 1996).

Overall, adults of *A. gigas* inhabit preferably the várzea ecosystem of the Amazon biome. They are described as solitary and non-migratory, and found normally in shallow depth areas (~3 m) with low currents (Castello, 2008a). Their optimal temperature ranges from 28-30 °C, with temperatures of 16 °C or less being lethal (Castello, 2008a; Lawson et al., 2015). Results from a capture-recapture experiment detected migrations distances of 64 km, and released animals were also recaptured in the same releasing site five years later suggesting site fidelity and territoriality in the species (Queiroz, 2000). Likewise, a study monitoring animals using radio-telemetry in a Peruvian lake of 38 km^2, found strong patterns of residency and territoriality, with a maximum home-range of 4 km^2 (Núñez et al., 2015). Using a different approach with molecular markers to estimate the dispersal capacity of *A. gigas,* a study showed high levels of genetic similarity among lakes separated by 25 km, moderated genetic differentiation in sites separated by 100 km and highest genetic differentiation among regions separated by more than 1,500 km, which taken together indicate *A. gigas* do not migrate long distances in wild environments (Araripe et al., 2013).

A. gigas is a dioecious and oviparous species that reaches sexual maturity after three to five years of age, when females are measuring 145-154 cm and males 115-124 cm in TL (Godinho et al., 2005). The reproductive period coincides with the rainy season,

ranging from November to May in the Amazon region (Queiroz, 2000). After mating and spawning on nests (details in section 3.1.), a characteristic male parental care phase involves a lateral migration towards the inundated areas where the fry can find abundant food items (Castello, 2008a). The larval stages of *A. gigas* have not been described yet. However, it is known that after consumption of the yolk sac, larvae feed mostly on zooplankton, and juveniles (26 cm; TL) prey actively on microcrustaceans, insects and gastropods (Oliveira et al., 2005; Halverson, 2013). During the early developmental stages, the fry and juveniles are predated mainly by birds such as the *Anhinga anhinga* and *Phalacrocorax brasilianus*, or by carnivorous fishes such as the Piranhas *Serrasalmus* spp. and the Traíra *Erythrinus erythrinus* (Neves, 1995). During parental care, the male's head becomes darkened, and this physiological change is believed to provide camouflage for the offspring which shoal over the male's head. Sex ratio appeared to be balanced at hatch, but given a higher vulnerability of males to fisherman during the parental care stage, there is a preponderance of females in natural populations (Queiroz, 2000). With the start of the dry season, parental care ceases after approximately three months when the offspring (0.5 m; TL) disperse towards the lateral canals and main rivers, while adults tend to remain in the várzea lagoons until the next flooding season (Castello, 2008a).

During the droughts (Sept-Nov), water levels are the lowest in the várzea lagoons reducing water quality (Castello, 2008a). As an air-breather, *A. gigas* is well adapted to anoxic waters, and physiological studies showed the species can tolerate high ammonia levels (Cavero et al., 2004). With adaptations for harsh drying lagoons, *A. gigas* also find few competitors for the high abundance of detritivorous fish species that comprise their main feeding items during the dry season (Castello, 2008a). This abundant nutritional provision during the droughts might be essential in the conditioning for reproduction in

the next flooding period. In the food web, *A. gigas* is an omnivorous species better defined as a secondary consumer (Watson et al., 2013). Studies analysing the stomach content found *A. gigas* (96 - 202 cm; TL) prey mostly on loricariids, callichthyids, pimelodids, crabs and prawns, and there is a piscivory trend along with growth (Watson et al., 2013; Carvalho et al., 2017b). In adults, the cycloid scales of *A. gigas* thickens massively being adapted to resist brutal attacks of Piranhas (Meyers et al., 2012). Therefore, the main natural predators of adult Arapaima will include the jaguar *Panthera onca* and the Amazon alligator *Melanosuchus niger* (Migdalski, 1957). Despite being often described as long-lived, longevity of *A. gigas* has not been reported so far but a method was proposed to estimate age by counting the number of scale rings (Arantes et al., 2010). However, such a method cannot be applied to captive broodstocks which are not submitted to the natural flooding rhythms.

1.1.3. Prospects for aquaculture

The culture in South America is influenced by prior Portuguese and Spanish colonisations, where the consumption of salted Atlantic Cod *Gadus morhua* is historically marked. For a short period during the 19th century, *A. gigas* became a cheaper substitute for the imported cod, being commercialised salted in most commercial centres and popularised as the “Brazilian cod” (Nunes et al., 2012). With an increased internal demand, the pressure on the natural stocks intensified and most populations became depleted and even extinct in many Amazon localities. Estimates indicated actual populations would only comprise 13 % of the original stocks (Castello et al., 2011; Cavole et al., 2015). Since 1986, *A. gigas* has been included in the Convention on International Trade of Endangered Species of Wild Fauna and Flora (CITES) and in Brazil, fishery is closed from 1st December until 31st of May with a minimum size of

capture limited to 150 cm (TL) (Arantes et al., 2010; Castello & Stewart, 2010). Even with such restrictions, there is an overall inability of enforcing the law in most remote Amazon areas where active fishing takes place, and it has been estimated that 77 % of commercialised *A. gigas* is coming from the wild illegally with size of captured fish being reduced over time (Cavole et al., 2015).

In aquaculture, attempts to control reproduction of *A. gigas* have been hindered, with several bottlenecks (detailed in section 1.3.3.) limiting the regular seed supply for a fledgling industry. Even in regions where spontaneous captive reproduction is considered successful such as Loreto and Ucayali in Peru, the mean annual production recorded was only 117,821 alevins from 2010 to 2016, most of them being commercialised in the international ornamental market (Alvan-Aguilar et al., 2016). Concerning the aquaculture sector, there is still lack of specific feed technologies to supply the grow out phase. Nevertheless, given the natural biology of the species, production systems such as earth ponds, circular systems and net cages are suitable for *A. gigas* which can have growth rates of 10 kg+ in 12 months (Wosnitza-Mendo, 1984; Imbiriba, 2001; Pereira-Filho et al., 2003; Bocanegra et al., 2004; Oliveira et al., 2012).

The meat of *A. gigas* lacks intramuscular bones, has low fat levels and the species has a fillet yield close to 58 % (Fogaça et al., 2011). Given these attributes and the meat colour, texture and taste, a market study concluded that Arapaima "offers without doubt great perspectives for French, UK and Swiss markets" (Mueller, 2006). Not only the meat of *A. gigas* has great value, but also its by-products (*i.e.* scales, skin). Analyses of Arapaima scales demonstrated biomaterial and biomechanical properties with promising industry applications, and its leather is already commercialised in the weaving sector within a luxurious niche market (Lin et al., 2011; Meyers et al., 2012; Yang et al., 2013; Yang et al., 2014; Sherman et al., 2016). Despite the strong market potential, the Brazilian

meat production of *A. gigas* remains low with only 2,300 metric tons (MT) produced in 2013 (Lima et al., 2015a) contrasting with Atlantic cod which remained from 2012 to 2016 the second most imported fish species into Brazilian markets with a mean annual importation of 36,592 MT (MICES, 2017). *A. gigas* has therefore a great potential, however reproduction in captivity and supply of large numbers of juveniles to the grow out sector is considered as the main limiting factor.

1.2. Physiology of fish reproduction

1.2.1. Oogenesis

The identification of gonadal developmental stages in individual broodstock is a key tool used in aquaculture to identify fish reproductive condition and characterise potential reproductive dysfunctions in captivity. The term oogenesis refers to the morphological processes by which primordial germ cells (oogonia) develop into fully mature eggs ready for fertilisation (Grier et al., 2009; Lubzens et al., 2017). During oogenesis in teleosts, there are key oocyte characteristics, common to most species, used to identify stage of development, and there are many different schemes proposed so far, each using different terminologies (Wallace & Selman, 1981; Patiño & Sullivan, 2002; Núñez & Duponchelle, 2009). Grier et al. (2009) proposed a classification scheme which aims to use putatively homologous characteristics common to teleosts for the classification of oogenesis. This classification divides oogenesis in four periods of germ cell development: mitosis, active meiosis, arrested meiosis and second period of active meiosis. Also, six steps subdivided into stages are used, illustrated in Figure 1.3.

Along oogenesis, initially oogonia proliferates (OP) through mitosis forming oogonia nests. Oogonia then divide by active meiosis comprising the step of chromatin-nucleolus (CN), when folliculogenesis occurs (Patiño & Sullivan, 2002). Folliculogenesis

corresponds to the differentiation of the somatic cells surrounding the oogonia into follicles known as theca and granulosa. This is a key event that marks the start of the vitellogenesis process allowing the egg to grow and accumulate yolk. At this stage, the hormonal communication between oocyte and follicle, and also between theca and granulosa cells is initiated (Lubzens et al., 2017). The follicles comprise the oocyte and the follicle cell layer, being separated from the germinal epithelium and stroma by a basement membrane (Grier et al., 2009). Oogenesis enters the third period when meiosis is arrested at late diplotene at the start of primary growth (PG). During PG, there is intense synthesis of RNA, oil droplets and cortical alveoli first appear in most species and oocyte size increases markedly (Babin et al., 2007b; Lubzens et al., 2017). During the secondary growth (SG) or true vitellogenesis, vitellogenin (VTG) produced in the liver is transported through the blood and sequestered by endocytosis into the oocytes, being the precursors of lipoprotein yolk which accumulates in the ooplasm along with a large deposition of oil droplets causing significant increase in oocyte volume (Babin et al., 2007b; Lubzens et al., 2017). At the end of SG, oil droplets start to coalesce and form oil globules in some species, and the vitelline envelope becomes thicker with the micropyle, usually one in most species being formed at the animal pole (Grier et al., 2009; Bian et al., 2014).

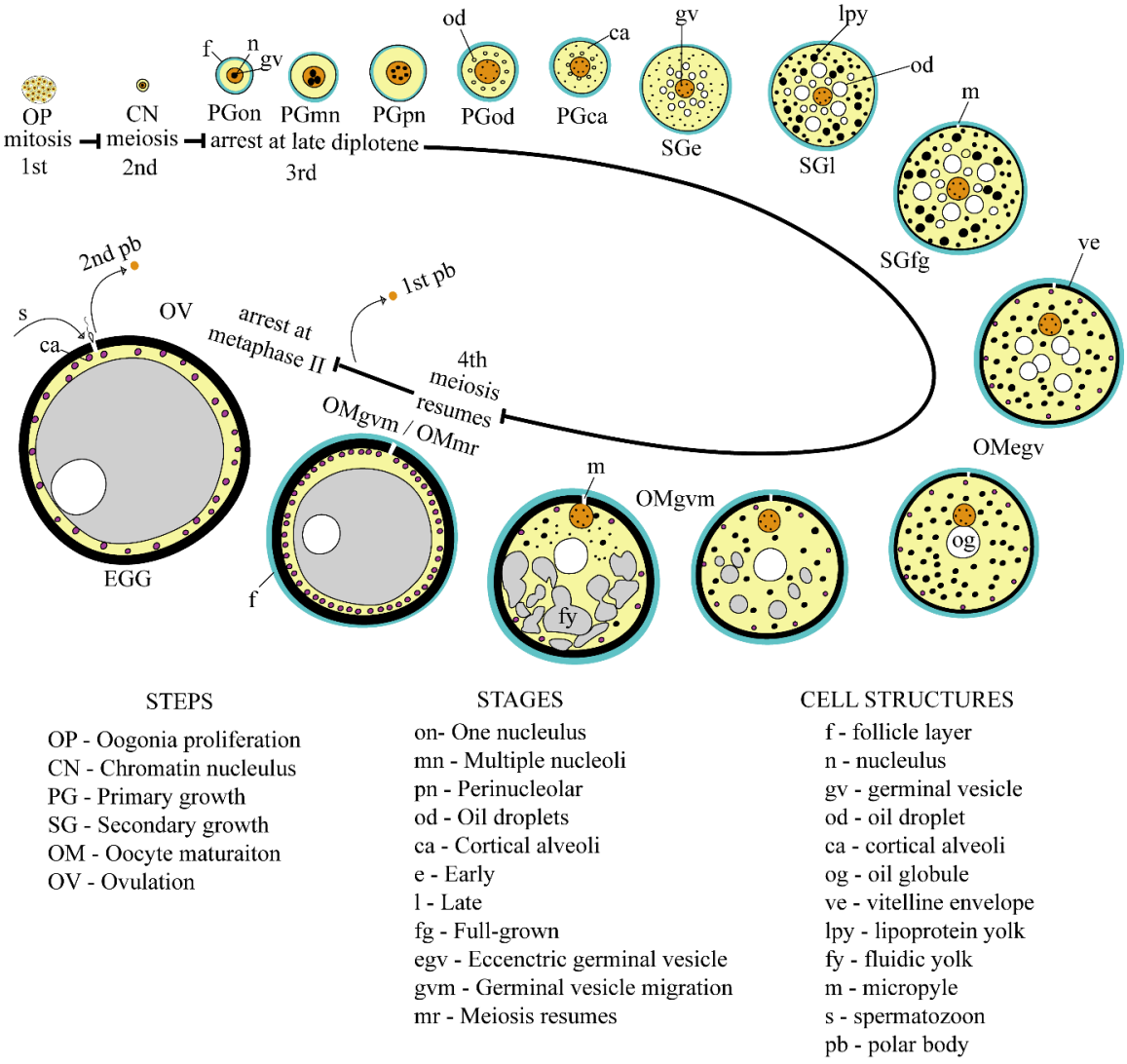

Figure 1.3. Schema illustrating oogenesis in teleosts along four periods of the cell cycle (mitosis, meiosis, arrest of meiosis at late diplotene and second period of active meiosis). These periods are divided into six steps, each step further subdivided into stages according to appearance of key cell structures. This schema has been proposed by Grier et al. (2009) and this illustration was modified from Rhody (2014).

During oocyte maturation (OM), the germinal vesicle initially moves slightly assuming an eccentric position in the ooplasm, hydrolysation of lipoprotein yolk is intense causing an increase in concentration of low molecular weight compounds, resulting in the oocyte hydration volume increase (Babin et al., 2007b; Grier et al., 2009;

Lubzens et al., 2017). The germinal vesicle then migrates towards the animal pole, breaks down, and meiosis is reinitiated producing the first polar body which degenerates (Grier et al., 2009). Oocytes become arrested again at metaphase II until reaching the ovulation step (OV). Egg activation is triggered by the contact with water, which promotes a rise in the intracellular calcium content that induce the break of dormancy and the completion of meiosis II, resulting in the release of the second polar body following fertilisation (Fig. 1.3) (Lubzens et al., 2017). At ovulation and spawning, the ovary of most species release eggs into an oviduct which channels eggs externally towards the gonopore (ovocoel type). Alternatively, salmonids, anguillids and also the osteoglossid *A. gigas* lack an ovarian capsule (gymnovarium condition) so eggs are released into the coelomatic cavity and reach the gonopore with the action of a funnel-shaped coelomic cavity (Colombo et al., 1984; Godinho et al., 2005; Grier et al., 2009).

1.2.2. Spermatogenesis

Spermatogenesis comprises the developmental process where diploid spermatogonial stem cells produce a large number of highly differentiated haploid spermatozoa carrying a recombined genome (Schulz et al., 2010). In teleosts, two groups of testis can be distinguished, lobular (*i.e.* guppy *Poecilia reticulata*) and tubular (*i.e.* trout *Oncorhynchus mykiss*), each with a particular physiology and morphology (Billard, 1986). Overall, the testis of teleosts is formed by somatic tissue which comprises the seminiferous tubules or lobules, supporting connective tissue and specialised somatic Leydig and Sertoli cells (Mañanós et al., 2008; Schulz et al., 2010). Spermatogenesis occurs within cysts, which are surrounded by Sertoli cells enclosing a single primary spermatogonium, and the array of cysts in the testis forms the germinal epithelium (Schulz et al., 2005; Mañanós et al., 2008). The Sertoli and Leydig cells are endocrine

cells which provide the physical, nutritional, and regulatory factors that support the developing germ cells along the three distinct phases of spermatogenesis (e.g. proliferative, meiotic and spermiogenic phases), illustrated in Figure 1.4 (Schulz & Nóbrega, 2011).

During spermatogenesis, spermatogonia type A proliferates through several cycles of mitosis forming spermatogonia type B, with number of cycles varying largely among species (Billard, 1986). In the second phase, spermatogonia type B undergo meiotic divisions to become spermatocytes I, spermatocytes II, and then haploid spermatids. During the third phase, also known as spermiogenesis, the spermatids differentiate into flagellated spermatozoa, when mitochondria concentration increase, and a midpiece and flagellum are formed (Mañanós et al., 2008; Schulz et al., 2010; Mylonas et al., 2017). In contrast with mammals, teleosts spermatozoa lack an acrosome and fertilisation occurs through the micropyle on the egg surface, whose diameter coincides with the spermatozoa width (Mañanós et al., 2008; Grier et al., 2009). Spermiation involves the release of mature spermatozoa from cysts after degeneration of Sertoli cells (Schulz et al., 2005). After rupture of the cysts, the spermatozoa are released into the testicular lumen, where they undergo maturation becoming capable of fertilizing eggs (motility) in the spermiduct (Mañanós et al., 2008; Schulz et al., 2010). Depending on the species, milt can be stored in the testis lumen or in vesicles, and fish release sperm spontaneously through the spermiduct which opens at the genital papilla (Mañanós et al., 2008; Mylonas et al., 2017).

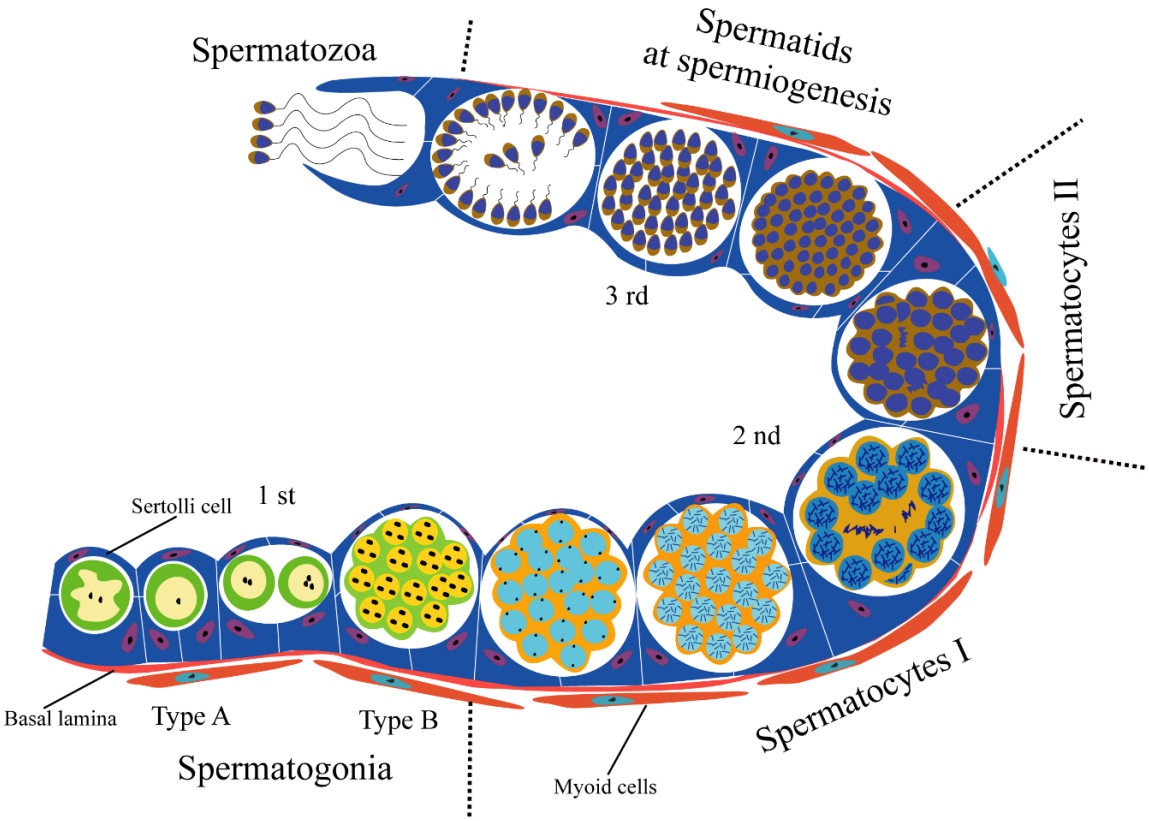

Figure 1.4. Schematic representation of spermatogenesis in teleosts. Proliferation of spermatogonia type A through mitosis becoming spermatogonia type B (1st phase), meiotic divisions of spermatogonia B becoming spermatocytes I, spermatocytes II and haploid spermatids (2nd phase), and differentiation of spermatids into spermatozoa during spermiogenesis (3rd phase). Schema modified from Schulz & Nóbrega (2011).

1.2.3. Neuroendocrine control of reproduction

Reproduction is one of the most important events in the life-history of a species, involving a high-energy expenditure to produce gametes, and an increased vulnerability to predation and diseases. Teleosts display an astonishing diversity of reproductive strategies and modes, and each species reproduces at the most appropriate time to maximise the chances of offspring survival. Environmental signals including photoperiod, water temperature, food availability and others vary seasonally, and fish

species (either temperate or seasonal) have adapted to use such cues to control and entrain reproduction involving the brain-pituitary-gonadal (BPG) axis (Migaud et al., 2010; Cowan et al., 2017). Based on a combination of gametogenic and environmental cycles which determine spawning frequency, fish species are classified into synchronous, group-synchronous or asynchronous (Grier et al., 2009; Lubzens et al., 2017). In synchronous species, one single cohort of oocytes undergoes vitellogenesis synchronously and matures together, resulting in a single batch of spawned eggs. Group-synchronous development refers to species with two or more oocyte cohorts at different developmental stages present in the ovary, thus resulting in more than one spawning along the reproductive season (*i.e.* batch spawners). Lastly, asynchronous species have all stages of oocyte development in the ovary but without a dominant oocyte population, and oocytes are recruited into maturation for ovulation and spawning several times (from daily to monthly) along the reproductive season (Lubzens et al., 2017).

Along the BPG axis (Fig. 1.5), the brain integrates environmental (*i.e.* photoperiod, temperature, pheromonal signals) and metabolic (*i.e.* nutrition, growth, stress factors) signals, conveying information through a cascade of hormones along the reproductive axis. This cascade can involve melatonin synthesised in the pineal, which interact with the BPG axis through dopaminergic inhibition of FSH and LH production (Maitra et al., 2013), and the kisseptin (KiSS) system, which is thought to be a key regulator of gonadotrophin releasing hormone (GnRH) in the hypothalamus of some species (Ogawa & Parhar, 2013; Chi et al., 2017). Hypothalamic GnRH is a decapeptide considered as the most potent stimulator of gonadotrophins synthesis and release in the pituitary (follicle stimulating hormone - FSH and luteinizing hormone - LH). Once released from the adenohypophysis, gonadotrophins reach target sites (steroidogenic cells) in the gonads eliciting the production of sex steroids that regulate gonadal

development (Mañanós et al., 2008; Zohar et al., 2010). Other neuropeptides and neurohormones can also play distinct roles modulating gonadotrophin production and release, these including neuropeptide Y, somatostatin, gamma-amino-butyric acid (GABA), and others reviewed in Levavi-Sivan et al. (2010) and Zohar et al. (2010). In addition, while GnRH stimulates production and release of gonadotrophins, dopamine can play an inhibitory effect on the pituitary of some species (Dufour et al., 2010; Zohar et al., 2010). Also, gonadotropin inhibitory hormone (GnIH) is a key neurohormone, regulating GnRH and GTH systems and locomotor and behavioural processes (Ubuka et al., 2012). The brain-pituitary axis also receives positive or negative feedback from sex steroids, when the gonads inform the brain about the stages of gonadogenesis along the reproductive cycle (Mañanós et al., 2008), and energy homeostasis messengers such as insulin-like growth factor (IGF), growth hormone (GH), ghrelin and leptin also interacting with the reproductive cycle through yet unclear pathways (Migaud et al. (2010).

In males and females, the reproductive cycle occurs in two phases, the first phase involving vitellogenesis and spermatogenesis and the second oocyte maturation and spermiogenesis. In species with synchronous ovarian development such as salmonids, it has been found that FSH stimulates vitellogenesis and spermatogenesis while LH stimulates oocyte maturation and spermiation (Mylonas & Zohar, 2001; Rosenfeld et al., 2007). However, this appears to be the exception to the rule as FSH and LH appear to play important roles during both phases in all other asynchronous and group-synchronous species studied (Fig. 1.5) (Rosenfeld et al., 2007; Mañanós et al., 2008). FSH and LH are heterodimeric proteins which share a common α subunit but have a species-specific β subunit, each subunit being encoded by a specific gene, and binding to specific receptors present in gonadal endocrine cell targets (FSH-R and LH-R) (Levavi-Sivan et al., 2010;

Golan et al., 2015). This explains the difficulty to measure gonadotrophin level in fish and the relatively limited assays available in only a few species (Levavi-Sivan et al., 2010). Molecular studies involving analyses of gene expression levels therefore became key to understand gonadotrophin regulation along the reproductive cycle. Very recently, the α and β FSH and LH subunits in *A. gigas* have been cloned and characterised (Faria et al., 2013; Bartolini et al., 2015), which will enable future investigations of gonadotrophin regulation in Arapaima although this will rely on invasive sampling.

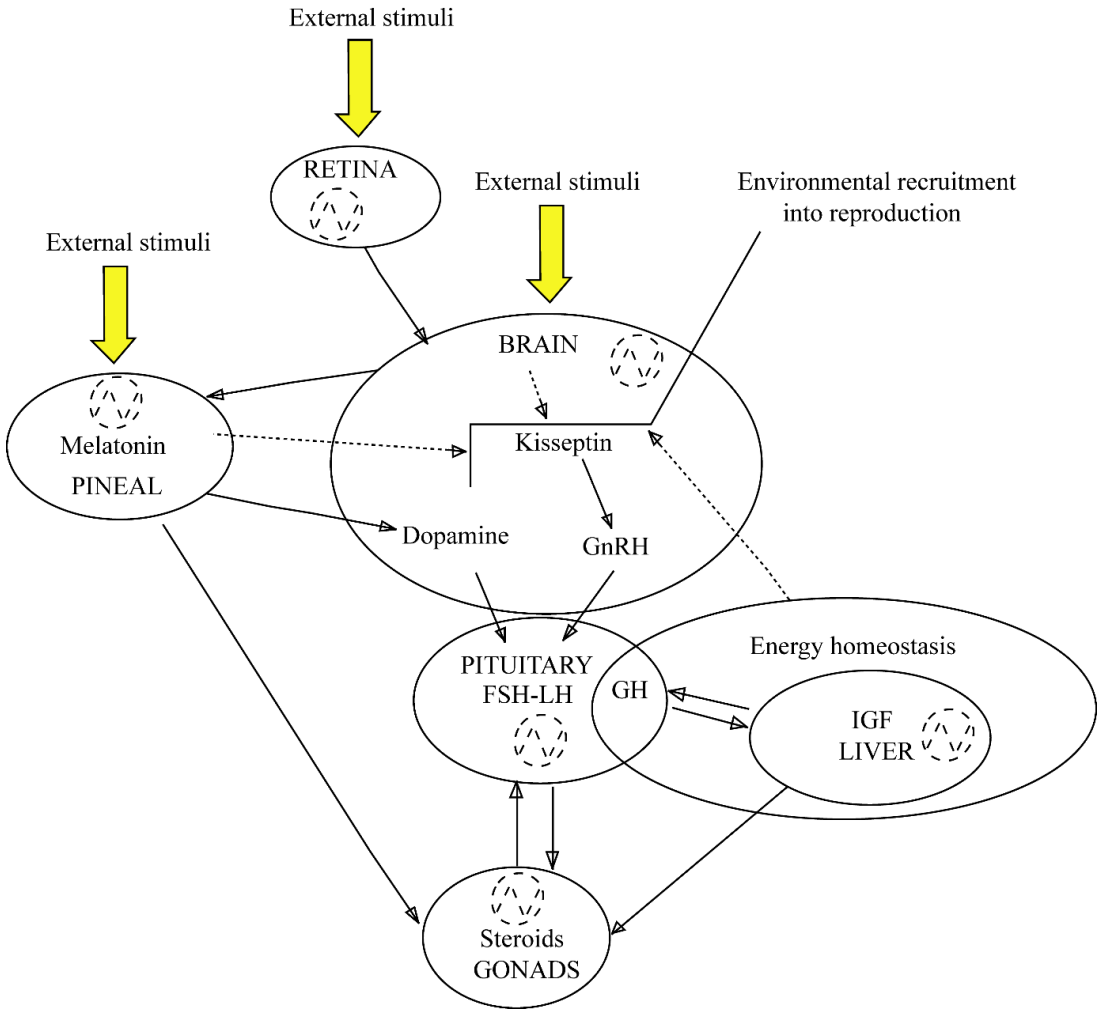

Figure 1.5. Schematic representation of the neuroendocrine control of reproduction in teleosts depicting photic stimuli perceived in the pineal, retina and deep brain photoreceptors (yellow arrows). Dashed line arrows correspond to still unclear pathways. At the brain level, kisseptin regulates gonadotrophin releasing hormone (GnRH) release

from the hypothalamus, which entrains the brain-pituitary-gonad (BPG) axis cascade where follicle stimulating hormone (FSH) and luteinizing hormone (LH) are released from the pituitary, controlling vitellogenesis and spermatogenesis through synthesis of sexual steroids. Synthesis of melatonin in the pineal can interact with the BPG axis through dopaminergic inhibition of FSH and LH production. Energy homeostasis messengers such as insulin-like growth factor (IGF), growth hormone (GH), ghrelin and leptin can interact with the reproductive cycle through unclear pathways. Figure extracted from Migaud et al. (2010).

During vitellogenesis in females, theca cells respond to FSH or LH by producing testosterone (T), which is converted into 17β-oestradiol (E_2) in the granulosa cells (Lubzens et al., 2017). In turn, E_2 regulates oocyte development and stimulates the production of VTG in the liver, which is released into the bloodstream and incorporated into the oocytes (Hiramatsu et al., 2015). At the end of vitellogenesis, germ cells develop the capacity to respond to maturation inducing steroids (MIS) (Mylonas et al., 2010). The secretion of LH (but also FSH in most species) from the pituitary elicits a shift in gonadal steroidogenesis, stimulating a rise in levels of MIS that promotes oocyte maturation (Mylonas et al., 2010). In fish, the main MIS include the 17α,20β,dihydroxy-4-pregnen-3-one (DHP) and 17α,20β,21-trihydroxy-4-pregnen-3-one (20β-S), and their action during oocyte maturation is mediated by factors such as prostaglandins, IGF and activin B (Mañanós et al., 2008; Mukherjee et al., 2017).

Similarly in males, spermatogenesis is initiated with the pituitary release of FSH or LH, which stimulates the Leydig cells to synthesise 11-ketotestosterone (11-KT), considered as the major sex steroid regulator of spermatogenesis (Schulz & Nóbrega, 2011). At the completion of spermatogenesis, pituitary release of gonadotrophins

promotes a shift in gonadal steroidogenesis stimulating production of MIS, which regulates spermiation, resulting in increased milt volume and inducing the motility capacity of spermatozoa (Mañanós et al., 2008; Mylonas et al., 2017).

1.2.4. Reproductive dysfunctions in captivity

The ability to control fish spawning to produce large numbers of high quality seeds is imperative to develop the aquaculture of a new species. In captive rearing systems, the environmental cues and optimal conditions for reproduction may not always be present, resulting in a range of reproductive dysfunctions (Mañanós et al., 2008; Mylonas et al., 2010; Migaud et al., 2013; Mylonas et al., 2017). In general, reproductive dysfunctions in males are less frequent, being associated with reduced milt volumes (*i.e*. common snook, *Centropomus undecimalis* (Rhody et al., 2014) and also lack of breeding behaviour in the first generation (F_1) such as in the Senegalese sole *Solea senegalensis* (Guzmán et al., 2011). In females, reproductive dysfunctions are more common in commercially important aquaculture species. In species like the Japanese eel, *Anguilla japonica*, vitellogenesis appears to not proceed spontaneously in captivity, due to a strong dopaminergic inhibition; this is unusual in fish and difficult to overcome in aquaculture (Ohta et al., 1997; Mylonas & Zohar, 2007). The most common problem with female fish is the failure to undergo oocyte maturation and spawn spontaneously, being associated with the lack of appropriate cues to induce the release of pituitary gonadotrophins (Mylonas et al., 2010). This is the reproductive dysfunction observed in the Atlantic Bluefin Tuna *Thunnus thynnus* (Rosenfeld et al., 2012), the meagre *Argyrosomus regius* (Mylonas et al., 2015), the greater amberjack *Seriola dumerili* (Fernandez-Palacios et al., 2015) and many other teleosts. Alternatively, there are species like Atlantic salmon *Salmo salar* and other salmonids, where females can undergo oocyte maturation and ovulation,

but fail to spawn, with their eggs remaining in the coelomatic cavity thus requiring stripping of females for artificial fertilisation (Mylonas & Zohar, 2001; Mylonas et al., 2010). This is also common in many flatfish species such as turbot *Scophthalmus maximus* (Nissling et al., 2013) or the halibut *Hippoglossus hippoglossus* (Mommens et al., 2013).

To address these reproductive bottlenecks in captive-reared fish, knowledge of the species' reproductive biology is key, especially to understand possible environmental cues controlling the reproductive cycle. Such cues can be related to photoperiod (Migaud et al., 2006), water properties (*i.e.* temperature, salinity, pH) (Migaud et al., 2002), tidal/moon cycles (Rhody et al., 2015), flooding cycles (Bailly et al., 2008) and rain (Kembenya et al., 2016). In addition, many other factors can impact directly on the success of reproduction including physical factors (lack of spawning substrates, inappropriate pond or tank dimensions), social factors involving pheromone communication to synchronise matings, nutritional factors with inappropriate nutrition of broodstocks and stress factors (Mañanós et al., 2008; Mylonas et al., 2010). Once the species requirements and optimal conditions for reproduction are known, environmental/hormonal manipulations can be used to entrain reproduction or induce ovulation and spawning. In some species, stimulation of the reproductive axis can also involve social manipulations where the release of pheromones is key to induce oocyte maturation and to synchronise spawning, as demonstrated for the African catfish *Clarias gariepinus* and the Mozambique tilapia *Oreochromis mossambicus* (Resink et al., 1989; Huertas et al., 2014).

Amongst the manipulations used to induce oocyte maturation and spawning in fish, hormonal treatments are the most frequently used (Mylonas & Zohar, 2001; Mylonas & Zohar, 2007; Mañanós et al., 2008; Mylonas et al., 2010; Mylonas et al.,

2017). Along the BPG axis, several hormones have been tested including KiSS, GnRH, gonadotrophins, carp pituitary extract (CPE), sex steroids and human chorionic gonadotrophin (HCG) (Mañanós et al., 2008; Mylonas et al., 2010; Selvaraj et al., 2015). From these, GnRH is considered as the most potent, reliable and safe hormone to use, for acting at the brain level stimulating production and release of the species' own gonadotrophins (Mylonas et al., 2010). To successfully stimulate final stages of gonadal development, the levels of gonadotrophins should be maintained high for long periods, with the main problems of GnRH treatments being the low residence time of the native hormone in the bloodstream (Mylonas & Zohar, 2001). Therefore, GnRH analogues (GnRHa) are required for having a higher resistance to cleavage (Mylonas & Zohar, 2001). Often multiple intraperitoneal injections spaced in time are necessary, or delivery-system technologies can be used to deliver hormone for longer periods (up to 8 weeks), reducing the need to handle fish multiple times (Mylonas & Zohar, 2001). Such delivery systems can be made with several matrices such as cholesterol-cellulose, ethylene-vinyl acetate and biodegradable microspheres (Mylonas & Zohar, 2001). Strong dopamine inhibition has been demonstrated in freshwater species (mainly cyprinids), and has also been suggested for marine fishes (Dufour et al., 2010). In some species (*i.e. Solea senegalensis*), dopaminergic inhibition on the reproductive axis can affect negatively on the effectiveness of the hormonal treatments, these requiring additional treatment with dopamine antagonists such as domperidone or pimozide (Mylonas et al., 2010; Guzmán et al., 2011).

To be effective, hormonal treatments rely on selecting females at the right stage of development (late vitellogenic stage), when their follicles are responsive to gonadotrophin or start of final oocyte maturation (Mylonas & Zohar, 2001). Therefore, tools to monitor gametogenesis *in situ*, especially in females, are essential in hatcheries

to induce fish at the most appropriate time. These tools often include the analysis of the genital papilla intumescence (Carosfeld et al., 2003), ultrasound monitoring of gonads (Albers et al., 2013), endoscopy procedures either surgical or accessing the gonoduct ("gonoductoscopy") to observe the gonad (Divers et al., 2009; Divers, 2010), or biopsy through cannulation to obtain gamete samples for wet mount or histological assessment (Samarin et al., 2015).

1.3. Reproduction of *Arapaima gigas*

1.3.1. General aspects

The proximate cues that control reproduction of *A. gigas* are still poorly understood, being possibly related to rain and flooding factors characteristic of the spawning season (Núñez et al., 2011). Pre-copulatory behaviour involves chasings and fights associated with territorialism, mating competition and possibly subordination (Fontenele, 1948; Rojas, 2005). The potential participation of pheromones during the mating process has been suggested (Amaral, 2009), although there is still very limited evidence and no experiment monitoring pheromone release along the reproductive season. Following the still poorly described courtship behaviour, mating couples move to shallow flooding areas with c. 1-1.5 m depth, where male and female dig a nest using their head and fins (Migdalski, 1957; Castello, 2008b). Nests in shallow areas are believed to facilitate fry transition to aerial breathing, and Arapaima appears to prefer sandy or clay soils without vegetation for nest building (Fontenele, 1953; Castello, 2008b). Building the nest can take from 3 to 5 days and copulation, spawning and external fertilisation occur very soon after (Fontenele, 1948; Fontenele, 1953). Interestingly, a study using 10 microsatellite markers for paternity analysis showed a single female can effectively copulate with more than one male suggesting potential sneaking behaviour (Farias et al., 2015). Following

external fertilisation, the male and female guard the nest for approximately nine days until the larvae start aerial breathing, and during the nest guarding period, when one parent is gulping air, the other is always protecting the nest (Fontenele, 1948; Fontenele, 1953). Mouthbrooding is a common behaviour observed in other osteoglossids (*i.e.* Arowanas) and was suggested for *A. gigas*, although this subject is still controversial with only anecdotal evidence (Halverson, 2013). When the larvae start breathing air, the male guides the offspring towards the food-rich inundated areas, while the female swims circling the male at longer distances (>3m) for approximately one month although this remains poorly understood (Fontenele, 1948; Castello, 2008b). Females have an asynchronous ovarian development, with a gonadosomatic index (GSI) ranging from 0.05-0.12 at the start of vitellogenesis and 0.60-1.00 at advanced vitellogenesis prior to oocyte maturation, and fecundity has been estimated at 47,000 full vitellogenic oocytes which are recruited and spawned up to 6 times along the reproductive season with an inter-spawning interval of one month (Fontenele, 1948; Queiroz, 2000; Godinho et al., 2005). Fecundity estimated from counted alevins range from 500 to 13,000 per spawning (Sawaya, 1946; Rojas, 2005; Núñez et al., 2011). To date, it is still not clear whether females have an annual or biannual gonadal cycle with both possibilities being suggested in the literature, and investigations are hindered by the lack of non-invasive tools to monitor reproductive function in live stocks (details in section 1.3.3.) (Queiroz, 2000).

1.3.2. The cephalic secretion and chemical communication

Chemical communication is known to mediate key aspects in the life history of teleosts, such as social organisation, territoriality, pair formation, courtship, parental care, identification of food items, habitats and predators (Derby & Sorensen, 2008; Keller-Costa et al., 2015). According to Wyatt (2010), a pheromone is defined as a “single

molecule or combination of molecules which can elicit an innate reaction such as a stereotyped behaviour or a developmental process in a conspecific individual". Because pheromones are coupled to the endocrine system, an individual can inform a conspecific about their physiological and reproductive condition at a given moment, helping individuals to synchronise spawning (Huertas et al., 2014). The main biochemical classes used as pheromones by fish include sex steroids, prostaglandins, bile acids and small peptides (Hara, 1994; Sorensen & Stacey, 1999; Stacey, 2003; Sorensen & Stacey, 2004; Keller-Costa et al., 2015). Routes of pheromone release into the water include the gills, urine, seminal and ovarian fluids (Sorensen & Wisenden, 2015). However, another important concept in chemical communication of aquatic organisms is a signature mixture, defined as "variable subsets of molecules of an animal's chemical profile which are learnt by other animals, allowing them to distinguish individuals or colonies" (Wyatt, 2010). Investigations on pheromone and signature mixtures require isolation of the released chemical(s) and specific experimentation to confirm either an innate or a learnt response (Sorensen & Wisenden, 2015). Indeed, it has been concluded the major bottleneck in studies with chemical communication is the lack of appropriate characterisations of the chemicals involved as pheromones and signature mixtures (Keller-Costa et al., 2015).

Chemical communication in osteoglossids including Arapaima has been so far little investigated. Interestingly, *A. gigas* secretes a whitish fluid from several pore-bearing features present on the head surface. These pores are part of the cephalic canals of the anterior lateral line system. The Amazon indigenous people named this secretion the Arapaima "milk", with the belief that it would provide nutrients to the developing offspring during the parental care period. Eigenmann & Allen (1942) were the first authors to mention the cephalic secretion in *A. gigas*, when they raised the possibility of

a nutritional role to the offspring. The literature reported observations where larvae "would suck the milk from the adult's head", thus offering nourishment (Sawaya, 1946; Fontenele, 1948). Alternatively, it has been suggested the possible use of this secretion is to cleanse the head from the mud after nest building (Menezes, 1951). Despite the lack of data supporting such hypotheses, the only consensus seems to be that the cephalic secretion is released in higher volumes during the reproductive season. Anatomical studies on the head of *A. gigas* showed the cephalic canals are well irrigated and innervated, containing polygonal, piston and mucous cells which could potentially source biochemicals into the secretion (Lüling, 1964; Lüling, 1969; Quay, 1972; Noakes, 1979). For the first time, biochemical analyses of the cephalic secretion of *A. gigas* revealed the presence of E_2 and 17α-hydroxyprogesterone (17α-OHP) in males and females (Amaral, 2009). Although not conclusive, such results suggested the cephalic secretion could be a source of pheromones, but the route(s) of transfer of these chemical substances into the secretion are not clear yet. The parallel analysis of steroids in the secretion and other fluids (*i.e.* blood, cerebrospinal fluid), or analyses of biomolecules such as proteins and peptides could help to better characterise this biological fluid and understand its functions. Similar investigation approaches have been used before to characterise the mucus eaten by the larvae of the Discus fish (*Symphysodon aequifasciata*) (Chong et al., 2005) or the pouch fluid where male seahorses (*Hippocampus* spp.) incubate their eggs (Melamed et al., 2005).

1.3.3. Issues with captive reproduction

In recent years, it has been demonstrated that pairing *A. gigas* in earthponds seems to promote reproduction during the rainy season, in contrast with the extensive stocking of broodstocks in large lagoons which were shown to be ineffective (Rebouças et al., 2014;

Lima et al., 2015a). Pond dimensions where pairs of *A. gigas* have successfully spawned range from 150-200 m^2 and 1-1.5 m deep (Lima et al., 2015b). Nevertheless, separation of pairs in earthen ponds is still made without informed decisions on fish gender, stage of maturity, social or size compatibilities, and these can often result in fish fights with losses of expensive broodstocks. There are therefore several gaps in knowledge that need research to optimise separation of pairs and captive reproduction in *A. gigas*.

Gender identification of *A. gigas* is usually done using the differential male pigmentation, which is unreliable for most populations and restricted to mature fish. The species lacks evident and reliable meristic sexual dimorphisms, so several studies aimed to develop methods to identify gender in *A. gigas.* Initially, karyotype analyses revealed the lack of evident chromosome dimorphisms and bulked-segregant analyses could not find sex-associated markers (Marques et al., 2006; Almeida et al., 2013). The development of an enzyme immune assay (EIA) to identify females through VTG blood detection was successful and is already used in farms over the Amazon (Dugue et al., 2008; Chu-Koo et al., 2009), although such method is still costly, laborious and requires access to a lab. A laparoscopy surgery to observe the gonad has been described, but this surgical method has not been adopted given welfare concerns due to the severity of the procedure and availability of instruments (Carreiro et al., 2011). Similarly, ultrasound analyses have been hindered by the thickness and reflection of fish scales which prevents gonad observation (Carreiro, 2012; Lima et al., 2015b). To date, next generation sequencing (NGS) technologies have not been tested in *A. gigas* to search for sex-linked markers, though such technology could bring promising results as demonstrated for several fish species studied so far (e.g. Palaiokostas et al., 2013; Brown et al., 2016).

Another issue hindering studies on the reproduction of *A. gigas* is the lack of tools to monitor reproductive function *in vivo*. In the Osteoglossidae, only the left gonad is

functional, and when present, the right gonad is vestigial (Godinho et al., 2005; Nelson et al., 2016). In *A. gigas*, the ovulated eggs released into the coelomic cavity reach the gonopore through an oviduct or "short ligament" (Godinho et al., 2005). Likewise, males have a cord-like testis whose diameter ranges from 1-1.5 cm, being connected to the genital papilla through the spermiduct (Godinho et al., 2005). Stripping of broodstock fish to obtain gametes for artificial fertilisation is not suitable due to the thickness of body wall and large size of the broodfish, and the gonopore position in the genital papilla in both genders is unknown. Therefore, cannulation procedure to obtain gonad biopsy has not been possible in *A. gigas* so far (Núñez et al., 2011). Given these limitations, it has not been possible to characterise the reproductive dysfunction of *A. gigas* in captivity. Moreover, success of hormonal and environmental manipulations to induce oocyte maturation and spawning rely on the correct gender identification and on the possibility to monitor reproductive function in broodstocks. For this reason, no studies on hormonal or environmental manipulation in *A. gigas* have been published so far. The collection of fertilised eggs from nests has been reported, however this remains anecdotal and it requires extensive observation of fish behaviour, and therefore is labour intense and can be dangerous as parents protect the nest (Halverson, 2013). In addition, when natural spawning has been obtained, the most appropriate moment to collect eggs, larvae or alevins from the parents is unknown given the lack of knowledge on the possible beneficial roles of parental care to offspring overall fitness. In particular to this subject, the potential benefits of the parental cephalic secretion to the nutrition, immunity and survival of the developing larvae and offspring are still very speculative, although empirical evidence suggested better growth and survival rates in alevins raised under parental care (Rojas, 2005).

1.4. Experimental aims

Aquaculture of *A. gigas* in the Amazon has involved the translocation of fish families between distinct river basins to constitute broodstocks, and therefore knowledge on the genetic basis of founding populations is a key factor to consider given a recent multispecies scenario proposed for *Arapaima* and potential side-effects of inbreeding since natural populations have been depleted over the past 200 years. Given the lack of knowledge on the species' reproductive physiology and technical limitations inherent to the species, reproductive function of Arapaima is not controlled in captivity, hampering a stable seed supply for the development of aquaculture. In the reproductive season, pairing fish in earthen ponds has been done so far without informed decisions on fish gender, reproductive condition, fish size pairing, limiting our understanding on the reproductive dysfunction(s) of the species in captivity. With such limitations, there is no report on environmental and hormonal manipulations, neither attempts to induce spawning and investigate the BPG axis with endocrinological tools. In addition, the cephalic secretion is another mystery in the reproductive physiology of *A. gigas*, with potential roles suggested in pheromone release and parental care, although very little biochemical data has been reported on this biological fluid in any teleost species.

The general aims of this thesis were to investigate the reproductive physiology of *A. gigas* to advance basic knowledge on the species and develop tools to identify gender, monitor gonad development, test hormonal therapies to induce spawning and characterise the cephalic secretion.

The specific objectives of this thesis were:

1. To investigate the degree of polymorphism in *A. gigas* using double digest restriction-site-associated DNA (ddRAD) sequencing, and identify a single nucleotide

polymorphism (SNP) panel that would prove useful to characterise genetic diversity and structure in wild and captive populations (Chapter 2);

2. To examine the effects of slow-release mGnRHa implants and different fish size pairings on sex steroid profiles measured in blood plasma and cephalic secretion (Chapter 3);
3. To test a non-invasive endoscopy procedure (gonoductoscopy) for gender identification and ovary assessment in adult broodstock (Chapter 4 - Part A);
4. To validate the use of the endoscopy procedure to identify gender in juveniles and in adults under field operational conditions, describe the anatomy of the female gonopore to enable cannulation, describe oocyte development and monitor gonadal development of a broodstock population (Chapter 4 – Part B);
5. To evaluate the effects of an injection of a combination of GnRH (mGnRHa and sGnRHa) on oocyte development and sex steroids after selecting maturing females through cannulation (Chapter 4- Part C);
6. To characterise the proteome and the peptidome of the cephalic secretion released by males and females of *A. gigas* during and outside the parental care phase (Chapter 5).

CHAPTER 2

2. Chapter 2

DDRAD SEQUENCING IN *ARAPAIMA GIGAS*: ANALYSIS OF GENETIC DIVERSITY AND STRUCTURE IN POPULATIONS FROM AMAZON AND ARAGUAIA-TOCANTINS RIVER BASINS.

Abstract

Arapaima gigas (Schinz, 1822) is considered to be the largest freshwater scaled fish in the world, and a promising candidate species for aquaculture development in Amazon countries. The development of tools to evaluate genomic polymorphism, genetic differentiation and identification of stocks is key for species conservation and support of aquaculture. This study aimed to investigate genomic polymorphism through ddRAD sequencing to identify a single nucleotide polymorphism (SNP) panel to assess genetic diversity and structure in populations from the rivers Amazon, Solimões, Tocantins, Araguaia, and also from a captive broodstock. Compared to many other teleosts, the overall extent of polymorphism in *A. gigas* was surprisingly low with only 2.3 % of identified RAD-tags (135 bases long) containing SNPs. Nevertheless, a panel with 393 informative SNPs was identified and screened across the five populations. Higher genetic diversity indices (number of polymorphic loci and private alleles, Shannon's Index and H_O) were found in populations from Amazon and Solimões, intermediate levels in Tocantins and Captive, and very low levels in the Araguaia population. These results likely reflect larger population sizes and less urbanized environments in the Amazon basin compared to Araguaia. Populations were significantly differentiated with pairwise F_{ST} values ranging from to 0.086 (Amazon x Solimões) to 0.556 (Amazon x Araguaia). Mean pairwise relatedness among individuals was significant in all populations ($P<0.01$), reflecting a degree of inbreeding possibly due to severe depletion of natural stocks, the species sedentary behaviour and/or sampling biases. Although Mantel test was not significant ($P=0.104$; $R^2=0.65$), Bayesian analysis in STRUCTURE and discriminant analysis of principal components (DAPC) showed populations of Amazon and Solimões to be genetically differentiated from Araguaia, with Tocantins comprising individuals from both identified stocks. Future analyses should explore and validate this SNP panel

in a larger number of samples, populations and morphotypes (species?) using cheaper genotyping techniques.

Keywords: conservation, fisheries, genomics, NGS, Osteoglossomorpha, Pirarucu.

2.1. Introduction

The Amazon giant *Arapaima gigas* (Schinz, 1822) is the largest freshwater scaled fish in the world with adults reaching up to 250 kg and measuring over 2.5 m in total length (Nelson et al., 2016). This emblematic air-breather fish is a promising candidate species for aquaculture development and has a high-valued market in South America (Oliveira et al., 2012). The natural geographical distribution of *A. gigas* includes the basins of the Amazon, Tocantins-Araguaia and Essequibo rivers which cover Brazil, Ecuador, Guyana and Peru, and the species has already been introduced into several non-indigenous systems (Castello & Stewart, 2010). *A. gigas* is a dioecious and iteroparous species with sexual maturity reached after three to five years of age (Godinho et al., 2005). Reproduction involves nest building by males and females in the sandy bottom of lentic habitats during the rainy season from November onwards (Castello, 2008a; Castello, 2008b). External fertilisation generally involves a single female, but often participation of more than one male, suggestive of sneaking behaviours which could help increase genetic diversity in the species (Farias et al., 2015). After spawning the nest is guarded by both parents until egg hatching, whereupon intense male parental care continues and a characteristic lateral migration towards flooded food-rich areas ensues (Castello, 2008a). Females normally reproduce multiple times within the reproductive season (Godinho et al., 2005) with a mean fecundity estimated to be c. 11,000 fingerlings per spawning event and an equal sex ratio at hatch (Neves, 1995; Queiroz, 2000). In the dry season, from June onward, water levels in the rivers decrease, marking the end of parental care (Castello, 2008a). At this stage, adults and offspring migrate back to the low lands reaching the river canals and várzea lagoons, when dispersal through the main rivers occurs (Castello, 2008a; Araripe et al., 2013). Overall, adults are not believed to migrate long distances, they are solitary and well adapted to hypoxic conditions during the

droughts (Castello, 2008a; Araripe et al., 2013). In some regions the dry season can be severe, resulting in mass mortalities of Arapaima within dried up lagoons. At such times rescuing operations are sometimes carried out for conservation reasons (Vitorino et al., 2015).

Given its obligate air-breathing behaviour, *A. gigas* is an easy target for fishermen and natural populations have been historically depleted or even eradicated close to the main cities (Hrbek et al., 2005). It is estimated that populations today represent only 13 % of historic levels (Castello et al., 2011) and since 1986 *A. gigas* has been included in the CITES red list (Castello & Stewart, 2010). While occasionally successful, reproduction of *A. gigas* in captivity is not a routine practice, this requiring further development of technologies for gender identification and control of spawning (Núñez et al., 2011). Therefore, fingerlings are high-priced in the aquaculture market, and their illegal capture from the wild is another challenge for conservation. Translocations of animals from one site to another has been regarded as another issue of concern because morphotypes and novel species have been described, suggesting patterns of allopatric differentiation across hydrographic basins (Stewart, 2013a; Stewart, 2013b; Watson et al., 2016).

To date, several studies have been conducted aiming to characterise the genetic diversity and structure of natural arapaima populations. These have involved the use of mitochondrial markers (mtDNA), microsatellite or inter-simple sequence repeats (ISSR) markers to study eight populations from the Amazon, Solimões and Tocantins river systems (Hrbek et al., 2005; Hrbek et al., 2007; Hamoy et al., 2008; Araripe et al., 2013), four populations from Tocantins and Araguaia (Vitorino et al., 2015) and five populations from Essequibo and Branco rivers (Watson et al., 2016). Overall, these studies found higher levels of genetic diversity within the large Amazon river basin compared to other

systems with the population structure suggesting minimal genetic flow and high genetic differentiation between populations. So far, molecular data has failed to confirm a multispecies scenario for *Arapaima*, which today is supported only by morphological analyses of very few specimens (Stewart, 2013a; Stewart, 2013b; Watson et al., 2016). Further genetic research on *A. gigas* has focused on acquiring tools for gender identification since this species is not sexually dimorphic, a factor that has impeded captive reproduction and aquaculture development to date. First, the species karyotype was characterised (2n=56) though no obvious sex chromosome dimorphism was observed (Marques et al., 2006; Rosa et al., 2009), while later a bulked segregant analysis failed to find sex-related markers (Almeida et al., 2013). Up-to-date, next generation sequencing (NGS) technologies have not yet been applied to investigate genomic variability and diversity in populations of *A. gigas.*

Technologies involving the sequencing of restriction-site-associated DNA (RAD) markers allows the simultaneous discovery and genotyping of thousands of single nucleotide polymorphism (SNP) markers (Baird et al., 2008). A range of genotyping by sequencing (GBS) technologies which can be readily applied to non-model organisms that lack extensive genomic resources have become ever more popular for SNP genotyping (Zhanjiang, 2011; Hohenlohe et al., 2012; Andrews et al., 2016). A variant of the original RAD-seq protocol, double digest RAD (ddRAD, Peterson *et al.*, 2012), provides a more scalable methodology which also reduces both time and costs of library preparation. In aquaculture, RAD sequencing has been exploited in a range of studies, including identification of sex-related markers(Palaiokostas et al., 2013; Palaiokostas et al., 2015; Brown et al., 2016), construction of linkage maps (Kai et al., 2014; Manousaki et al., 2016) and investigations of different populations and species within a genome-scale view of evolutionary processes (Saenz-Agudelo et al., 2015; Bian et al., 2016).

The appropriate RAD-seq approach to use to genotype a novel non-model species needs to be derived empirically, being dependent on many factors, including the number of informative markers desired, the quality of extracted DNA samples, the specific restriction sites present and the natural level of polymorphism within the genome, and the size of budget available for the project. A pilot ddRAD-based quantitative trait loci (QTL) study involving two families from a captive Arapaima stock, using the same restriction enzymes combination used for other studies (*Sbf*I and *Sph*I; *i.e*. Brown et al. (2016); Manousaki et al. (2016); Taslima et al. (2017), gave disappointing low numbers of SNPs (< 200) compared to the 1-2K SNPs identified in other fish species. It was not possible to establish from this one case whether the low level of polymorphism was due to selecting inbred captive individuals or was the norm for current *A. gigas* populations. Thus, this study aimed to investigate the degree of SNP polymorphism within and amongst natural populations of *A. gigas* from the Amazon, Solimões, Tocantins and Araguaia rivers, with the goal to identify a SNP panel that would prove useful to characterise genetic diversity and structure in these populations.

2.2. Material and Methods

2.2.1. Arapaima samples

This study analysed samples from the rivers Amazon (Iquitos, Perú), Solimões (Jarauá, AM, Brazil), Tocantins (Tucuruí, PA, Brazil), Araguaia (Lagoa da Confusão, TO, Brazil), and also a captive broodstock (Taipas, TO, Brazil). Twelve individuals per population / stock were sampled, comprising a total of 60 individuals (Fig. 2.1). Samples from Amazon, Solimões and Tocantins rivers were collected from wild captured animals sacrificed by fisherman during the legal fishing season, samples from these fish having been previously studied using other molecular markers (Hrbek et al., 2005; Araripe et al.,

2013). Animals from Araguaia were opportunistically sampled during a rescue operation where juvenile fish that were trapped in a small lagoon were sampled and later released into an adjacent larger lagoon. The Captive broodstock comprised adult without known hydrographic origin, filial generation or degree of relatedness, and were sampled in a farm located in Taipas-TO, Brazil. Fin clips or muscle samples were fixed in 95 % ethanol and kept at -20 °C until DNA extraction at the University of Stirling (Scotland-UK). Table 2.1 includes information on geographical coordinates and total number of originally sampled fish in each location. All samples were collected in accordance with Brazilian regulations (process number 02001.007554/2005-76 IBAMA/MMA).

Table 2.1. Studied populations of *Arapaima gigas*. Information on site locations and geographical coordinates to four wild (Amazon, Solimões, Tocantins and Araguaia) and one captive population (Captive). A reference is indicated if the material was previously analysed in other study. * Material originally collected and studied in Hrbek et al. (2005) or ** Araripe et al. (2013).

Site name	n° genotyped fish	n° sampled fish	Sampling year	Location	Coordinates
Amazon	12	16*	2000	Amazon River, Iquitos, Peru	-3.767454/-73.248425
Solimões	12	223**	2003	Solimões River, AM-Brazil	-2.807539/-65.076747
Tocantins	12	38**	2002	Tocantins River, Tucuruí, PA-Brazil	-4.111435/-49.744085
Araguaia	12	31	2013	Araguaia River, Lagoa da Confusão, TO-Brazil	-10.918678/-50.183229
Captive	12	24	2016	Taipas, TO-Brazil	-12.161331/-46.859444

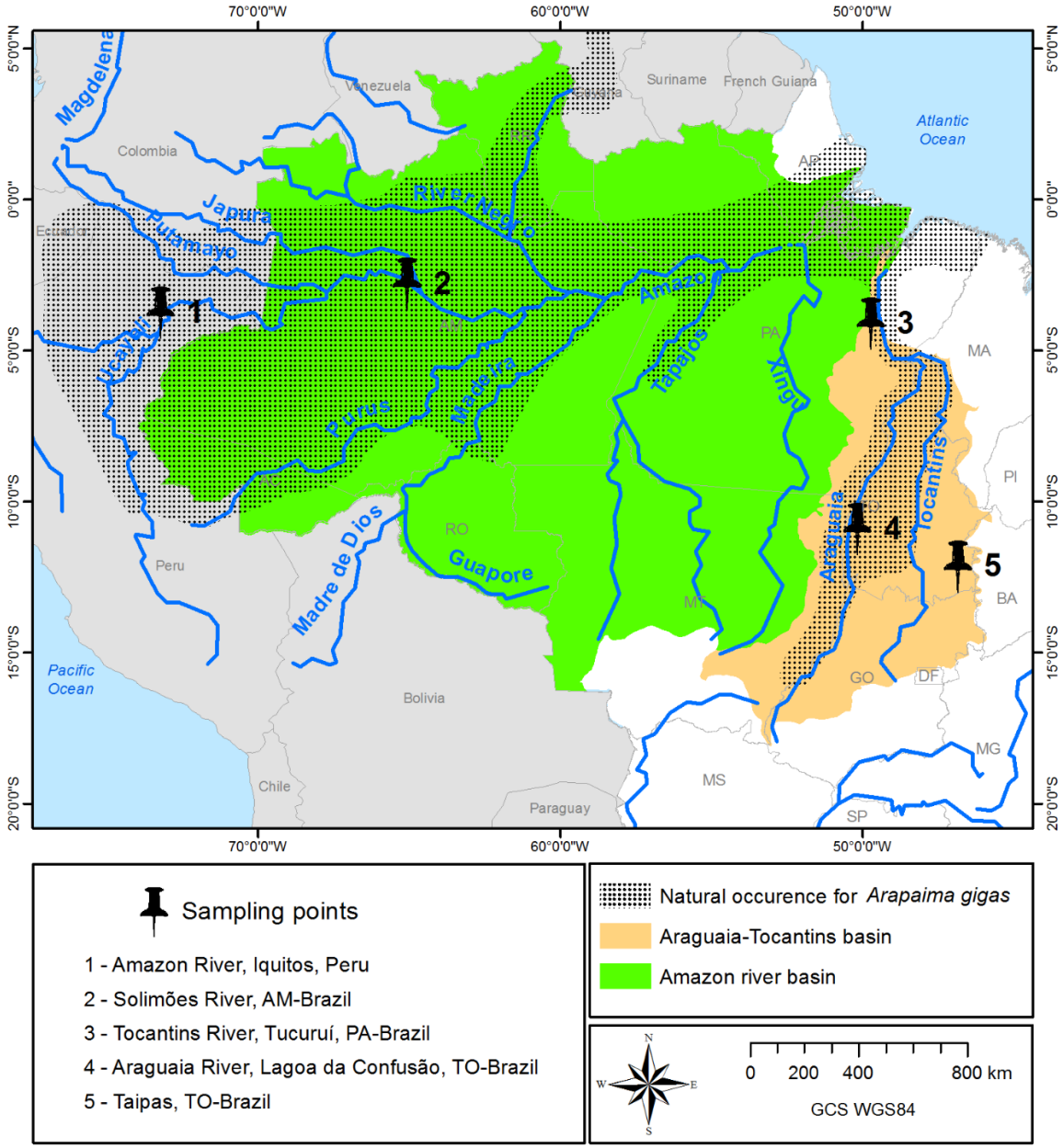

Figure 2.1. Map of Amazon region depicting approximate natural distribution of *Arapaima gigas* after Castello & Stewart (2010) (*by* Marta E. Ummus using software ArcGIS). Sampling points include two wild populations from Amazon river Basin (Amazon - 1 and Solimões -2), two wild populations from Tocantins-Araguaia Basin (Tocantins – 3 and Araguaia - 4), and one captive population (Captive - 5) from Tocantins state, Brazil.

2.2.2. DNA extraction

For genomic DNA extraction, a piece of fin clip or muscle sample from each of the 60 individuals was initially incubated for 16 hours at 55 °C in a lysis solution containing 200 µL of SSTNE buffer (Blanquer, 1990), 20 µL 10 % SDS, 5 µL proteinase K (10 $mg.mL^{-1}$) and 5 µL RNAse A (2 $mg.mL^{-1}$). Following, an incubation at 70 °C for 15 minutes was used to inactivate proteinase K. Then, 160 µL 5M NaCl (0.7 volumes) was added, mixed and cooled to 4 °C for 10 minutes to precipitate proteins. This mixture was centrifuged at 22,000 rcf for 10 minutes to pellet the proteins. The supernatant was then transferred to a new microfuge tube containing an equal volume of isopropanol, mixed to precipitate DNA and centrifuged at 22,000 rcf for 10 minutes to pellet the DNA. The supernatant was carefully removed and discarded and DNA pellet was then washed in 1 mL 70 % ethanol for 3 hours, after which the 70 % ethanol was replaced and washed for a further 16-20 hours. Finally, the samples were centrifuged at 22,000 rcf for 10 minutes to pellet down the DNA, the ethanol discarded and the DNA pellet air dried and reconstituted over a 24-hour period in 5 mM Tris (pH 8.0). The DNA quality and concentration was evaluated by spectrometry (NanoDrop; Thermo Scientific, USA). Agarose gel electrophoresis (0.8 %) was used to check the integrity of genomic DNA. Finally, a more accurate method, based on fluorescence (Qubit 2.0, Thermo Scientific, USA), was used to estimate DNA concentration prior to library preparation.

2.2.3. Library preparation and sequencing

A ddRAD library was prepared following method of Peterson et al. (2012) with modifications described by Brown et al. (2016). Briefly, 15 ng DNA was digested at 37 °C for 30 min with 0.3 U *Sbf*I (‘rare’ cutter, CCTGCA|GG motif) and 0.3 U *SphI* (‘common’ cutter, GCATG|C motif) high fidelity restriction enzymes (New England

Biolabs; NEB) in a 5 µL reaction volume that included 1× CutSmart™ buffer (NEB). After cooling the reactions to room temperature, 2.5 µL of a premade barcode-adapter mix was added to the digested DNA, and incubated at room temperature for 10 min. This adapter mix comprised individual-specific barcoded combinations of P1 (*Sbf*I-compatible) and P2 (*Sph*I-compatible) adapters at 6 nM and 72 nM concentrations respectively, in 1× reaction buffer 2 (NEB). Adapters were compatible with Illumina sequencing chemistry. The barcoded adapters were designed such that adapter–genomic DNA ligations did not reconstitute RE sites, while residual RE activity limited concatemerization of genomic fragments during ligation. The adapters included an inline five- or seven-base barcode for sample identification. Ligation was performed over 75 min at 22 °C by addition of a further 2.5 µL of a ligation mix comprising 4 mM rATP (Promega, UK), and 2000 cohesive-end units of T4 ligase (NEB) in 1× CutSmart buffer. The reactions were terminated by addition of 20 µL PB buffer (Minelute PCR Purification Kit; Qiagen UK). Each arapaima DNA sample was processed in duplicate (*i.e.* 120 separate digestion – ligation reactions) in order to lessen the impact of efficiency differences among reactions. All 120 reactions were pooled and column-purified (MinElute PCR Purification Kit) and eluted in 52 µL of EB buffer (Qiagen, UK). The purified sample was separated by agarose gel (1.1 %) electrophoresis and fragments ranging from approximately 450 bp to 650 bp were excised and purified using a MinElute Gel Extraction Kit (Qiagen, UK). The eluted size-selected template DNA (c. 54 µL in EB buffer) was amplified by PCR (11 cycles; 20 separate 20-µL reactions, each with 1.5 µL of DNA template) using a high fidelity *Taq* polymerase Q5 Hot Start High-Fidelity DNA Polymerase (NEB). The PCR reactions were combined (400 µL total) and column-purified with the MinElute PCR Purification Kit (Qiagen, UK), then eluted in 50 µL of EB buffer. The library was then re-purified, to ensure no carry over of PCR primer dimer

or other low size range amplicons (< 200 bp) using an equal volume of AMPure magnetic beads (Perkin-Elmer, UK) and eluted in 15 µL EB buffer. Finally, the library was sequenced at the Institute of Aquaculture (University of Stirling, Scotland) on a single run on the Illumina MiSeq platform (162 bp paired end reads, 300 cycle kit, v2 chemistry kit; Illumina, Cambridge, UK).

2.2.4. ddRAD genotyping and comparative datasets

After sequencing, reads were de-multiplexed, low quality reads with missing restriction sites and containing ambiguous barcodes were excluded and sequences trimmed to 135 bases long using Stacks version 1.42 (Catchen et al., 2013). Given lack of a reference genome for *A. gigas*, the RAD loci were assembled *de novo* using the following parameter settings: 4 as the minimum read depth to create a stack (m), 2 as the maximum number of mismatches in an individual (M) and 1 as the maximum mismatch between loci for building the catalogue (n). Finally a robust set of SNPs suitable for population analyses was exported from the catalogue in Genepop format (Rousset, 2008), using the 'populations' module within Stacks to select only those SNPs scored in at least 80% of individuals in each of the five populations and confined to RAD-tags that contained no more than 2 SNPs).

The levels of polymorphism within the *A. gigas* samples were compared to ddRAD data of other teleosts generated from other projects undertaken at the University of Stirling using the same methodologies. These species and numbers of samples were: *Cyprinus carpio* L. (n = 85), *Dicentrarchus labrax* (L.) (n = 26), *Oreochromis niloticus* (L.) (n = 6), *Clarias anguillaris* (L.) (n = 5), *Sprattus sprattus* (L.) (n = 8), *Ctenolabrus rupestris* (L.) (n = 20) and *Melanogrammus aeglefinus* (L.) (n = 16). For each species, the mean (± SD) of the following parameters were calculated: number of unique stacks,

number of polymorphic loci and number of SNPs obtained. Non-parametric Kruskal–Wallis one-way ANOVA followed by Dunn's pairwise *post hoc* tests were used to compare the ratio between polymorphic loci and unique stacks obtained between species. Statistical analyses were conducted using Minitab version 17.3.1 (Minitab, PA, USA) with significance set at $P<0.05$.

2.2.5. Analyses of genetic diversity and structure

Initially, tests for Hardy-Weinberg Equilibrium (HWE) and linkage disequilibrium (LD) were conducted for the SNP loci both within populations and across the global dataset (60 individuals, 5 populations) using Genepop version 4.6 (Rousset, 2008). To do so, the Markov chain Monte Carlo (MCMC) parameters used were 10,000 dememorizations, 20 batches and 5000 iterations per batch), with sequential Bonferroni (Rice, 1989) being used correct for type I error multiple tests. The software Arlequin version 3.5.2.2 (Excoffier & Lischer, 2010) was used to estimate F_{ST} values for each locus in the global population, to estimate pairwise genetic differentiation (F_{ST}) between populations (10,000 permutations; $P<0.01$) and to identify outlier loci (hierarchical island model, 20,000 simulations, 100 demes simulated per group with 10 groups simulated).

To investigate the hypothesis of isolation by distance, the shortest waterway path among sampling sites (Captive excluded) was measured using Google Earth version 7.1.8 (https://www.google.com/earth). The software GenAlEx version 6.5 (Peakall & Smouse, 2006; Peakall & Smouse, 2012) was used to perform a Mantel test in which geographical (km) and genetic (F_{ST}) distances were correlated. GenAlEx was also used to estimate percentage of polymorphic loci, number of private alleles, Shannon's Information Index (I), expected (H_E) and (H_O) observed heterozygosity, and coefficient of inbreeding (F_{IS}) from Weir & Cockerham (1984). Also, pairwise individual relatedness was estimated

using Lynch & Ritland (1999) method to calculate a square matrix of individual pairwise relatedness (r). The average of pairwise values was calculated using "Pops mean" option, in which significance was tested using 1000 permutations and 1000 bootstrap resamplings to estimate 95 % confidence intervals. Pearson Product Moment Correlations was used to correlate levels of relatedness (r) with Shannon's I and H_O using Minitab ($P<0.01$).

Population structure was first investigated using a model-based approach, which assumes HWE and linkage equilibrium, implemented in STRUCTURE version 2.3.4 (Pritchard et al., 2000). Initially, estimation of the K-value which maximizes the global likelihood of the dataset (50,000 burn-in; 100,000 MCMC; 10 independent runs *per* K; ranging K from 1 to 5) was made using an admixture model and frequencies were assumed correlated. Optimal K-value was determined by Evanno's method (Evanno et al., 2005) using the Best K pipeline of CLUMPAK program (Kopelman et al., 2015). Using the best K-value, a final analysis was conducted with a burn-in of 250,000 using 500,000 MCMC and 10 replications. Results were then averaged and displayed using main pipeline of CLUMPAK. The analysis described above was independently conducted for the global dataset (5 populations), and then for Amazon river basin (Amazon and Solimões) and Tocantins-Araguaia river basin (Tocantins, Araguaia and Captive) separately.

Population structure was explored using a discriminant analysis of principal components (DAPC), conducted in R version 3.3.2. (R-Core-Team, 2014) using the package Adegenet v. 2.0.1 (Jombart, 2008). DAPC analysis is not model-based, and optimizes the variance between groups while minimizing the differences within clusters. It requires prior identification of a cluster number, which was made using the Bayesian Information Criterion (BIC) (Jombart, 2008). The function *find.clusters* was used to

transform original data into principal components (PC), retaining 60 PCs in the analysis. Following, *dapc* function performed a discriminant analysis using 20 PCs (>80 % of variance explained) and 5 eigenvalues were retained and examined. Following, the *assign.per.pop* function evaluated how clear-cut the genetic clusters were in reassigning individuals into their original clusters. Finally, to investigate evolutionary relationships among populations, a Neighbor-Joining (NJ) analysis (Saitou & Nei (1987) was conducted in MEGA version 7.0.21 (Kumar et al., 2016) using a genetic distance (GD) matrix produced in GenAlEx as input file.

2.3. Results

2.3.1. ddRAD sequencing and degree of polymorphism

The SNP identification from the ddRAD sequencing for 60 individuals of *A. gigas* is summarised in Figure 2.2. In total, 33,932,300 raw reads comprising 16,966,150 paired-end reads were obtained from the single MiSeq run. After quality filtering, a total of 29,133,620 reads (85.8 %) were kept. Assembling loci (RAD-tags) into the 60 individuals identified 12,378 unique RAD-tags. This panel resolved down to 448 markers containing 1-2 SNPs present in more than 80 % of individuals in each of the five populations. A final filtering was used to select only one SNP (the most polymorphic one) when two SNPs were identified in a RADlocus, and also removed the few instances where the same SNP was identified at the 3' end of paired reads. This reduced the final panel used for population genetic analyses to 393 SNPs. Only 2.3 % of *A. gigas* ddRAD generated RAD loci (135 bases long) were identified to contain SNPs. This was 3-8 times lower than observed for other teleosts studied at the University of Stirling (Table 2.2). The levels of polymorphism detected among the *A. gigas* samples from different locations varied significantly ($P<0.05$); higher polymorphic levels were found in the Amazon (3.5 %) and

Solimões (2.7 %) rivers compared to the Araguaia river (1.0 %), whilst fish from the Tocantins river and the Captive stock showed intermediate levels (2.0 %) (P<0.05; Table 2.3).

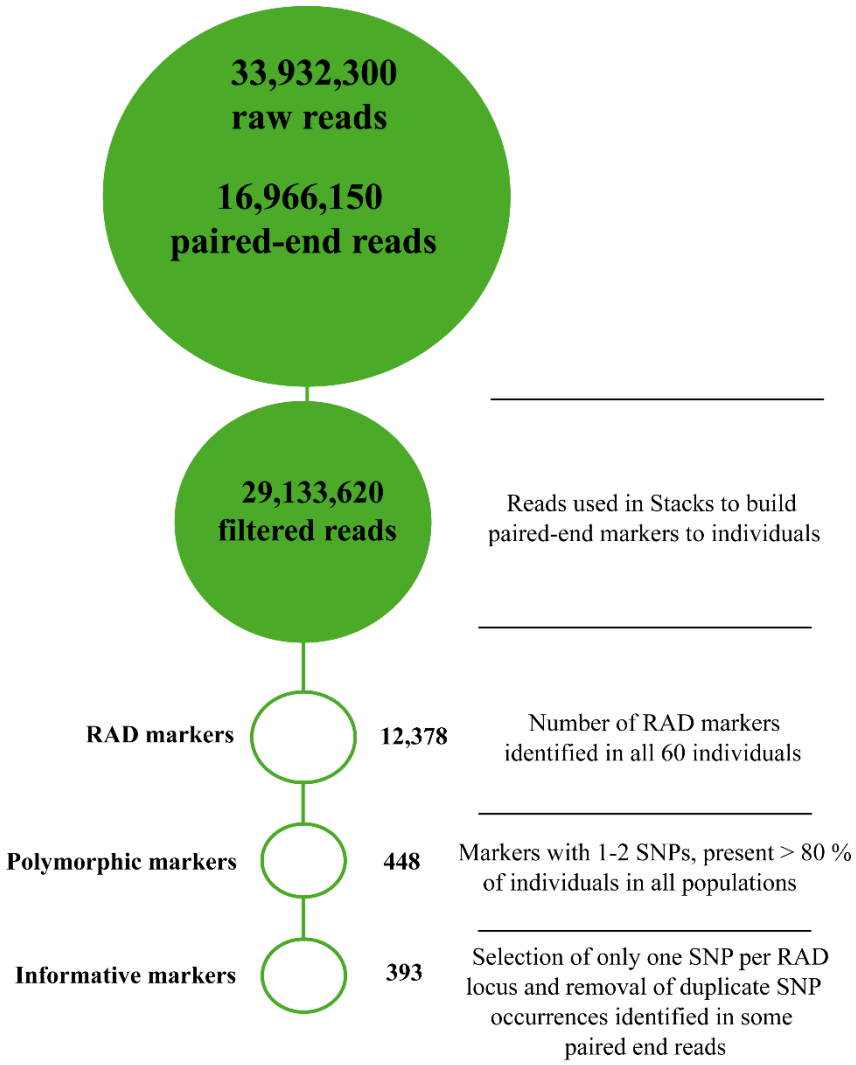

Figure 2.2. Summary of ddRAD sequencing for *Arapaima gigas*, scheme modified from Brown et al. (2016). Workflow of data processing from the obtained raw reads (upper disk) down to the markers used to investigate genomic diversity and structure in populations.

Table 2.2. Comparison of ddRAD sequencing results obtained after *de novo* assembly in Stacks (Catchen et al., 2013). Results represent the mean values in genotyped individuals for the number of unique stacks obtained, polymorphic loci, SNPs found and ratio of polymorphic loci *per* unique stacks obtained (%), different letters indicate significantly different (Kruskal-Wallis, P<0.05).

Species	Popular name	Family	n° fish	Unique stacks per individual	n° Polymorphic loci (%)	SNPs found
Arapaima gigas	**Pirarucu**	**Osteoglossidae**	**60**	**6667.9 ± 611.6**	**154.0 ± 103.8 (2.3) ^{F}**	**192.6 ± 141.4**
Dicentrarchus labrax	Seabass	Moronidae	26	6830.9 ± 201.5	519.2 ± 522.1 (7.6) ^{E}	1034.7 ± 784.0
Cyprinus carpio	Carp	Cyprinidae	85	7697.4 ± 1213.4	637.6 ± 416.8 (8.3) ^{DE}	868.6 ± 553.9
Oreochromis niloticus	Tilapia	Cichlidae	6	14612.8 ± 520.6	1796.2 ± 130.0 (12.3) ^{CDE}	2463.3 ± 161.9
Clarias anguillaris	Mudfish	Clariidae	5	9666.0 ± 79.2	1356.6 ± 43.2 (14.0) ^{BCD}	1945.4 ± 83.0
Sprattus sprattus	Sprat	Clupeidae	8	15140.6 ± 1561.5	2629.5 ± 319.8 (17.4) ^{ABC}	3444.5 ± 441.0
Ctenolabrus rupestris	Goldsinny	Labridae	20	11817.4 ± 3772.6	2217.9 ± 715.6 (18.8) ^{AB}	3129.3 ± 944.5
Melanogrammus aeglefinus	Haddock	Gadidae	16	12481.8 ± 3255.7	2529.1 ± 415.5 (20.3) ^{A}	3349.4 ± 550.2

Table 2.3. Comparison of ddRAD sequencing results for different populations of *Arapaima gigas*. Results represent the mean values in genotyped individuals for the number of unique stacks obtained, polymorphic loci, SNPs found and ratio of polymorphic loci *per* unique stacks obtained (%), different letters indicate significantly different (Kruskal-Wallis, P<0.05).

Population	n°	Unique stacks per individual	n° Polymorphic loci (%)	SNPs found
Amazon	12	6744.3 ± 330.9	241.3 ± 160.2 (3.5) ^{AB}	309.4 ± 241.8
Solimões	12	7029.3 ± 397.2	193.5 ± 25.8 (2.7) ^{A}	230.1 ± 31.3
Tocantins	12	6582.5 ± 416.6	131.1 ± 21.5 (2.0) ^{B}	157.7 ± 24.5
Araguaia	12	5838.9 ± 462.0	60.8 ± 13.1 (1.0) ^{C}	80.6 ± 20.3
Captive	12	7144.7 ± 437.0	143.1 ± 26.9 (2.0) ^{B}	185.2 ± 33.2

2.3.2. Analyses of genetic diversity

The 393 loci analysed were all found to be in HWE both within individual population groupings and across the global population sample (P<0.05; sequential Bonferroni corrected). Analyses of LD indicated significant association between locus 373_82 and 3408_78 in the global dataset after Bonferroni correction. Therefore, locus 373_82 (lower F_{ST}) was removed from further analyses resulting in a final dataset with 392 SNPs. For this dataset (392 SNPs, 60 individuals), the global F_{ST} calculated across all loci was 0.389, and individual locus F_{ST} values ranged from -0.04 to 0.97 with their frequency distribution depicted in Figure 2.3A. Analysis also detected 57 loci putatively under selection (outlier) (Fig. 2.3B). These were kept in the dataset in order to integrate possible adaptive information in further population analysis (Moore et al., 2014), while their removal did not significantly alter population structure results (data not shown).

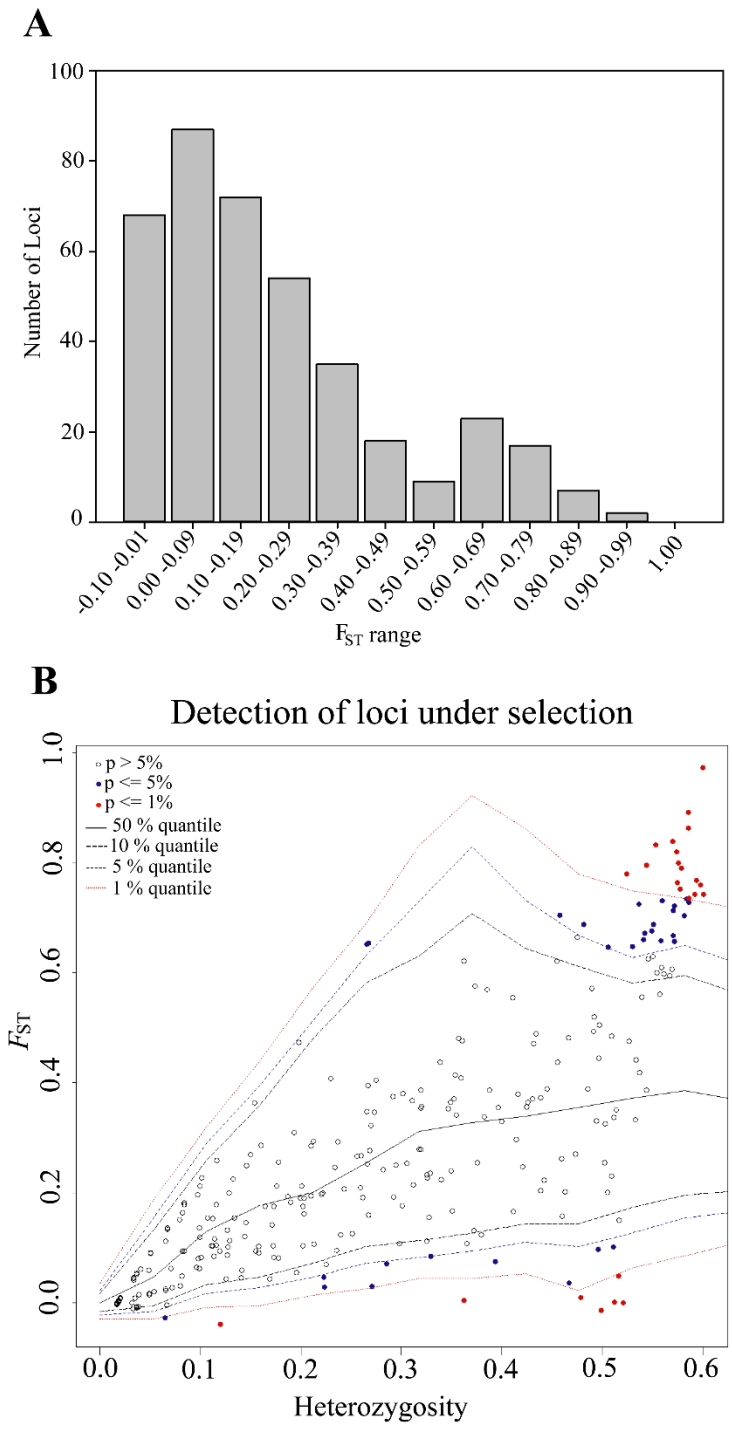

Figure 2.3. Loci analysis for 392 SNPs across 60 individuals analysed across the five populations of *Arapaima gigas.* A. Frequency distribution of F_{ST} values. B. Detection of loci under selection (outlier) using hierarchical structure model implemented in Arlequin v. 3.5.2.2. Outlier loci are indicated in red (P<0.01) and blue (P<0.05) dots.

All pairwise F_{ST} comparisons among population samples were significant (at least P<0.01; Table 2.4). Moderate degrees of differentiation were found between Amazon and Solimões (0.086) and between Captive and Araguaia (0.077). High degrees of differentiation were found between Tocantins and Captive (0.227), Tocantins and Araguaia (0.332), Solimões and Tocantins (0.336), Amazon and Tocantins (0.390), Solimões and Captive (0.423) and Amazon and Captive (0.459). Very high degrees of genetic differentiation were observed comparing Araguaia with Amazon (0.556) and with Solimões (0.523). Table 2.4 also indicates waterway distances calculated among sampling sites, and Mantel test based on these values indicated that the isolation by distance hypothesis was not supported (P = 0.104; R^2=0.65).

Table 2.4. Genetic differentiation (F_{ST}) and geographical distance (km) among populations of *Arapaima gigas*. Below diagonal values are pairwise F_{ST} comparisons made with Arlequin v. 3.5.2.2, performing 10,000 permutations. Above diagonal values depict waterway geographical distance measured among wild populations (*Captive* excluded) using Google Earth version 7.1.8 (https://www.google.com/earth). ** P<0.01; *** P<0.001.

Population pairwise F_{ST} / Geographical distance (km)					
	Amazon	**Solimões**	**Tocantins**	**Araguaia**	**Captive**
Amazon	—	1,323	3,613	4,708	—
Solimões	0.086***	—	2,290	3,385	—
Tocantins	0.390***	0.336***	—	1,095	—
Araguaia	0.556***	0.523***	0.332***	—	—
Captive	0.459***	0.423***	0.227***	0.077**	—

Overall, the population samples from the Amazon basin (Amazon and Solimões) were genetically more diverse than populations from those from the Araguaia and

Tocantins rivers in terms of percentage of polymorphic loci, number of private alleles, Shannon's Information Index (I) and observed heterozygosity (H_O) (Table 2.5). The mean pairwise relatedness (r) between individuals was significantly different from zero for all population samples ($P<0.05$), being lower in rivers Amazon (0.127) and Solimões (0.125), increasing in Tocantins (0.182) and highest in river Araguaia (0.323) and Captive broodstock (0.238) (Fig. 2.4). Relatedness correlated negatively with H_O and with I ($R^2 = 0.957$, $P<0.01$, and $R^2 = 0.956$, $P<0.01$ respectively), indicating inbreeding as a potential cause for the loss of genetic diversity in the studied populations.

Table 2.5. Genetic diversity statistics for five population samples of *Arapaima gigas.* Polymorphic loci (%), I = Shannon's Information Index, H_O = observed heterozygosity, H_E = expected heterozygosity, F_{IS} = coefficient of inbreeding of Weir & Cockerham (1984), SE = standard error. Indices calculated using 392 SNPs with GenAlEx v. 6.5.

Population	**Fish (n°)**	**Polymorphic loci (%)**	**Private alleles (n°)**	**I**	**H_O±SE**	**H_E±SE**	**F_{IS}±SE**
Amazon	12	71.4	62	0.316 ± 0.013	0.192 ± 0.010	0.203 ± 0.009	0.034 ± 0.016
Solimões	12	70.4	49	0.316 ± 0.013	0.220 ± 0.011	0.204 ± 0.009	-0.070 ± 0.013
Tocantins	12	39.0	9	0.190 ± 0.013	0.134 ± 0.011	0.125 ± 0.009	-0.061 ± 0.016
Araguaia	12	14.5	9	0.061 ± 0.008	0.042 ± 0.007	0.039 ± 0.006	-0.073 ± 0.013
Captive	12	34.7	5	0.148 ± 0.011	0.110 ± 0.009	0.092 ± 0.007	-0.174 ± 0.010
Overall	**60**	**46.0 ± 10.9**	**—**	**0.206 ± 0.006**	**0.140 ± 0.005**	**0.133 ± 0.004**	**-0.052 ± 0.007**

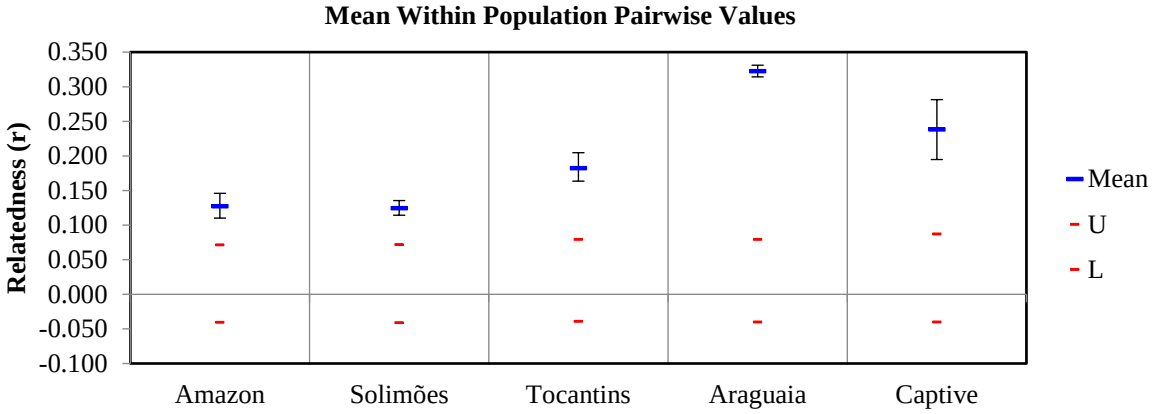

Figure 2.4. Within population pairwise mean relatedness (r) for *Arapaima gigas* (n=12 individuals each population). Calculations followed method of Lynch & Ritland (1999) with confidence intervals of 95% denoted by the U (upper) and L (lower) marks, calculated after 1,000 bootstrap resamplings and 1000 permutations in GenAlEx version 6.5.

2.3.3. Population structure

The optimal number of clusters (K) across the five studied populations was resolved, by the Evanno method, to be two (Fig. 2.5B). Analysis indicated the Amazon river basin (Amazon and Solimões) and Araguaia river constitute distinct genetic stocks, and the lower Tocantins river sample suggestive of a hybrid zone between these two groupings (Fig. 2.5A-C). Analyses also indicated eight individuals from the Captive broodstock being genetically similar to Araguaia population, and 4 individuals similar to Tocantins population (Fig. 2.5A-C). Analysing datasets split by hydrographic basin, the Evanno method again resolved two clusters (K=2) for Amazon river basin (Amazon and Solimões) and also two clusters (K=2) for Tocantins-Araguaia basin (Tocantins, Araguaia and Captive). In the Amazon river basin, two individuals from river Amazon were clearly differentiated from the remainder of the sample (Fig. 2.5D). In Tocantins-Araguaia basin, results depicted population of Tocantins as a differentiated stock

compared to Araguaia, and supported the previous interpretation of a mixed origin for the Captive broodstock (Fig. 2.5E).

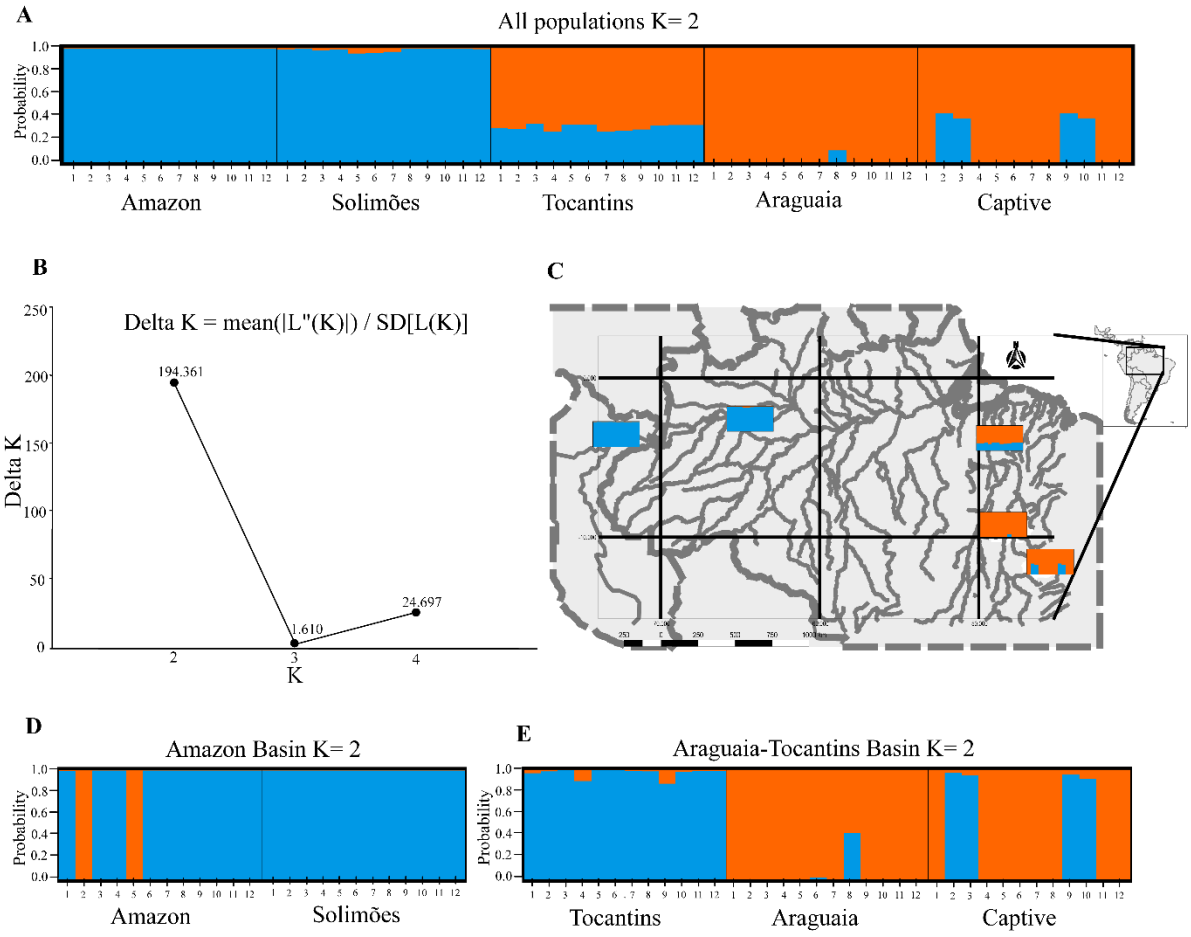

Figure 2.5. Bayesian clustering representation for populations of *Arapaima gigas* using 392 SNP markers in STRUCTURE v. 2.3.4 (Pritchard et al., 2000). A. Analysis of five populations (n=60 individuals; optimal Evanno's K = 2). B. Graphical representation of optimal number of clusters (K) across the five studied populations determined by Evanno's method (Evanno et al., 2005), indicated by delta K peaking at K=2. C. Geographical representation of structure results for the global population. D. Structure analysis for Amazon river basin (Amazon and Solimões, optimal Evanno's K = 2). E. Structure analysis for Araguaia-Tocantins basin (Tocantins, Araguaia and Captive, optimal Evanno's K = 2).

The DAPC analysis identified three groups (K=3) and generally supported the findings resolved by STRUCTURE. Populations from the Amazon basin (Amazon and Solimões) were clustered together, the genetically "hybrid" area formed by Tocantins was considered a distinct cluster in DAPC, which also included four Captive individuals. The third cluster grouped together Araguaia and the remaining eight Captive individuals (Fig. 2.6A). Selection of three clusters was suggested by the Bayesian Inference Criterion (BIC) (decreasing elbow; Fig. 2.6B), with all individuals being correctly reassigned into their original clusters with 100% membership probability (Fig. 2.6C).

Evolutionary reconstruction using an unrooted NJ tree also supported results from STRUCTURE and DAPC. Analysis showed separation of two main clades. One grouped populations from Solimões and Amazon, and the other populations from Tocantins, Araguaia and Captive. Indeed, NJ showed two individuals of Amazon genetically distant from the remaining clade, congruent with results found in STRUCTURE (compare Fig. 2.5C with Fig. 2.7). Also, eight Captive individuals were grouped close to Araguaia individuals and four formed a sister-group of Tocantins and Araguaia populations (Fig. 2.7).

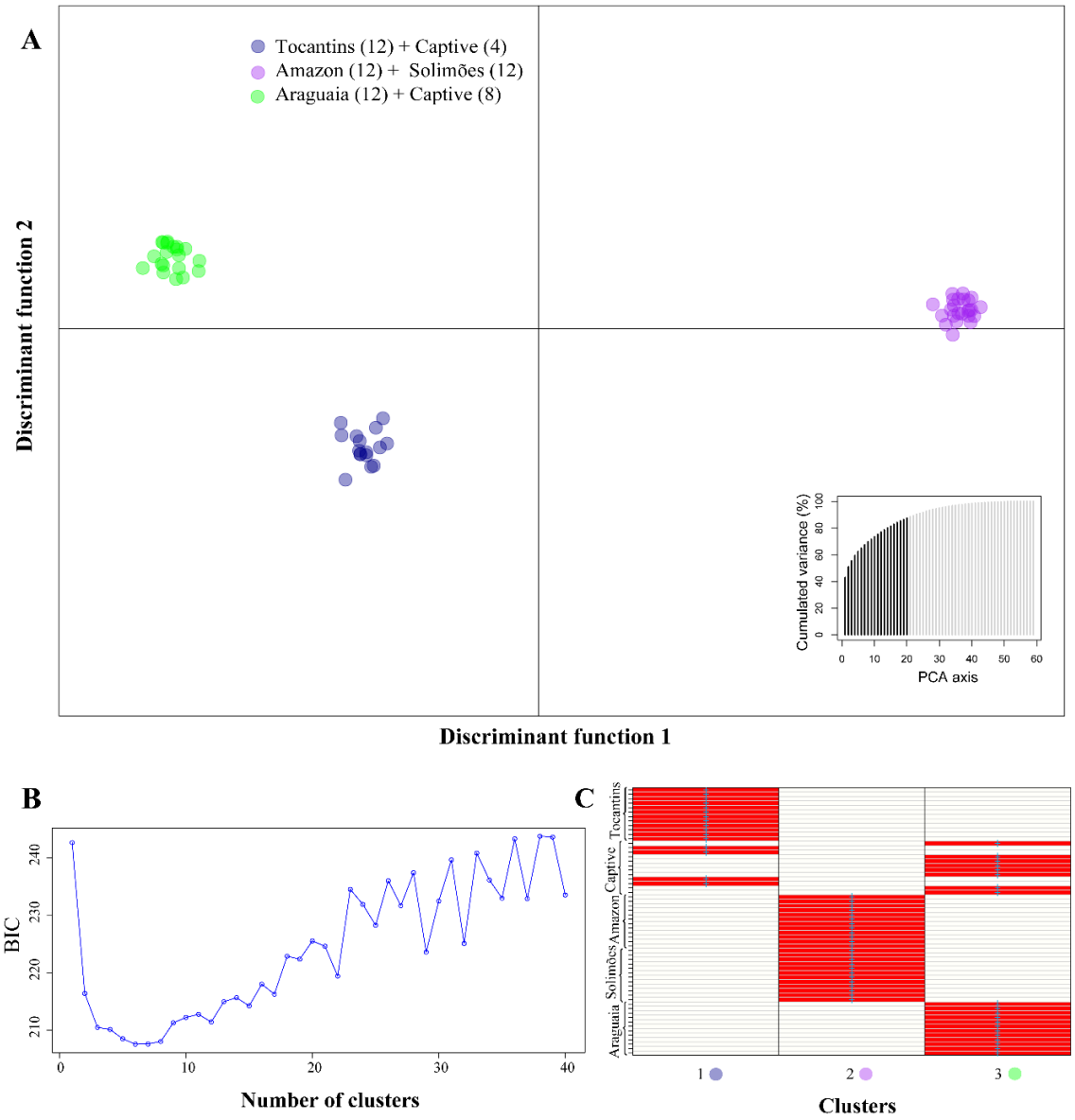

Figure 2.6. A. Discriminant analysis of principal components (DAPC) using 392 SNP markers in *Adegenet* v. 2.0.1 (Jombart, 2008) for five population samples of *Arapaima gigas* (n = 60 individuals). B. Selection of number of cluster was based on value of Bayesian Inference Criterion (BIC), which indicates 3 clusters for data summarization (elbow drop). C. Membership probabilities (red = 1, white = 0) for individuals into clusters, blue crosses indicate the prior cluster provided into DAPC.

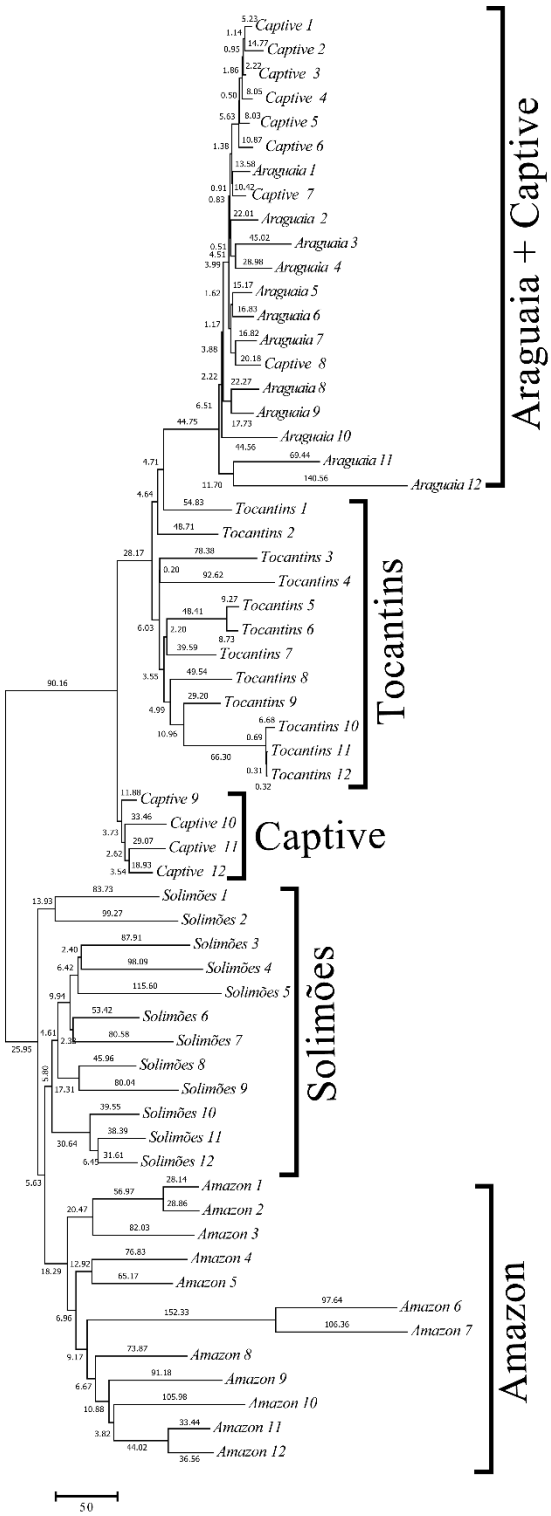

Figure 2.7. Neighbor-Joining phylogram showing relationship among 60 individuals of *Arapaima gigas* built in MEGA version 7.0.21. The branch length of the optimal tree is 3,574.9. The tree is drawn to scale, with branch lengths (next to the branches) based on a matrix of genetic distance (GD) among individuals produced in GenAlEx v 6.5.

2.4. Discussion

This study evaluated the degree of polymorphism in *A. gigas* using ddRAD sequencing and identified a nuclear SNP panel for population genetic inference. The work was intended as a pilot study to ascertain how useful the approach could be in the study of this iconic fish species. Access to samples was limited and the budget constrained. Thus, only a small number of individuals per river and only five population samples in total were available from the wide geographical range of the species. As such only tentative observations and conclusions can be drawn from the data generated, and sampling bias should always be borne in mind when interpreting population genetic data.

Samples from the Amazon, Solimões and Tocantins rivers were caught by individual fishermen which, given the solitary behaviour of adult *A. gigas* in the Amazon floodplains, is likely to have minimised sampling bias (Hrbek et al., 2005; Araripe et al., 2013). However, the samples from the Araguaia river were obtained during a rescue operation, when fish were trapped in a small lagoon during the dry season. These samples are likely to be more prone to bias, especially with regard to genetic relatedness. The origins of the captive broodstock are not documented, but likely were sourced from related families commercialised for aquaculture in North Brazil, which composed part of the broodstock analysed with endoscopy in Chapter 4 – Parts B and C.

Initially, this work confirmed the observation of low levels of genomic polymorphism in wild populations of *A. gigas*, which was apparent from an earlier attempt to generate a linkage map for this species which resulted in low numbers of SNPs (< 200) and prevented further QTL analysis for sex-associated markers. While the extent of polymorphism across a range of other species screened at the University of Stirling using the same approach were 3-8 times higher than that observed for *A. gigas*, this is not a unique situation for fish. Rondeau et al. (2014) reported very few SNPs being identified

(c. 1 % of RAD loci being polymorphic) in a RAD analysis of Northern pike (*Esox Lucius*), which undermined a planned linkage mapping study using RAD generated SNP data. The reasons for the relative paucity of genetic variability in *A. gigas* populations may be many and varied. It is likely that the lack of genetic variability within populations has been influenced by past bottleneck effects, both historic and those more recently documented, while genetic drift post population recovery has accentuated genetic differences among rivers and regions. The analysis of polymorphism levels in pristine populations from remote areas could help resolve this issue.

Although SNPs are considered less informative c.f. other markers used in previous population studies with *A. gigas* (*i.e.* microsatellites), hundreds of biallelic SNPs, even in reduced sample sizes, can be analysed to produce robust estimates of population diversity and structure, with the advantage that large numbers of SNPs invariably survey a large extent of the genome (Hohenlohe et al., 2012; Jeffries et al., 2016). The 393 SNP markers identified for *A. gigas*, were found to be in HWE and only two loci were found to be statistically associated (LD). Further population analyses identified 335 neutral and 57 potential outlier loci. Putatively outlier loci were incorporated in the dataset because they provide valuable information on local adaptation and their inclusion is recommended in analysis of threatened species for aiding define conservation units (Moore et al., 2014).

Population samples from the Amazon and Solimões rivers were more genetically variable (percentage of polymorphic loci, number of private alleles, Shannon's index (I) and observed heterozygosity (H_O)) compared to fish from the Araguaia. The river Tocantins population sample and Captive broodstock had intermediate diversity levels. These observations agree with previous studies using mtDNA and microsatellite markers which showed higher diversity in *A. gigas* sampled from the Amazon and Solimões

compared to river Tocantins (Hrbek et al., 2005; Hrbek et al., 2007), and very low levels of diversity in samples from the Araguaia river (Vitorino et al., 2015). The higher genetic diversity in populations from the Amazon basin have previously been explained as a consequence of accommodating larger population sizes of *A. gigas* in less urbanised environments compared to factors affecting the Araguaia, where the flooding regime is markedly different, the dry season being considered more severe and populations, for this reason, more depleted (Vitorino et al., 2015). Therefore, it would be desirable to analyse more populations from Araguaia-Tocantins and identify areas with higher diversity suitable for preservation as repositories of genetic diversity.

All five population samples showed significant levels of relatedness (r). Mean population relatedness was negatively correlated with diversity indexes such as observed heterozygosity (H_O) and Shannon's Index, which is indicative of a degree of inbreeding, similar to that reported for mudminnows (*Umbra krameri*) (Takács et al., 2015). While, to some extent, the relatively high relatedness within populations could partially explain the low level of genomic polymorphism detected by ddRAD sequencing, the low level of variation was detected across all 60 animals screened and five different sampling sites. A high degree of relatedness is also a characteristic of populations restocked with related individuals (Takács et al., 2015). Populations of *A. gigas* are believed to have been overexploited for over 200 years, with numbers of fish reduced to just 13 % of original population estimates, and have become extinct in many localities, with later restocking with related individuals being practised in some regions (Castello et al., 2011; Garcia-Dávila et al., 2011; Castello et al., 2015). Taken together, these biological particularities of *A. gigas* could explain the high levels of relatedness found in the wild population samples investigated in this study.

Based on F_{ST} pairwise comparisons, populations of *A. gigas* were demonstrated to be significantly differentiated from each other. The pattern of differentiation did not conform to an 'isolation by distance' model suggesting current population structure is likely to be mainly influenced by factors other than geographical separation. This mirrors findings from other genetic studies of *A. gigas* populations from Amazon, Tocantins and Araguaia rivers (Hrbek et al., 2005; Araripe et al., 2013; Vitorino et al., 2015), where past bottleneck events were considered to be the main factor underlying current population genetic structure (Hrbek et al., 2005; Vitorino et al., 2015). The relatively sedentary behaviour of *A. gigas* has been regarded as a major factor contributing to low genetic flow, resulting in population differentiation. Ecological studies using radio-telemetry to monitor wild individuals also found strong patterns of residency and territoriality in *A. gigas* (Núñez et al., 2015). Indeed, a study evaluating the dispersal capacity of *A. gigas* concluded the existence of high levels of genetic similarity among lakes separated by 25 km, moderate genetic differentiation in sites separated by 100 km and highest genetic differentiation among regions separated by > 1,500 km (Araripe et al., 2013), supporting previous ecological observations. In the current study, the high inter-population differentiation, the lack of support for isolation by distance, the elevated levels of mean population relatedness negatively correlated with diversity (I and H_O), together, corroborate the hypothesis of Hrbek et al. (2005) that past bottleneck effects and genetic drift led to a loss of genetic variability within and increased differentiation between populations of *A. gigas.*

An initial population genetic study of *A. gigas* within the Amazon basin, using two discontinuous mitochondrial DNA regions of 1204 base-pairs (bp) (NADH1 segment) and 1143 bp (ATPase segment) revealed minimal evidence of substructuring, which loosely fitted an isolation-by-distance model (Hrbek et al., 2005). A later study,

based on 7 microsatellite markers and using Bayesian analysis in STRUCTURE detected two distinct clusters, one comprising fish mostly from the lower Tocantins and lower Amazon rivers, the other comprising *A. gigas* predominantly from the mid- region of the Amazon river (Araripe et al., 2013). A more recent study of population samples from the Essequibo river basin employing 11 hypervariable microsatellite markers and mtDNA markers (NADH1 segment), identified patterns of allopatric differentiation within the species, the authors suggesting that this pointed to sympatric "species" being present in three sampled sites (Watson et al., 2016). This latter interpretation was influenced by recent proposals to consider *Arapaima* as a multispecies group based on morphological analyses (Stewart, 2013a; Stewart, 2013b). The current SNP-based analysis contrasts with these previous findings as it identified significant substructuring both within and between river basins. Though the data did not fit an isolation by distance model, this cannot be ruled out as only a few population samples were available for study. Though some pairwise F_{ST} values were high, there was no convincing genetic evidence to suggest that genetically distinct species were sampled. This was supported by the observation that > 96 % of RAD loci identified as polymorphic contained just one or two SNPs, a higher proportion than observed for all other fish species studied by ddRAD at the University of Stirling.

Though restricted by the unusually low level of polymorphism detected within the Arapaima genome, the screening of hundreds of SNPs using ddRAD technology is more than sufficient to generate reliable measures of genetic variability within and between arapaima population samples. Furthermore, these genetic markers have a role to play in identifying the origin of animals used in aquaculture, and in informing the maintenance of the genetic diversity of captive reared broodstocks. There is also the opportunity to develop a subset of the most informative SNPs for screening using

alternative platforms (qPCR assays; small scale SNP chips), which require less labour intensive and less expensive protocols. Such a panel would probably be better suited than RAD for the much more extensive survey of Arapaima populations that is required to better understand and manage the species. As a priority, surveys should focus on the different morphotypes of *A. gigas* (orange-fleshed and white-fleshed) and the contended species recently described for *Arapaima* (Stewart, 2013a; Stewart, 2013b) which currently lacks support from molecular data. More detailed genetic studies (*i.e*. genome scans, QTL analyses, linkage mapping) need much more dense marker panels. While this would be feasible using RAD approaches in arapaima, it is clear that it would require the selection of restriction enzymes that cut Arapaima DNA much more frequently, which would necessitate a 10-50 fold higher sequencing effort than currently used, with associated budget implications.

CHAPTER 3

3. Chapter 3

EFFECTS OF GNRHA IMPLANTS AND SIZE PAIRING ON PLASMA AND CEPHALIC SECRETION SEX STEROIDS IN *ARAPAIMA GIGAS* (SCHINZ, 1822).

Abstract

Arapaima gigas is one of the world's largest freshwater fish and is considered an emerging species for aquaculture development due to high growth rates and meat quality. However, the lack of reproductive control in captivity has limited the expansion of *Arapaima* farming. This study aimed to test the effects of hormonal induction using GnRHa implants and size pairing on broodstock reproduction through the analyses of sex steroids. To do so, broodstock of different sizes (large, small or mixed) were paired and implanted. Plasma and cephalic secretion profiles of testosterone (T), 11-ketotestosterone (11-KT) and 17β-oestradiol (E_2) were analysed. Compared to control (non-implanted), implanted broodstock showed a significant increase in plasma 11-KT (large and small males) and T (large and mixed females) post GnRHa implantation. In females, despite the significant increase in plasma T levels, E_2 remained unchanged after implantation. Despite the lack of clear spawning induction, this study showed the potency of GnRHa on sex steroid production regardless of pairing groups. Interestingly, significant correlations between blood plasma and cephalic secretion levels of 11-KT in males and T in females were observed, suggesting the possibility of hormonal pheromone release through the cephalic canals of *A. gigas*, a biological fluid still poorly investigated and a possible novel source of pheromone release in teleosts.

Keywords: Hormonal induction, lateral line, pheromones, Pirarucu, reproduction.

3.1. Introduction

Current knowledge on the biology of the Amazon Pirarucu *Arapaima gigas* (Schinz, 1822) remains scarce with regards to wild and captive populations, conservation and reproduction (Castello & Stewart, 2010). The Pirarucu has been considered as an emerging species for aquaculture diversification in South America with strong market demand due to its growth potential and meat quality. *A. gigas* is an obligate air-breather species, and is one of the largest scaled freshwater fish in the world (Nelson et al., 2016), reaching more than 250 kg in the wild and with a growth potential of 10kg+ in 12 months (Oliveira et al., 2012). However, achieving consistent spawning in captivity has remained the key challenge over the past decades that has prevented the expansion of the industry (Farias et al., 2015). Consequently, increased pressure on wild capture to meet market demands has resulted in *A. gigas* being placed on the CITES threatened list (Castello & Stewart, 2010). As such there is a clear need to develop protocols to induce spawning in captivity and understand the factors influencing the reproductive success of *A. gigas*.

Arapaima gigas is gonochoristic and iteroparous with fish reaching first sexual maturity after three to five years of age (Godinho et al., 2005). In their Amazonian habitat, spawning occurs year-round peaking in the rainy season from December to May (Castello, 2008b; Núñez et al., 2011). Breeding couples build nests in shallow flooded areas (*c.* 1-1.5 m depth) where mating, spawning and external fertilization occurs. After spawning, the couple guard the nest for approximately nine days and parental care is performed by the male for approximately three months (Castello, 2008b). During parental care the male's head and trunk becomes dorsally darkened providing camouflage for the offspring. Females can mate with several males and spawn multiple times during a reproductive season (Farias et al., 2015). This is made possible since ovary development is asynchronous with several batches of vitellogenic oocytes recruited for maturation

during the reproductive season (Godinho et al., 2005). On the other hand, males have a tubular cord-like left testis, which after spermiation will still contain lobules of spermatozoa and semen, allowing multiple spermiation during a reproductive season in cases when parental care is interrupted (Núñez & Duponchelle, 2009). Successful reproduction of *A. gigas* in captivity is problematic, and to date, isolating pairs in earth ponds during the rainy season appears to stimulate reproduction, although outcomes are often variable and with very limited success (Núñez et al., 2011).

Reproductive dysfunction in fish reared in captivity is common. In most cases, spawning can be induced through the use of hormonal therapies (Mylonas & Zohar, 2001; Mylonas et al., 2010). While there are several hormones associated with stimulation of the brain-pituitary-gonad axis (BPG) to artificially induce oocyte maturation, ovulation/spermiation and spawning in fish, hypothalamic gonadotropin releasing hormone (GnRH) is considered the most potent, safe and reliable hormone to use (Mylonas et al., 2010). GnRH is a decapeptide involved in the initiation of the BPG axis through the stimulation of gonadotropic cells in the pituitary to produce and release gonadotropins which in turn stimulate steroidogenic cells in the gonads to produce sex steroids regulating gametogenesis and spawning (Golan et al., 2015). To date, 25 forms of GnRH have been described in chordates and grouped into three main classes (GnRH 1, 2 and 3) based on molecular phylogeny (Fernald & White, 1999; Kim et al., 2011). The brain of teleosts normally express two or three GnRH forms with GnRH1 believed to regulate reproduction in most species whereas GnRH 2 and 3 act as neuromodulators of reproductive behaviours (Okubo & Nagahama, 2008). In fish hatcheries, GnRH analogues (GnRHa) are used to stimulate oocyte maturation and spermiation since they have an increased resistance to enzymatic cleavage compared to the native forms (Mylonas & Zohar, 2001). For asynchronous spawners such as *A. gigas*, the use of slow-

release implants is preferred rather than multiple injections, as it promotes a sustained elevation in gonadotropins reducing stress caused by repetitive handling (Mylonas et al., 2010). In osteoglossids and especially *A. gigas*, responsiveness of captive broodstock either to hormonal or environmental manipulations has not yet been examined.

Development of protocols to induce gonadal recruitment, gametogenesis and spontaneous spawning must consider technical limitations related to the biological features and reproductive strategy traits of a species (Mylonas et al., 2010). Varying widely among teleosts, paired mating systems are often associated with male body size or behavioural characteristics. To date, little is known about the social, behavioural and physiological factors controlling mating in *A. gigas.* This includes a lack of knowledge on body weight criteria for pairing fish in captivity, a critical component in the breeding system of many fish species (Lehtonen et al., 2015). When breeding pairs are isolated in captivity, knowledge on mating preferences and success rates are unknown and potential drawbacks may also exist regarding male-female agonistic interactions resulting in unsuccessful mating. When pairing couples, gender identification in *A. gigas* is another key limitation as the species is not sexually dimorphic. Several techniques have been used to sex fish including colour patterns, laparoscopy to visualise the gonads, vitellogenin measurement to identify females and sex-specific DNA markers but these can be unreliable, invasive or expensive (Chu-Koo et al., 2009; Carreiro et al., 2011; Almeida et al., 2013). Given the gonopore is not externally visible in *A. gigas*, cannulation to obtain ovarian biopsies is difficult and gonadal development can therefore not be assessed easily, which is a critical component to develop protocols for induced ovulation/spermiation and spawning (Mylonas et al., 2010; Torati et al., 2016). Likewise, stripping of gametes for artificial fertilisation, routinely done in many other species, is not suitable due to the species' thick abdominal body wall preventing artificial stripping

and collection of eggs and milt. In *A. gigas*, reproductive success in ponds cannot be confirmed through observation of spawning behaviour, oviposition or assessment of gonadal development. Proxy indicators of reproductive success such as a cessation of feeding behaviour and male darkening have been applied before (Monteiro et al., 2010) although these are not always reliable. Given these limitations, the profiling of sex steroids following GnRHa induction is therefore particularly important in this species to assess its impact on the BPG axis directly associated to gonad maturation.

In teleosts, the main sex steroids regulating spermatogenesis in males are testosterone (T) and 11-ketotestosterone (11-KT), whereas 17β-oestradiol (E_2) is directly related with vitellogenesis in females (Lubzens et al., 2010; Schulz et al., 2010). Fish also use sex steroids such as 11-KT, maturation inducing steroids (MIS) and prostaglandins as pheromones to convey information on gender and reproductive condition to conspecifics to help synchronise the release of gametes and external fertilization (Kobayashi et al., 2002). In most fish species, pheromones are released through the gills, urine, seminal and ovarian fluids (Sorensen & Wisenden, 2015). In *A. gigas*, the cephalic secretion released from the head sensorial cavities have been suggested to play a role in pheromone release, although supporting evidence is still limited (Lüling, 1964; Amaral, 2009).

The aims of the present study were to: 1) test the effects of GnRHa slow-release implants on sex steroid profiles measured in blood plasma and cephalic secretion; and 2) examine the effects of different size pairings on reproductive success and sex steroid profiles.

3.2. Materials and Methods

3.2.1. Experimental set up

This experiment was conducted at the *Rodolpho von Ihering* Station - DNOCS (3°48'09.54"S, 39°15'56.73"W) in Pentecoste-CE (Northeast Brazil). A total of 59 adult captive-reared broodstock of approximately the same age (over six-year-old) had been previously held (since 2013) in two large earth ponds: 8 females with 11 males in a 2300 m^2 pond and 19 females with 21 males in a 930 m^2 pond. In these ponds and with such a fish density, spontaneous spawning is normally not observed in the species, though some of the broodstocks used in this study reproduced in previous years (2010-2011) after isolation of pairs in smaller earthponds (Rebouças et al., 2014).

The experiment started on January 21st 2014 (day 0), when broodstock were measured for body weight (BW) (± 0.1 kg) and total length (TL) (± 0.1 cm), and Fulton's condition factor (K) was calculated as $K=(BW\times100)/TL^3$ (Froese, 2006). Fish were photographed and implanted with a passive integrated transponder (PIT; AnimallTAG®, São Carlos, Brazil) in the dorsal muscle to allow individual identification. Each fish was sexed using a specific *A. gigas* vitellogenin enzyme immune assay (EIA) kit using the manufacturer's protocol (Acobiom, Montpellier, France) based on the work of Dugue et al. (2008). Based on BW, each female was paired with a single male and pairs were allocated into 18 earthen ponds of 330 m^2 and depth of 1.95 ± 0.06 m (deepest point) (Fig. 3.1A). Four treatments were tested: control (no implant) with large fish couples (53.8 ± 3.3 kg; n=5), and three GnRHa implanted groups: large (58.8 ± 5.3 kg; n=5), small (29.8 ± 5.0 kg; n=3) or mixed size (large female: 56.1 ± 4.1 kg paired with small male: 21.5 ± 1.8 kg, n=5) couples. Means of fish BW, TL and K among sex and treatments are presented in Table 3.1.

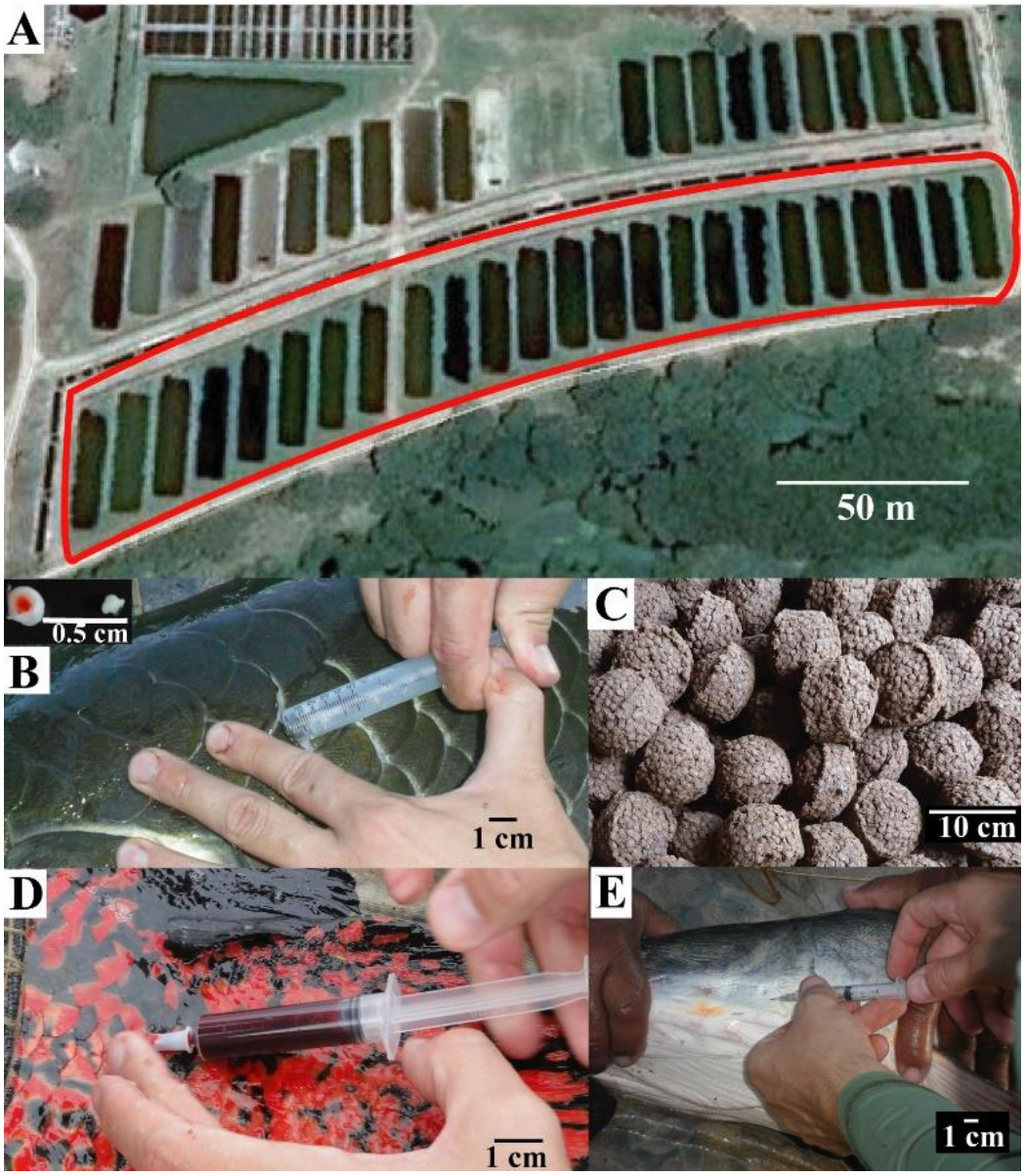

Figure 3.1. *Arapaima gigas* experimental details. A. Site indicating earthponds used for pair allocation (image from http://www.google.com/earth/index.html; accessed at 13.10.16), B. GnRHa slow-release implant and fish implantation in dorsal muscle, C. Feed pellet offered to fish during the trial, D. Sampling of blood from the caudal vein, E. Sampling of cephalic secretion from preopercle cavity.

Table 3.1. Body weight (BW, kg), total length (TL, cm) and Fulton's condition factor (K) in control, large, small and mixed size-pairing couples. Values are presented as mean (± SD).

	Sex	BW (kg)	TL (cm)	K
Control (*n* = 5)	Male	54.1±3.6	185.8±2.6	0.85±0.08
	Female	53.6±3.4	186.6±4.3	0.82±0.02
Large (*n* = 5)	Male	57.4±3.8	188.8±3.3	0.85±0.04
	Female	60.2±6.7	189.2±5.5	0.89±0.09
Small (*n* = 3)	Male	26.7±2.5	143.0±0.9	0.91±0.06
	Female	33.0±5.2	157.0±8.5	0.85±0.06
Mixed (*n* = 5)	Male	21.5±1.8	134.8±3.9	0.88±0.06
	Female	56.1±4.1	187.0±2.5	0.86±0.05

At day 62 post pairing and stocking into ponds, treated couples received a GnRHa slow-release implant (Center of Marine Biotechnology, Baltimore, MD, USA) with a dose of 84.7 ± 8.7 µg.kg^{-1} for females and 49.1 ± 6.7 µg.kg^{-1} for males (Fig. 3.1B). Implants were made with ethylene-vinyl acetate polymer (EVAc) delivering desGly^{10}, DAla^{6}, Pro^{9}-GnRH- NEthylamide for approximately 21 days (Mylonas et al., 2007). Each implant was inserted in the dorsal muscle using an implanter (Fig. 3.1B).

Fish were fed once a day *ad libitum* with 160 g floating balls made with a commercial ration (38 % crude protein, Aquamix, Brazil) mixed with 10 % tilapia flesh (*Oreochromis niloticus*) (Fig. 3.1C). Given water turbidity in the ponds hindered the possibility to directly observe spawnings, daily feed intake (*i.e.* number of floating balls consumed) per couple was recorded, and cessation in feeding and couples swimming at the same location for long periods were monitored as *proxy* indicators of mating and nest guarding behaviour (Fontenele, 1953). Also, colour pattern changes (darkening in males) and one eventual egg release during a sampling was recorded. With such limitations to infer reproductive activity, effects of GnRHa implantation will be mainly reported as hormonal profiles.

This experiment occurred under natural photo-thermal regimes. Figure 3.2 shows rainfall variation during the study. Climatic data was obtained from the *National Institute for Space Research* (INPE, *bancodedados.cptec.inpe.br*) and photoperiod from the R package "*StreamMetabolism*" (R-Core-Team, 2014). Maximum daily rainfall was 819 mm and air temperature ranged from 23.8 to 30.5 °C. The last meaningful rain (204.8 mm) occurred at day 112. Photoperiod ranged from 11.9 to 12.3 hours of photophase during the study.

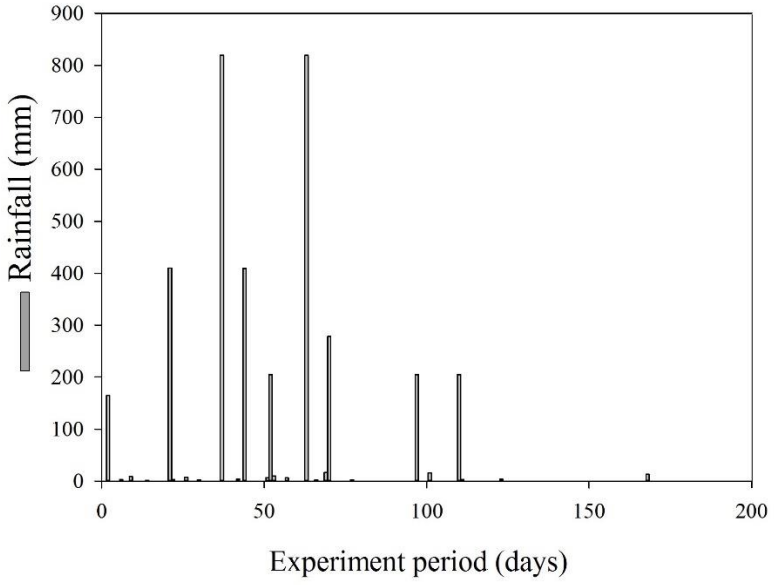

Figure 3.2. Rainfall (mm) recorded in Pentecoste-CE (Brazil) during the experiment (data obtained from the *National Institute for Space Research* -INPE; *bancodedados.cptec.inpe.br*).

3.2.2. Sampling procedures

All broodfish were sampled for blood and cephalic secretion at couple allocation (day 0), GnRHa implantation (day 62), two weeks post-implantation (day 76) and then monthly thereafter (days 111, 146 and 181). At each sample point, ponds were sampled in the same daily order and between 6:00 to 10:00am. Prior to each sampling, fish were fasted for 24 hours. Fish were netted from earthponds and contained in cylinder-shaped net, then

kept on a soft wet mat with eyes covered with a wet cloth for approximately 5-10 minutes. Anaesthetics were not applied during sampling as anaesthesia has been shown to compromise welfare and result in mortalities in *A. gigas* due to its air breathing behaviour (Farrel & Randall, 1978). Netting, handling procedures and equipment used in this study are described in Lima et al. (2015b). Fish breathing behaviour was closely monitored during each procedure (breathing at regular intervals of 4-6 minutes). Fish were photographed to analyse colour patterns. Approximately 4 ml of blood was sampled from the caudal vein using syringes (BD Precisionglide, New Jersey, USA) flushed with 560 $IU.ml^{-1}$ heparin ammonium salt solution (Sigma Aldrich, Saint Louis, MO, USA) (Fig. 3.1D). Plasma was collected by centrifugation at 1200 g at 4°C for 15 minutes, stored in cryovials and frozen in liquid nitrogen. Cephalic fluid (2-3 ml) was sampled from the dorsalmost lateralis cavity of the preopercle using a sterile syringe carefully inserted underneath the dermis sensorial cavity, then immediately frozen in liquid nitrogen (Fig. 3.1E). Fish were then returned to the ponds and monitored until normal breathing behaviour returned. Due to unknown reasons, one female from the small group and one male from the mixed group died after samplings on 13th May 2013 (day 111).

Samples were transported to EMBRAPA research centre in Fortaleza (Brazil) and stored at -80°C, and then shipped on dry ice to the University of Stirling (Stirling, Scotland) for analyses (Permit IBAMA/CITES n°14BR015849/DF and 14BR015850/DF). This research complied with the Brazilian guidelines for the care and use of animals for scientific and educational purposes (CEUA/CNPASA n°09).

3.2.3. Steroid analyses

Samples of plasma and cephalic secretion were thawed at room temperature, and extraction of sex steroids made in 1 ml of ethyl acetate using 50 µl of plasma of 100 µl

for cephalic secretion. The mixture was vigorously vortexed and spun for 1 hour at room temperature (16°C) using a rotary mixer, then centrifuged at 430 g for 10 min at 4°C. Just prior to assay, extracts were dried in a vacuum oven at 35°C for 40 minutes. Given sample variability in steroid concentration, different amounts of extract (50-100 µl for plasma; 50-250 µl for cephalic secretion) were used in assays. Levels of testosterone (T) and 17β-oestradiol (E_2) in plasma and cephalic secretion were quantified by radioimmunoassay (RIA), following methods developed and detailed in Duston & Bromage (1987). Tritiated radiolabels for T (GE Healthcare, UK) and E_2 (PerkinElmer, Boston, USA) were used with anti-T and anti-E_2 antisera (CER group, Marloie, Belgium). Radioactivity was measured using a Packard 1900 TR Liquid Scintillation Analyser (Pangbourne, UK). For 11-ketotestosterone (11-KT), an enzyme-linked immunosorbent assay (ELISA) kit was used (Cayman Chemical Inc., Michigan, USA) following the manufacturer's protocol using samples extracted in ethyl acetate. Given sample variability, dilutions in plasma samples ranged from 1:5 to 1:30 (v:v). For analysis of cephalic secretion, samples were either concentrated 4:1 or diluted 1:2. Microplates were read at 405 nm using an ELX808 reader (Biotek, Swindon, UK). T and E_2 RIA and 11-KT ELISA were validated for *A. gigas* through parallelism comparing serial dilutions of extracts to known concentrations of hormone standards as described in Sink et al. (2008) (data not shown). Concentration of steroids in the blood or cephalic secretion were calculated from the value yielded in the assay (pg. $tube^{-1}$) corrected for: (a) proportion of extract added to the assay tube and (b) volume of blood or cephalic secretion used for extraction. Samples were assayed in duplicate with intra- and inter- assay coefficients of variation being respectively 12.0 and 6.6 % for T (7 assays), 12.6 and 9.8 % for E_2 (7 assays) and 9.7 and 10.0 % for 11-KT (4 assays).

3.2.4. Statistical analysis

Statistical analyses were conducted with Minitab (version 17.3.1, Minitab, PA, USA). Non-parametric Kruskal–Wallis one-way ANOVA and Dunn's pairwise *post hoc* tests were used to compare GnRHa effect between treatments using the level change from implantation (day 62) as the primary measure. One-way repeated measures ANOVA followed by Tukey *post hoc* tests were used to describe time effects within treatments. Pearson Product Moment Correlations were used to correlate and compare steroid levels in blood plasma and cephalic secretion. Level of significance was set as $P \leq 0.05$ and data are presented as mean ± SE unless stated otherwise.

3.3. Results

3.3.1. Effects of GnRHa implants on sex steroid levels in plasma and cephalic secretion

3.3.1.1. Testosterone in males

Plasma T levels in implanted groups were not significantly different from control after 14 and 49 days post implantation (dpi) (Fig. 3.3A). Overall, no significant time effects were found in T plasma levels measured in control and small groups (Fig. 3.4A and C). In large couples, levels increased 2-fold (from 20.7 to 41.4 $ng.ml^{-1}$) after 14 dpi (Fig. 4B; $P<0.05$). This increase was significant compared to small couples at 14 and 49 dpi ($P<0.05$; Fig. 3.3A). In small males of the mixed group, levels increased significantly from the pair allocation into earthponds (day 0) until day 76 ($P<0.05$; Fig. 3.4D).

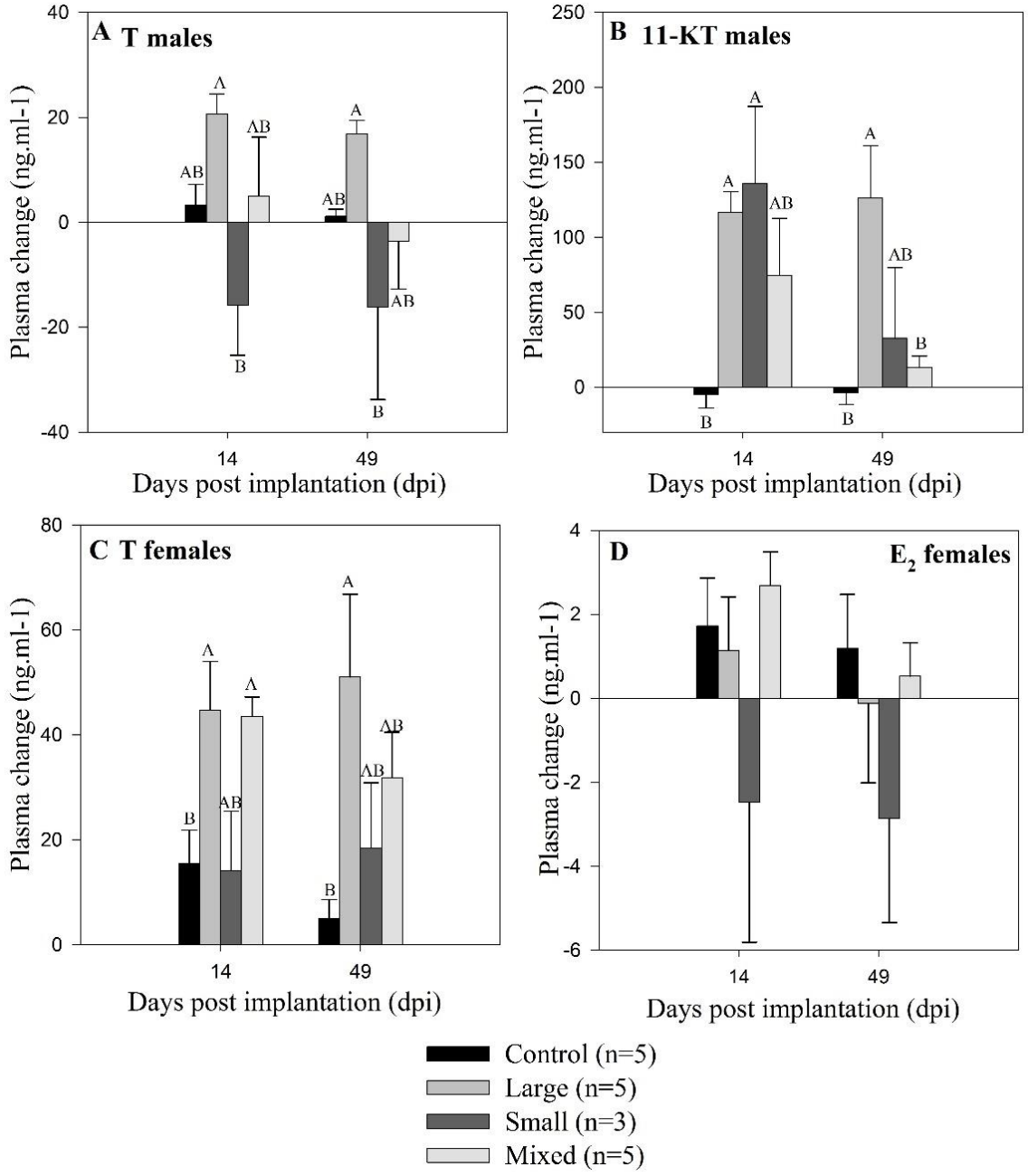

Figure 3.3. Post-implantation changes (at 14 and 49 days post implantation, dpi) in plasma sex steroid levels (ng.ml^{-1}). A and B. Testosterone (T) and 11-ketotestosterone (11-KT) in males, respectively. C and D. Testosterone and 17β-oestradiol (E_2) in females, respectively. Data presented as mean ± SE. Different uppercase letters denote statistical difference among groups at a given time ($P<0.05$).

In the cephalic secretion, no time effects were observed in T levels from all groups (Fig. 3.4). Correlation between T levels in blood and cephalic secretion was significant (r^2=0.05; P<0.05), although linearity assumption was not met (Fig. 3.5).

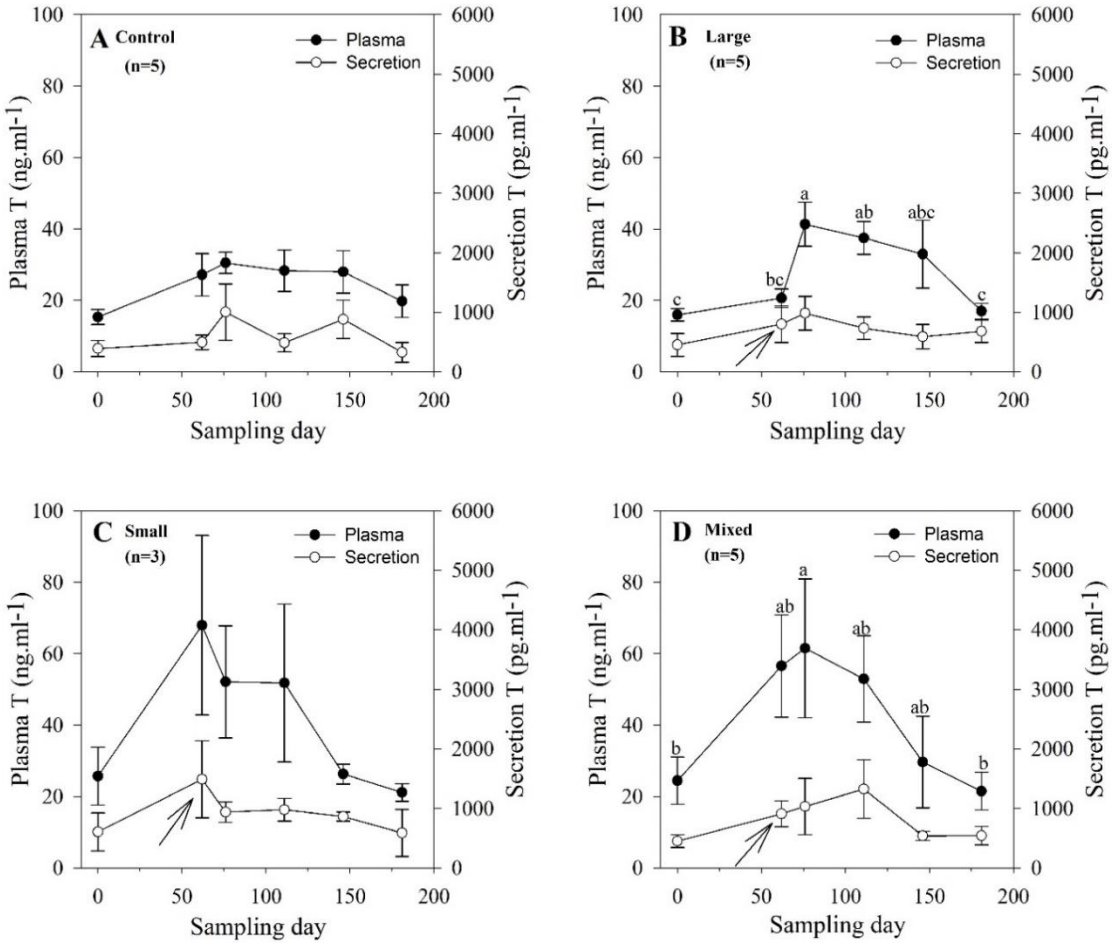

Figure 3.4. Testosterone (T) levels in plasma (ng.ml^{-1}) and cephalic secretion (pg.ml^{-1}) in *Arapaima gigas* males from day 0 (couple pairing and stocking into ponds) to day 181 (119 days post GnRHa implantation) in the four experimental groups: A. Control couples (n = 5), B. Large implanted couples (n = 5), C. Small implanted couples (n = 3) and D. Mixed size implanted couples (n = 5). Values are presented as mean ± SE. Lowercase superscripts denote time effect in blood plasma levels (P<0.05). Arrows indicate GnRHa implantation time.

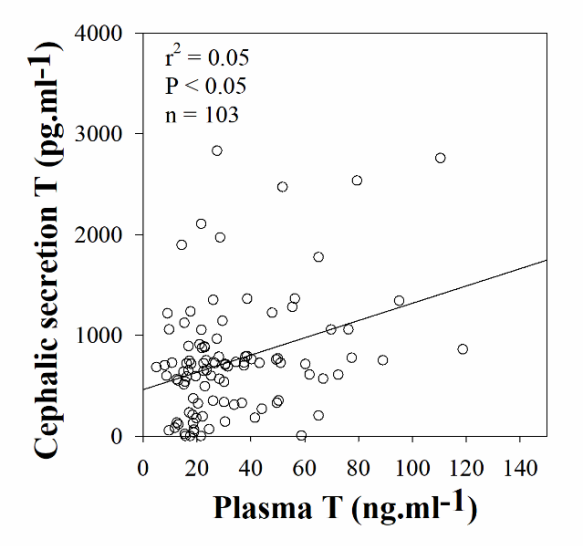

Figure 3.5. Testosterone (T) levels in male *Arapaima gigas* broodstock. Correlation between blood plasma (ng.ml^{-1}) and cephalic secretion (pg.ml^{-1}) levels. Pearson Product Correlation Coefficients were calculated on log-transformed data.

3.3.1.2. 11-ketotestosterone in males

Plasma 11-KT levels remained below 30 ng.ml^{-1} in males from the control couples throughout the experimental study. Plasma 11-KT levels increased in males from large (7.3-fold, from 13.6 to 135.2 ng.ml^{-1}) and small (2.3-fold from 106.8 to 242.9 ng.ml^{-1}) couples after 14 dpi ($P<0.05$; Fig. 3.3B, 3.6B). This increase was sustained until 49 dpi in large couples but not small, compared to control couples and returned to basal pre-implantation levels after 84 dpi (Fig. 3.3B, 3.6B). In small group and small males of mixed group plasma 11-KT levels increased significantly between day 0 (allocation to earthponds) and day 76 (Figs. 3.6C-D; $P<0.05$). No time effects were found in control couples (Fig. 3.6A).

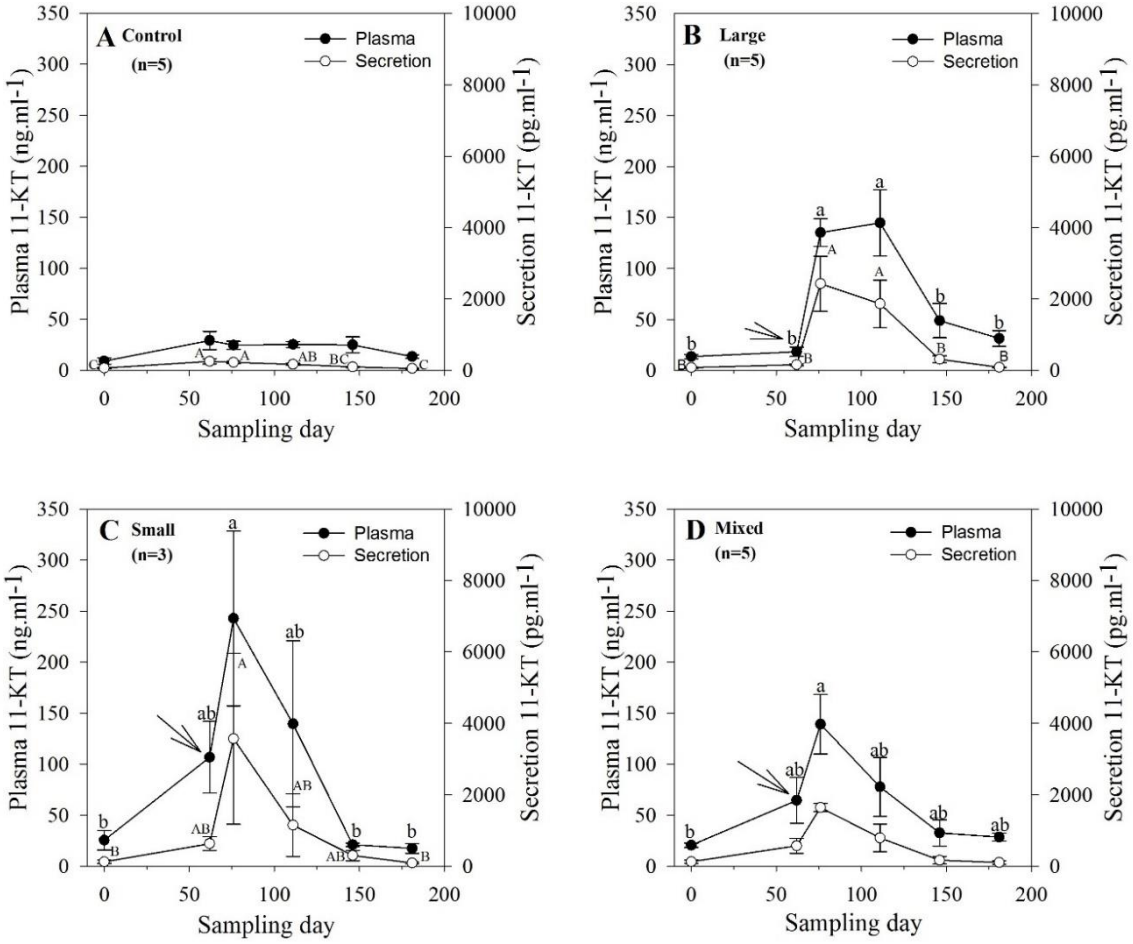

Figure 3.6. Levels of 11-ketotestosterone (11-KT) in the plasma (ng.ml^{-1}) and cephalic secretion (pg.ml^{-1}) in males of *Arapaima gigas* from day 0 (couple pairing and stocking into ponds) to day 181 (119 days post GnRHa implantation) in the four experimental groups: A. Control couples (n = 5), B. Large implanted couples (n = 5), C. Small implanted couples (n = 3) and D. Mixed size implanted couples (n = 5). Values are presented as mean ± SE. Lowercase and uppercase superscripts denote time effects in blood plasma and cephalic secretion levels, respectively ($P<0.05$). Arrows indicate GnRHa implantation time.

In the cephalic secretion, 11-KT levels of control group increased 4-fold (from 63.8 pg.ml^{-1} to 252 pg.ml^{-1}) between day 0 (pair allocation into earthponds) and day 62 ($P<0.05$, Fig. 3.6A). In large couples, levels increased 15-fold (from 162.4 to 2433.1

pg.ml^{-1}) between implantation and 14 dpi (P<0.05), returning to pre-implantation levels from 84 dpi onwards (Fig. 3.6B). In small couples, 11-KT levels increased from day 0 to day 76 (P<0.05, Fig. 3.6C). In small males of mixed group, levels remained unchanged over time (Fig. 3.6D).

Plasma T and 11-KT levels were positively correlated in males (r^2=0.62; P<0.001) (Fig. 3.7A). In addition, a significant positive correlation was found between 11-KT levels in blood plasma and cephalic secretion (r^2=0.74; P<0.001, Fig. 3.7B).

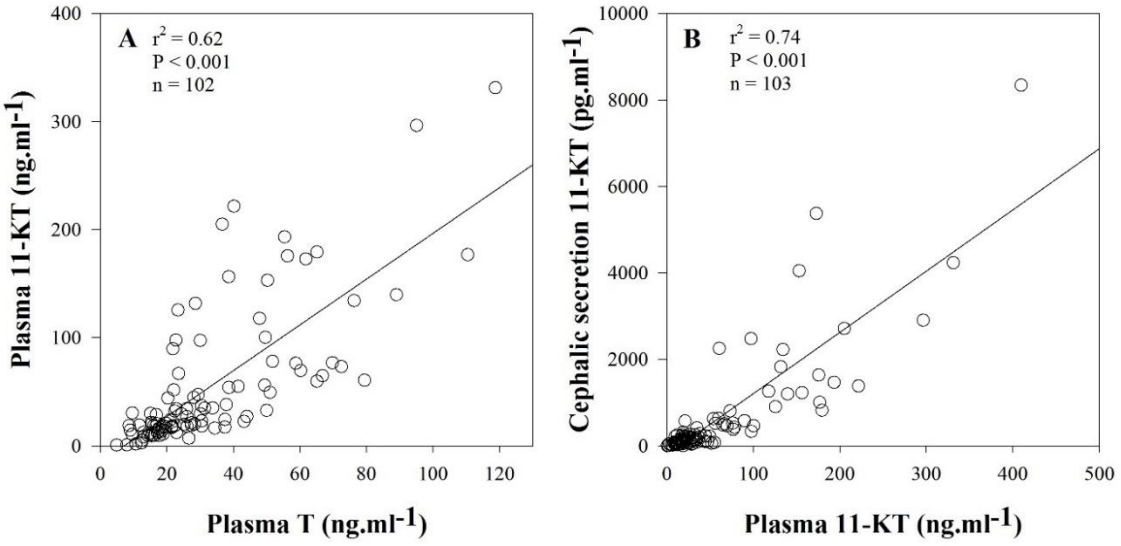

Figure 3.7. Sex steroid correlations in male *Arapaima gigas* broodstock. A. Correlation between blood plasma levels (ng.ml^{-1}) of Testosterone (T) and 11-ketotestosterone (11-KT). B. Correlation between 11-KT levels in blood plasma (ng.ml^{-1}) and cephalic secretion (pg.ml^{-1}). Pearson Product Correlation Coefficients were calculated on log-transformed data.

3.3.1.3. Testosterone in females

Plasma T levels remained below 40 ng.ml^{-1} in control couples throughout the experimental study with levels significantly higher at day 76 than 181 (Fig. 3.8A). Levels increased significantly in large (3.2-fold from 20.0 to 64.7 ng.ml^{-1}) and large females of

mixed couples (4.5-fold, from 12.4 to 56.0 ng.ml^{-1}) at 14 dpi (P<0.05; Fig. 3.3C; 3.8B-D). This increase was sustained until 49 dpi in both groups and resumed to basal pre-implantation levels by the end of the study (Fig. 3.3C; 3.8B-D). In small couples, no time effects were seen (Fig. 3.3C, 3.8C).

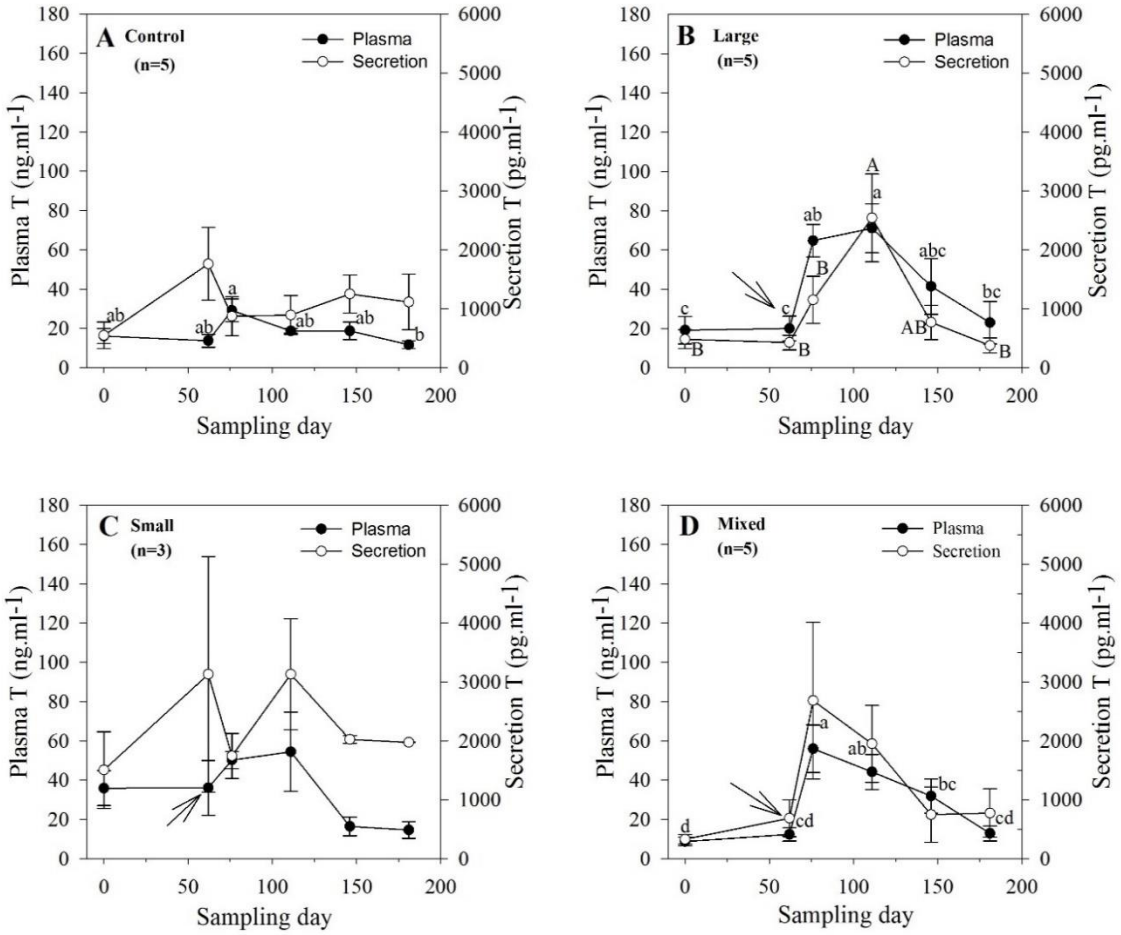

Figure 3.8. Levels of testosterone (T) in the plasma (ng.ml^{-1}) and cephalic secretion (pg.ml^{-1}) in females of *Arapaima gigas* from day 0 (couple pairing and stocking into ponds) to day 181 (119 days post GnRHa implantation) in the four experimental groups: A. Control couples (n = 5), B. Large implanted couples (n = 5), C. Small implanted couples (n = 3) and D. Mixed size implanted couples (n = 5). Values are presented as mean ± SE. Lowercase and uppercase superscripts denote time effects in blood plasma and cephalic secretion levels, respectively (P<0.05). Arrows indicate GnRHa implantation time.

No significant time effects were observed in female T levels measured in the cephalic secretion of control, small and small females of mixed group although levels appeared to increase in small females of mixed couples at 14 dpi and returned to pre-implantation basal levels by the end of the study (Fig. 3.8A, C, D). In large pairs, T levels raised by 6-fold (from 429.4 to 2544.0 $pg.ml^{-1}$) between implantation (day 62) and 49 dpi ($P<0.05$; Fig. 3.8B). A significant positive correlation was found between T levels in blood plasma and cephalic secretion in females ($r^2=0.25$; $P<0.001$; Fig. 3.10A).

3.3.1.4. 17β-oestradiol in females

No significant time effects were seen in E_2 levels measured in blood plasma and cephalic secretion along the study period for any of the GnRHa implanted groups (Fig. 3.3D, 3.9). No statistical correlation was found between E_2 levels in blood plasma and cephalic secretion ($r^2=005$; $P=0.453$, Fig. 3.10B). Plasma T levels showed a significant positive correlation with plasma E_2 levels ($r^2=0.36$; $P<0.001$, Fig. 3.10C).

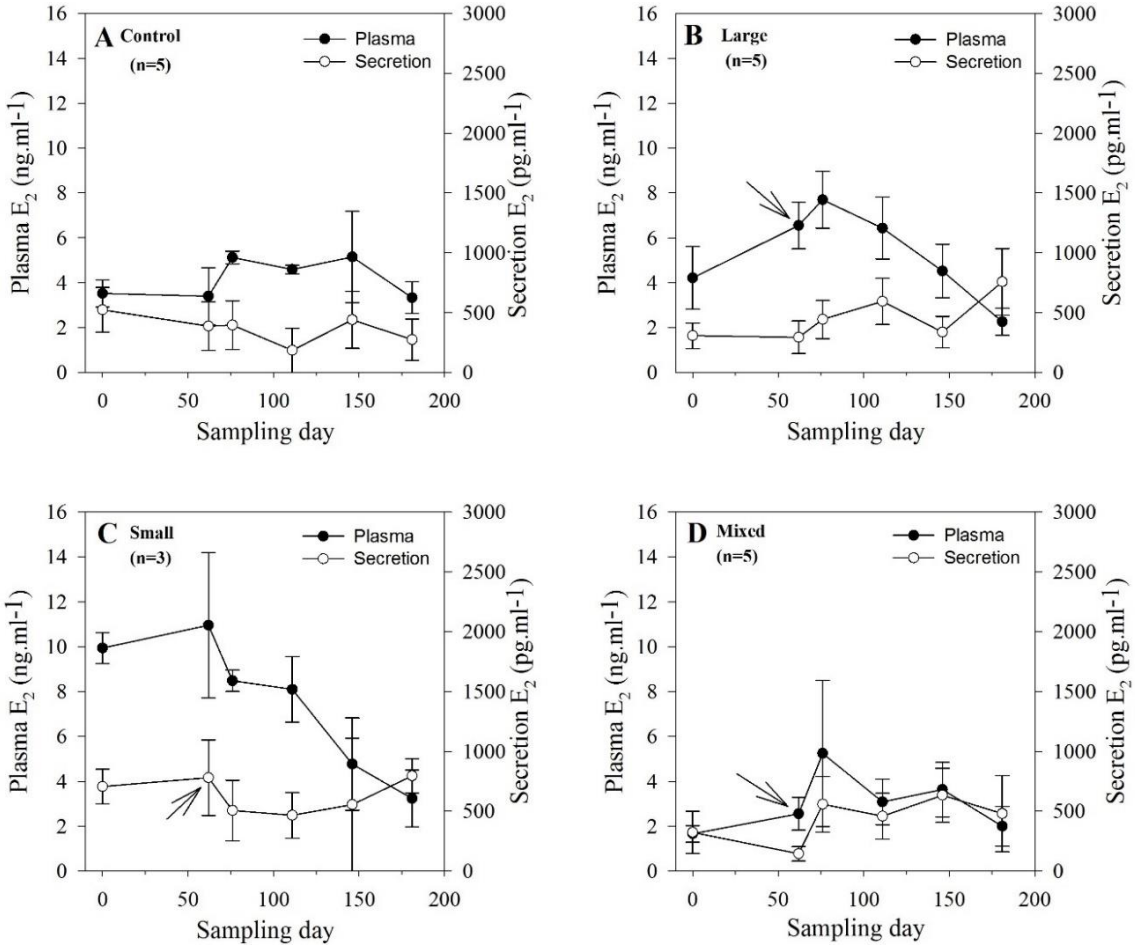

Figure 3.9. Levels of 17β-oestradiol (E_2) in the plasma ($ng.ml^{-1}$) and cephalic secretion ($pg.ml^{-1}$) in females of *Arapaima gigas* from day 0 (couple pairing and stocking into ponds) to day 181 (119 days post GnRHa implantation) in the four experimental groups: A. Control couples (n = 5), B. Large implanted couples (n = 5), C. Small implanted couples (n = 3) and D. Mixed size implanted couples (n = 5). Values are presented as mean ± SE. Arrows indicate GnRHa implantation time.

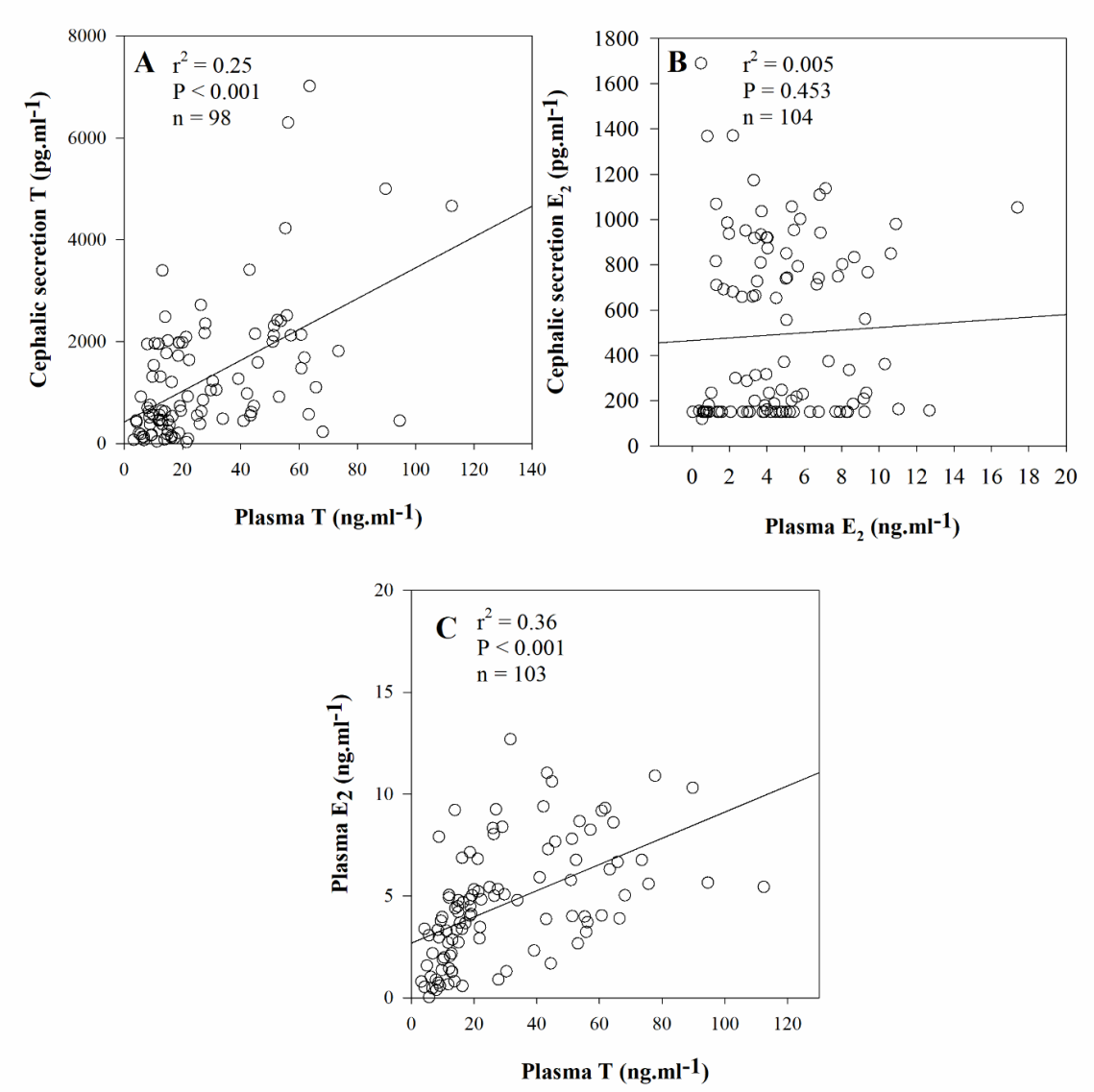

Figure 3.10. Sex steroid correlations in female *Arapaima gigas* broodstock. A. Correlation between Testosterone (T) levels in blood plasma (ng.ml^{-1}) and cephalic secretion (pg.ml^{-1}). B. Correlation between 17β-oestradiol (E_2) levels in blood plasma (ng.ml^{-1}) and cephalic secretion (pg.ml^{-1}). C. Correlation between blood plasma levels (ng.ml^{-1}) of T and E_2. Pearson Product Correlation Coefficients were calculated on log-transformed data.

3.3.2. Reproduction and sex steroids

No behavioural observations of reproductive activity (cessation in feeding, couples swimming in the same location and changes in colour pattern) were recorded in any of the experimental couples from day 0 until day 146. However, at day 146 (84 dpi), two of the large implanted couples (Fig. 3.11, couple a and c) displayed nesting behaviour and one female (couple b) from the small implanted couples spawned during sampling at day 146 (Fig. 3.11). Following GnRHa implantation, plasma 11-KT increased sharply in these males (2.3-10.6 folds, Fig. 3.11B) but not T (Fig. 3.11A). Similarly, plasma T increased sharply in females a and b (4.3-4.4 folds, Fig. 3.11C) although not E_2 (Fig. 3.11D). From day 146 onwards, T and 11-KT in males and T and E_2 in females decreased and returned to basal levels, spawning could not be confirmed and behavioural *proxy* of reproductive activity were not observed in any couple.

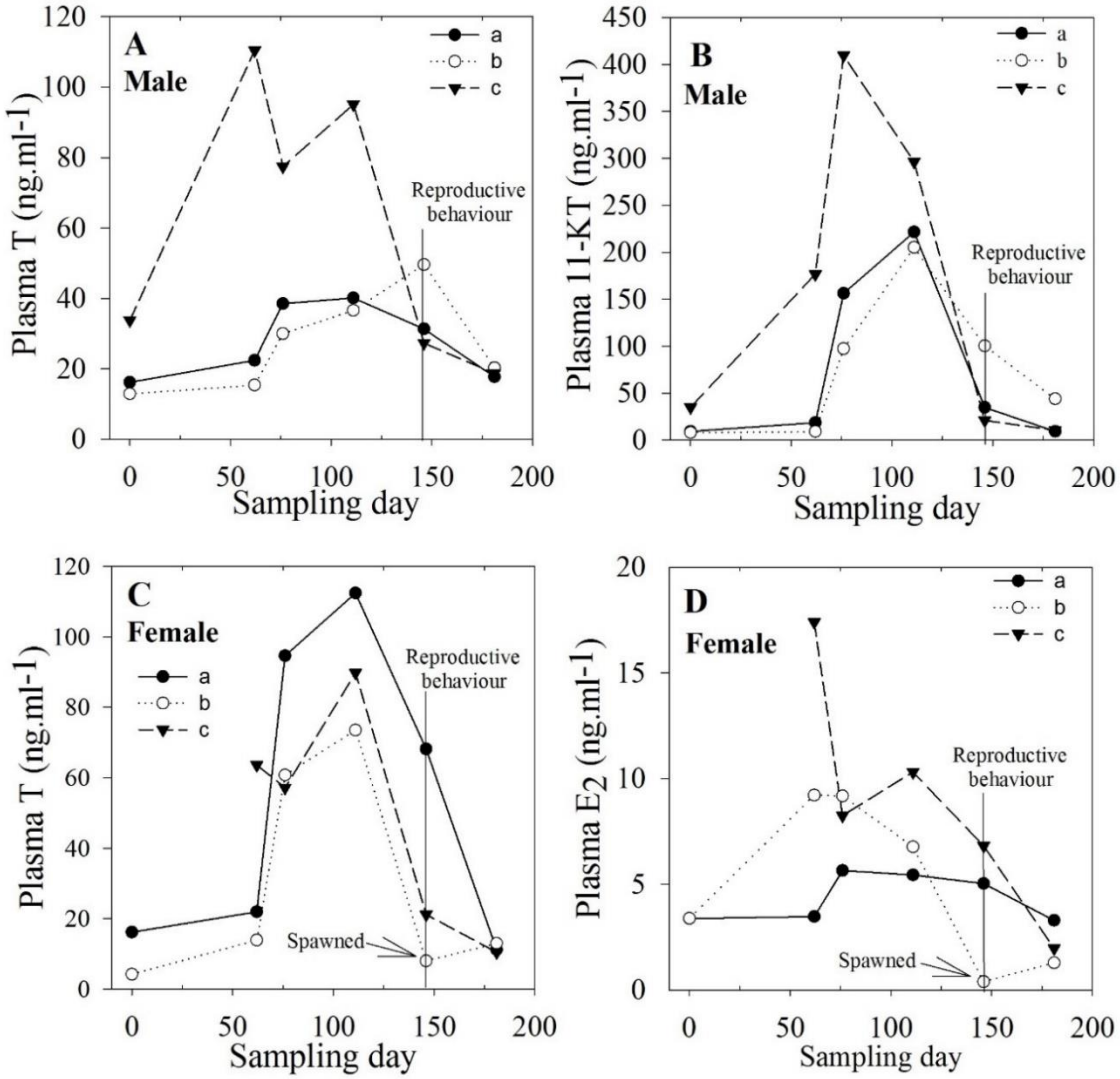

Figure 3.11. Plasma steroid levels (ng.ml^{-1}) for three couples (a, b and c) of *Arapaima gigas* showing reproductive behaviour at day 146. A and B. Levels of testosterone (T) and 11-ketotestosterone (11-KT) in males. C and D. Levels of T and 17β-oestradiol (E_2) in females.

3.4. Discussion

In recent years, a series of publications have studied the reproductive physiology of *A. gigas* reared in captive conditions or from the wild. These included the description of the adenohypophysis (Borella et al., 2009), the isolation of the pituitary gonadotrophic α-subunit hormone (Faria et al., 2013), the identification of gender through analysis of sex

steroid, blood vitellogenin levels and colour patterns (Chu-Koo et al., 2009) and the description of the gametogenesis and gonadogenesis in both sexes (Godinho et al., 2005; Núñez & Duponchelle, 2009). However, the reproductive dysfunction of *A. gigas* in captivity and the impact of hormonal induction on the control of gametogenesis and spawning has not yet been investigated. This is largely explained by the constraints to reliably identify genders, assess reproductive condition *in vivo* through ovarian biopsy and source implants that can deliver hormonal dosage for large broodstock, which are all essential for developing hormonal induction protocols (Mylonas et al., 2010). In addition, sampling of adult broodstock from the wild or in captivity for research purposes has been prohibitive due to ecological and economic factors. Trying to overcome these limitations and capitalise on the recent validation of sex identification techniques (Dugue et al., 2008; Chu-Koo et al., 2009) and availability of suitable implants (Mylonas et al., 2007), this study described the combined effects of couple size pairing and GnRHa implantation on reproductive function of captive *A. gigas*. Although spawning directly associated to the hormonal treatment could not be recorded, results showed effects of the implants on sex steroid secreted in the blood and cephalic secretion. This confirmed that the hormonal induction protocol used stimulated the BPG axis, however without clear influence of size pairing criteria. Also, this study showed higher basal steroid levels compared to previous studies on *A. gigas* (Chu-Koo et al., 2009; Monteiro et al., 2010), possibly related to the studied broodstock, for dealing with pairs separated in breeding ponds and given the very high steroid levels obtained after hormonal stimulation.

The current experiment was carried out during the rainy season when spawning has been previously reported at the experimental site (Rebouças et al., 2014), and therefore females were expected to be recruited into reproduction. This was supported by two spawnings observed prior to the start of the experiment with these couples excluded

from the study. However, no spawning or reproductive behaviour associated with nest building or mating (*i.e.* reduced feeding, alteration in swimming and air breathing patterns) were observed in the control and implanted couples post hormonal induction. The lack of clear observations of spontaneous spawning contrasts with previously published results obtained in other asynchronous spawners like Senegalese sole (*Solea senegalensis*) which released eggs after 4 dpi (Guzmán et al., 2009), Meagre (*Argyrosomus regius*) which spawned after 2-3 dpi (Mylonas et al., 2015) or the Atlantic Bluefin Tuna (*Thunnus thynnus*) which spawned after 6 dpi (Rosenfeld et al., 2012). A possible explanation for the lack of spawning following hormonal implantation in *A. gigas* could be the suppression of the normal mating and spontaneous spawning behaviour. This has been reported for other species and in such cases egg collection through stripping is often necessary (Mylonas & Zohar, 2001). If this was the case, the implants could have induced oocyte maturation and spawning in females which then released eggs without fertilization or lacking parental care provision. Alternatively, such reproductive dysfunction can either be hormonal with the involvement of other GnRH forms than the mGnRHa type tested herein, playing a neuromodulation role on the reproductive axis, or also behavioural due to a lack of pheromonal signalling and reproductive synchronization among partners. In such cases, a treatment with only mGnRHa might not be enough to induce oocyte maturation, ovulation and spawning despite the observed impact in 11-KT (males) and T (females). Indeed, this explanation finds support in the unchanged E_2 levels in all implanted females post-implantation, though since gonadogenesis could not be monitored, conclusions can not be made at this stage.

In species where GnRHa induction was reported to be successful, plasma sex steroid levels usually peak a few days after the implantation as reported in Atlantic

Bluefin Tuna (*Thunnus thynnus*) (Rosenfeld et al., 2012), Senegalese sole (*Solea senegalensis*) (Guzmán et al., 2011) and yellowtail flounder (*Pleuronectes ferrugineus*) (Larsson et al., 1997). In this study, a window of 14 days (days 62-76) was given between implantation and the following sampling point to minimise handling stress which is known to disrupt reproduction in many species and monitor potential mating and breeding behaviour in *A. gigas*. By day 76 (14 dpi), plasma T levels had raised significantly in females from large and mixed implanted couples contrasting with the lack of plasma E_2 response. The lack of increase in circulating E_2 suggests either an enzymatic deficiency in cytochrome P450 aromatase ($P450_{arom}$) activity converting precursor T into E_2 in the granulosa cells of the oocytes, or that the timing of the sampling would not have allowed to detect the E_2 peak usually slightly phase shifted from the T increase. However, while the main role of E_2 during oogenesis is to stimulate hepatocytes to produce vitellogenin that accumulates in the oocytes during vitellogenesis (Lubzens et al., 2010), it has also been suggested to play a role during final oocyte maturation (FOM) and ovulation (OV). While plasma E_2 levels during vitellogenesis can remain high during a prolonged window (from weeks to months in iteroparous spawners), increase during FOM and OV can be transient (hours to days depending on species) (Lubzens et al., 2010). Therefore, in further experiments, sampling schedules should be adapted to confirm this hypothesis and alternatively, *in vitro* experiments could be performed studying other sex steroids or hormone-like compounds involved in the later stages of oogenesis, such as the MIS or prostaglandins.

In hatcheries, problems with male reproduction are less common than for females and generally are associated with a reduced sperm volume and quality. Since possible dysfunctions in males of *A. gigas* are unknown, this study evaluated the effects of a lower GnRHa dose (49.1 ± 6.7 µg.kg^{-1}) compared to females, but intended to increase chances

of reproduction by synchronizing the couples and also evaluate GnRHa impact on male BPG axis. Given that T is the main precursor of 11-KT, levels of both androgens co-vary during most of the reproductive season (Mylonas & Zohar, 2001), with 11-KT being considered as the key hormone peaking during spermatogenesis and declining prior to the spermiation period (Mylonas & Zohar, 2001; Schulz et al., 2010). In the present experiment, a significant correlation between T and 11-KT was found, suggesting a positive impact on the BPG axis. Stripping of males after implantation was attempted without success due to the thick abdominal body wall of *A. gigas*, and it is therefore difficult to infer possible impact of the hormonal induction on spermiation or milt volumes. When compared to control males, all implanted groups showed a significant increase in 11-KT levels post GnRHa implantation (14 dpi). Though 11-KT is the main steroid regulating spermatogenesis in male teleosts, these results showed no evidence for implants inducing spermiation, courtship behaviour and synchronisation of gamete release.

During the present study, one female spawned while being sampled (from a small couple) after 84 dpi and outside the rainy season. At the same period, two other implanted pairs (large couples) were observed with nest building behaviour with the male displaying darkened external pigmentation. The link between these reproductive events and the GnRHa treatments is unclear, however spawning of *A. gigas* outside the rainy season is rare especially on the studied site. Analysis of steroid profiles for these three pairs clearly showed GnRHa contributed to stimulate the BPG axis and possibly spermatogenesis/vitellogenesis. The observation of ripe females spawning while being sampled has been reported in other samplings made in Pirarucu farms (*pers. comm.*), suggesting the involvement of stress factors in the induction of spawning in *A. gigas*. Spawning induced by stress is common among other fish species (Schreck et al., 2001)

but so far not reported for *A. gigas*. These novel observations suggests artificial fertilisation could be feasible in *A. gigas* if tools to monitor female ovary development could be developed (Torati et al., 2016). However, artificial fertilisation in *A. gigas* will require further characterisation of the unusual male gonadal anatomy especially regarding the position of the gonopore in the genital papilla, to develop non-invasive protocols for milt collection as done in the mudfish *Clarias anguillaris* (Idahor, 2014).

Sex steroids were also analysed in the cephalic secretion released from the head of *A. gigas* males and females. There is very limited available data on the biochemical nature of this fluid in the cephalic canals of the lateral line system in teleosts (Coombs et al., 2014). This cephalic fluid in *A. gigas* is known by the Amazonian indigenous people as the "Pirarucu milk", given its whitish colour especially during the parental care phase. However, the role(s) of this substance in the biology of the species is still unknown. In a recent study performed on wild specimens, steroids (*i.e.* T, E_2 and 17α-hydroxyprogesterone) were detected in the cephalic secretion (Amaral, 2009), however methodological details of this unpublished work are lacking. In the present experiment, positive correlations between plasma and cephalic secretion steroid levels were observed for 11-KT and T. This suggests the release of steroids as pheromones through the cephalic canals and the circulatory system. Interestingly, no correlation was found for E_2. Since the brain is also a site of T conversion into E_2 through brain aromatase (Forlano et al., 2001), these results suggest a possible metabolic link between the cephalic fluid and the cerebrospinal fluid in *A. gigas*. Given that the lateral line in osteoglossids is an open system, and the cephalic secretion is released externally, the results would indicate that the cephalic fluid could play an important role in pheromonal signalling. Further investigations are needed to characterise the nature and roles of this cephalic secretion in *A gigas*.

In conclusion, this study showed for the first time the impact of slow-release mGnRHa implants on the pituitary-gonad axis of *A. gigas*, eliciting significant increases in T and 11-KT levels in females and males, respectively. Couples paired with different sizes showed similar responses to GnRHa in terms of steroid levels, but impact on mating and spawning could not be assessed. Interestingly, positive correlations between plasma and cephalic secretion steroid levels suggest a link between the anterior lateral line and the circulatory systems. This is a possible new route of pheromone release in a teleost species. Further analyses are being carried out to describe other biochemical compounds in the cephalic secretion (*i.e.* secreted proteins and peptides) and their putative roles during parental care and in the reproductive biology of the species.

CHAPTER 4

4. Chapter 4

DEVELOPMENT OF TOOLS TO MONITOR FEMALE REPRODUCTIVE FUNCTION AND HORMONAL MANIPULATION USING SELECTED MATURE FEMALES

Abstract

The lack of tools for gender identification and to assess gonadal development are hindering our understanding of *Arapaima gigas* reproductive dysfunction in captivity. This chapter addresses these two issues and data are presented in three consecutive parts. Part A aimed to evaluate a non-surgical gonoductoscopy procedure to access the ovary in adults. In Part B, gonoductoscopy procedure was validated under field operational conditions in 12 adults, and used to identify gender in 20 juveniles. Cannulation assisted through the description of the genital anatomy made ovary biopsy sampling possible to describe oocyte development from primary growth to pre-ovulation, providing an oogenesis classification scheme for the species. Also, sampling of non-fertilised eggs confirmed that *A. gigas* has a single micropyle, whose morphology was described for the first time in an osteoglossid using scanning electron microscopy (SEM). Cannulation was also successfully performed without endoscopic guidance, and allowed to compare ovarian development along the reproductive season together with plasma 17β-oestradiol (E_2) profiling. Results showed females paired with males in earthponds during the spawning season underwent oocyte maturation, but eggs were either resorbed or released without male fertilisation in nests. Plasma E_2 levels remained high in females, explained by the constant presence of vitellogenic oocytes being recruited into oocyte maturation, characteristic of asynchronous species during the spawning season. In Part C, eight maturing females were selected, treated with injections of GnRHa (types 1 and 3) and paired with males in earthponds aiming to induce ovulation and spawnings. Handling fish multiple times likely disrupted reproduction: spawnings were not seen and mortalities were recorded. From baseline until 12 days post injections, the leading cohort oocyte diameter advanced significantly in 80 % of injected females. Levels of E_2 in females and 11-ketotestosterone (11-KT) in males remained unchanged post-treatment. Although

induced spawnings were not obtained, this study advanced our knowledge on the reproductive biology of *A. gigas* in captivity, with novel data on gonad anatomy and ovarian development. With the information provided herein, gonoductoscopy and cannulation are tools which can be promptly applied for gender identification and monitoring of reproductive function in wild and captive stocks.

Keywords: biopsy, endoscopy, hormonal manipulation, oogenesis, sex steroids.

CHAPTER 4 - PART A

A. Chapter 4 - Part A

ENDOSCOPY APPLICATION IN BROODSTOCK MANAGEMENT OF *ARAPAIMA GIGAS* (SCHINZ, 1822)

SHORT COMMUNICATION

Submitted: May, 2015

Accepted: August, 2015

Published in: Journal of Applied Ichthyology 32 (2016), 353-355

Torati, L.S.; Varges, A.P.S.; Galvão, J.A.S.; Mesquita, P.E.C. & Migaud, H.

Breeding and Physiology team, Institute of Aquaculture, School of Natural Sciences, University of Stirling, FK9 4LA, Stirling, Scotland, UK.

Contributions: The present manuscript was compiled in full by the author of this thesis (LT). Experiment was conceived and designed by LT and HM. Examinations were carried out by LT, AV, JG and PM. All authors approved the final version of the manuscript.

4.A.1. Introduction

The air-breathing fish *Arapaima gigas* (Schinz, 1822) is an emblematic species of the Amazon with adults reaching up to 2.5 m in total length. It has long been considered as a promising new candidate species for aquaculture, with reported growth rates of 10 kg over one year, no intramuscular bones and suitability to different production systems. However, the control of reproduction in captivity is a key bottleneck limiting the expansion of the aquaculture sector mainly due to the lack of fingerling supply, which also increases the pressure on the natural stocks.

Sexual maturity is reached after three to five years, when only the left gonad becomes functionally developed (Godinho et al., 2005). The oocyte development is asynchronous and spawning can occur several times in a same reproductive period, usually during the rainy season from November to April in the Amazon (Núñez et al., 2011). During this period, couples seek shallow flooded regions for nest building and mating (Castello, 2008b). In captivity, reproduction is stimulated by isolating couples in earth ponds, but this strategy is ineffective with only few spawns usually obtained. One of the main reasons for such unreliable reproductive control is the absence of consistent sexual dimorphisms between genders and the lack of tools to assess fish reproductive status. Broodstock pairing is therefore done without accurate information on gender and reproductive condition.

In many fish species, cannulation is routinely used for gonad biopsy and assessment. However, in some species like sturgeon, specialised endoscopic equipment has been successfully used with minimal body incisions (Hernandez-Divers et al., 2004). Interestingly, endoscopic techniques can also be used to navigate through the gonoduct (gonoductoscopy) (Divers et al., 2009). In *A. gigas*, cannulation for gonad assessment is impossible due to its atypical gonadal morphology (Núñez et al., 2011). Females have a

left gymnovarium lacking a capsule, therefore mature oocytes are released directly into the coelomic cavity, which guides them into an elongated structure called ligament ("oviduct") that conducts oocytes externally to the genital papilla (Godinho et al., 2005). The aim of this study was to test a non-invasive gonoductoscopy *in A.* gigas broodstock, taking advantage of modern endoscopes which have reduced gauge, for gender confirmation and assessment of ovary development.

4.A.2. Materials and Methods

4.A.2.1. Examined fish, anaesthesia and management

Two adult females (Fig. 4.A.1A and 4.A.1B) and two adult males were fasted for 24 hours and then examined on 23rd October 2014 in the DNOCS pisciculture station (Pentecoste-CE, Brazil). Fish were previously PIT tagged and sex evaluated *a priori* by colour pattern and blood vitellogenin detection (Dugue et al., 2008) and later by endoscopic analysis. Fish total length and weight were 1.53 ± 0.12 m and 32.8 ± 8.5 kg, respectively. Each fish was anaesthetised by spraying MS222 buffered solution onto the gills (200 $mg.L^{-1}$, Sigma Aldrich), as previously demonstrated in Honczaryk & Inoue (2009) for the air breather *A. gigas*. For safety reasons and to facilitate the procedure, each fish was contained lying on the left side on a wet soft table. Aerial breathing was stimulated at intervals of 4-6 minutes by repositioning the fish belly down. After the procedure, each fish was held in a recovery tank and water was forced through the gills to flush out anesthetic residues. The study complied with the Brazilian guidelines for the care and use of animals for scientific and educational purposes – DBCA – Concea.

4.A.2.2. Endoscopy

Analyses were done with a medical uretero-renoscope (34 cm x 2.6 mm, 6 ° of angular field, 8-13.5 Charr operating sheath, model 27001L/K, Karl Storz Endoscopy, Tuttlingen, Germany), equipped with a Telecam camera of 1 chip, a 50 watt Hi-Lux light source and a 15″ LCD monitor (200450 01-PT, Tele Pack X, Karl Storz, Tuttlingen, German). During each examination, a 0.5 L sterile saline bag (0.9 % sodium chloride) was used to flush the system allowing image acquisition and secure endoscope navigation. The saline was connected to the operating sheath with a medical intravenous line. Telescopes and sheaths were sterilized with 70 % ethanol for at least 3 min between each procedure.

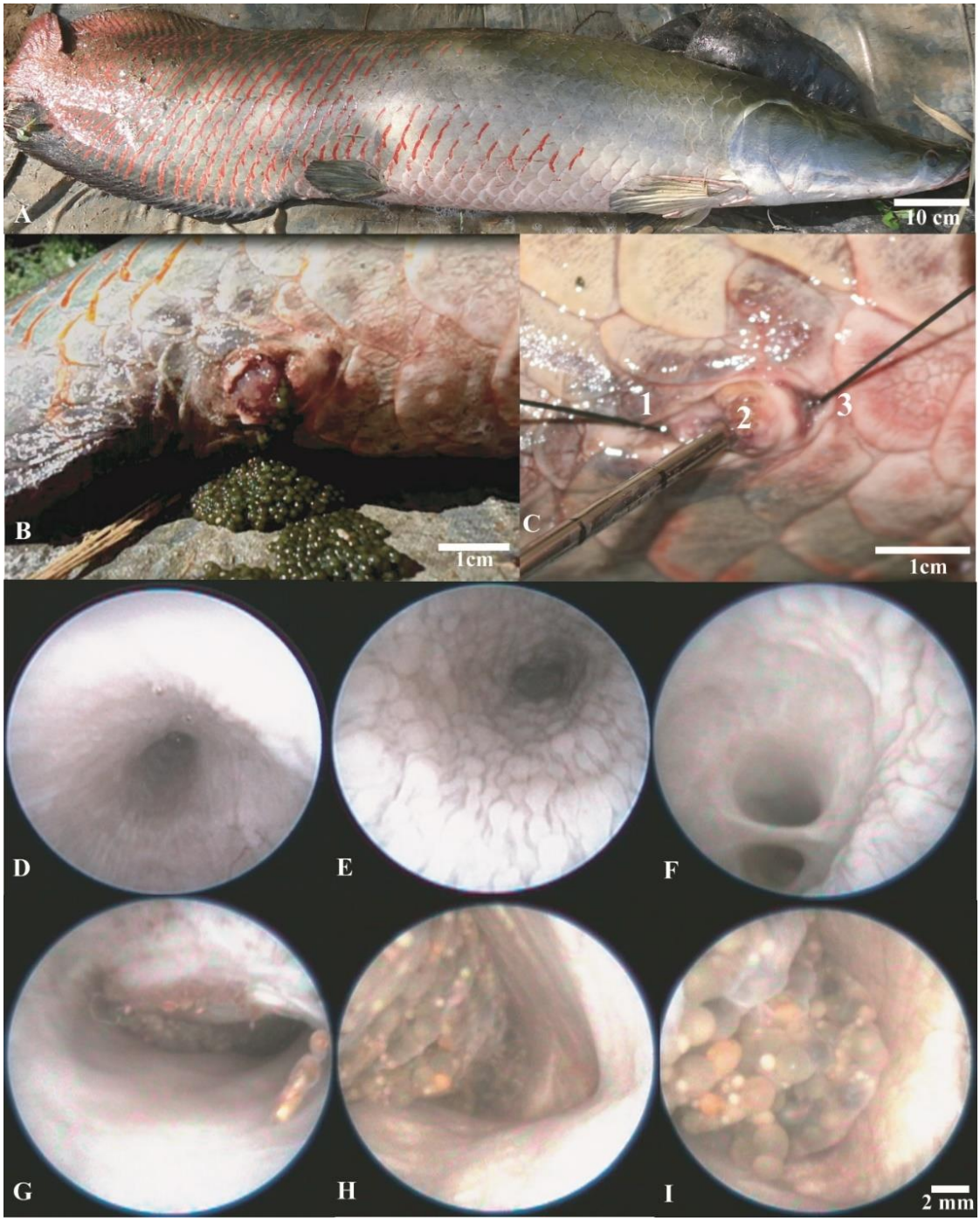

Figure 4.A.1. *Arapaima gigas*. A. Adult female. B. Genital papilla at spawning. C. Genital papilla at endoscopy: 1-Urinary opening; 2-Gonopore and 3-Anus. Endoscopic images of: D. Urinary channel, E. Urinary bladder, F. Ureter openings. G and H. Left ovary in the coelomic cavity. I. Mature ovary depicting green vitellogenic oocytes (stage III).

4.A.3. Results

Examinations were followed on the monitor screen and recorded for later analyses. As the gonopore is hardly distinguishable from the urinary pore without a guided inspection, the endoscope was essential to reveal their relative position on the genital papilla (Fig. 4.A.1C), and also their different angulation for endoscope navigation. For future guidance, images of the urinary canal, bladder and ureter openings are also presented (Fig. 4.A.1D, E and F). Accessing through the gonopore, the endoscope reached the straight short ligament ("oviduct") gaining access to the coelomic cavity (Fig. 4.A.1G). Internal organs could then be observed as well as the left functional ovary (Fig. 4.A.1H and I). Ovary assessment was made within 1-3 min by a trained operator. Following a macroscopic ovary developmental scale proposed in Núñez & Duponchelle (2009), the two females could have their ovary determined at stage 3, with the lamellae containing oocytes at stages II and III (vitellogenic). The fish analysed fully recovered after the procedure with no injury observed in the internal tissues.

4.A.4. Discussion

Given the impossibility of cannulating *A. gigas* (Núñez et al., 2011), the only method successfully tested so far for maturity identification was through laparoscopy, although it was initially developed for juvenile sex identification (Carreiro et al., 2011). However, such technique is invasive and involves surgery. Due to welfare concerns, it is not applicable to broodstock. The progress made in medical and veterinary endoscopic equipment and their use as a diagnostic tool for sexing and/or assessment of gonadal development is significant for many fish species (Hurvitz et al., 2007; Swenson et al., 2007; Divers, 2010). It is particularly true in *A. gigas* where traditional methods for

reproductive assessment are not suitable. Such a technique was proved to be useful in this preliminary trial.

The preliminary endoscopic test made so far could have a direct application for farmers and breeders as it enables the identification of maturing females (e.g. late vitellogenesis/final oocyte maturation) within a population. Broodstock management of the species would then be based on informed decisions on couple pairing and it could increase mating and spawning success in farms. This adapted technique could also improve fishery management practices of the species across the Amazon. When dealing with males, the unknown position and gauge of the spermiduct prevented assessing the testis, and further analyses are still required. At this stage, despite the many applications this technical adaptation can have, it will be essential that a more thorough study, under practical operational conditions, be undertaken to evaluate the reliability of the technique before it comes to use by the industry.

CHAPTER 4: PART B

B. Chapter 4 - Part B

GENDER IDENTIFICATION AND MONITORING OF FEMALE REPRODUCTIVE FUNCTION USING ENDOSCOPY AND CANNULATION TECHNIQUES IN THE GIANT *ARAPAIMA GIGAS*

4.B.1. Introduction

The Pirarucu *Arapaima gigas* (Schinz, 1822) is a dioic and iteroparous species without evident sexual dimorphisms (Chu-Koo et al., 2009). Like salmonid and anguillid species, the ovary of *A. gigas* lack an external capsule (gymnovarium) meaning mature oocytes are directly released into the coelomatic cavity before reaching the gonopore at spawning (Colombo et al., 1984; Godinho et al., 2005; Grier et al., 2009). However, the gonopore morphology and position have not been described yet and procedures to obtain ovarian biopsies by cannulation are not available. This is a major constraint to the study of reproductive function in the species for both wild and captive stocks (Núñez et al., 2011; Torati et al., 2016), as assessment of reproductive condition can only rely on sacrificing animals to sample gametes which contrasts with conservation efforts (Godinho et al., 2005; Núñez & Duponchelle, 2009; Faria et al., 2013). Consequently, data on gonadal development and spawning rhythms are scarce which limits greatly advances made to optimise reproduction in captivity.

A. gigas breeding and spontaneous spawning has been reported in captivity in many different farms in Brazil, however it remains uncontrolled and unpredictable. Empirical evidence suggested that nutrition and rainfall would act as proximate factors for the control of final oocyte maturation and spawning (Migaud et al., 2010; Núñez et al., 2011). However, little is known about the reproductive dysfunction of *A. gigas* in captivity. Current practices for breeding the species consist in isolating pairs randomly into earthen ponds without any scientific and validated assessment of gender and reproductive condition (Núñez et al., 2011). Available methods to identify gender in *A. gigas* include colour patterns which display large intra and inter population variability and can therefore be considered as unreliable (Chu-Koo et al., 2009). Other methods to identify gender were investigated using genetic karyotyping and bulked segregant

analyses (Rosa et al., 2009; Almeida et al., 2013), however these were not successful (Almeida et al., 2013). Hormonal analyses demonstrated gender could be identified using the ratio between plasma 17β-oestradiol (E_2) and 11-ketotestosterone (11-KT) in maturing but also immature individuals (Chu-Koo et al., 2009). In addition, a vitellogenin (VTG) enzyme immune assay (EIA) was developed and can be used to identify adult females (Dugue et al., 2008; Chu-Koo et al., 2009). Alternatively, surgical laparoscopy was used to visualise gonads in juveniles (Carreiro et al., 2011), however the method is laborious, not fully reliable, requires availability of equipment and expertise and is very invasive, with associated welfare concerns.

Recently, a non-surgical endoscopy procedure showed promising results and allowed images to be taken of adult *A. gigas* females ovaries (Torati et al., 2016). With this technique, it became possible to stage ovary development. However, this preliminary proof of concept study only analysed a limited number of adult *A. gigas* with endoscopy (Torati et al., 2016), and the method requires further validation on farms including the screening of juveniles. If such a technique was proven feasible, it could then provide a gender identification tool for broodstock pairing during captive reproduction in earth ponds. Importantly, endoscopic examinations of urogenital papilla would also provide data on gonadal stage of development through cannulation and biopsy collection followed by wet or histological analysis of oocytes to assess reproductive condition of stocks (Grier et al., 2009; Mylonas et al., 2010).

The aims of this study were: 1) to test endoscopy as a tool for gender identification and ovary assessment in juveniles of *A. gigas*, 2) validate the endoscopy procedure (described in Chapter 4 – Part A) in adults of *A. gigas* under field operational conditions, 3) describe the anatomy of the female gonopore to enable cannulation and biopsy samples of oocytes, 4) describe oocyte development in *A. gigas* including details of micropyle

number and morphology, and 5) monitor gonadal development in a captive broodstock population during the rainy season.

4.B.2. Materials and Methods

4.B.2.1. Studied fish, endoscopy and sampling procedures

From 20th to 21st Feb 2016, 20 juveniles of *A. gigas* measuring 106.8 ± 4.6 cm in total length (TL) and weighting 10.2 ± 1.1 kg (body weight, BW) were fasted for 24 hours, sacrificed with a blow into the head and examined with endoscopy for gender determination and morphological analyses of the urogenital papillae. These juveniles were 19 months of age, and were held in the facilities of Embrapa Fisheries and Aquaculture, Palmas (TO), Brazil.

Adult broodstock females were studied in a farm located in Taipas-TO (12°09'40.79" S, 46°51'34.00" W), North Brazil. Broodstock were known to be over 6 years of age, and reproduction had never been recorded in the study site. Twelve adult females (n=12, TL: 166.6 ± 7.2 cm and BW: 43.0 ± 5.6 kg), and twelve adult males (n=12, TL: 159.6 ± 9.8 cm, BW: 37.5 ± 6.0 kg) were monitored. Prior to each sampling, fish were fasted for 24 hours, netted from earthen ponds, and kept contained outside on a soft wet mat for approximately 5-10 minutes. Anaesthetics were not applied during sampling since it has been shown to compromise welfare and result in mortalities in *A. gigas* due to its air breathing behaviour (Farrel & Randall, 1978). Fish breathing behaviour was closely monitored (breathing at regular intervals of 4-6 minutes). After each procedure, fish were returned into the ponds and monitored until they return to normal breathing behaviour.

Endoscopy procedures in juveniles and adult fish were performed according to Torati et al. (2016) (also described in Chapter 4-Part A) with the exception of the

equipment used in the current study which was a cistouretroscope (13 cm x 2.6-3.6 mm, 6 ° of angular field, 8-13.5 Charr operating sheath, model 27030KA, Karl Storz Endoscopy, Tuttlingen, Germany). The main difference with the uretero-renoscope previously used in chapter 4 Part A is that it is shorter and had better manoeuvrability. A hydrophilic guidewire (B|Braum, Barcarena-Portugal) was used to inspect the urogenital papilla region and search for the gonopore position (Suppl. video I). The endoscope was equipped with a Telecam camera of 1 chip, a 50 watt Hi-Lux light source and a 15″ LCD monitor (200450 01-PT, Tele Pack X, Karl Storz, Tuttlingen, German) and all examinations were recorded for later analyses. Ovary visualisation was used to identify juvenile females, and ovary colour pattern was observed in adults to estimate overall developmental condition. In sacrificed juveniles, a 4-mm-thick slice was dissected from the gonad median region for histology. In adults, ovary biopsies were taken by inserting a flexible silicon tubing catheter (3 mm internal diameter) into the gonopore and applying a gentle suction using a 25-ml syringe. Collected gonadal samples were fixed in Bouin's solution (Sigma Aldrich, Saint Louis, MO, USA) for 24 hours, washed in distilled water and stored in 70% ethanol until histological analysis.

Broodstock were sampled three times during this study. Initially on 7th Jan 2016, fish were tagged with a passive integrated transponder (PIT; AnimallTAG®, São Carlos, Brazil) inserted in the dorsal muscle to allow individual identification in subsequent samplings. Blood (4 ml maximum) was collected from the caudal vein of males and females using heparinised syringes (BD Precisionglide, New Jersey, USA) flushed with 560 IU.ml^{-1} heparin ammonium salt solution (Sigma Aldrich, Saint Louis, MO, USA). Blood plasma was extracted after centrifugation at 1200 g for 15 minutes and stored at -20 °C until steroid analysis. Gender was determined with a VTG enzyme immune assay (EIA) kit (Acobiom, Montpellier, France) developed specifically for *A. gigas* (Dugue et

al., 2008). On 8th Jan 2016, females were randomly paired with males into 250 m^2 breeding earthen ponds. By 16th March 2016, the 12 adult females were examined with endoscopy under practical operational field conditions (Fig. 4.B.1A-B), and cannulated to obtain biopsy samples (Fig. 4.B.1C). Adult males were not examined under field conditions given prior results obtained for juvenile males and known impossibility to assess spermatic duct with endoscopy. On 17th May 2016, the cannulation biopsy procedure was repeated on the same females without endoscopy guidance. No mortalities were recorded following the procedure during the 30 days study. This study complied with the Brazilian guidelines for care and use of animals for scientific and educational purposes (CEUA/CNPASA n°09).

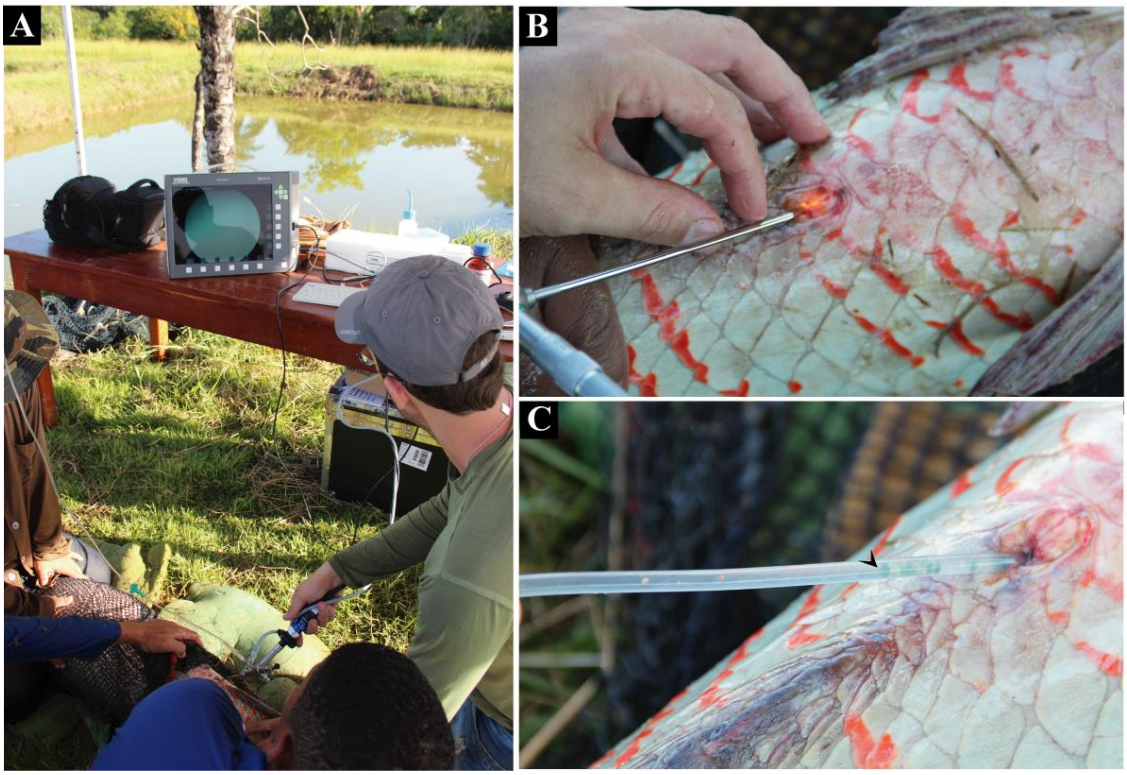

Figure 4.B.1. A. Details of field endoscopy procedure in adult females of *Arapaima gigas.* B. Angulation for cistouretroscope access into gonopore, to reach coelomatic cavity and observe ovary. C. Cannula inserted through gonopore suctioning greenish oocytes (arrow).

4.B.2.2. Gonad histology

Prior to histological analysis, all adult oocytes (n = 2101) were photographed on a scaled dish plate and individually measured for two perpendicular diameters (OD) using ImageJ v. 1.49 software (Schneider et al., 2012). The mean OD was calculated and used to estimated oocyte volume using the formula $V=1/6\pi\ OD^3$ according to Jones & Simons (1983). Measurements made before histological processing aimed to avoid effects of shrinkage and membrane collapsing which occur after oocyte dehydration in histological protocols (Kennedy et al., 2011). Oocytes analysed for histology ranged in diameter from 354.5 to 2466.0 µm (n = 363 oocytes). Samples were gradually dehydrated in ethanol series (80-90-100%), then infiltrated in glycol methacrylate resin and blocked following the manufacturer's protocol (Leica Microsystems™, Wetzlar, Germany). Blocks were sectioned at 3-5 µm thick using a rotary microtome (Leica RM 2235, Heidelberg, Germany). Alternatively, oocytes were also infiltered in paraffin, and sectioned in a Shandon Finesse microtome (Thermo Fisher Scientific, Waltham, MA, USA) aiming to improve issues faced with oocyte crumbling during the sectioning step. Slides were stained with haematoxylin – eosin (HE) (Fischer et al., 2008) for optical microscopy.

Aiming to provide a detailed characterisation of oocyte development for *A. gigas*, we adopted the staging terminology from Grier et al. (2009), in which stages are indicated by uppercase letters (*i.e.* PG=primary growth), and steps within a stage by lowercase letters (*i.e.* PGl = late step in primary growth stage). Gonadogenesis stages followed classification of Núñez & Duponchelle (2009). Gonad sections were analysed using an Olympus BX51 microscope equipped with a Zeiss AxioCam MRc camera system. To describe female reproductive condition, leading cohort (LC) oocyte diameter was defined as the mean of the 10 largest oocytes. For juvenile females, LC was measured from pictures taken after histological sections using the 10 largest oocytes *per* juvenile female.

4.B.2.3. Micropyle analysis with scanning electron microscopy (SEM)

Given egg release during sampling is not common in *A. gigas*, non-fertilized eggs could only be kept frozen in liquid N_2 instead of being fixed with appropriate fixatives for SEM analysis. Aiming to prevent egg damage at defrosting given lack of cryoprotectants, eggs were thawed at room temperature while immersed in a 0.5 M sucrose solution prepared with NaCl PBS buffer (pH 7.2), and then immediately fixed in Karnovsky solution (Karnovsky, 1965) for five hours. Fixed eggs were washed three times for 10 min in 0.1 M phosphate buffer. A second fixation (2 hours) was made in 1% osmium tetroxide followed by a quick (2 sec) immersion in distilled water. A gradual dehydration in ethanol was made, increasing 10% grade until 100% (15 min each grade). Eggs were dried in a Critical Point K850 (Quorum Technologies, Lewes, UK) following equipment protocol, then carefully mounted onto stubs, coated with gold, and observed in a Zeiss DSM 940A SEM. Egg surface was visualised in 20 eggs. Given the immovable position of eggs on stubs, a single micropyle was visualised in 5 eggs, and measurements of micropylar cone diameter undertaken in 3 of them.

4.B.2.4. Steroid analysis

Plasma samples were thawed at room temperature, and extraction of sex steroids made mixing 50 µl of plasma with 1 ml of ethyl acetate using 1 ml polypropylene assay tubes. The mixture was vigorously vortexed and spun for 1 hour at room temperature (16°C) using a rotary mixer, then centrifuged at 430 g for 10 min at 4°C. Levels of 17β-oestradiol (E_2) and 11-ketotestosterone (11-KT) were quantified in the blood plasma of females and males, respectively, using an enzyme-linked immunosorbent assay (ELISA) (Cayman Chemical Inc., Michigan, USA) previously validated for *A. gigas* (Chu-Koo et al., 2009). Just prior to assay, extracts were dried in a vacuum oven at 35°C for 40 minutes. Given

sample variability, dilutions in samples ranged from 1:5 to 1:15 (v:v) for 11-KT, and for E_2 samples had to be concentrated 6:1 (v:v). The manufacturer's protocol was followed and microplates were read at 405 nm using a BioEasy microplate reader (Belo Horizonte, Brazil). Steroid concentrations in blood were calculated from the assay value (pg. $tube^{-1}$) corrected for the proportion of extract used in the assay and the volume of blood used for extraction. Intra-assay coefficient of variation was 6.5 % for 11-KT (1 assay) and 1.7 % for E_2 (1 assay).

4.B.2.5. Statistics

Statistical analyses were conducted in Minitab (version 17.3.1, Minitab, PA, USA). One-way repeated measures ANOVA was used to compare steroid levels ($ng.ml^{-1}$) during the study period. Paired t-tests were used to compare LC (mm) for each female between biopsy samplings points (16th March and 17th May 2017). The level of significance was set as $P \leq 0.05$, and values are presented as mean ± SE.

4.B.3. Results

4.B.3.1. Gonopore position, endoscopy and cannulation in females

Internally, the ovary was found to be suspended in the coelomatic cavity by the mesovarium, whilst intestine and urinary ducts were located caudally (Fig. 4.B.2A-B). Externally, the intestine opened cranially, and there was no evident sexual dimorphism in the urogenital papilla, where a single external opening was evident in both genders (Fig. 4.B.2C). Endoscope insertion, approximately through 4-5 mm of the female urogenital papilla pore, and using a guiding wire inserted through the endoscope working channel, allowed identification of a membranous septum dividing the gonopore (cranial) from the urinary opening (caudal) (Fig. 4.B.2D). The septum identification was recorded

in Suppl. Video I, and its recognition was key to visualise the gonopore and reach the oviduct and coelomatic cavity where the internal organs and ovary could be visualised (Fig. 4.B.2E). In practice, the urinary duct (Fig. 4.B.2F) was more easily assessed caudally if the septum was not noticed or inspected, especially because the septum was normally found collapsed cranially enclosing the gonopore. In juvenile males, neither a septum nor the spermatic duct opening could be observed after endoscopic analyses.

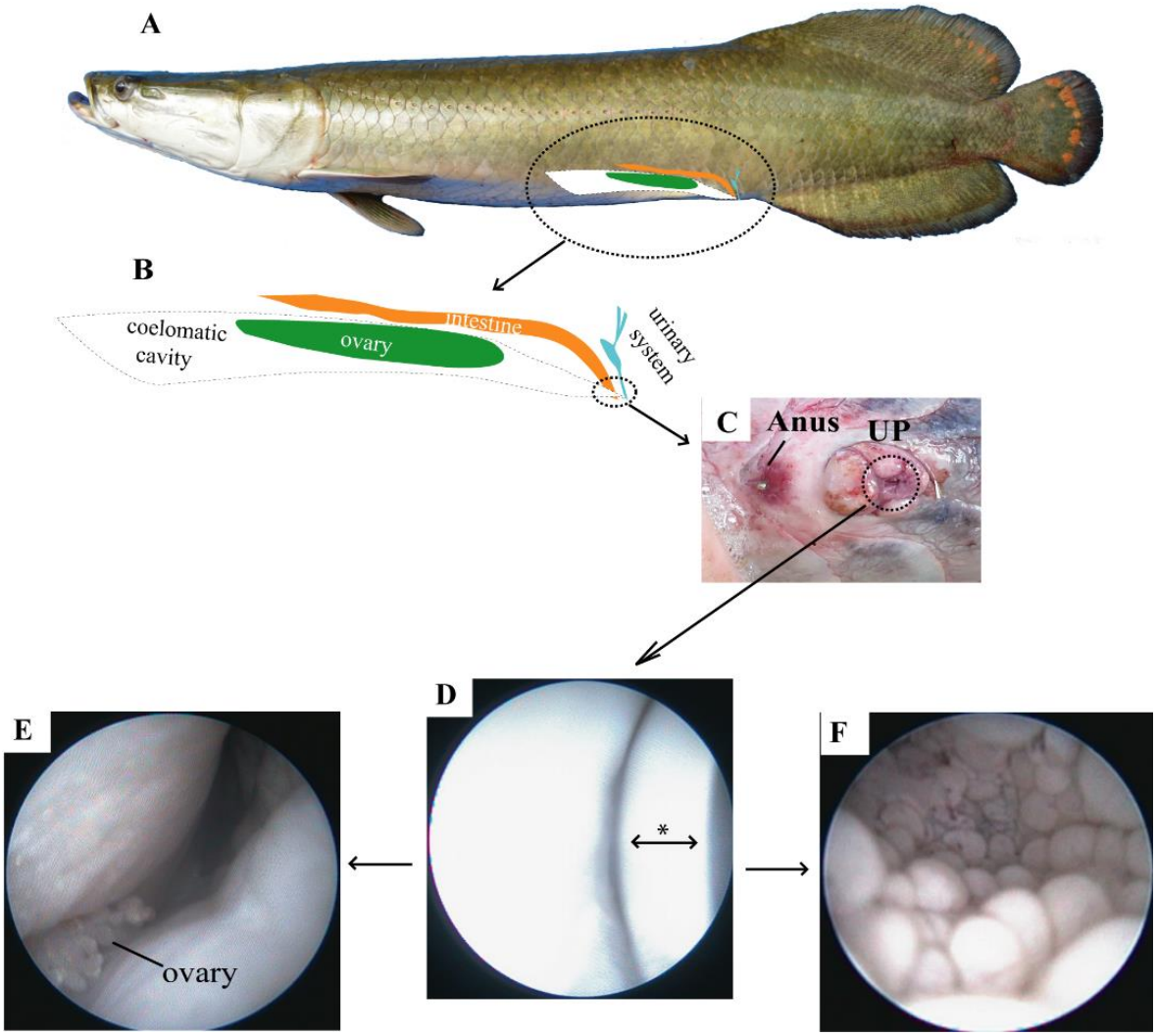

Figure 4.B.2. Illustration of urogenital anatomy in females of *Arapaima gigas*. A. B. Lateral view of juvenile illustrating coelomatic cavity (white), left ovary (green), intestine (orange) and urinary system (blue). C. External morphology depicting relative position of anus and urogenital papilla (UP). D. Endoscopic image of the septum inside the UP,

separating cranial access to coelomic cavity where the ovary is found (E) from caudal access to urinary bladder (F). * septum.

Out of 20 juveniles examined with endoscopy, 11 had the gonopore assessed and left ovary observed after approximately 2.9 ± 2.7 minutes of examination. After gonad dissection and histological analyses, these fish were confirmed to be females. Among the remaining 9 fish whose gonopore was assessed, 7 were males and 2 were females which couldn't be identified after examinations (Table 4.B.1; Suppl. Video I). With this method, male identification remained indirect, and dependent on operator ability to locate the septum to assess gonopore, and also on fish morphological variability. After histological analyses, the leading cohort (LC) in the ovary of juvenile females was classified at the primary growth perinuclear step (PGpn), characteristic of immature fish (Stage I, Table 4.B.1). Externally, their ovaries had a pale colour when observed during endoscopy and its foliaceous shape could also be noted (Fig. 4.B.2E; Suppl. Video I).

Table 4.B.1. Endoscopy procedure in juveniles of *Arapaima gigas*. Fish body weight (BW; kg), total length (TL; cm), time elapsed to observe ovary (minutes), gender confirmation after histology, diameter and stage of leading cohort (LC, µm) and ovary stage *according to Núñez & Duponchelle (2009). PGpn – Primary growth perinucleoli.

Juvenile n°	BW (kg)	TL (cm)	Procedure time (min)	Gender	LC (µm)	LC Stage	Ovary Stage *
1	12.1	110.5	4.0	Female	220.9	PGpn	I
2	9.5	107.0	2.6	Female	161.9	PGpn	I
3	10.2	107.2	1.6	Female	159.1	PGpn	I
4	12.2	100.0	0.3	Female	208.9	PGpn	I
5	10.1	107.0	—	Male	—	—	—
6	8.9	106.0	—	Male	—	—	—
7	10.2	107.5	—	Male	—	—	—
8	11.7	112.5	1:6	Female	215.4	PGpn	I
9	9.8	106.0	1.1	Female	191.2	PGpn	I
10	11.6	115.0	7.1	Female	167.4	PGpn	I
11	10.6	109.5	—	Male	—	—	—
12	9.9	107.7	—	Male	—	—	—
13	9.8	106.0	No access	Female	209.2	PGpn	I
14	10.9	112.0	—	Male	—	—	—
15	8.9	101.5	9.0	Female	206.1	PGpn	I
16	9.6	103.0	—	Male	—	—	—
17	11.3	110.0	0.7	Female	220.1	PGpn	I
18	9.6	98.5	2.3	Female	246.5	PGpn	I
19	10.4	110.5	1.9	Female	229.5	PGpn	I
20	8.0	98.5	No access	Female	170.3	PGpn	I

Field endoscopy in adult females was successfully done without complications (Fig. 4.B.1A-B). The coelomatic cavity of 12 adult females was assessed and the ovary observed after approximately 1.24 ± 1.53 minutes (Table 4.B.2; Suppl. Video I). Following endoscopy, a cannula was introduced into the gonopore and oocyte biopsies were sampled (Fig. 4.B.2C). With the morphological knowledge gained after endoscopic analyses, cannulation was repeated in the same females two months later without use of endoscopy guidance. For each biopsy, from 3 to 206 oocytes (mean of 87.5 ± 53.4) were obtained per individual female with a total of 2101 oocytes staged and measured over the two sampling times. Endoscopy and oocyte analyses combined showed greenish colour ovaries at Stage IV with LC oocytes classified either at the germinal vesicle migration stage (OMgvm, n =10) or ovulation (OV, n=1). One female had a yellowish ovary at Stage II with LC oocytes at the full-grown stage (SGfg, n=1) (Suppl. Video I, Table 4.B.2).

Table 4.B.2. Endoscopy procedure in adult broodstock females of *Arapaima gigas*. Females are numbered as in presented in Figure 4.B.7, and differently from Table 4.B.1 which correspond to examinations in juveniles. Female body weight (BW; kg), total length (TL; cm), time elapsed to observe ovary (minutes), ovary predominant colour, diameter and stage of leading cohort (LC, µm) and ovary stage *according to Núñez & Duponchelle (2009). OMgvm = Oocyte maturation germinal vesicle migration, SGfg = Secondary growth full-grown, OV = Ovulation.

Adult female n°	BW (kg)	TL (cm)	Procedure time (min)	Ovary colour	LC (µm)	LC Stage	Ovary Stage *
1	37.0	155.0	1.3	green	2391.5	OMgvm	IV
2	31.0	156.0	0.7	green	2152.5	OMgvm	IV
3	45.0	162.0	1.2	green	1740.9	OMgvm	IV
4	-	-	0.6	green	2256.4	OMgvm	IV
5	48.5	175.0	0.7	green	2399.5	OMgvm	IV
6	41.0	167.0	0.6	green	2064.5	OMgvm	IV
7	48.0	173.0	0.5	green	2232.6	OMgvm	IV
8	51.0	175.0	0.7	green	2299.1	OMgvm	IV
9	41.0	160.0	5.8	yellow	1324.0	SGfg	II
10	45.0	168.0	2.1	green	2225.1	OMgvm	IV
11	42.0	166.0	0.5	green	2258.0	OMgvm	IV
12	44.0	171.0	0.3	green	2435.4	OV	IV

4.B.3.2. Oocyte development

Primary growth stage. Multi nucleoli oocytes (**PGmn**) were identified by the presence of several nucleoli scattered within the germinal vesicle. Balbiani bodies were present in the ooplasm, and oocyte diameter ranged from 32.2 to 113.1 µm (72.0 ± 17.3 µm; Fig. 4.B.3A). PGmn oocytes were present in all juvenile females analysed, but not in cannulated adults. Perinucleoli oocytes (**PGpn)** were the leading cohort observed in all juvenile females. These oocytes were identified by the presence of several spherical nucleoli positioned close to the inner germinal vesicle membrane. Balbiani bodies were still present throughout the ooplasm. PGpn oocyte diameter ranged from 64.7 to 320.8 µm (105.2 ± 103.9 µm; Fig. 4.B.3B). The oil droplet step (**PGod**) step was initiated with the appearance of few oil droplets at the ooplasm periphery identified on the micrographs as they are cleared out during histological preparation. Also during the PGod step, other ooplasm inclusions were firstly seen. Cortical alveoli appeared centripetally and large lipoprotein yolk granules at the circumnuclear area. Oocytes at the PGod step were not observed in the juveniles analysed, and were collected through cannulation in all adult females. In PGod oocytes, the zona pellucida appeared completely formed and visible, and oocyte diameter ranged from 354.5 to 599.0 µm (473.4 ± 82.1 µm; Fig. 4.B.3C). At the cortical alveoli step (**PGca)**, the multiple layers of cortical alveoli had completely migrated towards the ooplasm periphery. Granules of lipoprotein were still present although fewer and smaller compared to the PGod step. PGca oocyte diameter ranged from 627.0 to 777.0 µm (697.5 ± 57.6 µm; Fig. 4.B.3D).

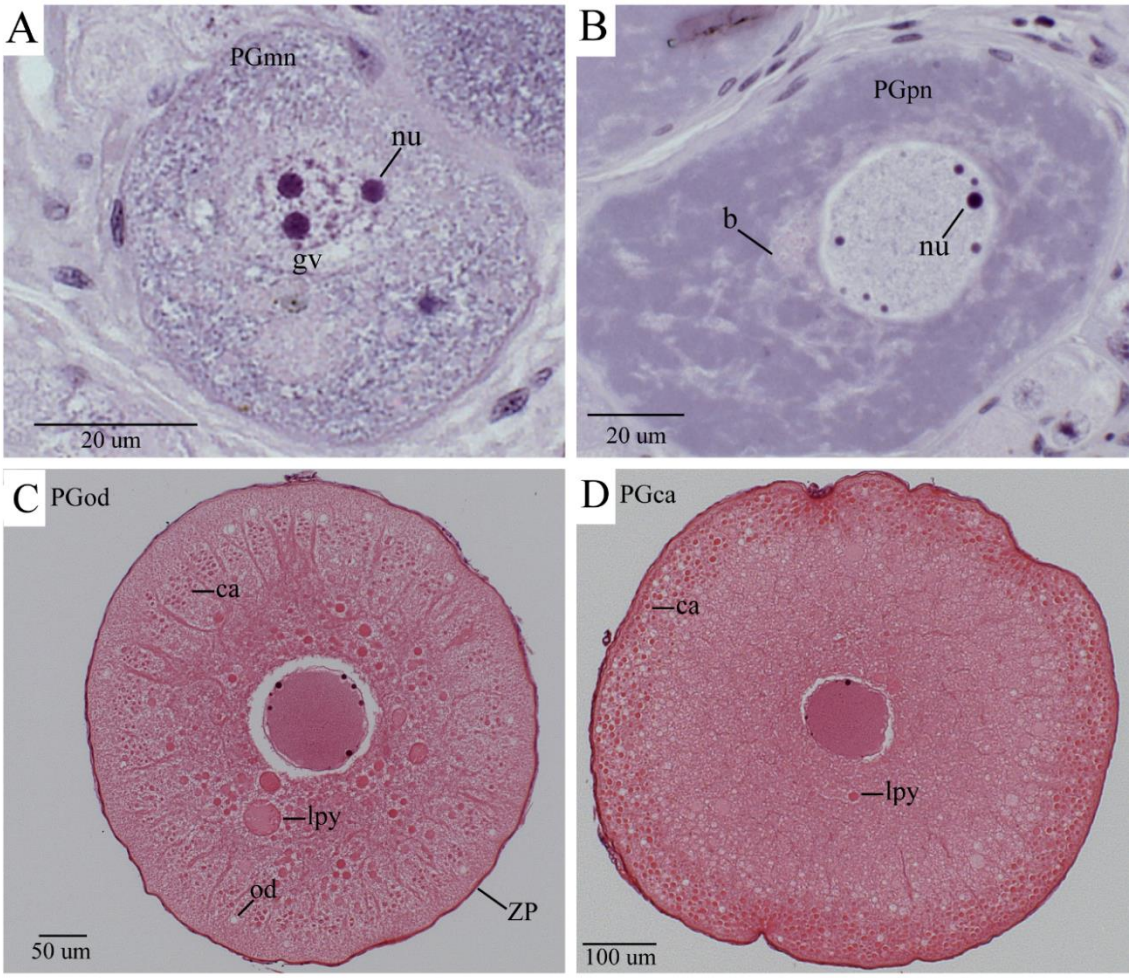

Figure 4.B.3. Light micrographs of primary growth (PG) step in oocytes of *Arapaima gigas*. A. Multiple nucleoli stage (PGmn), depicting several nucleoli inside the germinal vesicle. B. Perinucleoli (PGpn) stage, nucleoli at the internal margins of germinal vesicle and Balbiani bodies scattered throughout the ooplasm. C. Oil droplets stage (PGod), depicting early appearance of oil droplets at the ooplasm periphery, large globules of lipoprotein yolk and cortical alveoli. D. Cortical alveoli stage (PGca) depicting the cortical alveoli layer migrated towards ooplasm periphery and presence of lipoprotein yolk. b = Balbiani body, gv = germinal vesicle, nu = nucleoli, ca = cortical alveoli, od = oil droplet, lpy = lipoprotein yolk, ZP = zona pellucida.

Secondary growth stage. The follicle layer was more evident lining the zona pellucida from secondary growth onwards. Early secondary growth (**SGe**) oocytes were distinguished from previous PGca step by a marked increase in the number of small yolk globules (lipoprotein) and oil droplets throughout the ooplasm. SGe oocyte diameter ranged from 783.5 to 946.5 µm (892.6 ± 46.5 µm; Fig. 4.B.4A). In the late secondary growth (**SGl**), larger yolk granules became evident surrounding the germinal vesicle. In SGl oocytes, the multiple layers of cortical alveoli became also more compacted lining the zona pellucida, and oocyte diameter ranged from 1007.0-1111.5 µm (1066.1 ± 39.8 µm; Fig. 4.B.4B). In full-grown oocytes (**SGfg**), an extensive accumulation of oil droplets was observed and these were larger compared to SGl oocytes. Oil droplets were found between the nuclear region and the cortical alveoli layer. At SGfg, oocytes reached their maximum diameter during vitellogenesis ranging from 1139.5 to 1276.0 µm (1214.9 ± 52.9 µm; Fig. 4.B.4C-D).

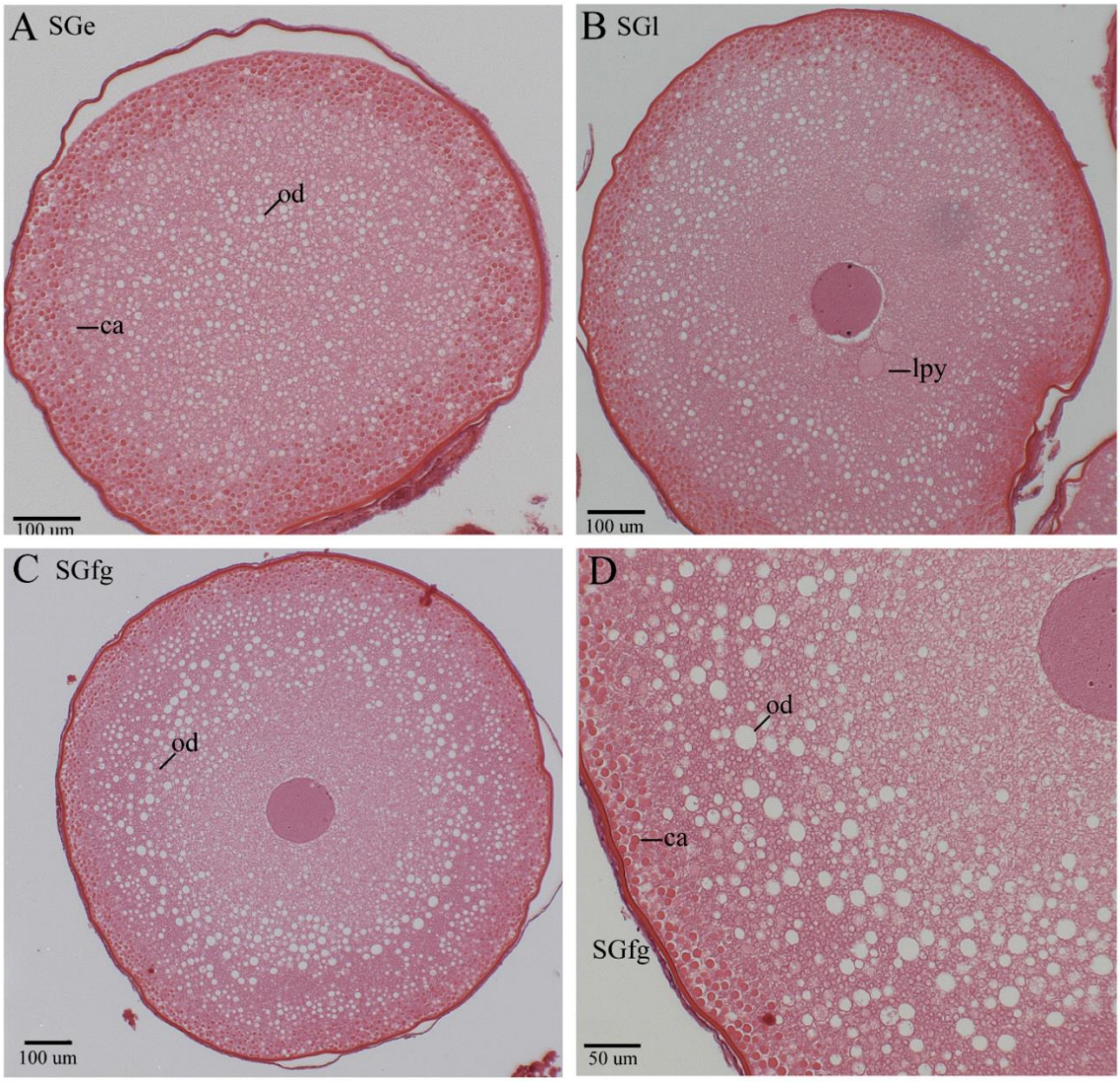

Figure 4.B.4. Light micrographs of secondary growth (SG) step in *Arapaima gigas*. A. Early secondary growth (SGe) depicting wide layer of cortical alveoli at the ooplasm periphery and increase in oil droplet number throughout the ooplasm. B. Late secondary growth (SGl) depicting appearance of lipoprotein yolk globules at the central region. C. D. Full-grown (SGfg) oocyte depicting enlarged oil droplets and thin layer of cortical alveoli. ca = cortical alveoli, od = oil droplet, y = lipoprotein yolk.

Oocyte maturation and pre-ovulation steps. The germinal vesicle started its migration towards the animal pole marking the start of the **OMgvm** stage. During OMgvm, most of the oocyte hydration occurred, at a period when oocytes became greener

when observed externally with endoscopy or after biopsy. At OMgvm, oil droplets started to coalesce becoming larger compared to the SGfg stage, although oil globules were not observed. At OMgvm, yolk became more fluid and nucleo-cytoplasmic ratio decreased. OMgvm oocyte diameter ranged from 1302.5 to 2354.0 µm (1923.3 ± 275.6 µm; Fig. 4.B.5A-D). The ovulation (**OV**) step was observed in one female only which released eggs during sampling. Ovulated oocytes measured from 2394.0 to 2466.0 µm (2427.0 ± 21.4 µm) in diameter. Since ovulated eggs are detached from the follicular layer, only the zona pellucida could be observed at this step. In OV oocytes, the germinal vesicle has reached the animal pole and was found broken down and only a lightly stained region could be observed (Fig. 4.B.6A). SEM analyses on unfertilized eggs revealed the presence of a single micropyle. The micropyle in *A. gigas* has a series of radially arranged ridges leading into the micropylar canal, whose diameter measured 11.3 ± 0.02 µm (Fig. 4.B.6B-C).

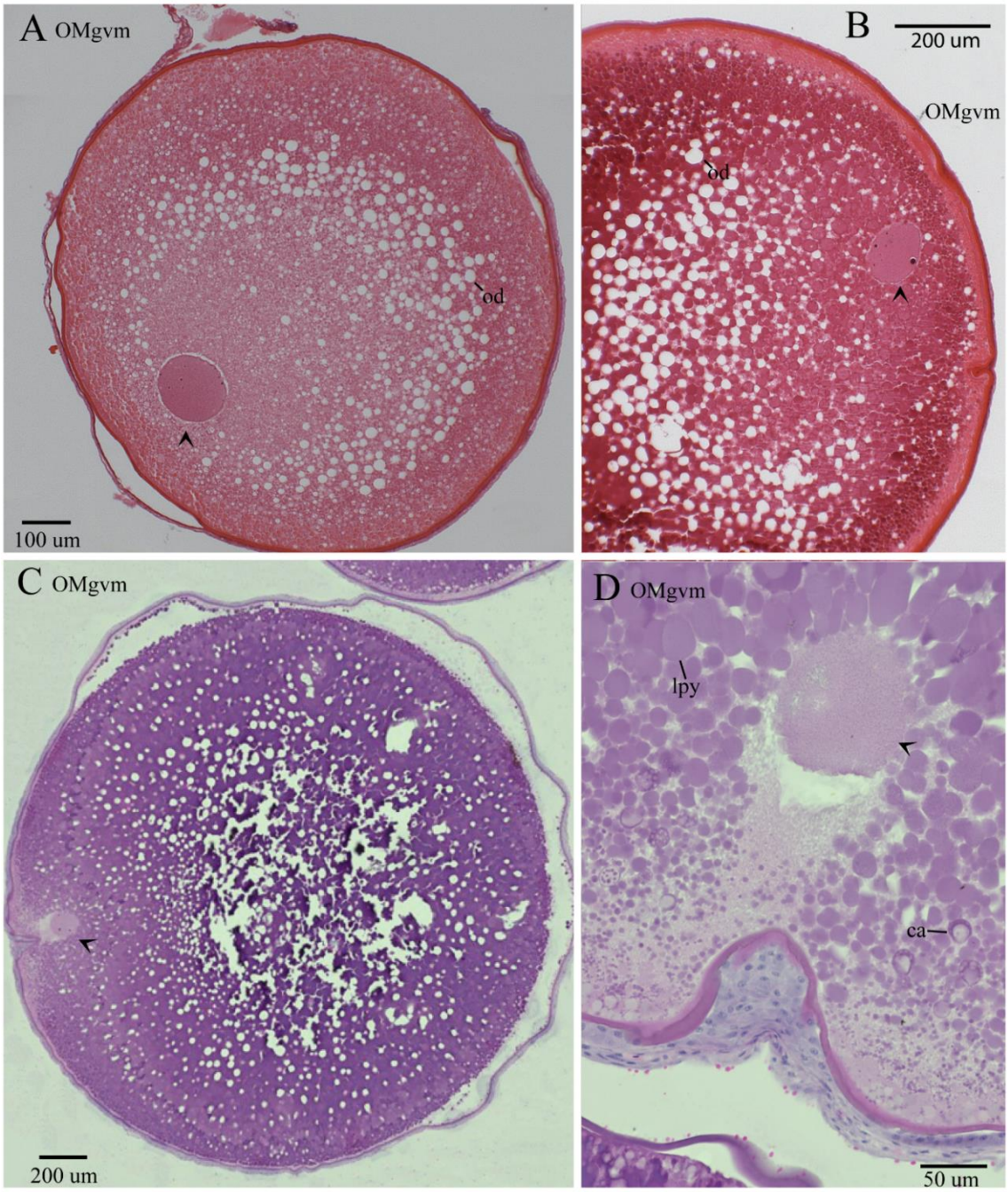

Figure 4.B.5. Light micrographs of oocyte maturation (OM) step in *Arapaima gigas* at different hydration moments. A. Early migration of germinal vesicle towards the animal pole, depicting oil droplets coalescing. B. Germinal vesicle closer to the oocyte periphery and with a reduced nucleo cytoplasmic ratio. C. D. Germinal vesicle close to micropyle area at the maximum hydration volume observed prior to germinal vesicle breakdown.

Arrow = germinal vesicle, od = oil droplet, lpy = lipoprotein yolk, ca = cortical alveoli, gvm = germinal vesicle migration.

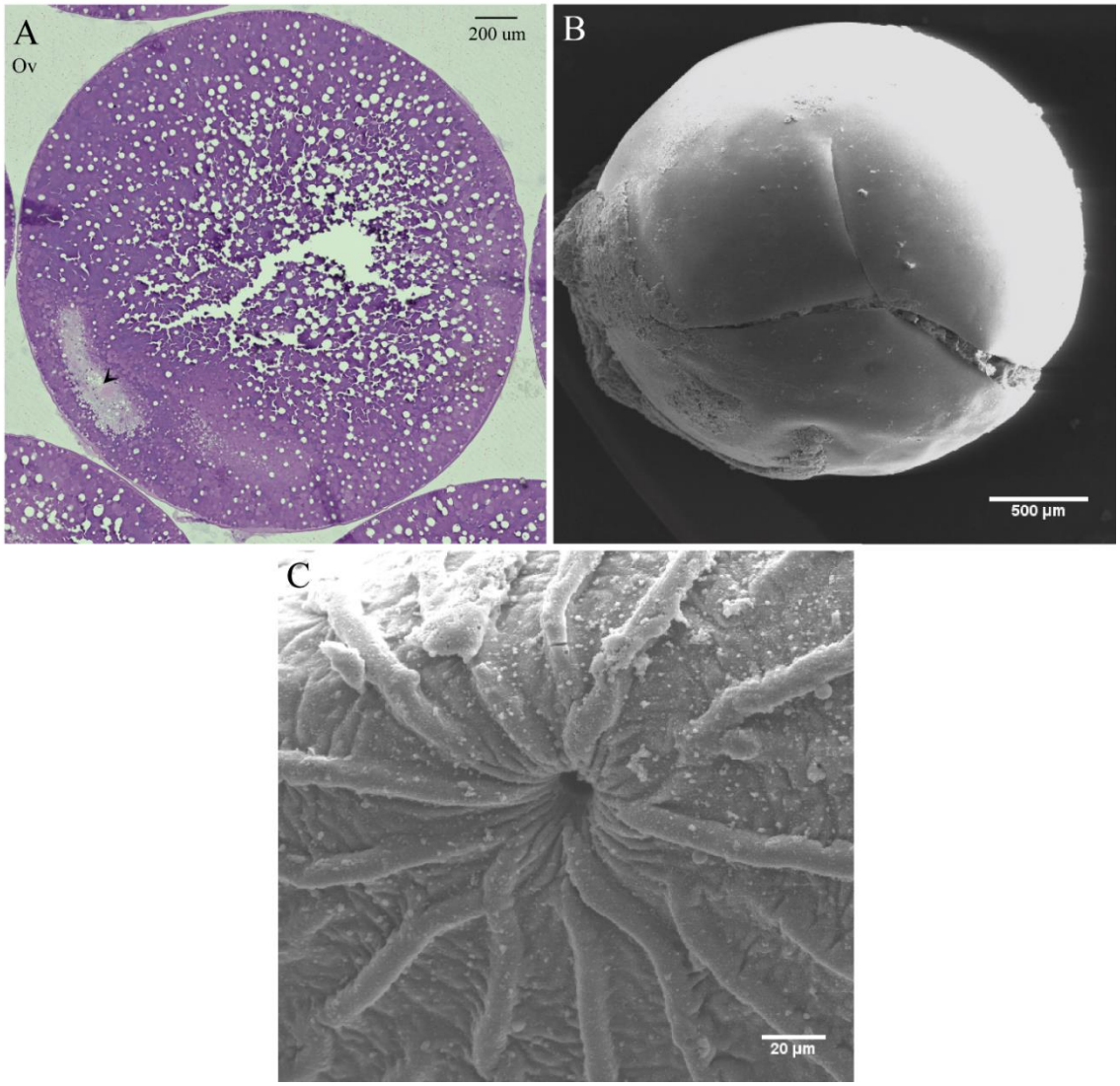

Figure 4.B.6. Non-fertilized eggs of *Arapaima gigas* at the post ovulatory step (OV). A. Light micrograph depicting region where germinal vesicle broke down (arrow). B. Scanning electron microscopy (SEM) depicting egg surface and (C) details of concentric ridges of micropyle external morphology.

4.B.3.3. Monitoring of the reproductive function of a captive broodstock

When the 12 couples were first paired in earthponds on 7th January 2016, the cannulation technique had not been developed yet. These females were first sampled on 16th March 2016, when 11 of them had their LC oocytes at OMgvm or OV (Table 4.B.2, Fig. 4.B.7). In comparison with the following sampling undertaken on 17th May 2016, 9 out of these 11 females displayed reduced diameters of their LC oocyte ($P<0.01$; Fig. 4.B.7), revealing ovarian development regressed. The ovary of only a single female (♀9) showed advanced oocyte maturation during the monitoring period ($P<0.01$; Fig. 4.B.7).

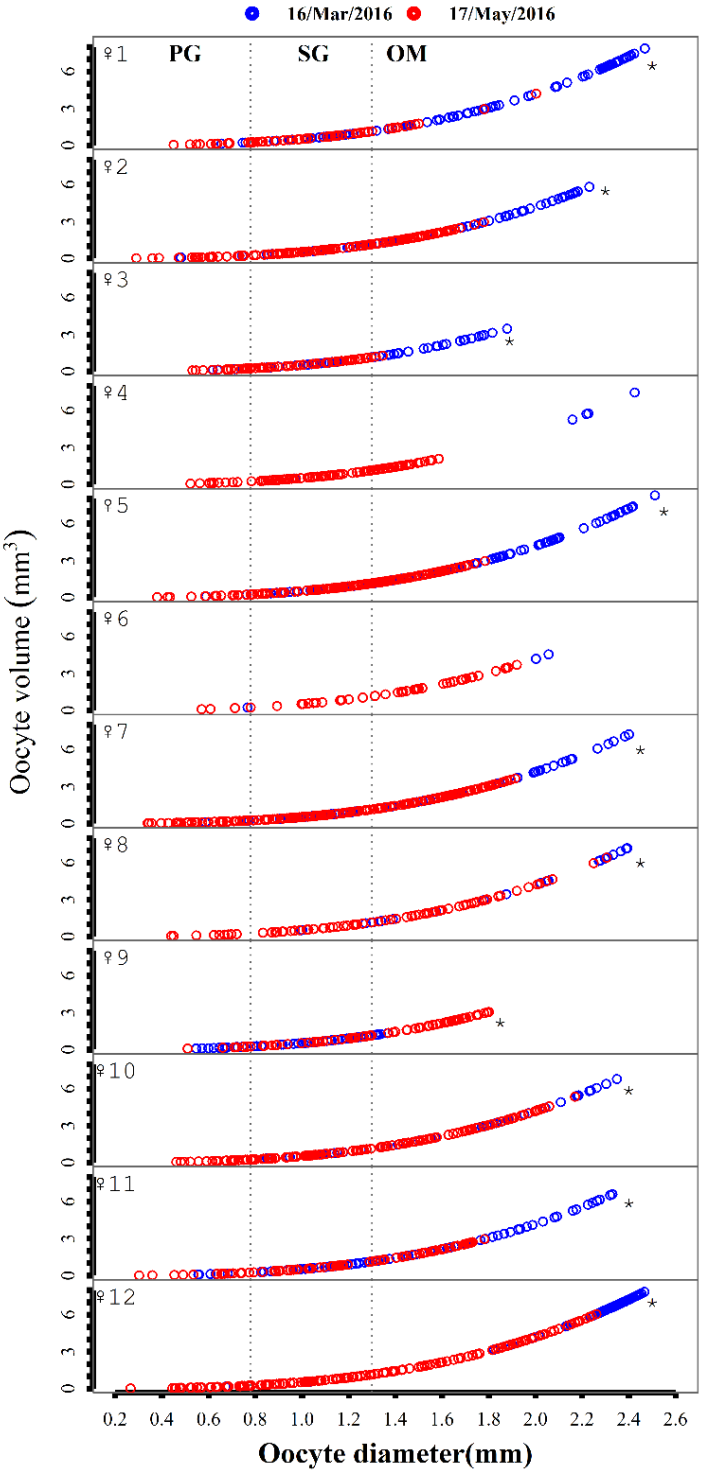

Figure 4.B.7. Oocyte diameter (mm) *vs* volume (mm^3) for 12 adult *Arapaima gigas* females sampled on 16th Mar (blue) and 17th May (red) 2016. Dashed line indicates stage of oocyte development for a given oocyte diameter range (PG = primary growth; SG = secondary growth and OM = oocyte maturation). * indicate significant difference in the leading cohort (LC) between different sampling dates.

Levels of E_2 remained steady throughout the rainy season (Fig. 4.B.8A, P>0.05), ranging from 0.6 to 16.0 ng.ml^{-1} (5.1 ± 4.5 ng.ml^{-1}). In males, levels of 11-KT also remained unchanged with a decreasing trend towards the end of the rainy season (Fig. 4.B.8B, P>0.05). It ranged from 4.5 to 59.0 ng.ml^{-1} (27.0 ± 17.2 ng.ml^{-1}).

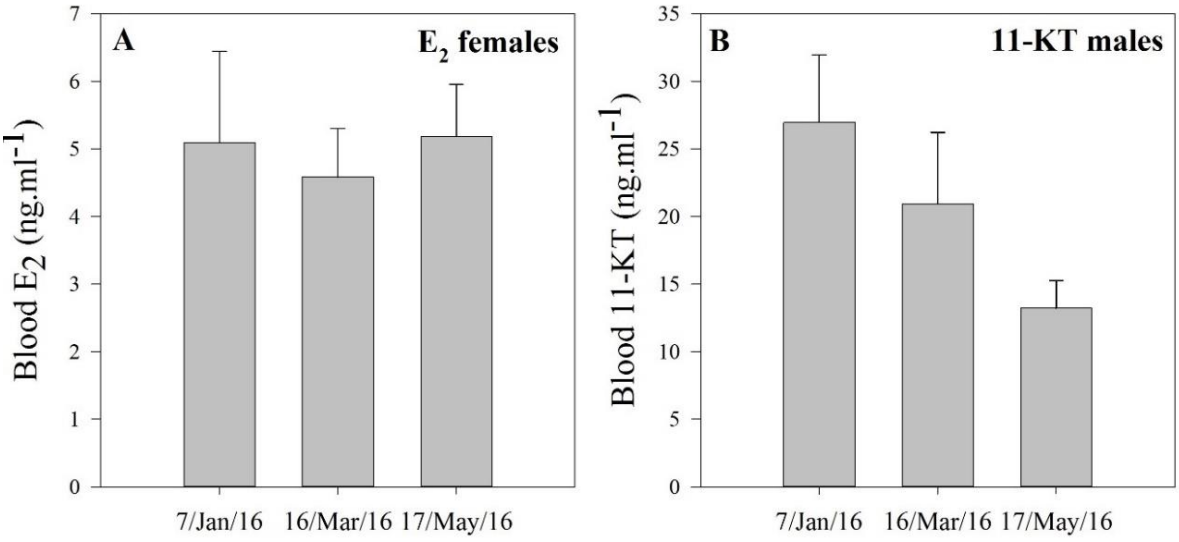

Figure 4.B.8. Plasma sex steroid levels in studied broodstock of *Arapaima gigas*. No significant time effects were seen (P>0.05). A. 17β-oestradiol (E_2) in females (12 individuals). B. 11-ketotestosterone (11-KT) in males (12 individuals).

4.B.4. Discussion

Methods for gender identification and assessment of fish reproductive condition are key to broodstock management to estimate maturity, to understand reproductive dysfunctions and to employ hormonal manipulations (Mylonas & Zohar, 2001; Mylonas et al., 2010). In practice, such methods vary according to the species morphology, ranging from observations of the urogenital papilla intumescence (Carosfeld et al., 2003), biochemical analyses of ovarian fluids (Rime et al., 2004), cannulation and/or massaging out oocytes and sperm (West, 1990; Rhody et al., 2013) or through examinations with ultrasound, endoscopy, boroscopy and laparoscopy (Kynard & Kieffe, 2002; Wildhaber et al., 2005;

Divers et al., 2009; Albers et al., 2013). The description of the urogenital anatomy is critical to perform cannulation and allow gonad assessment (Rasotto & Shapiro, 1998; Carlisle et al., 2000; Siqueira-Silva et al., 2015). In *A. gigas*, there is no evident secondary sexual dimorphism in the urogenital papilla. However, endoscopy examination in females showed the urinary canal and the gonoduct reach a common aperture inside the urogenital papilla, with their openings separated by a membranous septum. To observe and displace this septum, a guiding wire was used through the endoscope working channel, allowing to assess the gonopore and observe *in situ* the ovary. Given male identification could only be done indirectly, further studies should better describe morphology and position of the spermiduct opening with anatomical and histological analyses. These would enable to collect semen and carry out artificial fertilisation and cryopreservation, techniques already routinely used in several other Amazon species such as the catfish *Pseudoplatystoma corruscans* or the characid *Brycon orbignyanus* (Carosfeld et al., 2003).

Endoscopy analyses in adult females under farm conditions made the observation *in situ* of the ovary possible, allowing rapid and non-invasive validation of preliminary results described in Chapter 4 – Part A (Torati et al., 2016). Thereafter, through a better understanding of the urogenital anatomy, adult *A. gigas* females could be cannulated and oocytes sampled by biopsy without endoscopic guidance. This led to the description of the oogenesis in *A. gigas* with an oocyte development scheme. This was a major breakthrough for the study of the species reproduction. According to other teleost species, oogenesis includes six distinct periods of development: mitotic division of oogonia, chromatin nucleolus, primary growth, secondary growth, oocyte maturation and ovulation (Grier et al., 2009; Grier, 2012). During these periods, the appearance of ooplasm inclusions (*i.e.* cortical alveoli, oil globules) varies between species, justifying

the need for a staging scheme of oocyte development in *A. gigas* (Mañanós et al., 2008; Grier et al., 2009; Grier, 2012). Previous studies performed on wild-caught fish have described oocyte development in *A. gigas* (Bazzoli & Godinho, 1994; Godinho et al., 2005), however without details of final oocyte maturation and ovulation which are key stages to understand potential reproductive dysfunctions and environmental signalling of reproduction in *A. gigas*.

Overall, oogenesis in *A. gigas* is similar to oogenesis described for most freshwater species, and particularities of oocyte maturation and micropyle morphology in Arapaima has not been studied before possibly due to sampling limitations (Godinho et al., 2005). During primary growth in *A. gigas*, Balbiani bodies were seen in oocytes at the multi nucleoli and perinucleoli stages only seen in juveniles (PGmn and PGpn, from 32.2 - 320.8 µm). Balbiani bodies comprise a complex association of mitochondria, multivesicular bodies and Golgi elements of still unclear functions (Tyler & Sumpter, 1996; Grier et al., 2009). Oil droplets in *A. gigas* were marked from the PGod stage (354.5 – 599.0 µm) and increased in number and size until the full-grown stage of secondary growth step (SGfg), observed only in adult fish. Also known as lipid yolk, oil droplets are formed by very low density lipoproteins (VLDL) and their nutritional composition have been studied in several teleosts but excluding *A. gigas* for which there is still no information (Wiegand, 1996; Le Menn et al., 2007). In this study, cortical alveoli appeared during the PGod step (354.5 – 599.0 µm), and were found migrated at the ooplasm periphery from the PGca (627.0 – 777.0 µm) stage onwards. They remain at the periphery until ovulation, and their content is discharged upon fertilisation preventing polyspermy (Patiño & Sullivan, 2002; Grier et al., 2009), though polyspermy is very unlikely in species with a single micropyle. Herein, cortical alveoli of *A. gigas* appears continuous with multiple layers of small vesicles, supporting previous histological

observations from Bazzoli & Godinho (1994), who also found these are composed by neutral glycoproteins and sialic acid-rich glycoproteins.

In teleosts, the onset of secondary growth is marked by a significant uptake of VTGs, which are synthesized in the liver under control of E_2, and then largely transported into the oocytes (Babin et al., 2007a; Lubzens et al., 2010; Lubzens et al., 2017). This process results in visible accumulation of yolk in the ooplasm (Grier et al., 2009). In *A. gigas*, yolk accumulation was accompanied by an intense oil droplet increment, which is in accordance with overall patterns observed for other fish species (Patiño & Sullivan, 2002). During the reproductive season, levels of E_2 remained steady (*cc*. 5 $ng.ml^{-1}$), consistent with observation of vitellogenic cohorts in the ovary of all females analysed at both samplings. In species with asynchronous ovarian development, plasma E_2 levels tend to increase and remain high during the spawning season (Rinchard et al., 1993; Barcellos et al., 2001), when several oocyte batches are recruited for final oocyte maturation, resulting in batch spawning. When oocytes reach the full-grown stage (in *A. gigas* from 1139.5 to 1276.0 µm), they become maturationally competent and their membrane receptors become capable of binding maturation-inducing steroids, which initiates oocyte maturation (Patiño & Sullivan, 2002; Brown-Peterson et al., 2011). Females with full-grown oocytes are capable of spawning and often selected for hormonal manipulation in hatcheries (Mylonas et al., 2010; Brown-Peterson et al., 2011). In the wild, full-grown oocytes indicate the start of the breeding period and the moment when fecundity can be better estimated (Núñez & Duponchelle, 2009).

Oocyte maturation in teleosts is characterised by a sequence of morphological changes in the ovarian follicle prior to ovulation, when the germinal vesicle migrates towards the animal pole, breaks down and meiosis resumes completing the first meiotic division (Patiño & Sullivan, 2002; Grier et al., 2009; Lubzens et al., 2017). During oocyte

maturation, hydrolysis of lipoprotein yolk occurs, in a period when the oocyte hydrates intensely resulting in external colour changes which can be species-specific (Cerdá et al., 2007). A change in colour from yellow to green was reported during oocyte development in *A. gigas* (Chu-Koo et al., 2009; Núñez & Duponchelle, 2009), and based on histological analyses, the results showed that this colour change appeared to be associated with final oocyte maturation. As such, ovary/oocyte colour could be used as a good indicator of maturation to interpret oocyte biopsies in the field. During OM, oocyte volume increased 6-fold (diameter ranging from 1.30 to 2.35 mm). This is considered high compared to other freshwater species which release benthic eggs (typically 1-3 fold) (Cerdá et al., 2007). High hydration levels and presence of oil globules often provide buoyancy capabilities in fish species (Cerdá et al., 2007). In *A. gigas*, the apparent lack of oil globules may explain the benthic nature of its eggs, since eggs are deposited in nests built in shallow areas with low water flow (Castello, 2008b).

Histological analyses of ovary biopsies in *A. gigas* showed most females had oocytes ranging from primary growth (PGod) to final oocyte maturation (OMgvb), typical of species with asynchronous ovarian development, where heterogeneous populations of developing oocytes subsequently undergo ovulation in several batches during the spawning season (Tyler & Sumpter, 1996). Previous works with-wild caught *A. gigas* have reported up to six batches of eggs being recruited and spawned along the reproductive season from October to May, with an inter-spawning period of approximately 1 month and in some cases spawning can occur outside the rainy season (Queiroz, 2000). For the first time, this study confirmed females of *A. gigas* reared in captivity can undergo final oocyte maturation without spawning when paired with a single male in earthen ponds. However, it must be acknowledged that it was not possible to confirm whether fish released eggs in earthen ponds. Females either released eggs

which were not fertilised in the nests, implying a male reproductive dysfunction or lack of synchronisation between spawning in both genders, or oocytes were resorbed, since most of the screened females showed atretic oocytes in the ovaries at the end of the spawning season. In species reared in captivity, females can fail to release all ovulated eggs, with a portion undergoing resorption in the ovary, yet it can happen that pre- or post- ovulation atretism occur (Migaud et al., 2002). 11-KT is a key indicator of spermatogenesis in teleosts (Schulz et al., 2010). In males, the impossibility to monitor spermatogenesis and limited sampling points restrict conclusions. Levels of 11-KT remained unchanged with a tendency to decrease, suggesting the end of the spawning season, although this remains speculative. Follow up studies should monitor reproductive function over a longer period and with a higher resolution to better describe oocyte recruitment, spermatogenesis and spawning rhythms in *A. gigas*. Hereupon, research efforts should go towards enabling cannulation in males.

Interestingly, the present study also describes the micropyle in the species taking advantage of the collection of fully mature eggs spontaneously released by one female during one of the samplings. The micropyle is a funnel-shaped canal through the zona pellucida at the animal pole which allows passage for a single spermatozoon upon fertilization (Wallace & Selman, 1981). Non-fertilized eggs of *A. gigas* were found to bear a single micropyle. Comparative studies have shown a single micropyle is a basal condition present in most Actinopterygii, although more derived groups can bear 3 to 15 micropyles (*i.e.* Sturgeon *Acipenser* spp.) as a secondary change (Bartsch & Britz, 1997). To the best of our knowledge, this is the first micropyle description for an osteoglossid species, and as such it can be useful in future comparative studies (Isaú et al., 2011). Together with molecular markers, the possibility to analyse micropyle morphology in

Arapaima after eggs biopsy could be powerful to clarify current taxonomic issues under debate for *Arapaima* (Stewart, 2013a; Stewart, 2013b).

In conclusion, this study advanced our knowledge of the reproductive biology of *A. gigas* in captivity, with novel data on gonad anatomy and ovarian development. With the information provided herein, endoscopy and cannulation are clearly useful tools which can be promptly applied for gender identification and monitoring of reproductive function in wild and captive stocks. Eggs of *A. gigas* have a single micropyle, whose morphology was described for the first time in an osteoglossid, being useful not only to systematics/taxonomy, but also for possible future studies on fertilisation processes in this ancient group of fishes. As such, the deployment of these tools should help to better understand the environmental, social and hormonal cues that are required to promote or induce spawning in the species. In addition, while no spawning and/or mating was clearly observed, almost all females under investigation displayed oocytes at final oocyte development or ovulation stages which confirmed that the difficulty to breed *A. gigas* is not due to reproductive dysfunctions in females during oogenesis but rather at spawning, either due to lack of spawning, males not spermiating or not being synchronised. As such, despite previous results obtained in Chapter 3, further optimisation of hormonal therapies should be done to promote *A. gigas* spawning together with investigations into male reproductive development and behaviour.

CHAPTER 4: PART C

C. Chapter 4 - Part C

EFFECTS OF GNRHA INJECTION ON OVARY DEVELOPMENT AND PLASMA SEX STEROIDS IN *ARAPAIMA GIGAS*

4.C.1. Introduction

In a previous experiment, the effects of slow-release GnRHa implants were tested in different size matching pairs of *A. gigas* (Chapter 3). A key limitation in Chapter 3 was the lack of monitoring of ovary development and therefore the random selection of broodstock for hormonal implantation (Núñez et al., 2011). The selection of females at advanced stages of oogenesis (*i.e.* full-grown stage and onset of final oocyte maturation) is considered as one of the main factor of success for hormonal induction in fish (Mylonas et al., 2010; Brown-Peterson et al., 2011). Therefore, in the previous experiment adult *A. gigas* were implanted during the rainy season ranging from November to May in the studied site, when females are most likely to have oocytes in advanced vitellogenic stage and possibly final oocyte maturation. The experiment involved the use of slow release (up to 21-days) GnRHa implants, which suits well species with asynchronous ovarian development such as *A. gigas* (Godinho et al., 2005; Mylonas et al., 2007; Mylonas et al., 2010). However, following hormonal induction, no spontaneous spawnings were observed. Steroid analyses showed increased 11-ketotestosterone (11-KT) and testosterone (T) levels in males and females respectively, however plasma 17β-oestradiol (E_2) levels remained unchanged. Lack of E_2 increase following implantation may indicate a reduced potency of mGnRHa to induce ovulation and spawning. Since this experiment was performed, advances were made with a non-surgical gonoductoscopy technique in *A. gigas* (Chapter 4 - Part A; Torati et al. (2016)), which made possible to perform ovarian biopsies and better characterise reproductive condition in captivity (Chapter 4 - Part B). The results also confirmed that captive *A. gigas* females screened were at advanced stages of oogenesis. The validation of cannulation as a technique to biopsy *A. gigas* ovary *in situ* is a major step forward as it means broodstocks to be induced hormonally can be

selected based on their gonad developmental stage and the effects of the induction can be monitored post-treatment.

Another critical factor of any hormonal therapy is the choice of the most appropriate form of GnRHa or mixture of GnRHs for inducing the hormonal cascade along the brain-pituitary-gonadal axis leading to ovulation and spawning (Mylonas & Zohar, 2001; Mylonas et al., 2010). Based on phylogenetic and neuroanatomical data, there are three main classes of GnRH (1, 2 and 3) grouping 25 different GnRH forms described for Chordata (Fernald & White, 1999; Kim et al., 2011; Chang & Pemberton, 2017). Any given teleost usually displays brain expression of two or three GnRH forms, with GnRH 1 often controlling reproduction (Chang & Pemberton, 2017). GnRH 2 is not considered hypophysiotropic, being involved in the neuromodulation of reproductive behaviours (Okubo & Nagahama, 2008), whilst GnRH 3 is restricted to teleosts being the hypophysiotropic form in species where GnRH 1 is absent (Roch et al., 2014; Chang & Pemberton, 2017). Therefore, analogues of GnRH 1 and 3 are the main forms used to induce spawning in Teleosts (Mylonas & Zohar, 2001) although species-specific differences in potency of the different GnRHa forms have been demonstrated in the literature (Forniés et al., 2003). In the previous experiment (Chapter 3), *A. gigas* were implanted with mGnRHa which appeared to not be effective at inducing spawning. While other factors may have been involved (*i.e*. stage of oogenesis which could not be assessed, timing of induction not optimal, social factors due to random pairing of broodstocks, dopaminergic inhibition in the species), one potential factor is the use of mGnRH. A previous study identified sGnRH in the adenohypophysis of *A. gigas* (Borella et al., 2009), and a proteomic investigation on the cephalic secretion identified mGnRH and sGnRH through GeLC-MS/MS analyses (Chapter 5; results not shown). These findings

could indicate that both GnRH types may contribute to stimulate the gonadotroph cells in the pituitary and induce *A. gigas* reproduction.

The present experiment aimed to test the dual effects of sGnRHa and mGnRHa to induce final oocyte maturation, ovulation and spawning in couples of *A. gigas*, taking advantage of the ability to assess stage of oocyte development through biopsy validated earlier in this chapter.

4.C.2. Materials and Methods

Fish and methods used for handling, gender identification, sample collection (blood and oocyte biopsy), steroid analysis, oocyte development terminology and measurement methods are described in Chapter 4 Part B.

4.C.2.1 Experimental design

On 17th May 2016 (day 0), 12 adult females (TL: 166.6 ± 7.2 cm and BW: 43.0 ± 5.6 kg) randomly paired with similar sized males (n=12, TL: 159.6 ± 9.8 cm, BW: 37.5 ± 6.0 kg) in 250 m^2 breeding earthen ponds were sampled for blood and females cannulated to determine stage of ovary development. All oocytes were measured and leading cohort (LC), defined as the mean of the 10 largest oocytes, was used to characterise female reproductive condition. The eight females with largest LC (from 1.69 to 2.20 mm; corresponding to final oocyte maturation and germinal vesicle migration step) were selected for hormonal stimulation. The remaining four females with LC ranging from 1.28 to 1.67 mm (full-grown and germinal vesicle migration steps) were not injected but monitored for comparison. Males were also injected with a lower dose and pairs kept in their same ponds throughout the experiment. Injections were given intraperitoneally (IP) inserting the needle at the base of the right pelvic fin (Fig. 4.C.1). The GnRHa forms (D-

Ala⁶)-LHRH (*mammalian*) and (Des-Gly¹⁰,D-Arg⁶,Pro-NHEt⁹)-LHRH (*salmon*) (Bachem n° H4020 and H9205, Switzerland) were dissolved in a saline solution resulting in a 50:50 (v:v) mixture containing both forms. A first small dose of 9.9 ± 0.7 $\mu g.kg^{-1}$ was given in males and 19.9 ± 1.1 $\mu g.kg^{-1}$ in females, followed by a higher dose of 37.0 ± 0.2 $\mu g.kg^{-1}$ in males and 60.2 ± 0.3 $\mu g.kg^{-1}$ in females spaced by 24 hours (Almendras et al., 1988). Following the dual injection, the behaviour of the couples was monitored as a proxy for reproduction (*i.e.* positioning in earthponds). Mortalities were recorded at days 4, 5, 6 and 7 (3 females and 1 male) only in the injected group suspected to be due to handling. Therefore, ovary biopsies and blood sampling were not performed post-implantation. When mortalities ceased, a follow up sampling (biopsy and blood) was performed on day 15 (31st May 2016) to evaluate the effect of hormonal injections on oocyte development and plasma sex steroid levels.

Figure 4.C.1. Intraperitoneal (IP) injection of mGnRHa and sGnRHa 50:50 (v:v) at the base the of right pelvic fin in *Arapaima gigas* broodstock.

4.C.2.2. Statistics

Statistical analyses were conducted in Minitab (version 17.3.1, Minitab, PA, USA). A 2-sample *t*-test was used to compare steroid change ($ng.ml^{-1}$) and LC mean change (mm)

from days 0 and 15 comparing injected and non-injected groups. Paired *t*-tests were used to compare LC (mm) among days 0 and 15 for each individual female. The level of significance was set as $P \leq 0.05$ and data presented as mean ± SE unless stated otherwise.

4.C.3. Results

Comparing injected and non-injected groups for the LC oocyte diameter growth (days 0 to 15), no significant difference was observed (2-sample *t*-test; $P>0.05$). Comparing LC oocyte diameter growth from day 0 to 15, 4 females (80 %) of injected group and 3 females of non-injected group (75 %) showed progression in LC oocyte diameter (paired *t*-test; $P<0.05$; Fig. 4.C.2). However, ovulation was not observed in any female of either groups.

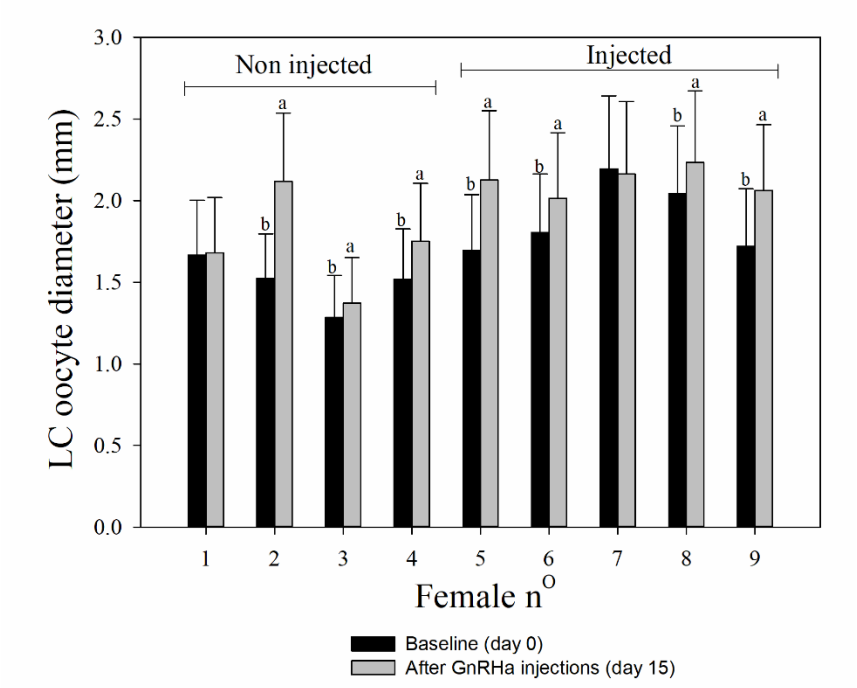

Figure 4.C.2. Changes in mean oocyte diameter of leading oocyte cohort (LC) between baseline (day 0) and 12 days post GnRHa injections (day 15) from injected and non-injected females of *A. gigas* treated with a combination of mGnRHa and sGnRHa 50:50

(v:v). Different letters denote significant change in female LC between sampling dates (Paired t-tests; P<0.05).

Baseline plasma E_2 levels in females were similar for injected (n = 5; 4.7 ± 1.6 $ng.ml^{-1}$) and non-injected (n = 4; 5.6 ± 2.7 $ng.ml^{-1}$) groups. No significant differences in plasma E_2 level variation between pre- (baseline day 0) and post-hormonal injections (day 15) were observed due to large individual variability (2-sample t-test; P>0.05; Fig. 4.C.3A). Baseline plasma levels of 11-KT in males were similar in injected (n = 7; 12.6 ± 5.4 $ng.ml^{-1}$) and non-injected groups (n = 4; 11.3 ± 4.3 $ng.ml^{-1}$). After 12 days post-injection (day 15), levels did not change significantly (2-sample t-test; P>0.05; Fig. 4.C.3B) between groups although levels appeared to increase in the injected group.

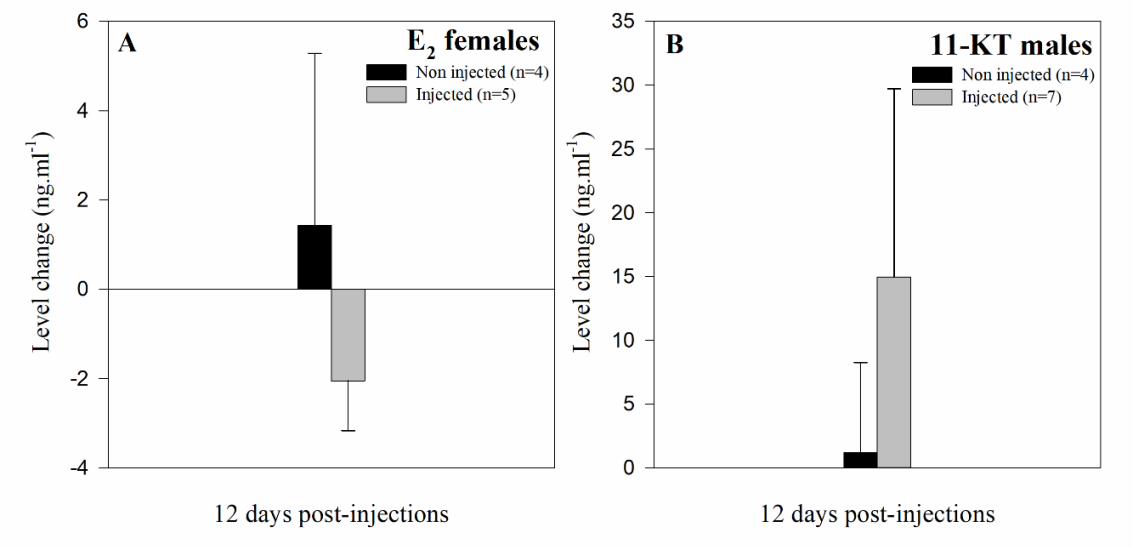

Figure 4.C.3. Change in plasma sex steroid levels ($ng.ml^{-1}$) after 12 days post GnRHa injections in *A. gigas*. No significant difference (2-sample *t*-test; P>0.05) was seen comparing non injected and injected groups treated with a combination of mGnRHa and sGnRHa 50:50 (v:v). A. 17β-Oestradiol (E_2) in females. B. 11-ketotestosterone (11-KT) in males.

4.C.4. Discussion

The present study aimed to test injection of different GnRH forms to induce ovulation and spawning in maturing females *A. gigas*. In other fish species with asynchronous ovarian development like in *A. gigas*, spawning would be expected within 2-5 days following injection (Fernandez-Palacios et al., 2015; Mylonas et al., 2015). Therefore, injected couples were monitored for nest building and guarding behaviours which are typical of the species reproductive behaviour characterised by fish remaining at a same position in the pond for extended periods of time. However, no such behaviour was observed following the hormonal induction protocol and mortalities prevented scheduled biopsy sampling following injections. In some commercially important fish species including salmonids like Atlantic salmon (*Salmo salar*), cyprinids like carp (*Cyprinus carpio*) or flatfish like turbot (*Scophthalmus maximus*), ovulation can be induced hormonally but spawning does not occur spontaneously and requires stripping fish to collect eggs for artificial fertilisation. This may have been the case also for hormonally induced *A. gigas* in the present study however due to unexpected mortalities following injections, fish could not be handled to prevent further stress and stripping is considered impossible in *A. gigas*.

Potential causes for the mortalities remain speculative. GnRHa is considered the safest hormone to use in hatcheries and the dose used was within the range usually tested in other teleosts (5 – 5,000 $\mu g.kg^{-1}$), though this data include other delivery systems (Mylonas & Zohar, 2001). Cannulation itself is also safe, employed in many species, and did not result in any mortality when performed in *A. gigas* broodstock (Chapter 4-Part B). The most likely cause of mortality was the repeated handling for 3 consecutive days involving fasting, netting the fish out of the ponds and sampling. Mortalities in broodstock are commonly reported in hatcheries following multiple handling for health

checks and disease treatment, however not following experimental manipulations. This clearly indicates slow-release implants are the best choice when attempting to induce spawning, although implants have limited market availability and high costs especially to suit large-sized species.

A follow-up sampling was made 12 days post-injection, when mortalities had ceased and fish had recovered. The results suggested oocyte maturation was advanced in LC oocytes of GnRHa-injected and control groups, without significant differences between treatments. Analysis of sex steroid change was not very informative given the interval of time between injection and sampling. Spontaneous spawning or ovulated females were also not seen, indicating the treatment had not been effective. Lack of effectiveness of GnRHa injection protocols can be related to the short-lived hormonal stimulation with GnRHa in the blood circulation only detected for a short window of time following injection (23 min *in vivo*) and negatively correlated to temperature (Mylonas & Zohar, 2001). In many species, the process of final oocyte maturation may require several days to be completed, requiring prolonged GnRHa stimulation of the gonadotroph cells and secretion of gonadotrophins in the blood circulation (Mylonas & Zohar, 2001; Mylonas & Zohar, 2007; Mylonas et al., 2010). In *A. gigas*, the inter-spawning period was suggested to be as long as 30 days (Queiroz, 2000), indicating the selection of females at a more advanced gonad developmental stage (LC > 2.3 mm) may be required to hormonally induce ovulation and spawning. In addition, stress is a well-known factor of reproductive arrestment and gonad resorption through atresia. The multiple handlings may have caused stress-induced suppression of normal advanced oogenesis and spawning behaviour, a bottleneck frequently reported in treatments involving multiple injections (Mylonas & Zohar, 2001). Future studies to characterise gonadal development in *A. gigas* following environmental and hormonal protocols will be key to the control of captive

reproduction. Taken together, despite limited and disappointing results, this experiment strongly advocated the development and use of less invasive induction methods with reduced fish handling (*i.e.* environmental, social, slow-release implants) and welfare impact. The control of spawning and the reliable supply of eggs remains a critical bottleneck that limits the expansion of the *A. gigas* aquaculture sector.

CHAPTER 5

5. Chapter 5

COMPARATIVE PROTEOME AND PEPTIDOME ANALYSIS OF THE CEPHALIC FLUID SECRETED BY *ARAPAIMA GIGAS* (TELEOSTEI: OSTEOGLOSSIDAE) DURING AND OUTSIDE PARENTAL CARE.

ORIGINAL ARTICLE

Submitted: April, 2017

Accepted (Minor revision): June, 2017

Accepted in: PlosOne

Torati, L.S.; Migaud, H; Doherty, M.K.; Siwy, J.; Mullen, W.; Mesquita, P.E.C. & Albalat, A.

Breeding and Physiology team, School of Natural Sciences, Institute of Aquaculture, University of Stirling, FK94LA, Stirling, Scotland, UK.

Contributions: The present manuscript was compiled in full by the author of this thesis (LT). The experiment was conceived and designed by LT, AA, PM and HM. Laboratory and data analyses were carried out by LT, MD, JS, WM and AA. All authors approved the final version of the manuscript.

Abstract

Parental investment in *Arapaima gigas* is pivotal for the species survival, which includes nest building, guarding and male parenting during which a cephalic fluid is released from the parent's head. This fluid presumably has important functions for the offspring but so far its composition has not been characterised. In this study, the proteome and peptidome of the cephalic secretion was studied in parental and non-parental fish using capillary electrophoresis coupled to mass spectrometry (CE-MS) and GeLC-MS/MS analyses. Multiple comparisons revealed 28 peptides were significantly different between males (M) and parental males (PM), 126 between females (F) and parental females (PF), 51 between M and F and 9 between PM and PF. Identification revealed peptides were produced in the inner ear (*pcdh15b*), eyes (*tetraspanin* and *ppp2r3a*), central nervous system (*otud4*, *ribeye a, tjp1b* and *syn1*) among others. A total of 422 proteins were also identified and gene ontology analysis revealed 28 secreted extracellular proteins. From these, 2 hormones (*prolactin* and *stanniocalcin*) and 12 proteins associated with immunological processes (*serotransferrin*, *α-1-antitrypsin homolog*, *apolipoprotein A-I*, and others) were identified. This study provides novel biochemical data on the lateral line fluid which will enable future hypothesis-driven experiments to better understand the physiological roles of the lateral line in chemical communication.

Keywords: CE-MS, GeLC/MS-MS, hormones, immunity, lateral line, Pirarucu.

5.1. Introduction

Freshwater oviparous teleosts lay eggs in rivers and lakes where food resources are not always abundant and predators or pathogens can significantly impact numbers of descendants. During the evolution of teleosts, several mechanisms have arisen comprising parental investment, which include parental actions to increase offspring survival (Trivers, 1974). Parental care behaviour in teleosts involves building nests with appropriate substrates, defending and guarding the eggs, embryos and larvae against predators, and in some species parents provide protection for extended periods during juvenile development (Balshine & Sloman, 2011). An important component of parental care strategies is the provision of nutrients to the offspring either through egg yolk or at post-hatch stages (*i.e.* mucus feeding) (Sloman & Buckley, 2011). In some species, passive immunity or stimulated growth can occur via parental biochemical interaction with the offspring (Schütz & Barlow, 1997). The provision of nutrients, growth and immunity factors at early developmental stages is particularly important to reduce the time offspring spend under more vulnerable conditions, therefore increasing likelihood of survival.

Parental investment is a critical component in the life-history of the obligate air-breather *Arapaima gigas* (Schinz, 1822) in the Amazon (Castello et al., 2011). Despite being an emblematic species of the Neotropical ichthyofauna (adults can reach up to 2.5 m in total length) (Nelson et al., 2016), the species is threatened due to overexploitation of natural stocks combined with the lack of knowledge on its basic biology (Castello & Stewart, 2010). Reproduction of *A. gigas* in natural environments begins with the rainy season generally from December to May (Núñez et al., 2011). With the start of the flooding, adults build nests for mating in temporary lagoons (Núñez et al., 2011). Courtship behaviour involves chasings and fights with marked territorialism, and

spawning is followed by external fertilization on the nest (Fontenele, 1948). After spawning, the male and female initiate nest guarding behaviour: while one parent is at the surface breathing, the other is always protecting the brood on the nest. The occurrence of mouth incubation or egg transportation has been suggested (Farias et al., 2015) but no systematic observation supporting oral incubation in *A. gigas* has been reported. Nine days after spawning, the eggs hatch and larvae start aerial breathing. The male provides intensive parental care which can last up to 3 months, guiding the offspring above its darkened head into zooplankton-rich areas for feeding (Castello, 2008a). Fry shoaling at this stage is remarkably organized, for instance the male darkening is believed to provide camouflage for the offspring against predators (Castello, 2008b). Female participation in parental care at this stage seems less relevant due to female unchanged body colour (Fontenele, 1953). However, the female swims around the male and offspring at longer distances (>2m) for some period after the nest guarding phase (Farias et al., 2015). This behaviour is still poorly understood, but could involve territorial inspection aiming at predator avoidance or location of food-rich areas (Fontenele, 1953). The female normally leaves the male and offspring after a poorly documented period (*c*. 1 month), and can reproduce again with other males during the same reproductive season (Núñez et al., 2011).

The mechanisms by which parental investment occur in *A. gigas* are not fully understood, and a particularity has called attention of many ichthyologists. On the head surface of adult males and females, the cephalic canals of the lateral line system are well developed forming a series of cavities covered by a pore-bearing integument, from which a whitish fluid is released into the water (Lüling, 1964). The indigenous refer to this fluid as “Arapaima milk” since after fertilization on the nest, eggs, developing larvae and the growing offspring are in constant interaction with the male and female’s head, being in

close contact with this secretion. Several roles for this fluid have been suggested since its production is intensified during reproduction and parental care. However, whether the cephalic secretion contributes to fingerling nutrition (Eigenmann & Allen, 1942; Sawaya, 1946; Quay, 1972; Noakes, 1979) or enhances the fry immunological system, increasing survival rates (Rojas, 2005), are hypothesis never truly investigated given the lack of basic information on the biochemical nature of the cephalic secretion. Lack of biochemical information hinders also our understanding of fish chemical communication, which involves release of pheromones and signature mixtures, most commonly peptides, that convey varied messages to conspecifics including the offspring (Wyatt, 2010).

In this context, the un-targeted study of the proteome and peptidome is an ideal "search space" for molecules which can help us to understand the functions behind biological fluids, with examples found in proteomic studies of cerebrospinal fluids (CSF), fish mucus, plasma, tear fluid, urine and others (Hayakawa et al., 2013; Holtta et al., 2015; Jurado et al., 2015). With the analytical –omics advances of recent years, proteomes and peptidomes have been studied in different species, generating protein databases that are available for protein identification, phylogenetic comparisons and gene ontology analyses based on homology-driven approaches (Shevchenko et al., 2009). Consequently, proteomic methods are being used in a wide range of disciplines such as behavioural ecology, aquaculture and food sciences (Zhou et al., 2012; Valcu & Kempenaers, 2014; Farias et al., 2015). In most cases, the initial mapping of proteins and peptides is necessary to allow later mechanistic-driven hypothesis. Such examples are illustrated in studies with other parental fish species, such as the mapping of the mucus proteome in the Discus fish (*Symphysodon aequifasciata*) (Chong et al., 2005), which later allowed better understanding on the roles of prolactin (PRL) in the parental care of that species (Khong et al., 2009). Similarly, after an initial mapping of the mucus

proteome of the Atlantic Cod (*Gadus morhua*), important proteins related to fish immune response were elucidated (Rajan et al., 2011). In general for teleosts, there is a lack of basic knowledge on the proteins present in the cephalic fluid comprising the anterior lateral line (Sorensen & Wisenden, 2015). Having access to few of these valuable samples during parental care of *A. gigas* enabled this study to characterise the proteome and peptidome comparing parental and non-parental males and females, thus generating novel data to increase our understanding of parental care processes in this species and possible roles of the cephalic secretion in teleosts.

5.2. Materials and Methods

5.2.1. Animal sampling

This study was carried out in the *Rodolpho von Ihering* research station- DNOCS (3°48'09.54"S, 39°15'56.73"W) at Pentecoste (Brazil). Adult broodstock of *A. gigas* known to be over 6 years of age received a passive integrated transponder (PIT; AnimallTAG®, São Carlos, Brazil) in the dorsal muscle to allow individual identification. Fish gender was identified with a vitellogenin enzyme immune assay (EIA) kit (Acobiom©, Montpellier, France) developed specifically for *A. gigas* (Dugue et al., 2008) and females were paired with males in 300 m^2 breeding earthen ponds. The sampled males measured 162.8 ± 25.9 cm in total length (TL) and weighed 39.9 ± 16.8 kg, whilst females measured 180.5 ± 14.5 cm in TL and weighed 52.3 ± 12.3 kg. Before sampling, fish were fasted for 24 hours, netted from earthen ponds and kept contained on a soft wet mat for 5 to10 minutes. Anaesthetics were not applied during sampling as anaesthesia has been shown to compromise welfare and result in mortalities in *A. gigas* due to its obligate air-breathing behaviour (Farrel & Randall, 1978). Fish breathing behaviour was closely monitored during sampling (breathing at regular intervals of 4-6 minutes). Tags were read

and 1-2 ml of cephalic fluid was sampled from the dorsalmost lateralis sensorial cavity of the preopercle using a sterile needle inserted inside the integument pocket (BD Precisionglide 22G, New Jersey, USA). The collected fluid was entirely from the fish sensorial cavity and samples did not include visible blood contamination. Samples were immediately frozen in liquid nitrogen and then stored at -80°C. After sampling, fish were returned to the ponds and monitored until return to normal breathing behaviour, and no mortality was recorded following this sampling procedure.

Non-parental and parental fish were sampled at different time points. A total of 11 non-parental males (M) and 12 non-parental females (F) were sampled on March 24th 2014. Additionally, two couples spawned twice from March to June 2014, and from them a total of 10 samples were taken during the parental care phase. Since it was not possible to observe spawning of *A. gigas* due to water turbidity in breeding ponds, the fertilization time in *A. gigas* was calculated by subtracting 9 days from the day when offspring were first observed air-breathing (Fontenele, 1953). Therefore, samples from parental males (PM) and females (PF) correspond to approximately 13, 25, 35 and 62 days post spawning (dps). All samples were individually analysed using capillary electrophoresis (CE-MS) after peptide extraction. For the protein analysis (GeLC-MS/MS), equal protein amounts of each sample were pooled after protein quantification. These protein pools corresponded to: M (n=6 males at 1-time point), F (n=6 females at 1-time point), PM (n=5 samples, 1 male at 2-time points and 1 male at 3-time points) and PF (n=5 samples, 1 female at 2-time points and 1 female at 3-time points).

This research complied with the Brazilian guidelines for the care and use of animals for scientific and educational purposes – DBCA – Concea (CEUA/CNPASA n°09). Samples were shipped in dry ice for protein and peptide extraction at the University of Stirling (Stirling, Scotland) (Permit IBAMA/CITES n° 14BR015850/DF).

Extracted peptides were analysed at the University of Glasgow and proteins at the University of the Highlands and Islands.

5.2.2. Analysis of peptides

5.2.2.1. Extraction

All buffers and solutions used were freshly prepared, following the method from Albalat et al. (2013). Cephalic fluid samples were thawed on ice and 800 µl was used for peptide extraction. An initial centrifugation at 2,000 g (5 min; 4 °C) was used to settle debris and impurities. The supernatant was removed and centrifuged at 16,000 g (10 min 4 °C) for delipidation. Following this, 750 µL of the supernatant was removed and mixed with 750 µL of urea buffer (2 M urea, 100 mM NaCl, 10 mM NH_4OH containing 0.01% SDS). This mixture was filtered in a 20 kD Centrisart tube (Sartorius Stedim, Geottingen, Germany), and centrifuged at 2,000 g to obtain 1.1 mL of filtrate. This filtrate was applied into an NAP™ -5 Saphadex™ desalting column (GE Healthcare Life Sciences, Buckinghamshire, UK) previously equilibrated with 25 ml of NH_4 buffer (0.01% NH_4OH in HPLC-grade H_2O, pH 10.5 - 11.5), and eluted with 2 ml of NH_4 buffer, collected in clean Eppendorfs, lyophilized and kept at 4 °C until CE-MS analysis was performed.

5.2.2.2. Capillary electrophoresis (CE-MS)

Peptides were analysed with a P/ACE™ MDQ capillary electrophoresis system (Beckman Coulter, Brea, CA, USA) equipped with a capillary PicoTip™ Emitter TaperTip™ of 90 cm, 360 µm OD and 50 µm ID (New Objective, Woburn, USA) online coupled to a MS micro-TOF (Bruker Daltonics, Leipzig, Germany). Prior to analysis, the capillary was conditioned for 10 minutes at 50 psi with 1 M NaOH, and later for 20 minutes with running buffer (20% acetonitrile (ACN), 0.94% formic acid and 79.05%

HPLC graded H_2O). The MS was calibrated with a tuning solution containing lysozyme (14,303 Da), ribonuclease (13,681 Da), aprotinin (6,513 Da) and the following synthetic peptides: ELMTGELPYSHINNRDQIIFMVGR (2,832 Da), TGSLPYSHIGSRDQIIFMVGR (2,333 Da), GIVLYELMTGELPYSHIN (2,048 Da) and REVQSKIGYGRQIIS (1,733 Da) and analysis was made in reverse mode. Lyophilised peptide samples were reconstituted in 9 µl of HPLC grade water (Roth, Karlsruhe, Germany) and centrifuged at 12,000 g for 10 min at 4 °C. Samples were injected at a pressure of 2 psi for 99s loading a volume of 290 nL. The capillary temperature was set at 25 °C with separation made at 25 kV for 30 minutes. Sheath flow liquid (30 % 2-propanol, 0.4% formic acid and 69.6% deionized water) was applied at the capillary end with a running speed of 0.02 ml/h. The ESI sprayer (Agilent Technologies, Palo Alto, CA, USA) was grounded (0 V), so the ion spray interface potential was set in -4.5 kV. The MS spectra were accumulated every three seconds, over a range of m/z of 50 to 3000 along 60 minutes. This method has been comprehensively described in Zürbig et al. (2009) and Good et al. (2010).

5.2.2.3. LC-MS/MS analysis

The extracted peptides were also analysed with a Dionex Ultimate 3000 RSLS nano-flow system (Dionex, Camberly UK). The samples (5 µL) were loaded onto a Dionex 100 µm × 2 cm 5 µm C18 nano-trap column at a flowrate of 5 µL/min by an Ultimate 3000 RS autosampler (Dionex, Camberley UK). The composition of the loading solution was 0.1 % formic acid and ACN (98:2). Once loaded onto the trap column the sample was washed off into an Acclaim PepMap C18 nano-column 75 µm × 15 cm, 2 µm 100 Å at a flowrate of 0.3 µL/min. The trap and nano-flow column were kept at 35 °C in a column oven in the Ultimate 3000 RSLC. The samples were eluted with a gradient of solvent A: 0.1 %

formic acid and ACN (98:2) versus solvent B: 0.1 % formic acid and ACN (20:80) starting at 5 % B rising to 50% B over 100 min. The column was washed using 90 % B before being equilibrated prior to the next sample being loaded.

The eluant from the column was directed to a Proxeon nano-spray ESI source (Thermo Fisher Hemel UK) operating in positive ion mode then into an Orbitrap Velos FTMS. The ionisation voltage was 2.5 kV and the capillary temperature was 200 °C. The mass spectrometer was operated in MS–MS mode scanning from 380 to 2000 amu. The top 20 multiply charged ions were selected from each full scan for MS–MS analysis, the fragmentation method was CID at 35% collision energy. The ions were selected for MS2 using a data dependent method with a repeat count of 1 and repeat and exclusion time of 15 s. Precursor ions with a charge state of 1 were rejected. The resolution of ions in MS 1 was 60,000 and 7500 for HCDMS2.

5.2.2.4. Data processing, peptide identification and statistical analysis

CE-MS raw data was processed for peak peaking, deconvolution and deisotoping using MosaiquesVisu software v. 2.1.0 (Mosaiques Diagnostics GmbH, Hannover, Germany). The threshold of signal to noise (SNR) was set at 4, and only signals present in 3 consecutive spectra were accepted. A matched filtering algorithm assigned charge based on the isotopic distribution. It generated a list containing all the interpretable signals with their *m/z*, migration time, charge, and signal intensity (ion counts). Signals representing the same compound with different charge states were then combined, generating a final list with compound identities defined by mass, migration time and relative abundance (ion count). To allow comparisons among different samples, CE migration time and ion signal intensity were normalized, resulting in a list with molecular mass (Da) and

normalized CE migration time (min) for each feature. The normalized signal intensity was used to measure and compare the relative abundance among samples.

LC-MS/MS data was processed initially uploading the raw spectra data into Thermo Proteome Discoverer 1.2 Thermo Scientific (Hemel Hempstead, UK). Peak picking was performed under default settings for FTMS analysis such that only peptides with signal to noise ratio higher than 1.5 and belonging to precursor peptides between 700-8,000 Da were considered. Peptide and protein identification was performed with SEQUEST algorithm. An in-house compiled database containing proteins from the latest version of the UniProt SwissProt database was compiled to include only *Danio rerio* entries. No enzyme cleavage was selected and oxidation of methionine and proline was chosen as variable modifications. Precursor tolerance was set at 20 ppm and 0.1 Da for MS/MS fragment ions. Resulting peptides and protein hits were further screened by excluding peptides with an error tolerance higher than 10 ppm and by accepting only those hits listed as high confidence by Proteome Discoverer software. Theoretical migration times in CE–MS for any resulting peptides were calculated so that sequences obtained with LC-MS/MS could be subsequently assigned to a position in the CE–MS analysis.

Aiming to discriminate peptide peaks and compare different groups, a non-parametric Wilcoxon signed rank test was conducted in R version 2.15.3 (R-Core-Team, 2014), and corrected for false-discovery rate with Benjamini and Hochberg (BH) procedure (Benjamini & Hochberg, 1995). Based on Wilcoxon tests, significantly different ($P<0.05$) peptides among groups were identified using LC-MS/MS dataset (Table S5.1) following Zürbig et al. (2006). In this process, peptides data obtained with the CE-MS had their theoretical migration times, MW and number of basic aa matched with LC-MS/MS peptide list, enabling identification using a threshold of 80 ppm on MW.

This procedure has been successfully used in several studies linking specific CE-MS-identified peptides to sequences obtained with LC-MS/MS [19].

Using the complete catalogue of peptides detected through the LC-MS/MS, a Gene Ontology (GO) analyses were performed using STRAP *v 1.5*. software (Bhatia et al., 2009). Also, aiming to identify secreted (extracellular) peptides, a GO analysis was also made for those peptides significantly different among groups after unadjusted Wilcoxon testing.

5.2.3. Analysis of proteins by GeLC-MS/MS

5.2.3.1. Extraction and quantification

Initially, the remaining cephalic secretion was centrifuged at 2,000 g for 5 minutes at 4 °C to remove any debris. Given low protein levels in samples, a protein precipitation was made at 1:3 v/v secretion/acetone (HPLC grade, Fisher Scientific). The mixture was cooled down to -20 °C for 2 hours, then centrifuged at 5,000 g for 40 minutes at 4 °C. The supernatant was discarded, the protein pellet dried under a N_2 stream and reconstituted in non-reducing buffer (13.1 mM Tris - pH 6.8, 2.63 % v/v Glycerol, 0.42 % v/v SDS). An aliquot was taken for protein quantification whilst the extracts were kept at -80 °C until inGeLC-MS/MS analysis.

The total protein content was quantified using a bicinchoninic acid assay kit (BCA) (Uptima UP40840B, Interchim, France). Reactions were prepared in 96 well plates using 2 µl of extracted sample in 40 µl of working solution (1:20 reaction) following the kit protocol. The plate was incubated for 30 minutes at 37 °C, cooled at room temperature and read at 562 nm using a ND-1000 NanoDrop® Spectrophotometer (Thermo Fisher Scientific, DE, USA). A serial dilution of bovine serum albumin (10 –

2,000 µg/ml) was used to prepare a standard curve against which the sample readings were interpolated.

5.2.3.2. In-gel digestion

Aiming to substantially expand protein identifications following Paulo (2014), each pool of samples (M, F, PM and PF) was analysed with three technical replicates. Samples (10 µg of total protein per pool) were mixed in reducing buffer (13.1 mM Tris - pH 6.8, 2.63 % v/v Glycerol, 0.42 % v/v SDS, 0.243 % v/v bromophenol blue and 163.5 mM DTT), heated up to 95° C for 5 minutes on a heating block and centrifuged at 2,000 g for 30 seconds. Reduced lysates were loaded into a 1D SDS polyacrylamide gel (4-20 %, Mini-PROTEAN® TGX, BIO-RAD™) with a protein ladder reference (5 µl, BenchMark™, 10-220kDa, ThermoFisher Scientific). Gels were run in a BIO-RAD Mini-PROTEAN® Vertical Electrophoresis apparatus immersed in tris-glycine running buffer (24.7 mM Tris-Base, 191.8 mM glycine, 3.46 mM SDS, milliQ water, pH adjusted to 8.3) at 200V (400 mA) for 25 minutes until the frontline reached 2.5 cm (incomplete run). A compete run of these pools was made with the purpose of band visualization (graphical abstract). The run gels were fixed in 200 ml of fixation solution (7 % v/v acetic acid, 40 % v/v methanol) for 1 hour at room temperature. Following, fixation solution was discarded and gel washed 3 times (5 minutes each) with milliQ water (200 ml). Gels were stained with 25 ml of Comassie G-250 (SimplyBlue™ Safe Stain, Thermo Fisher Scientific) for 45 minutes at room temperature. After discarding the stain, gels were washed and rinsed in milliQ water for 1 hour. Under a laminar flow cabinet, each gel lane was sliced in 10-12 cuts (plugs) of approximately 2.0-2.5 mm each, then processed independently in 1.5 ml reaction Eppendorfs. For destaining, plugs were incubated twice (10 minutes each) in 100 µl of destaining solution (50 mM ammonium bicarbonate and ACN 50 % v/v) at 37

°C. Samples were then incubated in 50 µl of reducing solution (10 mM dithiothreitol-DTT and 100 mM ammonium bicarbonate) at 37 °C for 30 minutes. Following, plugs were incubated in the dark with 50µl of alkylation solution (55.1 mM iodoacetamide, 100 mM ammonium bicarbonate) at 37 °C for 30 minutes, and then dehydrated in 100 µl of ACN at 37 °C for 15 minutes. After complete ACN evaporation on a heat block (37 °C), plugs were digested with 50 µl of trypsin (Promega, Madison, USA) prepared in a solution of 5 mM acetic acid and 45 mM ammonium bicarbonate. Initially, samples were incubated at 37 °C for 1 hour. Then, 20 µl of 100 mM ammonium bicarbonate was added and reaction allowed for 14 hours at 37 °C. Finally, the digests were transferred into sterile Eppendorfs containing 2 µl of 10 % formic acid, and frozen at -20 °C until analysis.

5.2.3.3. LC-MS/MS analysis

Tryptic digests were analysed with a LTQ-Orbitrap XL LC−MS^n mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a nanospray source and coupled to an Ultra High Pressure Liquid Chromatographer (UPLC) system (Waters® nanoAcquity, Manchester, U.K.). Initially, 5 µL of sample were loaded, desalted and concentrated in a BEH C18 trapping column (Waters, Manchester, U.K.) with the instrument operated in positive ion mode. The peptides were then separated on a BEH C18 nanocolumn (1.7 µm, 75 µm × 250 mm, Waters) at a flow rate of 300 nL/min using an ACN/water gradient; 1% ACN for 1 min, followed by 0−62.5 % ACN over 21 min, 62.5− 85 % ACN for 1.5 min, 85 % ACN for 2 min and 1 % ACN for 15 min.

MS spectra were collected using data-dependent acquisition in the range m/z 400−2,000 using a precursor ion resolution of 30,000, following which individual precursor ions (top 5) were automatically fragmented using collision induced dissociation

(CID) with a relative collision energy of 35 %. Dynamic exclusion was enabled with a repeat count of 2, repeat duration of 30 s and exclusion duration of 180 s.

5.2.3.4. Protein identification

The ion spectra readings were converted into peak list text files for database search using Proteome Discoverer™ Software and analysed with two search algorithms: MaxQUANT v1.1.1.3624 Andromeda and MASCOT (http://www.matrixscience.com), against the Actinopterygii SwissProt database 57.15. The use of multiple search engines has been shown to expand the number of identified proteins (union) and validates protein identifications (intersection) (Paulo, 2014). The initial search parameters allowed for a single trypsin missed cleavage, carbamidomethyl modification of cysteine residues, oxidation of methionine, acetylation of N-terminal peptides, a precursor mass tolerance of 10 ppm, a fragment mass tolerance of ±0.5 Da, and a FDR of 0.01. After MASCOT search, only protein hits with score above ≥ 18 were accepted. The Exponentially Modified Protein Abundance Index (emPAI) (Ishihama et al., 2005) was also calculated and used for relative quantification between groups.

5.2.3.5. Bioinformatics and Gene Ontology (GO) analysis

Initially, lists obtained for each group and from the different search engines were combined and redundancies were manually curated. The mean emPAI values, frequency of each protein hit were calculated, and a table containing MW (kDa), highest MASCOT score, matches, highest MaxQUANT score was produced (Table S5.2). This list was analysed for gene ontology (GO) in STRAP *v 1.5*. (Bhatia et al., 2009), using UniProtKB and EBI GOA databases, after which secreted (extracellular) proteins were revealed.

Venny 2.1. (Oliveros, 2007) was used to generate a Venn diagram depicting all detected proteins and secreted (extracellular) proteins between different groups.

5.3. Results

5.3.1. Peptide analysis on the cephalic secretion of *Arapaima gigas*

Based on CE migration time and molecular weight (MW) from individual samples, a compilation of peptide patterns was generated for each studied group (Fig 5.1). The range and number of detected peptides (mean ± SD) ranged from 187 to 3005 (1374.1 ± 1017.8; n = 9) in M, from 402 to 1397 (682.8 ± 446.2; n = 5) in PM, from 170 to 2243 (755.3 ± 639.9; n = 11) in F and from 202 to 2560 (1008.0 ± 935.2; n = 5) in PF. From these peptides, 28 peptides were significantly different between M and PM (Wilcoxon $P<0.05$; BH>0.05) (17 identified). Comparing F and PF, 126 peptides were found significantly different (Wilcoxon $P<0.05$; BH>0.05) (76 identified); 51 peptides were significantly different between M and F (Wilcoxon $P<0.05$; BH>0.05) (41 identified) and finally, 9 peptides were found significantly different between PM and PF (8 identified). Table S5.1depicts all identified peptides, MW, primary gene name, sequences and group differences.

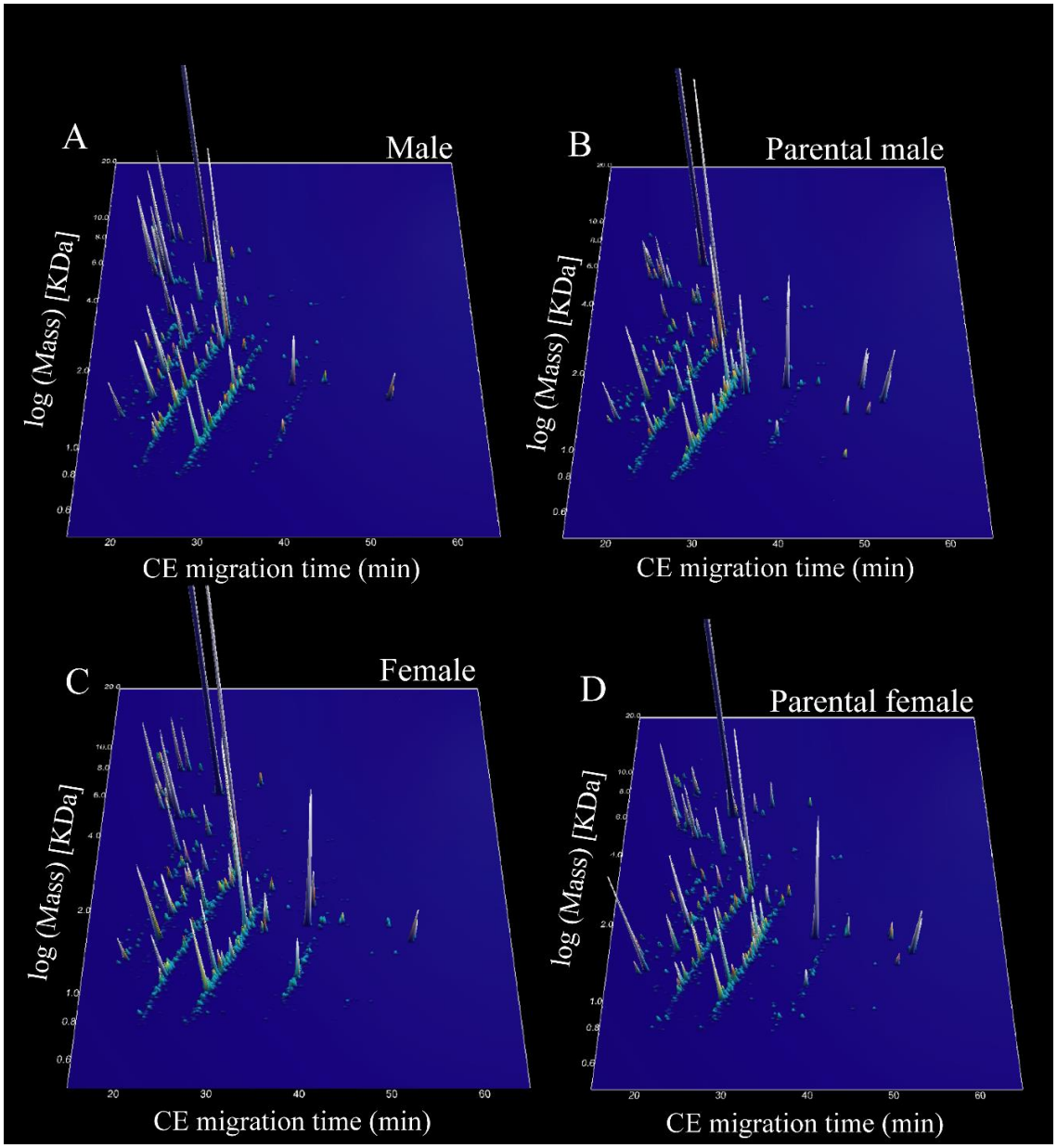

Figure 5.1. Compiled CE-MS peptide fingerprints from the cephalic secretion of *Arapaima gigas.* A. Non-parental males (M; n = 9). Parental males (PM; n = 5). C. Non-parental females (F; n = 11). D. Parental females (PF; n = 5). Migration time of capillary electrophoresis (CE) is shown in X axis whereas Y is the logarithmic scale of molecular mass (kDa) and Z is the mean signal intensity.

A list with 7009 unique peptide sequences was obtained compiling LC-MS/MS data from all the studied groups. Regarding their GO biological processes, most were related to regulation (1462; 27.8 %), cellular process (1461; 27.8 %), developmental process (698; 13.3 %), localization (345; 6.6 %), interaction with cell and organisms (298; 5.7 %) and others (625; 11.9 %) (Fig 5.2A). Considering their molecular functions, main GO categories were binding (2364; 47.4 %), catalytic activity (1485; 29.8 %), molecular transducer activity (292; 5.8 %), structural molecule activity (127; 2.5 %), antioxidant activity (7; 0.1 %) and others (709; 14.2 %) (Fig 5.2B). Considering their GO cellular components, most belonged to the nucleus (767; 17.6 %), plasma membrane (438, 10 %), cytoplasm (421; 9.6 %), cytoskeleton (228, 5.2 %), macromolecular complex (219, 5 %), extracellular (207, 4.7 %) and others (Fig 5.2C).

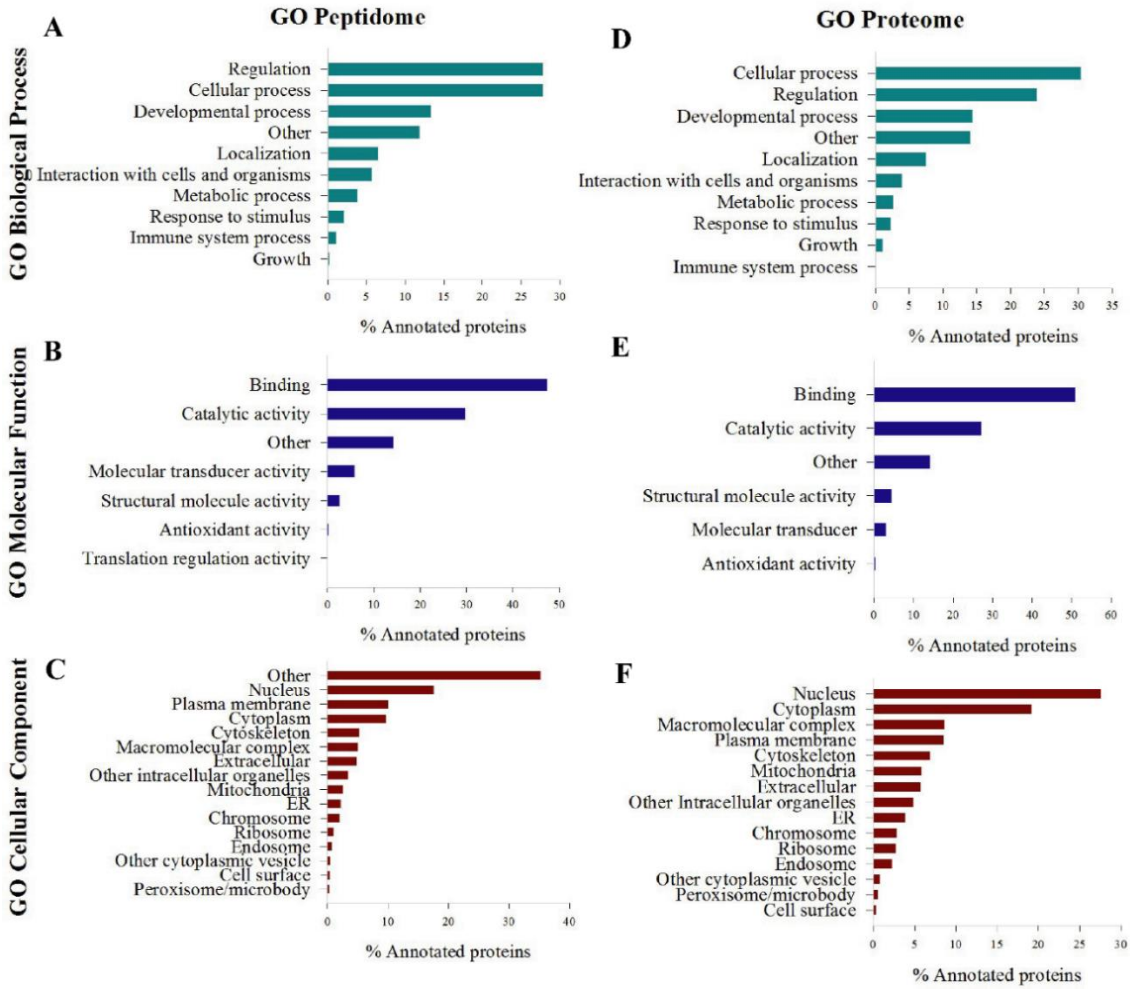

Figure 5.2. Gene ontology (GO) comparison of peptidome (7009 peptides) and proteome (422 proteins) identified in the cephalic secretion of *Arapaima gigas*. A, B and C. Peptidome GO for biological process, molecular function ad cellular component, respectively. D, E and F. Proteome GO for biological process, molecular function ad cellular component, respectively. Analyses were conducted in STRAP *v. 1.5* (Bhatia et al., 2009).

5.3.2. Protein analysis on the cephalic secretion of *Arapaima gigas*

The total protein content was not significantly different among M, F, PM and PF (one-way ANOVA; P>0.05). Total protein ranged from 0.25 to 4.29 µg/µL (1.80 ± 1.35 µg/µL; n = 6) in F, from 0.60 to 16.26 µg/µL (6.23 ± 6.29 µg/µL; n = 5) in PF, from 0.39 to 6.88 µg/µL (1.71 ± 2.55 µg/µL; n = 6) in M and from 1.82 to 7.89 µg/µL (4.08 ± 2.59 µg/µL; n = 5) in PM. In total, 422 proteins were identified combining all studied groups. A complete list that includes protein names, family, species, molecular weight (MW), scores, matches, mean emPAI values, frequency is catalogued in Table S5.2. Aiming to compare the number of proteins unique and shared between groups, a Venn diagram was generated (Fig 5.3A). Parental males had 35 exclusive proteins whereas parental females had 45 with both groups sharing other 35 proteins at a parental care condition. Males had 21 exclusive proteins and females 23 with both groups sharing 18 proteins. A total of 111 proteins were common to all groups.

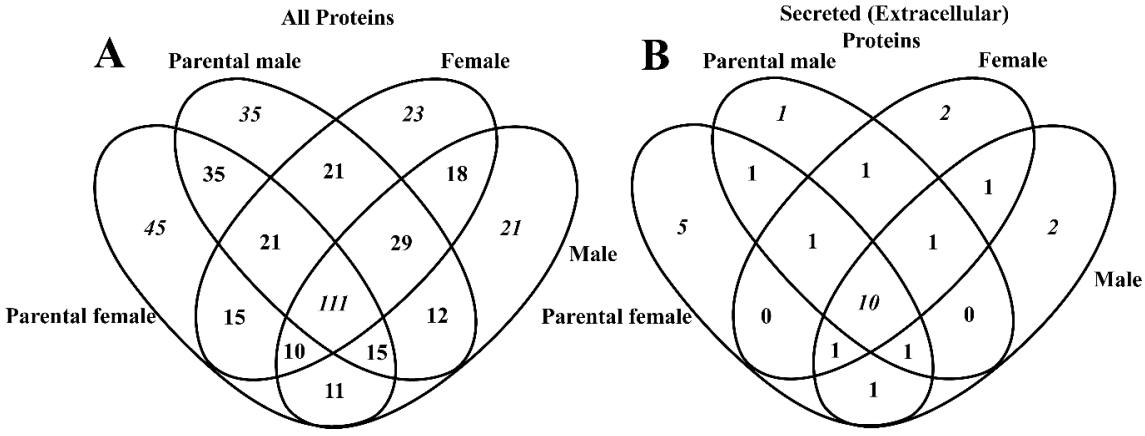

Figure 5.3. Number of unique and shared proteins in the cephalic fluid of *Arapaima gigas* comparing parental (PM) and non-parental males (M), and parental (PF) and non-parental females (F). A. Total of 422 proteins catalogued after GeLC-MS/MS and the 28 secreted (extracellular) proteins (B) revealed after GO analysis. Venn diagrams were produced in Venny (Oliveros, 2007).

GO analyses were conducted to classify the complete catalogue with 422 proteins, resulting in very similar pattern compared to the peptidome GO analysis (Fig 5.2). Based on annotations obtained for biological processes, most of the identified proteins are associated with cellular processes (194; 30.5 %), regulation (147; 23.1 %), developmental process (97; 15.3 %), localization (46; 7.2 %), interaction with cell and organisms (24; 3.8 %) and others (92; 14.5%) (Fig 5.2D). Considering molecular functions, proteins were mainly associated with binding (245; 51.8 %), catalytic activity (116; 24.5 %), structural molecule activity (22; 4.7 %), molecular transducer (12; 2.5 %), antioxidant activity (2; 0.4 %) and others (76; 16.1 %) (Fig 5.2E). Considering cellular component GO analysis, most of the proteins belonged to the nucleus (133; 22.9 %), cytoplasm (96; 16.5 %), macromolecular complex (46, 7.9 %), plasma membrane (36, 6.2 %), mitochondria (24, 4.1 %), extracellular (28, 4.8 %) and others (86; 14.8 %) (Fig 5.2F).

5.3.2.1. Secreted (extracellular) proteins in the cephalic fluid of *Arapaima gigas*

GO analysis revealed 28 secreted (extracellular) proteins in the different groups (Fig 5.3B), which are catalogued in Table 5.1. Their putative functions in the biology of *A. gigas* were retrieved from available literature on teleost species and emPAI values used as a measure of relative quantification between studied groups. Two hormones were detected in PF: *Prolactin* (PRL) and *Stanniocalcin* (STC). In all studied groups, proteins related to immunological system processes were identified (Table 5.1). Among them, proteins with known antibacterial roles and other defence proteins are listed. *Serotransferrin* (TF) was found to be up-regulated in PM (3.8-fold in relation to M group), *α -1-antitrypsin homolog* was detected in all groups, *Apolipoprotein A-I* (apoA-I) was present in F, PM and PF, *Complement C3* (C3) and *Complement component C8 beta chain* (C8B) were both present in all groups. The remaining secreted proteins are

putatively involved in growth (*Fibroblast growth factor 3* found in PF and *Growth/differentiation factor 6-A* in PM) brain regulation and development (different *Wnt* proteins), embryonic development in parental and non-parental fish (*Chordin*, *Olfactomedin-like protein 3B, Laminin subunit gamma-1*) and others.

Table 5.1. List of 28 secreted proteins (extracellular) present in the proteome of *Arapaima gigas*. Putative functions, detected groups and relative measure of concentration (emPAI) are given for pools of males (M), females (F), parental males (PM) and parental females (PF).

UniProt n°	Name	Hormones	Detected group	emPAI	Reference
O93337	Prolactin	Several roles in parental fish including mucous production and growth	PF	0.16	(Ogawa, 1970; Khong et al., 2009; Whittington & Wilson, 2013)
Q08264	Stanniocalcin	Prevention of hypercalcemia	PF	0.13	(Bonga & Pang, 1991; Wagner et al., 1992)
		Immune system			
P84122	Thrombin	Functions in blood homeostasis, inflammation and wound healing	M	1.22	(Manseth et al., 2004)
P79819	Serotransferrin	Role in stimulating cell proliferation. Known to forbid bacterial colonization in fish	M, F, PM, PF	0.14, 0.15, 0.53, 0.09	(Røed et al., 1995; Denovan-Wright et al., 1996)
P80426	Serotransferrin-1	Role in stimulating cell proliferation. Known to forbid bacterial colonization in fish	M, F, PM	0.25, 0.29, 0.18	(Røed et al., 1995; Denovan-Wright et al., 1996)
P80429	Serotransferrin-2	Role in stimulating cell proliferation. Known to forbid bacterial colonization in fish	PM, PF	0.21, 0.21	(Røed et al., 1995; Denovan-Wright et al., 1996)

Table 5.1. (Continued).

UniProt n°	Name	Immune system	Detected group	emPAI	Reference
P32759	α-1-antitrypsin homolog	Identified in carp perimeningeal fluid	M, F, PM, PF	0.14, 0.09, 0.19, 0.09	(Huang et al., 1995; Iq & Shu-Chien, 2011)
Q8JFG3	Tumor necrosis factor	Important mediator in resistance against parasitic, bacterial and viral infections	M, F, PM, PF	0.13, 0.13, 0.13, 0.14	(García-Castillo et al., 2002)
O42363	Apolipoprotein A-I	Participates in the reverse transport of cholesterol from tissues to the liver	F, PM, PF	0.13, 0.13, 0.13	(Babin et al., 1997)
Q92079	Serotransferrin	Role in stimulating cell proliferation. Known to forbid bacterial colonization in fish	M, F, PM, PF	0.11, 0.13, 0.11, 0.07	(Røed et al., 1995; Denovan-Wright et al., 1996)
P98093	Complement C3	Central role in the activation of the complement system	M, F, PM, PF	0.05, 0.07, 0.07, 0.07	(Lambris et al., 1993)
Q9PVW7	Complement component C8 beta chain	Play a key role in the innate and adaptive immune response	M, F, PM, PF	0.06, 0.06, 0.06, 0.06	(Katagiri et al., 1999)
Q3B7P7	Ubiquitin-60S ribosomal protein L40	Ubiquitin A-52 residue ribosomal protein	M, F, PM, PF	–	–
Q9W686	Semaphorin-3ab	Influence pathway choice of extending motor axons along development	M, F, PM, PF	0.04, 0.04, 0.04, 0.04	(Torres-Vázquez et al., 2004)

Table 5.1. (Continued).

UniProt n°	Name	Growth	Detected group	emPAI	Reference
P48802	Fibroblast growth factor 3	Regulation of cell proliferation, differentiation and embryonic development	PF	0.13	(Léger & Brand, 2002)
P85857	Growth/differentiation factor 6-A	Growth factor that controls proliferation and cellular differentiation in the retina	PM	0.08	(Sidi et al., 2003)
		Brain regulation/development			
P24257	Protein Wnt-1	Involved in neurogenesis	F	0.09	(McFarland et al., 2008)
P47793	Protein Wnt-4a	Probable brain developmental protein	M, PF, F	0.09, 0.09, 0.09	(Toyama et al., 2013)’
P43446	Protein Wnt-10a	Signalling molecule important in CNS development	M, PF	0.07, 0.07	(Kelly et al., 1993)
		Embryonic development			
Q0P3W2	Olfactomedin-like protein 3B	Secreted scaffold protein with essential role in dorsoventral patterning in early development	PF	0.08	(Zeng et al., 2005)
O57472	Chordin	Developmental protein, dorsalizing factor, somitogenesis	M, F, PM, PF	0.03, 0.03, 0.03, 0.04	(Miller-Bertoglio et al., 1997)
Q1LVF0	Laminin subunit gamma-1	Mediate attachment, migration and organization of cells into tissues in embryonic development	M, PF, PM	0.02, 0.02, 0.03	(Van Eeden et al., 1996)

Table 5.1. (Continued).

UniProt n°	Name	Others	Detected group	emPAI	Reference
Q6NWB6	Unique cartilage matrix-associated protein	Control of osteogenic differentiation	M	0.24	(Kessels et al., 2014)
B9TQX1	Unique cartilage matrix-associated protein	Control of osteogenic differentiation	PF	0.23	(Viegas et al., 2008)
B0JZP3	Protein THEM6	Thioesterase superfamily member 6	M, F, PM, PF	0.16, 0.16, 0.22, 0.16	(Yoshida et al., 2014)
Q7T297	Protein FAM172A		F	0.08	
O93484	Collagen alpha-2(I) chain	Type I collagen is a member of group I collagen (fibrillar forming collagen)	F, M	0.05, 0.03	(Saito et al., 2001)
Q90X49	Coiled-coil domain-containing protein 80	Promotes cell adhesion and matrix assembly	F, PM	0.04, 0.04	(Noce et al., 2015)

5.4. Discussion

The lateral line of *A. gigas* is an open system like in most teleosts, contrasting closed systems present in coryphaenoids, notopterids and mormyrids (Fänge et al., 1972; Hilton, 2003). On the head surface of *A. gigas*, many pore-bearing sculptures can be easily recognized, and these are the external openings of the interconnected cephalic canals composing the anterior lateral line system. Internally, the canals bear neuromast organs responsible for sensing mechanical stimuli through displacements in the cephalic fluid. Being an open system, water inflow and fluid outflow occur within these canals (see Coombs et al. (2014) for revision). The biochemical composition of the fluids in contact with the mechanoreceptors of the lateral line vary from one species to another (Fänge et al., 1972), and to the best of our knowledge, there has been no proteome investigation on this type of biological fluid published to date. In *A. gigas,* this fluid is secreted from the head into the water and its release is intensified during parental care (Lüling, 1964), when the secretion remains within range of developing eggs, larvae and offspring. Given this behavioural strategy, the cephalic secretion would appear to have significant roles in the life-history of the species. Previous authors have suggested that offspring of *A. gigas* would feed from this secreted fluid (Lüling, 1964; Noakes, 1979; Holbrook, 2011). This comparable to several cichlid parents that feed their fry with mucous which provides nutrients, hormones and passive immunity (Schütz & Barlow, 1997; Melamed et al., 2005; Buckley et al., 2010).

The cephalic fluid of *A. gigas* has a complex biological origin. It involves a fluid interaction with the endolymph from the inner-ear, which in this study was supported by the presence of the peptide *pcdh15b,* and with the CSF, supported herein by the presence of peptides related to synapse assembly and neurotransmitter secretion (*i.e. otud4*, *ribeye a protein*, *tjp1b* and *syn1*). Also, the canals are composed by the integument, skull bones

and cartilage, which has been shown to be highly vascularized and innervated in *A. gigas* and also containing integumentary glands potentially secreting peptides and proteins into the cephalic fluid (Lüling, 1964; Noakes, 1979). Ultimately, there is also an influx of environmental water into the cephalic canals since the system is opened and environmental water is necessary to supply the mechanical function of the neuromasts.

Given this multiple tissue interactions and the opened nature of the cephalic canal system, studies on the cephalic fluid of *A. gigas* are technically difficult. For the giant *A. gigas* it is also difficult to obtain precise information on fish age, maturity and reproductive condition (Torati et al., 2016), and these taken together explain the large variability in peptide number and total protein content observed in the cephalic secretions in the present study. Investigation of the parental care condition in *A. gigas* is also challenging. Collection of samples in the wild is prohibitive since the species is endangered (Castello & Stewart, 2010) and captive reproduction is still the main bottleneck for aquaculture development, hindering further investigations on the possible roles of the cephalic fluid in parental care (Núñez et al., 2011). In this study, two couples reproduced in the research station allowing the collection of valuable samples from parental fish for the first time in the species.

After CE-MS analysis, significant differences in peptide abundance among the studied groups were detected, and surprisingly, a higher number of peptides were found in PF compared to F. This suggests a marked physiological change in PF reflected at the peptide level in the secretion, especially given that females are regarded to play a minimal role in parental care after the nest guarding phase (Fontenele, 1953). Given that peptides are processed by olfaction and are widely used by fish in several chemical communication processes (Døving & Lastein, 2009; Wyatt, 2010), this data suggests PF of *A. gigas* could use peptides as pheromones and/or signature mixtures to convey information to

conspecifics (*i.e.* alarm signals, food-rich locations, among others), displaying roles going beyond behavioural observations. In this study, several candidate peptides and proteins were catalogued that will support future comparative studies involving chemical communication and behaviour in parental *A. gigas* and other teleosts.

Since the collection of cephalic fluid samples is a relatively non-invasive method, and the fluid has biochemical components from the lateral line system, inner ear endolymph, blood plasma, CSF and skin mucus, application of CE-MS would potentially be a suitable method for health diagnosis as already applied in humans and other mammals if sample collection could be standardised (Mansor et al., 2013; Pontillo et al., 2015). Other applications could include the search for gender and sexual maturity biomarkers which are considered as critical problems for the captive reproduction of *A. gigas* (Núñez et al., 2011), or markers for behavioural investigations. In fish, studies to evaluate physiological condition normally focus on mucus samples (Cordero et al., 2015; Cordero et al., 2016) which carry important immune components as being the first body barrier against pathogens and parasites. The mucus and skin are also sites for alarm signals and odorants in teleosts (Døving & Lastein, 2009). The present study on *A. gigas* showed that the cephalic fluid not only carried immune components, but also proteins found in the CSF (*otud4, ribeye a*, *tjp1b* and *syn1*), inner ear endolymph (*pcdh15b*) and eyes (*tetraspanin* and *ppp2r3a*).

In the present study, a total of 422 proteins were identified from the cephalic secretion of parental and non-parental *A. gigas*. From these, GO analysis revealed 28 extracellular products of secretion. The pituitary hormone *prolactin* (PRL) was detected exclusively in PF. PRL is considered a key hormone that has been detected previously in mucus of several teleost females exhibiting parental care behaviour (Schütz & Barlow, 1997; Khong et al., 2009). It also agrees with the increased mucin production seen in PF.

The detection of PRL only in PF agrees with a pattern observed in teleosts displaying parental care. In the Midas cichlid (*Cichlasoma citrinellum*), PRL can be actively transferred into the young through mucus feeding, increasing fry growth (Schütz & Barlow, 1997). In the Amazon Discus fish (*Symphysodon aequifasciata*), PRL is also present and upregulated in the mucus of parental females (Khong et al., 2009). Other studies on teleosts showed the association of PRL with female reproductive behaviour such as nest building, mouth brooding and parental care (Whittington & Wilson, 2013). Given that PRL is produced by the pituitary, the likely routes into the cephalic secretion are through the CSF or alternatively, PRL, as a mucus component, could be derived from the circulatory system as suggested in other teleosts.

Stanniocalcin (STC) is a potent hypocalcemic or antihypercalcemic from bony fishes which in this study was detected in PF. STC is synthesized and secreted from the kidney Stannius corpuscles (Wagner et al., 1992), so its route into the cephalic fluid is very likely through the circulatory system. After spawning in *A. gigas*, females undergo a period of rapid ovarian growth leading to repeat spawning during the same reproductive period, since the inter-spawning period appears to be approximately one month in the species (*pers. observation*). The presence of STC in PF may reflect oocyte recruitment during which the liver produces vitellogenin under the influence of oestrogens, which is transported into the ovary and incorporated into the growing oocytes (Lubzens et al., 2010). This process causes an elevation of plasma levels of protein-bound calcium used as a carrier in the blood circulation, so the production of SCT during parental care could be a response to these high protein-bound calcium levels which need to be counterbalanced during this phase, as suggested in teleosts (Bonga & Pang, 1991). Whether these changes would impact larvae and fry development remains to be investigated.

Being an open system, the cephalic canals of *A. gigas* are a vulnerable door for pathogens and parasites, which could significantly impact adult health. Parasitism and infections in the inner ear and lateral line are indeed serious diseases for many teleosts (Barber, 2007). Therefore, the presence of immune-competent proteins to defence against bacteria, viruses and parasites would be expected in the cephalic secretion, and their presence during parental care could be enhanced. Out of the 28 secreted proteins detected, 12 have immunological functions described in teleosts, some of which have been related to parental care. From these, four *Serotransferrin* (TF) species were identified with two present in all studied groups and one with higher concentration (emPAI-based) in PM. TF is known to suppress iron and therefore prevents bacterial colonization, playing also an important role in macrophage activation (Stafford et al., 2001). The importance of TF in immunological defence has been reported in the mucus of gilthead seabream (*Sparus aurata*) (Jurado et al., 2015; Sanahuja & Ibarz, 2015), seabass (*Dicentrarchus labrax*) (Cordero et al., 2015), and also within the internal pouch where male seahorses (*Hippocampus* spp.) incubate their fry during early pregnancy (Melamed et al., 2005). Therefore, the presence of TF in the cephalic secretion indicates protective functions similar to the role of mucus in teleosts, and the increase observed in PM suggests potential roles for developing offspring of *A. gigas*, with mechanisms already studied in other parental teleosts.

Another immune active protein detected in the cephalic fluid of all studied groups was the glycoprotein *α-1-antitrypsin homolog* (A1AT). A1AT belongs to the serpin family of glycoproteins, involved in the control of blood coagulation, fibrinolysis, complement activation, and inflammation processes (Huang et al., 1995). This protein has a marked protective function in carp (*Cyprinus carpio*) perimeningeal fluid (Huang et al., 1995), and is upregulated in the mouth mucus where eggs are incubated in parental

Tilapia females (*Oreochromis* spp.) (Iq & Shu-Chien, 2011). A similar protective role for developing fry and offspring of *A. gigas* could be at work, although further studies are required to test this hypothesis. Similarly, *apolipoprotein A-I* (apoA-I) was also detected in the cephalic fluid of PM and PF. ApoA-I has antibacterial properties previously demonstrated *in vitro* in the striped bass (*Morone saxatilis*) (Johnston et al., 2008) and its activity against pathogens has been suggested in the mucus of gilthead seabream (Sanahuja & Ibarz, 2015; Cordero et al., 2016), seabass (Cordero et al., 2015) and Atlantic cod (*Gadus morhua*) (Rajan et al., 2011). Two other proteins involved in complement activation were also detected in all groups: c*omplement C3* (C3) and *complement component* C8 (C8B). The complement activation system results in the formation of the membrane attack complex (MAC), which kills bacteria by disrupting their membranes (Katagiri et al., 1999). C3 has a central role in phagocytic and immunoregulatory processes (Lambris et al., 1993). The presence of C3 has been shown in the mucus of seabass (Cordero et al., 2015) and C3 is upregulated in seabream after probiotic intake (Cordero et al., 2016). The importance of such a varied number of immune proteins in immunological response of adult *A. gigas* seems consistent as these were detected in all studied groups.

This study detected two secreted proteins in parental fish related to growth regulation: *fibroblast growth factor 3* (FGF-3) in PF and *growth/differentiation factor 6-A* (GDF6) in PM. FGF-3 is a protein required for inner-ear induction, patterning and maintenance as demonstrated in early larval stages (Léger & Brand, 2002), whilst GDF6 plays a role in later eye development in *D. rerio* (Asai-Coakwell et al., 2013). Therefore, the presence of these proteins in PM and PF likely reflect inner-ear and eye metabolisms, and the absence in M and F groups suggests metabolism changes during parental care which could have association with offspring development in *A. gigas*. However, several

of these detected proteins have known functions in developing embryos. The presence of these proteins in parental care groups therefore remains to be understood. Future studies aimed at the study of the CSF in teleosts should provide a better understanding of their origin and possible roles within the cephalic secretion. This is the case also for several proteins previously reported to be expressed exclusively in developing embryos. These include reported roles in somitogenesis (*i.e. Chordin*), mesoderm segmentation (*i.e. Wnt inhibitory factor 1*) and embryonic brain development (Protein Wnt-10b). If these proteins cannot be produced and secreted from the brain of adult parents, as the literature suggests, an alternative parsimonious explanation would involve the embryonic production followed by a later inflow into the open cephalic canals of the adults during parental care. This is plausible since during nest guarding, parents remain with their head constantly inside the nest for egg mass fanning and guarding (Fontenele, 1953) during which surrounding embryo proteins could be absorbed by the open cephalic canals, although further behavioural studies are needed. Alternatively, the existence of mouthbrooding in *A. gigas* could explain the presence of embryonic proteins in the cephalic fluids. Mouth incubation and transport of eggs and larvae has been reported in *A. gigas* (Farias et al., 2015), although not systematically. Therefore, if not mouthbrooding, this study suggests a very close interaction among parent's head with brood along parental care, and this is in strong agreement with behavioural observations available for the species.

In conclusion, for the first time, this investigation applied proteomic and peptidomic techniques to survey cephalic fluids from the anterior lateral line of a teleost species. Analyses in *A. gigas* showed sample variability, and a proteomic composition influenced by components from the cephalic canals, inner-ear endolymph, CSF and circulatory system. This study enhances information on the biochemistry of the lateral

line system, opening research possibilities in fish physiology and chemical communication. Data from this study highlights the complex role that the cephalic secretion of *A. gigas* may play not only in the adults but also on the development of the fingerlings. Previous work suggested fingerlings raised under parental care condition would have higher survival rates and enhanced growth performance compared to indoor reared ones (Rojas, 2005). Although the present study does not confirm such observations, it does indicate the importance of parental care strategies not only on survival *per se* but also on fingerling condition which could be positively improved by being in contact with the cephalic secretion.

Supporting information

Table S5.1. List of peptides. Peptides identified with CE-MS and LC-MS/MS and comparisons among studied groups.

Peptide ID	CE time (min)	Peptide mass (Da)	Sequence peptide	UniProt assession n°	Protein identity	Primary Gene Name	Difference between groups	Wilcox P-value
376	22.99	894.53	YKRKSTL	E9QI87	Uncharacterized protein	zgc:162971	PF>F	0.039
652	28.78	956.54	LDAVVSPRT				PF>F	0.039
689	28.85	964.50	KANANLTSF				PF>F	0.008
725	29.05	974.55	LPLQDVYK	Q7ZWA1	Elongation factor 1-alpha (Fragment)	si:dkey-37o8.1	PF>F	0.039
784	36.13	986.54	LTAQLSQLN	E7FC83	Uncharacterized protein	LOC101886519	PF>F	0.039
879	35.85	1006.55	TISTSTVSVL	F8QXP5	Rank (Fragment)	tnfrsf1b	PF>F	0.039
887	28.78	1008.51	EAFSAICRL				PF>F	0.039
896	23.45	1010.52	KGPQKYGSF	X1WH84	Uncharacterized protein	ptprnb	PF>F	0.039
940	29.04	1019.56	TIKCVICQL	E7F8D5	Uncharacterized protein	ldb3a	PF>F	0.039
985	29.99	1030.57	DVNSAIATIK	Q6NWJ5	Tubulin alpha 6	tuba8l4	PF>F	0.039
1073	29.54	1053.61	VPEEIKQVI	B0V153	Methylenetetrahydrofolate reductase	mthfr	PF>F	0.039
1194	30.07	1086.59	KDQISPAISL	X1WEW9	Uncharacterized protein	si:ch211-167j9.4	PF>F	0.037
1266	30.29	1104.60	ISSASPVKSTL	F1QNB3	Uncharacterized protein (Fragment)	ank3b	PF>F	0.008
1472	39.34	1160.60	ELTAVPFVNGV				PF>F	0.039
1488	21.62	1163.74	IIKALKEPPR	Q5BKW5	Ribosomal protein L12	rpl12	PF>F	0.039
1543	29.52	1177.64	KYLGINSEGVV				PF>F	0.039
1565	26.62	1183.60	TVPPLAANHQH				PF>F	0.039
1809	31.17	1244.67	TRTPEELTILG	E7F798	Chloride channel protein	clcn6	PF>F	0.039
1885	30.75	1262.66	LSENTSAMIIGK	F1QCY7	Uncharacterized protein (Fragment)	pak7	PF>F	0.039
1900	31.48	1266.70	LEGGDIPLQGLR	F1QBE7	Uncharacterized protein (Fragment)	sorbs1	PF>F	0.039
1901	26.43	1266.69	KSLEIIDFFR	E7F541	Uncharacterized protein	atg9b	PF>F	0.039
2006	31.46	1290.70	TQTVTLDMAAKI				PF>F	0.039

Table S5.1. List of peptides. (Continued).

Peptide ID	CE time (min)	Peptide mass (Da)	Sequence peptide	UniProt assession n°	Protein identity	Primary Gene Name	Difference between groups	Wilcox P-value
2277	31.71	1355.70	EIDIQASDKGIAP	E7FF57	Uncharacterized protein	LOC100535010	PF>F	0.039
2653	32.38	1442.80	LSQRENLLFLNP	E7F8M6	Uncharacterized protein		PF>F	0.008
2719	32.40	1459.76	TEALMALTAAPSLR				PF>F	0.039
2749	33.17	1466.84	VEENANVVVRLLI	F1QEX9	Uncharacterized protein (Fragment)	ryr1b	PF>F	0.039
3021	26.75	1535.81	LKGEIDPHIALDLC	Q6NYI7	Zgc:77563	pla2g7	PF>F	0.039
3044	33.09	1542.86	KPSWTSALPLVCIL	B1N746	Interleukin 15 receptor alpha isoform 11	IL-15RA	PF>F	0.039
3125	33.20	1564.79	ESAHPPAFVDLQAAL	E9QFI8	Uncharacterized protein	si:dkey-83m22.7	PF>F	0.039
3237	53.12	1593.79	LPTTTVSPTMTTVMV	X1WG18	Uncharacterized protein	muc5.2	PF>F	0.039
3251	33.91	1594.85	SPGTIEIVTVTQVVH				PF>F	0.039
3332	26.99	1616.81	MEKPNVRWSDVAGL	Q08BZ6	Vacuolar protein sorting 4a (Yeast)	vps4a	PF>F	0.039
3359	33.54	1624.92	IGGIGTVPVGRVETGTL	A2VCX2	Elongation factor 1-alpha	eef1a1l2	PF>F	0.046
3399	52.98	1635.77	PAQQQQQPQSQPQQ	U3JAU2	Uncharacterized protein	tnrc6b	PF>F	0.039
3441	33.33	1645.74	PGRPGEEGPQGPPGEPG	Q8JGL8	Type IX collagen alpha 2	col9a2	PF>F	0.039
3476	33.19	1654.84	SPAPSSPTPRPPSPPPA	E9QG28	Uncharacterized protein	ppp2r3a	PF>F	0.039
3484	53.24	1656.70	NCGEPLEAVGSAYFGQ				PF>F	0.039
3493	33.46	1658.94	DLPALIKDLVYTVAL	Q1LUJ4	Uncharacterized protein	fuca1.2	PF>F	0.039
3513	33.49	1664.81	QGKPGAPGEPGLAGVPGE	Q1L863	Uncharacterized protein (Fragment)	col8a1b	PF>F	0.026
3515	42.14	1664.84	AEAVAAPAEAVAAPPEAV	F1QJI9	Histone-lysine N-methyltransferase	setd2	PF>F	0.039
3529	33.28	1668.84	VRSQGPAVSPQPISQS	B0UXI4	Uncharacterized protein	syn1	PF>F	0.039
3544	33.30	1672.88	CKVSIIPPTALNSIDS	F1RC67	Diacylglycerol kinase (Fragment)	dgkh	PF>F	0.039
3568	33.67	1678.90	TPTPLPSGADLVVPQR	E0YA53	Calcineurin binding protein 1	cabin1	PF>F	0.008
3812	34.12	1747.94	TPGVKGPAGIPGVQGPVGL	E7FE47	Uncharacterized protein	col5a2b	PF>F	0.008
3910	34.11	1775.01	LPPNLLALFAPRDPIP	Q503W8	Snrnp70 protein (Fragment)	snrnp70	PF>F	0.025

Table S5.1. List of peptides. (Continued).

Peptide ID	CE time (min)	Peptide mass (Da)	Sequence peptide	UniProt assession n°	Protein identity	Primary Gene Name	Difference between groups	Wilcox P-value
3918	33.83	1776.88	KPNLNAGATPTSYVIEA	E7FDW5	Uncharacterized protein	robo1	PF>F	0.039
3947	34.44	1783.94	LSGTPLPSLQSLSLCPR	E7FE22	Uncharacterized protein	inpp5ja	PF>F	0.039
3965	28.30	1787.91	RSLDTGSDDINIIKNQ	F1QM62	Uncharacterized protein (Fragment)	snx13	PF>F	0.039
3976	19.01	1789.99	GAHQNIIPASTGAAKAVGK	F1R3D3	Glyceraldehyde-3-phosphate dehydrogenase	gapdhs	PF>F	0.008
3987	34.04	1792.93	PSPSVPPHPAPPPPPPLP	F1QVE2	Uncharacterized protein (Fragment)	pcdh15b	PF>F	0.046
3998	34.06	1796.94	DKIIESSQPEVEVIPV				PF>F	0.039
4012	33.92	1799.88	DNPPVAFKSPVPVGIAE	I3IT71	Uncharacterized protein (Fragment)	zgc:152948	PF>F	0.016
4072	28.21	1817.89	PSVEVFFQSVKSSSMK				PF>F	0.011
4115	34.18	1831.95	KPLPTPIGPIPTDTPQP	A7MC91	Csnk1g2b protein	csnk1g2b	PF>F	0.039
4293	28.73	1882.00	KPAPVRYEEPPPPPPPV	X1WHN9	Uncharacterized protein	tjp1b	PF>F	0.034
4420	34.40	1921.03	LAAARALELAPPIPADSTP				PF>F	0.039
4544	28.97	1963.99	EEGKNAINAPLLHTSPEL	X1WCE0	Uncharacterized protein	parvb	PF>F	0.013
4585	42.51	1978.00	QPVASPNPIYTPVPPVAPA	Q5BU19	Ribeye a protein	ctbp2a	PF>F	0.039
4661	34.88	2001.98	QSKDLCLSGTPSSAAAPAVSP	E7FE43	Uncharacterized protein	lrfn2b	PF>F	0.039
4765	35.31	2037.98	KAVFVMVNPQTESSASAAGE	F1QS95	Uncharacterized protein	prx	PF>F	0.039
4934	29.28	2092.10	LKNIGNNFFKAQNWQSAI	Q6IQL7	Peptidylprolyl isomerase D (Cyclophilin D)	ppid	PF>F	0.039
4942	35.00	2095.00	SPGQTPLGGPHTLSSPPTSLT	E7FFV5	Uncharacterized protein	si:ch211-114m9.1	PF>F	0.039
5100	33.83	2148.01	EFSDTLFSTLASHQPVAWP	F1QKC3	Uncharacterized protein (Fragment)	mlxip	PF>F	0.008

Table S5.1. List of peptides. (Continued).

Peptide ID	CE time (min)	Peptide mass (Da)	Sequence peptide	UniProt assession n°	Protein identity	Primary Gene Name	Difference between groups	Wilcox P-value
5132	30.15	2155.15	KFKDLDLTNLYAPPPPPLT	A5D6U6	Akap1b protein	akap1b	PF>F	0.039
5212	29.87	2176.20	IALKGINRVLAQDPMGVPPGM	E7F2F7	Histone-lysine N-methyltransferase	kmt2d	PF>F	0.039
5213	25.66	2176.14	ADEINVWREKDGRLSIYI				PF>F	0.039
5343	25.01	2216.18	KGLKGFQGPPGPPGTRGLPGPPG	X1WCX6	Uncharacterized protein	si:ch211-196i2.2	PF>F	0.039
5565	30.27	2277.21	ALPAAASGPGAPRTPPMPKPPSPP	E7F092	Uncharacterized protein	cdk13	PF>F	0.046
5590	25.63	2286.20	RLSTASSVTKPPHFILTTEW	B5DDZ3	Parp12a protein	parp12a	PF>F	0.039
6032	31.71	2473.35	PAENIIAIAATNNLYIFQDRVK				PF>F	0.008
6317	28.73	2602.44	VKNISTQIPTILGAAESPGAVVLHVG	A9JRX8	LOC797572 protein	zgc:77327	PF>F	0.039
6703	24.82	2768.49	HVPRAVFVDLEPTVIDEVRTGTYR	Q6NWJ5	Tubulin alpha 6	tuba8l4	PF>F	0.039
1045	29.76	1044.58	QTLVDGKVVS	B2GQG8	Krt18 protein	krt18	F>PF	0.019
1479	30.36	1161.61	KQFLDNLGVE	Q5TYQ3	Cytochrome c oxidase assembly factor 7	coa7	F>PF	0.038
1766	31.13	1234.62	RLWLTSYPSP	E7FGS8	Uncharacterized protein	dnah12	F>PF	0.009
6258	26.60	2568.33	HSLPSLVTPISNKQRMSILVMAD				F>PF	0.047
921	29.53	1014.58	DVNAAIATIK	B8A516	Uncharacterized protein	tuba1b	F>M	0.049
946	29.39	1021.53	KDGNPGLPGPA	F1Q593	Uncharacterized protein	si:dkey-27a13.3	F>M	0.039
989	29.54	1031.55	KADTTSTVPI	F1QUP3	Uncharacterized protein	ehmt1b	F>M	0.049

Table S5.1. List of peptides. (Continued).

Peptide ID	CE time (min)	Peptide mass (Da)	Sequence peptide	UniProt assession n°	Protein identity	Primary Gene Name	Difference between groups	Wilcox P-value
1045	29.76	1044.58	QTLVDGKVVS	B2GQG8	Krt18 protein	krt18	F>M	0.026
1076	29.92	1054.60	LALIGKTADGP	R4GEK7	Uncharacterized protein	spint2	F>M	0.049
1104	39.39	1061.53	SPDEIPFII	A4FVN1	Hmha1 protein (Fragment)	hmha1b	F>M	0.049
1697	24.69	1214.69	TPTSDNKIVLK	Q1LUA6	Triple functional domain protein	trio	F>M	0.039
2465	32.16	1396.69	LSGGALGHVTSNTPS	E7FC98	Uncharacterized protein		F>M	0.040
2548	51.29	1417.60	PAGPTGQGSELDGGSG	E7FBN5	Uncharacterized protein		F>M	0.028
2685	32.94	1449.87	APAVVPVVVPITRT	F1R9S8	Uncharacterized protein	ktn1	F>M	0.049
3618	27.85	1692.93	LLPQLPPEGRDRALS	E7F9F7	Sodium channel protein		F>M	0.031
4518	34.36	1953.96	EPEQESLRPPAGLSWLN	I3IT38	Uncharacterized protein	march7	F>M	0.002
4890	29.71	2079.07	QASGVPHIPVQNEVFARAQ	R4GDU6	Uncharacterized protein	si:ch211-207i1.2	F>M	0.049
107	27.57	830.45	VIGDSGVGK	G1K2X3	Rab-like protein 3	rabl3	M>F	0.019
170	27.68	845.45	ARPVMSVS	A1A5I3	LOC557935 protein (Fragment)	zmp:0000001082	M>F	0.019
424	23.10	905.44	FEFAQKH	E9QFS4	Uncharacterized protein	zgc:113274	M>F	0.031
860	29.05	1002.53	ERDTLVSAL	E7FAN6	Kinesin-like protein	kif20ba	M>F	0.031
904	23.41	1011.60	KVESLRAPL	Q5TYP2	Uncharacterized protein	zp3a.2	M>F	0.007
940	29.04	1019.56	TIKCVICQL	E7F8D5	Uncharacterized protein	ldb3a	M>F	0.019
1025	30.65	1041.49	AMVGYCGTLK	Q7T2G0	Tetraspanin	tspan12	M>F	0.006
1108	29.52	1062.48	TQDEGDAIKS	Q8QGB8	DNA (cytosine-5)-methyltransferase	dnmt1	M>F	0.031
1541	24.98	1176.67	AELYLEALKK	A9JTH4	Zgc:175226 protein	slc26a6l	M>F	0.003
1560	30.62	1181.60	SIGAGASVSLSHP	E7FCA0	Uncharacterized protein	snphb	M>F	0.019
1791	31.26	1239.61	ETYLIVGCSQK				M>F	0.007

Table S5.1. List of peptides. (Continued).

Peptide ID	CE time (min)	Peptide mass (Da)	Sequence peptide	UniProt assession n°	Protein identity	Primary Gene Name	Difference between groups	Wilcox P-value
2318	25.62	1361.78	ISAVPVKTQPSVH	Q1LYF7	Novel protein similar to vertebrate cell adhesion molecule-related/down-regulated by oncogenes (CDON)	DKEY-79N12.1-001	M>F	0.028
2717	26.31	1458.73	ENRFTVRPNVDP	E7F418	Uncharacterized protein	msh5	M>F	0.019
2719	32.40	1459.76	TEALMALTAAPSLR				M>F	0.019
2829	26.25	1487.79	EMINRELQISQK	X1WGW7	Uncharacterized protein	ints1	M>F	0.019
2956	33.00	1519.83	PPPLPVQKAVGSTPT	Q501U6	Zgc:113078	sh3d19	M>F	0.019
3024	32.89	1536.79	DDIVNLQNPPLRQ	F2Z4W6	Uncharacterized protein	LOC799523	M>F	0.019
3239	33.18	1593.79	LAVDGRWVLYGTMGG	E7FB99	Uncharacterized protein	LOC796473	M>F	0.042
3316	27.29	1610.85	ATPTRKIASMPAPAAP	Q4KMK8	POU domain protein	pou2f1a	M>F	0.019
3613	27.33	1691.85	LTVYARDGGSPPNFAK				M>F	0.019
3695	33.50	1716.89	ICSIPGLPGAPGKPGPPGA	E7F3N3	Uncharacterized protein	CABZ01048675.1	M>F	0.044
3765	33.91	1736.88	SIALAAEDPVWPESPR	E9QJF6	Uncharacterized protein	cacna1aa	M>F	0.030
3878	33.79	1766.84	PMATPSLPLSPAPDRDA				M>F	0.016
4629	34.62	1992.00	PAADFFLVEKPSISPQNT	Q0P4D6	Transcription elongation factor, mitochondrial	tefm	M>F	0.025
4650	29.08	1998.06	QVEDTISKALNTLVNEPK	E7F3Q4	Uncharacterized protein	wu:fc34e06	M>F	0.019
5212	29.87	2176.20	IALKGINRVLAQDPMGVPPGM	E7F2F7	Histone-lysine N-methyltransferase	kmt2d	M>F	0.019
6191	30.11	2539.15	DEEAPTQIVPDITKTDYYHSTT	A5X6X5	Titin a	ttna	M>F	0.049
6962	27.75	2905.52	NKGVQPLLDAITAYLPAPNERNHDLV	A0JMI9	Ribosome-releasing factor 2, mitochondrial	gfm2	M>F	0.025
1431	30.26	1152.63	KELYVFGGLQ	E7FFD1	Uncharacterized protein	zgc:163014	PM>M	0.045

Table S5.1. List of peptides. (Continued).

Peptide ID	CE time (min)	Peptide mass (Da)	Sequence peptide	UniProt assession n°	Protein identity	Primary Gene Name	Difference between groups	Wilcox P-value
2045	31.79	1301.69	PTPSFPKSPALL	I3ISD8	Uncharacterized protein	ahdc1	PM>M	0.007
2465	32.16	1396.69	LSGGALGHVTSNTPS	E7FC98	Uncharacterized protein		PM>M	0.045
3066	19.06	1546.83	KSKGSVENSDKPKK				PM>M	0.045
3173	21.12	1576.71	IMGSSRHRMTPDGT	F1R7R3	Uncharacterized protein	hmcn1	PM>M	0.023
1025	30.65	1041.49	AMVGYCGTLK	Q7T2G0	Tetraspanin	tspan12	M>PM	0.026
1045	29.76	1044.58	QTLVDGKVVS	B2GQG8	Krt18 protein	krt18	M>PM	0.032
1541	24.98	1176.67	AELYLEALKK	A9JTH4	Zgc:175226 protein	slc26a6l	M>PM	0.026
1561	39.75	1182.59	TIGFGDYVALQ	E7F9D1	Uncharacterized protein	kcnk3b	M>PM	0.037
2031	25.42	1298.75	KRSALPLSISDI	E7F348	Uncharacterized protein	vwa5b1	M>PM	0.032
2238	31.85	1345.73	YTQGPNLDLVVK				M>PM	0.032
2475	32.52	1398.68	EAAGPSDLGGVPWK	F1QH17	Protein-methionine sulfoxide oxidase	mical3a	M>PM	0.032
3128	33.25	1565.88	VPPIRDPLSEGLII	F1R8Z8	Uncharacterized protein (Fragment)	ncor2	M>PM	0.007
3370	33.48	1627.87	RIFNIIADPPSPVAS	E7F372	Uncharacterized protein	rasa1b	M>PM	0.032
3423	33.54	1640.81	TVTDPVHDIAFAPNL	I3ISK9	Nucleoporin seh1 (Fragment)	seh1l	M>PM	0.037
3822	34.02	1749.91	KPPAEPSPLLLNVAEQ	F1Q897	Uncharacterized protein (Fragment)	magi2	M>PM	0.032
6246	26.93	2562.26	QQPSANHSTPSAPAPPTFDRSLKP	F1QW91	Uncharacterized protein	stambpb	M>PM	0.032
847	29.14	999.54	RAVSVSTPAP	F8W3X0	Uncharacterized protein (Fragment)	otud4	PF>PM	0.025
1045	29.76	1044.58	QTLVDGKVVS	B2GQG8	Krt18 protein	krt18	PF>PM	0.025
3370	33.48	1627.87	RIFNIIADPPSPVAS	E7F372	Uncharacterized protein	rasa1b	PF>PM	0.025
4012	33.92	1799.88	DNPPVAFKSPVPVGIAE	I3IT71	Uncharacterized protein (Fragment)	zgc:152948	PF>PM	0.025
4072	28.21	1817.89	PSVEVFFQSVKSSSMK				PF>PM	0.007
4293	28.73	1882.00	KPAPVRYEEPPPPPPPV	X1WHN9	Uncharacterized protein	tjp1b	PF>PM	0.025
6451	27.08	2646.45	TGVAEVPPGTPNVLRASHLLNTLYK	F1QBQ4	Uncharacterized protein	tubgcp5	PF>PM	0.045
2045	31.79	1301.69	PTPSFPKSPALL	I3ISD8	Uncharacterized protein	ahdc1	PM>PF	0.036

Table S5.2. List of proteins. Information on the 422 proteins identified through GeLC-MS/MS in all studied groups.

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
A0JMD2	Lateral signaling target protein 2 homolog (Zinc finger FYVE domain-containing protein 28)	109.2	22	8(1)	22.14	1	M
A0JMF1	CST complex subunit CTC1 (Conserved telomere maintenance component 1)	136.7	30	14(2)	30.27	5	PF, F, M
A0JMK5	Myotubularin-related protein 2 (Phosphatidylinositol-3,5-bisphosphate 3-phosphatase) (EC 3.1.3.95) (Phosphatidylinositol-3-phosphate phosphatase) (EC 3.1.3.64)	71.7	24	5(2)	23.30	2	PM, F
A0JML8	DALR anticodon-binding domain-containing protein 3	62.7	19	3(2)			PF, M
A0JPF5	Protein FAM160B2 (RAI16-like protein) (Retinoic acid-induced protein 16)	84.4	34	15(4)	32.21	28	PF, PM, F, M
A0MGZ7	Heparan-sulfate 6-O-sulfotransferase 3 (HS 6-OST-3) (EC 2.8.2.-)	50.1	23	19(3)			PM, F
A1A5V7	Transmembrane protein 41B	31.5	28	5(3)	27.96	3	PF, F
A1L1F1	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex assembly factor 3	19.3	20	9(2)			PF, F
A1L1F4	Sister chromatid cohesion protein PDS5 homolog A	150.3	18	23(2)			PF, PM, F, M
A1L1R6	Zinc finger protein 423	155.7	20	3(1)			PF, PM
A1L252	Ran-binding protein 9 (RanBP9)	66.4	18	2(1)			PF, PM
A2BFP5	Solute carrier family 12 member 9	99.7	22	15(3)			PF, PM, F, M
A2BGA0	Transcription factor RFX4 (Regulatory factor X 4)	82.8	20	17(1)	20.01	2	PF, PM
A2BGP7	Coiled-coil domain-containing protein 125	54.5	24	12(1)	24.28	2	PF, PM
A2BGT0	Ubiquitin carboxyl-terminal hydrolase 30 (EC 3.4.19.12) (Deubiquitinating enzyme 30) (Ubiquitin thioesterase 30) (Ubiquitin-specific-processing protease 30) (Ub-specific protease 30)	56.3	23	15(2)	22.49	5	PF, PM

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
A2BGU8	Neuronal-specific septin-3	40.6	18	11(2)			PF
A2BID5	Ventricular zone-expressed PH domain-containing protein (Protein melted homolog)	92.3	21	17(1)	21.33	3	PM
A2BIL7	Tyrosine-protein kinase BAZ1B (EC 2.7.10.2) (Bromodomain adjacent to zinc finger domain protein 1B) (Williams syndrome transcription factor homolog)	177.2	21	66(1)	21.19	2	PM, F
A2T928	Retinoic acid receptor gamma-B (RAR-gamma-B) (Nuclear receptor subfamily 1 group B member 3-B)	56.4	31	7(1)	31.24	15	PM, F, M
A3KNA5	Filamin A-interacting protein 1-like	130.3	20	29(2)			PF, F, M
A3KP77	Oxidoreductase NAD-binding domain-containing protein 1 (EC 1.-.-.-)	31.1	18	24(1)			PM, F, M
A3KQI3	Protein MEF2BNB (MEF2B neighbor gene protein homolog) (mef2b neighbor gene transcript)	14.1	20	3(1)			PM
A4IGF3	Mitochondrial inner membrane protease ATP23 homolog (EC 3.4.24.-)	29.5	24	1(1)	24.01	1	PM, F
A4QNR8	Leukocyte receptor cluster member 8 homolog	93.4	19	9(1)			PF, F, M
A5PLK2	Phospholysine phosphohistidine inorganic pyrophosphate phosphatase (EC 3.1.3.-) (EC 3.6.1.1)	29.9	18	11(2)			PF
A5WUT8	Ectopic P granules protein 5 homolog	286.6	27	55(3)			PF
A7UA95	Ras-associating and dilute domain-containing protein	126.5	19	12(1)			F, M
A8E7C5	Phosphatidylinositide phosphatase SAC2 (EC 3.1.3.-) (Inositol polyphosphate 5-phosphatase F) (Sac domain-containing inositol phosphatase 2) (Sac domain-containing phosphoinositide 5-phosphatase 2)	127.1	18	10(2)			PF, PM, M

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
A8E7G4	Protein FAM45A (Protein FAM45)	40.2	19	15(3)			PM, F, M
A8WE67	KIF1-binding protein homolog	73.6	21	2(1)	20.73	1	M
A8Y5U1	UPF0600 protein C5orf51 homolog		18	4(1)			PF
A9C3W3	Kazrin-A	89.7	20	7(2)			PM
A9LMC0	Zinc finger protein DPF3	45.8	32	15(4)	31.87	18	PM, F, M
B0JZP3	Protein THEM6	23.7	29	8(3)	22.87	5	PF, PM, F, M
B0R0I6	Chromodomain-helicase-DNA-binding protein 8 (CHD-8) (EC 3.6.4.12) (ATP-dependent helicase CHD8)	281.5	30	144(8)	28.96	16	PF, PM, F, M
B0R0T1	von Willebrand factor A domain-containing protein 8	213.2	23	24(1)	22.60	2	F, M
B0S6R1	Nck-associated protein 1 (NAP 1)	130.1	25	18(8)			PF, PM, F, M
B0V3F8	Interferon-inducible double-stranded RNA-dependent protein kinase activator A homolog	31.1	18	13(1)			PM, F, M
B3DK56	Receptor-type tyrosine-protein phosphatase U (R-PTP-U) (EC 3.1.3.48) (Receptor-type protein-tyrosine phosphatase psi) (R-PTP-psi)	164.7	19	3(1)			PM
B5XDD3	Partner of Y14 and mago A (Protein wibg homolog A)	22.4	31	10(6)			PF, PM, M
B5XG19	Partner of Y14 and mago B (Protein wibg homolog B)	21.2	20	6(2)	22.50	2	PF, M
B7ZC77	Catenin alpha-2 (Alpha N-catenin)	96.5	19	17(1)			PM
B8A4F4	Putative ATP-dependent RNA helicase TDRD9 (EC 3.6.4.13) (Tudor domain-containing protein 9)	152.1	84	44(19)	83.67	145	PF, PM, F, M
B9TQX1	Unique cartilage matrix-associated protein [Cleaved into: Unique cartilage matrix-associated protein C-terminal fragment (Ucma-C) (Gla-rich protein) (GRP)]	17.0	19	2(1)			PF
C6KFA3	G-protein coupled receptor 126	132.6	20	9(1)			F

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
O12990	Tyrosine-protein kinase JAK1 (EC 2.7.10.2) (Janus kinase 1) (JAK-1)	134.3	29	33(2)	28.28	6	PM, F, M
O13008	Fatty acid-binding protein, heart (Fatty acid-binding protein 3) (Heart-type fatty acid-binding protein) (H-FABP)	14.5	22	3(1)	22.10	1	M
O42099	Mitogen-activated protein kinase 8B (MAP kinase 8B) (MAPK 8B) (EC 2.7.11.24) (Stress-activated protein kinase JNKb) (c-Jun N-terminal kinase B)	48.7	46	15(8)			PM, F
O42356	Retinal homeobox protein Rx1	37.1	33	7(4)			PM
O42357	Retinal homeobox protein Rx2	37.1	22	8(3)	30.34	4	PF, PM, F, M
O42363	Apolipoprotein A-I (Apo-AI) (ApoA-I) (Apolipoprotein A1) [Cleaved into: Proapolipoprotein A-I (ProapoA-I)]	30.2	21	7(2)	25.18	4	PF, PM, F
O42368	Homeobox protein Hox-B3a (Hox-B3)	44.8	18	3(1)			F
O42467	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily B member 1	42.8	21	6(2)			PF, PM, F
O57472	Chordin (Protein chordino)	107.5	21	6(2)			PF, PM, F, M
O57656	Glycerol-3-phosphate dehydrogenase [NAD(+)], cytoplasmic (GPD-C) (GPDH-C) (EC 1.1.1.8)	38.7	20	5(1)			PM
O93267	Trypsinogen-like protein 3	29.2	19	2(2)			M
O93337	Prolactin (PRL)	23.3	18	3(1)			PF
O93484	Collagen alpha-2(I) chain (Alpha-2 type I collagen)	127.4	19	16(3)			F, M
O93542	L-lactate dehydrogenase A chain (LDH-A) (EC 1.1.1.27)	36.3	222	7(6)			PF, PM, F, M
P02016	Hemoglobin subunit alpha (Alpha-globin) (Hemoglobin alpha chain)	15.4	1585	131(112)			PF, PM, F, M
P02017	Hemoglobin subunit alpha (Alpha-globin) (Hemoglobin alpha chain)	15.5	1422	172(117)	1510.55	1149	PM, F, M

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
P02019	Hemoglobin subunit alpha-1 (Alpha-1-globin) (Hemoglobin alpha-1 chain)	15.3	24	10(1)			PF, PM, F, M
P02139	Hemoglobin subunit beta-A/B (Hemoglobin beta-A/B chain) (Hemoglobin subunit beta-1/2)	16.4	3175	211(176)	3100.49	1492	PF, PM, F, M
P02141	Hemoglobin subunit beta-4 (Beta-4-globin) (Hemoglobin beta-4 chain) (Hemoglobin beta-IV chain)	16.1	527	69(49)			PF
P02204	Myoglobin	15.8	47	64(10)	44.64	46	PF, PM, F, M
P0C218	Probable ATP-dependent RNA helicase DDX20 (EC 3.6.4.13) (DEAD box protein 20)	85.6	20	12(2)			PF, F
P0C237	Hemoglobin subunit alpha (Alpha-globin) (Hemoglobin alpha chain)	15.7	48	11(5)			F, M
P0C928	Regulator of telomere elongation helicase 1 (EC 3.6.4.12)	134.1	18	8(1)			F, M
P10777	Hemoglobin subunit alpha-1 (Alpha-1-globin) (Hemoglobin alpha-1 chain)	15.5	48	10(5)			PF, PM, F, M
P12115	Adenylate kinase isoenzyme 1 (AK 1) (EC 2.7.4.3) (EC 2.7.4.6) (ATP-AMP transphosphorylase 1) (ATP:AMP phosphotransferase) (Adenylate monophosphate kinase) (Myokinase)	21.5	25	6(2)	25.47	2	PF
P13104	Tropomyosin alpha-1 chain (Alpha-tropomyosin) (Tropomyosin-1)	32.8	113	35(15)	111.53	133	PF, PM, F, M
P14520	Hemoglobin subunit alpha (Alpha-globin) (Hemoglobin alpha chain)	15.3	78	15(8)	74.20	81	PF, PM, F, M
P14527	Hemoglobin subunit alpha-4 (Alpha-4-globin) (Hemoglobin alpha-4 chain)	15.9	133	5(3)	133.39	3	PF, F
P16058	Estrogen receptor (ER) (ER-alpha) (Estradiol receptor) (Nuclear receptor subfamily 3 group A member 1)	69.0	31	4(2)	30.98	2	PF, PM
P18520	Intermediate filament protein ON3	57.8	415	40(24)	400.53	163	PF, PM, F, M

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
P19180	Ig heavy chain V region 3	13.3	18	6(1)			PM, F, M
P19181	Ig heavy chain V region 5A	13.0	268	9(7)	311.09	60	PF, PM, F, M
P19618	NADPH--cytochrome P450 reductase (CPR) (P450R) (EC 1.6.2.4) (Fragments)	68.7	20	5(1)	20.38	1	PF, M
P20373	L-lactate dehydrogenase B chain (LDH-B) (EC 1.1.1.27)	36.2	222	8(6)			PF, PM, F
P24257	Protein Wnt-1	42.4	18	5(1)			F,
P25127	Metallothionein (MT)	7.1	321	40(25)	452.05	118	PF, PM, F, M
P26325	Alcohol dehydrogenase 1 (EC 1.1.1.1)	40.9	19	6(1)			M
P30436	Tubulin alpha chain	49.3	46	5(3)	45.97	8	PF, PM
P31579	Low choriolytic enzyme (EC 3.4.24.66) (Choriolysin L) (Hatching enzyme zinc-protease subunit LCE)	31.3	21	24(1)	20.93	5	PF
P32759	Alpha-1-antitrypsin homolog	42.0	35	10(2)			PF, PM, F, M
P34205	Deoxyribodipyrimidine photo-lyase (EC 4.1.99.3) (DNA photolyase) (Photoreactivating enzyme)	63.8	26	8(1)			PF
P43446	Protein Wnt-10a	50.0	19	7(1)			PF, M
P47793	Protein Wnt-4a	40.8	22	11(3)			PF, F, M
P48672	Vimentin A2 (Fragment)	28.1	20	8(1)			M
P48802	Fibroblast growth factor 3 (FGF-3) (Heparin-binding growth factor 3) (HBGF-3)	29.1	19	7(1)			PF
P49696	Valine--tRNA ligase (EC 6.1.1.9) (Valyl-tRNA synthetase) (ValRS)	139.8	19	18(2)			PF,
P49947	Ferritin, middle subunit (Ferritin M) (EC 1.16.3.1)	20.7	49	4(3)	49.48	10	PF, PM, F, M
P51112	Huntingtin (Huntington disease protein homolog) (HD protein homolog)	352.7	23	37(3)			PF, PM, F, M

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
P51958	Cyclin-dependent kinase 1 (CDK1) (EC 2.7.11.22) (EC 2.7.11.23) (Cell division control protein 2 homolog) (Cell division protein kinase 1) (p34 protein kinase)	34.6	20	24(1)			F
P53483	Actin, alpha anomalous	42.4	300	27(19)	347.18	208	PF, PM, F, M
P56250	Hemoglobin subunit alpha (Alpha-globin) (Hemoglobin alpha chain)	15.7	66	8(5)			PF, PM, F, M
P58312	Sodium/potassium-transporting ATPase subunit alpha-3 (Na(+)/K(+) ATPase alpha-3 subunit) (EC 3.6.3.9) (Na(+)/K(+) ATPase alpha(III) subunit) (Sodium pump subunit alpha-3)	112.9	22	13(2)			PF, PM, M
P59015	Vacuolar protein sorting-associated protein 18 homolog	114.8	18	6(1)			PM, F
P61207	ADP-ribosylation factor 3	20.6	104	4(2)			PF, PM, F, M
P62797	Histone H4	11.4	202	9(8)			PF, PM, F, M
P68199	Ubiquitin-40S ribosomal protein S27a	8.6	194	11(8)			PF, PM, F, M
P68200	Ubiquitin-40S ribosomal protein S27a (Ubiquitin carboxyl extension protein 80) [Cleaved into: Ubiquitin; 40S ribosomal protein S27a]	8.6	66	5(3)			PM
P70066	40S ribosomal protein S15 (RIG protein)	17.0	18	7(2)			M
P70083	Sarcoplasmic/endoplasmic reticulum calcium ATPase 1 (SERCA1) (SR Ca(2+)-ATPase 1) (EC 3.6.3.8) (Calcium pump 1) (Calcium-transporting ATPase sarcoplasmic reticulum type, fast twitch skeletal muscle isoform) (Endoplasmic reticulum class 1/2 Ca(2+) ATPase)	110.7	19	18(1)			PF, PM
P79819	Serotransferrin	76.4	894	141(90)	882.58	785	PF, PM, F, M
P80270	Hemoglobin subunit alpha (Alpha-globin) (Hemoglobin alpha chain)	15.8	1565	132(111)	1512.94	1158	PF, PM, F, M
P80426	Serotransferrin-1 (Serotransferrin I) (STF I) (sTF1) (Siderophilin I)	76.4	959	74(44)	946.58	517	PM, F, M
P80429	Serotransferrin-2 (Serotransferrin II) (STF II) (sTF2) (Siderophilin II)	76.5	959	74(44)			PF, PM
P80945	Hemoglobin anodic subunit alpha (Hemoglobin anodic alpha chain)	16.1	66	13(5)			PF, PM, F, M

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
P82159	Myosin light chain 1, skeletal muscle isoform (LC-1) (LC1) (Myosin light chain alkali 1) (Myosin light chain A1)	20.1	30	2(2)			PF, PM, F
P82160	Myosin light chain 3, skeletal muscle isoform (LC-3) (LC3) (Myosin light chain alkali 2) (Myosin light chain A2)	16.4	30	2(2)	30.12	4	F
P82990	Hemoglobin subunit alpha (Alpha-globin) (Hemoglobin alpha chain)	15.6	66	10(5)			PF, PM, F
P83613	Hemoglobin subunit alpha-2 (Alpha-2-globin) (Hemoglobin alpha-2 chain)	15.7	21	5(1)			PF, PM, F
P84122	Thrombin (EC 3.4.21.5) [Cleaved into: Thrombin light chain; Thrombin heavy chain] (Fragments)	3.8	18	1(1)			M
P84206	Hemoglobin anodic subunit beta (Hemoglobin anodic beta chain)	16.7	643	78(61)	610.18	653	PF, PM, F, M
P84611	Hemoglobin subunit beta-2 (Beta-2-globin) (Hemoglobin beta-2 chain)	16.8	643	75(61)	610.18	657	PF, PM, F, M
P85001	Centrosomal protein of 290 kDa (Cep290)	283.8	28	40(2)	27.50	12	PF, PM, F, M
P85281	Nucleoside diphosphate kinase B (NDK B) (NDP kinase B) (EC 2.7.4.6) (Fragments)	14.4	38	5(2)	37.48	3	M
P85312	Hemoglobin subunit beta-A (Beta-A-globin) (Hemoglobin beta-A chain) (Fragment)	4.8	83	8(7)	82.53	31	PF, PM, F, M
P85836	Ferritin, liver middle subunit (Ferritin M) (EC 1.16.3.1)	20.5	93	6(4)	92.52	29	PF, PM, F, M
P85838	Ferritin, heavy subunit (Ferritin H) (EC 1.16.3.1)	20.6	65	4(2)	65.19	10	PM, F, M
P85857	Growth/differentiation factor 6-A (GDF-6-A) (Growth differentiation factor 6A) (Protein radar)	46.9	18	16(1)			PM

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
P98093	Complement C3 [Cleaved into: Complement C3 beta chain; Complement C3 alpha chain; C3a anaphylatoxin; Complement C3b alpha' chain; Complement C3c alpha' chain fragment 1; Complement C3dg fragment; Complement C3g fragment; Complement C3d fragment; Complement C3f fragment; Complement C3c alpha' chain fragment 2] (Fragment)	183.6	702	32(27)	700.33	237	PF, PM, F, M
Q00IB7	Pleckstrin homology domain-containing family H member 1 (PH domain-containing family H member 1) (Protein max-1 homolog)	161.7	62	41(10)	62.28	19	PF, PM, F, M
Q03357	Homeobox protein MSH-A	29.3	18	4(2)			PF, PM
Q06176	L-lactate dehydrogenase C chain (LDH-C) (EC 1.1.1.27)	36.6	222	8(6)			PF, PM, F
Q07217	Cholesterol side-chain cleavage enzyme, mitochondrial (EC 1.14.15.6) (CYPXIA1) (Cholesterol desmolase) (Cytochrome P450 11A1) (Cytochrome P450(scc))	59.4	18	19(2)			PF, PM, F, M
Q07342	Forkhead box protein A2 (Axial protein)	45.3	19	10(2)			PF, F, M
Q08264	Stanniocalcin (STC) (Corpuscles of Stannius protein) (CS) (Hypocalcin) (Teleocalcin)	28.1	34	3(1)			PF
Q08BL3	Glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase 1-A (EC 2.4.1.122) (Core 1 O-glycan T-synthase A) (Core 1 UDP-galactose:N-acetylgalactosamine-alpha-R beta 1,3-galactosyltransferase 1-A) (Core 1 beta1,3-galactosyltransferase 1-A) (C1GalT1-A) (Core 1 beta3-Gal-T1-A)	47.1	21	17(3)			PM, M
Q08BR4	Histone-lysine N-methyltransferase SETDB1-B (EC 2.1.1.43) (SET domain bifurcated 1B)	137.5	18	7(2)			PF, F

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q08BV2	tRNA wybutosine-synthesizing protein 5 (EC 1.14.11.42) (tRNA(Phe) (7-(3-amino-3-carboxypropyl)wyosine(37)-C(2))-hydroxylase)	38.3	21	5(2)			PM
Q08C72	Pre-mRNA-splicing factor CWC22 homolog (Nucampholin homolog)	113.9	30	17(7)	29.76	10	PF, PM, F, M
Q08CF3	Coiled-coil domain-containing protein 61	57.7	21	7(2)			PM, M
Q08CH3	MPN domain-containing protein (EC 3.4.-.-)	51.5	28	18(4)			PM, F, M
Q08CH7	Radial spoke head 10 homolog B	84.0	24	4(2)	23.66	2	PF, PM, F, M
Q08CL8	CCR4-NOT transcription complex subunit 10	68.2	20	5(2)			PF
Q0P3W2	Olfactomedin-like protein 3B	45.2	18	4(1)			PF
Q0P485	Coiled-coil domain-containing protein 85C-A	44.4	29	5(3)			PF, PM
Q0P4C5	UPF0602 protein C4orf47 homolog	34.9	19	9(1)			PF, M
Q0V967	F-box only protein 5	44.1	18	5(1)			PF
Q1ECV4	Tubulin monoglycylase TTLL3 (EC 6.3.2.-) (Tubulin--tyrosine ligase-like protein 3)	90.1	18	9(1)			F, M
Q1ECX4	Serine/threonine-protein kinase tousled-like 2 (EC 2.7.11.1) (PKU-alpha) (Tousled-like kinase 2)	79.9	18	5(1)			F, M
Q1ECZ4	Protein CASC3 (Cancer susceptibility candidate gene 3 protein homolog) (Metastatic lymph node protein 51 homolog) (DrMLN51) (Protein MLN 51 homolog)	82.8	24	12(3)			PF, PM, F
Q1JPT7	THAP domain-containing protein 1	26.5	29	22(8)	27.39	24	PF, PM, F, M
Q1KKR9	Protein lunapark-B	40.5	22	11(1)	22.37	2	F
Q1KKT1	Homeobox protein Hox-D11a	30.7	40	9(4)	35.52	7	PF, PM, F, M
Q1KKU7	Homeobox protein Hox-C5a	23.8	18	2(1)			PM

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q1KKZ5	Homeobox protein Hox-A11b	31.7	26	7(2)	25.47	2	PF, F
Q1KL12	Homeobox protein Hox-A3a	45.3	18	8(1)			F
Q1L8G6	Ankyrin repeat and IBR domain-containing protein 1	121.5	22	8(1)	22.47	1	PM, F
Q1L994	Protein phosphatase 1 regulatory subunit 37 (Leucine-rich repeat-containing protein 68)	101.1	18	14(2)			PM, F, M
Q1L9A2	Optic atrophy 3 protein homolog	17.4	22	7(2)	21.29	2	PM, M
Q1LUS5	UPF0711 protein C18orf21 homolog	23.5	23	9(1)	23.34	2	PF, F
Q1LV19	A-kinase anchor protein SPHKAP (SPHK1-interactor and AKAP domain-containing protein)	176.9	23	8(1)	23.48	2	PF, PM, F, M
Q1LVF0	Laminin subunit gamma-1	176.1	23	22(2)			PF, PM, M
Q1LVK9	NEDD4-binding protein 1 (N4BP1)	96.9	24	45(4)			PF, PM
Q1LVQ2	SH3 domain-binding protein 4-A	109.6	22	7(1)	21.77	1	M
Q1LWG3	39S ribosomal protein L36, mitochondrial (L36mt) (MRP-L36)	13.3	20	4(1)	20.27	1	PM, M
Q1LWM5	Uncharacterized protein KIAA1522 homolog	146.0	33	16(2)	33.25	3	PF, PM
Q1LX59	Cholesterol 25-hydroxylase-like protein 1, member 2 (EC 1.14.99.-)	32.3	18	1(1)			F
Q1LXZ7	IQ motif-containing protein H	120.3	19	15(1)			PF
Q1LY84	UPF0524 protein C3orf70 homolog A	30.3	18	4(1)			PM
Q1LYG0	SH3 and PX domain-containing protein 2A	124.8	18	17(1)			PF
Q1LYM3	Protein spire homolog 1	86.9	19	9(3)			PF, F, M
Q1LYM6	Kelch-like protein 38	68.0	25	27(2)	24.36	3	PM, F
Q1MTI4	Triosephosphate isomerase A (TIM-A) (EC 5.3.1.1) (Triose-phosphate isomerase A)	27.2	664	18(16)	659.38	138	PF, PM, F, M
Q1RLX4	Trafficking protein particle complex subunit 11 (Protein foie gras)	130.3	18	3(1)			PM, M

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q1RM03	Trichoplein keratin filament-binding protein (Protein TCHP)	62.3	20	15(2)			PM
Q1RMA6	PCNA-interacting partner (PARI) (PARP-1 binding protein) (PARP1-binding protein) (PARPBP)	62.4	22	6(1)	21.72	1	PF, M
Q20JQ7	Sodium channel protein type 4 subunit alpha B (Voltage-gated sodium channel subunit alpha Nav1.4b)	203.9	21	16(2)			PF, PM, F, M
Q2KN94	Cytospin-A (SPECC1-like protein) (Sperm antigen with calponin homology and coiled-coil domains 1-like)	123.7	23	13(2)	22.62	2	PF, PM, F, M
Q2KNE5	Melanopsin-A (Mammalian-like melanopsin) (Melanopsin-M) (Opsin-4-A) (Opsin-4M)	66.2	32	222(7)			PF, PM, F, M
Q2PW47	Mediator of RNA polymerase II transcription subunit 24 (Mediator complex subunit 24) (Protein lessen) (Thyroid hormone receptor-associated protein 4 homolog) (Trap100 homolog)	112.4	19	13(2)			M
Q2QCI8	Mediator of RNA polymerase II transcription subunit 12 (Mediator complex subunit 12) (Protein kohtalo) (Protein motionless) (Thyroid hormone receptor-associated protein complex 230 kDa component) (Trap230)	246.3	19	35(1)			PF, PM
Q2YDS1	DNA damage-binding protein 2 (Damage-specific DNA-binding protein 2)	56.1	361	21(20)	359.55	130	PF, PM, F, M
Q32PN7	Zinc finger protein DZIP1L (DAZ-interacting protein 1-like protein)	86.7	42	14(8)	40.67	28	PF, PM, F, M
Q32PW3	Transcription initiation factor TFIID subunit 2 (TBP-associated factor 150 kDa) (Transcription initiation factor TFIID 150 kDa subunit) (TAF(II)150) (TAFII-150) (TAFII150)	137.8	23	10(1)	23.14	1	PM
Q3LGD4	Rab11 family-interacting protein 4A (FIP4-Rab11) (Rab11-FIP4-A) (zRab11-FIP4-A)	71.6	23	2(1)	23.24	3	PF, PM

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q3ZB90	Kelch repeat and BTB domain-containing protein 12 (Kelch domain-containing protein 6)	71.1	19	13(2)			PF, PM
Q499B3	KAT8 regulatory NSL complex subunit 3 (NSL complex protein NSL3) (Non-specific lethal 3 homolog)	90.2	18	8(1)			PM
Q49GP3	Phosphatidylinositol 4-kinase beta (PI4K-beta) (PI4Kbeta) (PtdIns 4-kinase beta) (EC 2.7.1.67)	94.6	20	8(2)			PM, F
Q4H4B6	Protein scribble homolog (Scribble1)	190.2	21	22(3)			PM, M
Q4KMC8	Fructose-bisphosphate aldolase C-A (EC 4.1.2.13) (Aldolase C-like) (Brain-type aldolase-A)	40.4	78	8(4)	76.73	33	PF, PM, F, M
Q4QRC6	Iron-sulfur cluster assembly 1 homolog, mitochondrial (HESB-like domain-containing protein 2) (Iron-sulfur assembly protein IscA)	14.4	28	10(2)	28.05	2	PM, F, M
Q4RSM6	Probable glutathione peroxidase 8 (GPx-8) (GSHPx-8) (EC 1.11.1.9)	24.0	19	5(1)			PM
Q4RWX9	Phosphotriesterase-related protein (EC 3.1.-.-) (Parathion hydrolase-related protein)	39.3	76	4(4)	75.70	4	PF, PM
Q4TVV3	Probable ATP-dependent RNA helicase DDX46 (EC 3.6.4.13) (DEAD box protein 46)	115.6	18	29(1)			F, M
Q4V8R6	Transcription factor E4F1 (EC 6.3.2.-) (Putative E3 ubiquitin-protein ligase E4F1) (Transcription factor E4F)	81.7	46	10(6)	41.35	16	PF, PM, F, M
Q4V8W7	Ubiquitin domain-containing protein 1	25.7	27	13(4)			PF, PM
Q4V921	Small integral membrane protein 19	11.5	23	1(1)	23.30	1	PM
Q4VBI7	Syntaxin-18	36.4	22	10(2)	21.28	4	PF, PM, M

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q502K3	Serine/threonine-protein phosphatase 6 regulatory ankyrin repeat subunit C (PP6-ARS-C) (Serine/threonine-protein phosphatase 6 regulatory subunit ARS-C)	116.5	116	34(12)	106.63	109	PF, PM, F, M
Q502M5	Enhancer of mRNA-decapping protein 3 (LSM16 protein homolog) (YjeF domain-containing protein 1)	55.4	34	19(3)	34.09	9	PF, PM, F, M
Q502M6	Ankyrin repeat domain-containing protein 29	32.3	23	7(2)			PF, M
Q50D79	Homeobox protein unc-4 homolog (Homeobox protein Uncx4.1)	52.2	18	3(1)			F
Q566Y8	Protein mago nashi homolog	17.3	19	5(2)			PF
Q567J8	Protein odd-skipped-related 2 (zOsr2)	27.1	27	5(4)	24.37	4	PF, PM, F, M
Q568D5	BRCA1-A complex subunit BRE (BRCA1/BRCA2-containing complex subunit 45) (Brain and reproductive organ-expressed protein)	44.4	19	29(1)			PF, PM, M
Q568E2	Zinc finger protein 750	68.8	27	10(3)			PF, PM, F, M
Q568K9	Protein FAM50A	40.6	18	9(1)			F
Q568R1	Nuclear speckle splicing regulatory protein 1 (Coiled-coil domain-containing protein 55)	61.2	21	22(1)	20.95	6	PF, PM
Q58EE9	Glial fibrillary acidic protein (GFAP)	51.3	102	20(11)	96.69	47	PF, PM, F, M
Q58EG3	Nectin-3-like protein (Poliovirus receptor-related protein 3-like)	64.4	23	7(1)	23.11	1	PF
Q58EL2	Transmembrane protein 35	17.9	41	1(1)	41.03	5	PF, F
Q5BJA5	Histone H2B 1/2	13.6	169	18(15)	167.35	101	PF, PM, F, M
Q5BJJ8	39S ribosomal protein L51, mitochondrial (L51mt) (MRP-L51)	14.8	18	3(1)			F

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q5BKW9	Isoaspartyl peptidase/L-asparaginase (EC 3.4.19.5) (EC 3.5.1.1) (Asparaginase-like protein 1) (Beta-aspartyl-peptidase) (Isoaspartyl dipeptidase) (L-asparagine amidohydrolase) [Cleaved into: Isoaspartyl peptidase/L-asparaginase alpha chain; Isoaspartyl peptidase/L-asparaginase beta chain]	33.2	20	5(1)	20.00	1	PM
Q5D013	Protein-lysine N-methyltransferase mettl10 (EC 2.1.1.-) (Methyltransferase-like protein 10)	54.3	24	4(2)			PF, PM, F
Q5D016	Prefoldin subunit 1	14.0	24	17(5)			PM, F, M
Q5EB20	Protein FAM65C	110.7	38	12(3)	37.54	9	PF, PM, F, M
Q5MJ86	Glyceraldehyde-3-phosphate dehydrogenase 2 (EC 1.2.1.12)	36.1	125	8(6)	123.76	35	PF, PM, F, M
Q5PRB3	PIH1 domain-containing protein 2	36.9	54	12(9)	47.12	13	PF, PM, F, M
Q5RG44	Uncharacterized protein KIAA1211 homolog	123.8	19	21(1)			PF, PM, F, M
Q5RGJ5	CWF19-like protein 1	62.3	23	10(2)	22.11	2	PF, PM, F
Q5RGJ8	N-acetylglucosamine-1-phosphotransferase subunits alpha/beta (EC 2.7.8.17) (GlcNAc-1-phosphotransferase subunits alpha/beta) (Stealth protein gnptab) (UDP-N-acetylglucosamine-1-phosphotransferase subunits alpha/beta) [Cleaved into: N-acetylglucosamine-1-phosphotransferase subunit alpha; N-acetylglucosamine-1-phosphotransferase subunit beta]	140.1	32	17(4)	31.55	10	PF, PM, F, M
Q5RID7	Sorting nexin-17	54.0	19	3(1)			M
Q5RKN4	Putative oxidoreductase GLYR1 (EC 1.-.-.-) (Glyoxylate reductase 1 homolog) (Nuclear protein NP60)	51.1	19	7(2)			F
Q5SNQ7	Protein SERAC1 (Serine active site-containing protein 1)	75.1	18	7(1)			PF

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q5SPP0	Clavesin-2 (Retinaldehyde-binding protein 1-like 2)	38.2	18	5(1)			F, M
Q5SPR8	N-acetyltransferase ESCO2 (EC 2.3.1.-) (Establishment of cohesion 1 homolog 2) (ECO1 homolog 2)	68.7	22	4(1)	22.37	1	F
Q5TZ07	Presqualene diphosphate phosphatase (EC 3.1.3.-) (Phosphatidic acid phosphatase type 2 domain-containing protein 2)	31.6	18	26(2)			PM, F, M
Q5TZ18	Neuron navigator 3	244.3	29	27(1)	29.05	4	PM, F
Q5TZ80	Protein Hook homolog 1	83.6	18	14(1)			PM, M
Q5U3F2	Pre-mRNA-splicing factor SLU7	66.7	18	14(1)			PM, F
Q5U3H2	Histone-lysine N-methyltransferase SUV420H1 (EC 2.1.1.43) (Suppressor of variegation 4-20 homolog 1) (Su(var)4-20 homolog 1) (Suv4-20h1)	91.6	19	11(1)			F
Q5U3H9	Protein spire homolog 2 (Spir-2)	69.1	46	7(3)	44.89	7	PF, PM, M
Q5U3I0	Sorting and assembly machinery component 50 homolog B	52.2	19	5(2)			PF
Q5U3U4	Cytidine and dCMP deaminase domain-containing protein 1	53.6	38	13(4)	35.52	20	PM, F, M
Q5U9X3	G-protein coupled receptor family C group 6 member A (Odorant receptor ZO6)	98.6	20	15(2)			PM, F
Q5XJ54	Glutaredoxin 3	36.5	25	2(2)	24.70	4	PF, PM, F
Q5XJA2	Coiled-coil domain-containing protein 149-A	58.8	29	10(4)			PF, PM, F
Q5XJQ7	Protein odd-skipped-related 1 (zOsr1)	30.1	27	5(4)			F
Q5XLR4	Borealin (Cell division cycle-associated protein 8) (Dasra-B) (DasraB) (DrDasraB)	30.5	19	2(1)			PM
Q66HV9	Leucine-rich repeat and immunoglobulin-like domain-containing nogo receptor-interacting protein 1-B	70.4	47	6(3)	46.55	4	PF, PM

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q696W0	Striated muscle preferentially expressed protein kinase (EC 2.7.11.1)	333.5	25	33(3)	23.16	9	PF, PM, F
Q6AZW2	Alcohol dehydrogenase [NADP(+)] A (EC 1.1.1.2) (Aldehyde reductase-A) (Aldo-keto reductase family 1 member A1-A)	37.1	121	9(2)	120.72	11	PF, F, M
Q6DBW0	Zinc finger protein Pegasus (Ikaros family zinc finger protein 5)	47.0	18	14(2)			F
Q6DC03	Polycomb protein suz12-A (Suppressor of zeste 12 protein homolog A)	75.8	20	5(1)			PF
Q6DC61	Meiotic nuclear division protein 1 homolog		20	2(1)			PF
Q6DC93	Peflin (PEF protein with a long N-terminal hydrophobic domain) (Penta-EF hand domain-containing protein 1)	29.5	21	4(1)	20.53	1	PM, F, M
Q6DG43	Ubiquitin domain-containing protein 2	26.7	27	15(4)	25.18	14	PF, PM
Q6DG88	Cysteine protease ATG4B (EC 3.4.22.-) (Autophagy-related protein 4 homolog B)	45.2	19	16(1)			F, M
Q6DH66	Mitochondrial import receptor subunit TOM20 homolog B (Mitochondrial 20 kDa outer membrane protein B) (Outer mitochondrial membrane receptor Tom20-B)	16.4	19	13(1)			PM, F
Q6DHI2	Coiled-coil domain-containing protein 65	58.1	47	8(4)	46.91	25	PF, PM, F, M
Q6DHL7	Coiled-coil domain-containing protein 85C-B	45.1	30	13(2)	29.63	7	PF, PM
Q6DHQ3	N-acyl-aromatic-L-amino acid amidohydrolase (carboxylate-forming) B (EC 3.5.1.114) (Aminoacylase-3.2) (ACY-3.2) (Aspartoacylase-2B)	35.6	22	12(2)			PM, F, M
Q6DHR3	Ras-GEF domain-containing family member 1B-A	60.4	18	6(1)			PF
Q6DHU8	Lymphokine-activated killer T-cell-originated protein kinase homolog (EC 2.7.12.2) (PDZ-binding kinase)	38.2	21	1(1)	21.48	1	PF
Q6DI40	Tetratricopeptide repeat protein 33 (TPR repeat protein 33)	30.9	36	18(5)	35.53	6	PF, PM, F, M
Q6DRC3	UPF0769 protein C21orf59 homolog	33.7	19	10(1)			F

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q6DRG7	Protein phosphatase 1 regulatory subunit 12A (Myosin phosphatase-targeting subunit 1) (Myosin phosphatase target subunit 1) (Protein phosphatase myosin-binding subunit)	117.7	36	20(3)	32.70	10	PF, PM
Q6DRJ4	Protein KRI1 homolog	91.9	18	15(1)			PF, PM
Q6DRL5	Myb-binding protein 1A-like protein	144.9	25	5(1)	24.73	1	PF, PM, F
Q6DRP2	Guanine nucleotide-binding protein-like 3 (Nucleostemin-like protein)	63.4	26	6(1)	26.47	3	PF, PM, F
Q6E2N3	E3 ubiquitin-protein ligase TRIM33 (EC 6.3.2.-) (Ectodermin homolog) (Protein moonshine) (Transcription intermediary factor 1-gamma) (TIF1-gamma) (Tripartite motif-containing protein 33)	130.2	23	7(1)	22.93	1	PM, M
Q6GMH0	Pre-mRNA-splicing factor 18 (PRP18 homolog)	39.7	22	3(2)			PF, PM
Q6GMI9	UDP-glucuronic acid decarboxylase 1 (EC 4.1.1.35) (UDP-glucuronate decarboxylase 1) (UXS-1)	47.8	25	32(3)	25.28	6	PF, PM, F, M
Q6GML7	Putative deoxyribonuclease TATDN1 (EC 3.1.21.-)	34.2	18	53(2)			PF, PM, F, M
Q6IQX0	Lysine-specific demethylase 5B-B (EC 1.14.11.-) (Histone demethylase JARID1B-B) (Jumonji/ARID domain-containing protein 1B-B)	174.5	24	9(1)	23.89	1	PF, PM, F, M
Q6JAN0	G-protein coupled receptor 98 (Monogenic audiogenic seizure susceptibility protein 1 homolog) (Very large G-protein coupled receptor 1)	676.1	18	14(1)			PF, F, M
Q6NUT7	Sortilin	92.0	21	12(1)	20.86	1	F, M
Q6NUV0	Rab3 GTPase-activating protein catalytic subunit	110.1	26	10(2)	25.97	9	PF, PM, F, M
Q6NV04	Acetolactate synthase-like protein (EC 2.2.1.-) (IlvB-like protein)	67.7	155	16(12)	151.31	32	PF, PM
Q6NV38	Transmembrane protein 41A-B	32.0	33	2(2)	33.10	3	PF, PM, F, M

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q6NVC5	Mitochondrial Rho GTPase 1-A (MIRO-1-A) (EC 3.6.5.-) (Ras homolog gene family member T1-A)	71.2	20	16(2)			PF, PM, F, M
Q6NVC9	Junction-mediating and -regulatory protein	96.8	21	17(1)	21.23	6	PF, PM, F, M
Q6NVJ5	Disco-interacting protein 2 homolog B-A (DIP2 homolog B-A)	171.6	48	28(5)			PF, PM, F
Q6NW58	Spastin (EC 3.6.4.3)	63.3	26	17(1)	25.91	1	M
Q6NW59	Cyclic AMP-dependent transcription factor ATF-4 (cAMP-dependent transcription factor ATF-4) (Activating transcription factor 4)	36.9	18	2(1)			PM
Q6NWB6	Unique cartilage matrix-associated protein [Cleaved into: Unique cartilage matrix-associated protein C-terminal fragment (Ucma-C) (Gla-rich protein) (GRP)]	16.1	19	2(1)			M
Q6NWF1	Solute carrier family 2, facilitated glucose transporter member 12 (Glucose transporter type 12) (GLUT-12)	66.4	20	8(1)	20.49	2	PF, PM
Q6NWG4	Protein arginine N-methyltransferase 6 (EC 2.1.1.-) (Histone-arginine N-methyltransferase PRMT6) (EC 2.1.1.125)	39.7	18	6(1)			M
Q6NX08	Ribosome biogenesis protein wdr12 (WD repeat-containing protein 12)	47.3	22	6(1)	22.14	1	PF
Q6NXA4	Interleukin enhancer-binding factor 3 homolog	90.9	19	1(1)			PM
Q6NXD2	Charged multivesicular body protein 2b (Chromatin-modifying protein 2b) (CHMP2b)	24.1	21	38(2)			PF, F
Q6NYE2	Protein RCC2 homolog	54.9	28	8(4)	27.09	6	PF, PM
Q6NYE4	Alkylglycerol monooxygenase (EC 1.14.16.5) (Transmembrane protein 195)	51.9	25	9(1)	24.59	2	M
Q6NYU2	Probable helicase with zinc finger domain (EC 3.6.4.-)	209.6	20	15(2)			F, M
Q6NYW6	CLIP-associating protein 2 (Cytoplasmic linker-associated protein 2)	141.2	20	11(2)			PF, PM, M

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q6NYY9	Vacuole membrane protein 1 (Transmembrane protein 49)	46.1	22	2(2)			PM
Q6NYZ6	Calcium-binding mitochondrial carrier protein SCaMC-2-A (Small calcium-binding mitochondrial carrier protein 2-A) (Solute carrier family 25 member 25-A)	52.9	26	9(2)	25.94	5	PM, F, M
Q6NZT2	Cerebellar degeneration-related protein 2-like	53.6	23	4(2)			PM
Q6P026	Barrier-to-autointegration factor	10.4	18	2(1)			PF, PM, M
Q6P0D0	Protein quaking-A (zqk)	37.9	18	19(2)			PF, F
Q6P0D7	MAGUK p55 subfamily member 7 (Protein humpback)	65.2	23	3(2)			PF
Q6P0H7	Carbonyl reductase family member 4 (EC 1.-.-.-) (3-oxoacyl-[acyl-carrier-protein] reductase) (EC 1.1.1.-) (Quinone reductase CBR4)	25.1	18	2(1)			M
Q6P132	Tax1-binding protein 1 homolog B	95.4	18	7(1)			PM, F
Q6P5L3	60S ribosomal protein L19	23.6	18	4(1)			PM
Q6P5L7	Arginine and glutamate-rich protein 1-A	30.4	18	5(1)			PF
Q6PBM8	Probable RNA-binding protein 18 (RNA-binding motif protein 18)	21.6	18	3(1)			PM, F
Q6PBT9	Zinc finger protein 385B (Zinc finger protein 533)	53.3	20	13(2)			PF, PM
Q6PC14	60S ribosomal protein L23	15.0	27	2(1)	26.90	2	PM
Q6PC29	14-3-3 protein gamma-1	28.4	448	35(27)	576.66	265	PF, PM, F, M
Q6PC91	Transcription factor BTF3 homolog 4 (Basic transcription factor 3-like 4)	17.4	21	2(1)	20.69	1	PF, F
Q6PCR6	Zinc finger RNA-binding protein	117.4	43	53(8)	42.13	28	PF, PM, F, M
Q6PCR7	Eukaryotic translation initiation factor 3 subunit A (eIF3a) (Eukaryotic translation initiation factor 3 subunit 10) (eIF-3-theta)	151.7	19	18(1)			PF, PM, F

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q6PFJ7	Sphingomyelin phosphodiesterase 4 (EC 3.1.4.12) (Neutral sphingomyelinase 3) (nSMase-3) (nSMase3) (Neutral sphingomyelinase III)	90.4	24	15(2)	22.72	2	PF, PM, F, M
Q6PFL8	Thymocyte nuclear protein 1	26.7	18	8(1)			PF, PM
Q6PG31	RNA-binding protein with serine-rich domain 1	31.7	18	4(1)			PF
Q6PGY6	E3 UFM1-protein ligase 1 (EC 6.3.2.-)	89.4	23	6(1)	22.56	1	PM
Q6PI20	Histone H3.3	15.4	134	7(4)			PF, PM, F, M
Q6QQT1	Zinc transporter ZIP1 (Solute carrier family 39 member 1) (Zrt- and Irt-like protein 1) (ZIP-1)	33.2	20	3(1)			PM
Q6R520	Calmodulin (CaM)	16.8	113	10(3)	113.37	21	PF, PM, F, M
Q6T499	Cellular retinoic acid-binding protein 1 (Cellular retinoic acid-binding protein I) (CRABP-I)	15.9	19	4(1)			PM
Q6TEN1	Ras-related protein Rap-1b	21.0	24	1(1)	23.74	1	PF, PM
Q6TEN6	WD repeat-containing protein 91	81.6	20	4(1)			PF
Q6UFZ9	14-3-3 protein beta/alpha-1 (Protein 14-3-3B1)	27.7	769	24(21)	765.59	296	PF, PM, F, M
Q6VEU3	RNA-binding protein PNO1	28.0	29	3(1)	28.66	3	PF, PM
Q6ZM89	Rho GTPase-activating protein 42 (Rho GTPase-activating protein 10-like) (Rho-type GTPase-activating protein 42)	92.0	18	15(1)			PM
Q708S7	Acid-sensing ion channel 1 (ASIC1) (Acid-sensing ion channel 1.2) (Amiloride-sensitive cation channel 2-B, neuronal) (ZASIC1.2)	58.6	20	3(2)			PF, PM, M
Q71PD7	Histone H2A.V (H2A.F/Z)	13.5	116	3(2)			PM, F
Q7SXA9	Crossover junction endonuclease MUS81 (EC 3.1.22.-)	68.8	18	8(2)			PF, PM
Q7SXE4	Golgin subfamily A member 5 (Golgin-84)	84.6	19	16(1)			F, M

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q7SXI6	Nuclear distribution protein nudE-like 1-A	38.7	20	1(1)	20.13	1	PM
Q7SXM7	U4/U6 small nuclear ribonucleoprotein Prp31 (Pre-mRNA-processing factor 31)	56.7	19	4(1)			PF, PM, F
Q7SXN5	Dynamin-1-like protein (EC 3.6.5.5)	77.7	19	27(1)			PM, M
Q7SXP4	Serine/arginine-rich splicing factor 1A (Splicing factor, arginine/serine-rich 1) (Splicing factor, arginine/serine-rich 1A)	28.5	21	7(1)	20.76	1	M
Q7SXU0	Eukaryotic translation initiation factor 3 subunit J-A (eIF3j-A) (Eukaryotic translation initiation factor 3 subunit 1-A) (eIF-3-alpha-A) (eIF3 p35-A)	28.7	19	12(1)			PM
Q7SXV2	Zinc finger protein AEBP2 (Adipocyte enhancer-binding protein 2 homolog) (AE-binding protein 2 homolog)	45.8	30	3(1)	29.89	1	PM, F
Q7SY23	Delta-1-pyrroline-5-carboxylate dehydrogenase, mitochondrial (P5C dehydrogenase) (EC 1.2.1.88) (Aldehyde dehydrogenase family 4 member A1) (L-glutamate gamma-semialdehyde dehydrogenase)	61.9	18	16(2)			PF
Q7SYB5	Differentially expressed in FDCP 6 homolog	72.5	39	22(7)	38.26	9	PF, F
Q7SYC9	Choline transporter-like protein 2 (Solute carrier family 44 member 2)	80.7	27	7(1)	26.84	1	PM, F, M
Q7SZS1	Transcription factor Sox-21-A (SRY-box containing gene 21a)	26.6	18	22(1)			PF, PM, M
Q7T047	Gap junction Cx32.2 protein (Connexin-32.2)	32.4	27	11(2)	33.42	18	PF, PM, F, M
Q7T297	Protein FAM172A	48.4	18	12(1)			F
Q7T320	RAB6-interacting golgin (N-terminal kinase-like-binding protein 1) (NTKL-BP1) (NTKL-binding protein 1) (SCY1-like 1-binding protein 1) (SCYL1-BP1) (SCYL1-binding protein 1)	38.0	41	17(3)	39.76	3	PF, PM, F
Q7T3P8	Protein fem-1 homolog C (FEM1c) (FEM1-gamma)	69.5	21	13(2)			PM, F, M

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q7ZTU9	T-box transcription factor TBX2b (T-box protein 2b)	74.5	19	9(1)			PF
Q7ZUQ3	Serine/threonine-protein kinase 3 (EC 2.7.11.1) [Cleaved into: Serine/threonine-protein kinase 3 36kDa subunit (MST2/N); Serine/threonine-protein kinase 3 20kDa subunit (MST2/C)]	56.3	29	14(6)	25.90	10	PM, F, M
Q7ZUW2	Hypoxia up-regulated protein 1	110.5	18	16(1)			PF
Q7ZUX1	DNA-directed RNA polymerase III subunit RPC3 (RNA polymerase III subunit C3) (DNA-directed RNA polymerase III subunit C)	60.5	29	12(5)			PF
Q7ZUY3	Histone H2AX (H2a/x) (Histone H2A.X)	15.0	150	5(5)			PF, F, M
Q7ZVA0	WD40 repeat-containing protein SMU1 (Smu-1 suppressor of mec-8 and unc-52 protein homolog)	58.1	24	16(1)	23.76	3	PM, F, M
Q7ZVI7	Actin, cytoplasmic 1 (Beta-actin-1) [Cleaved into: Actin, cytoplasmic 1, N-terminally processed]	42.1	1683	85(63)	1663.60	653	PF, PM, F, M
Q7ZVJ6	Queuine tRNA-ribosyltransferase subunit qtrtd1 (EC 2.4.2.29) (Queuine tRNA-ribosyltransferase domain-containing protein 1)	46.8	20	9(1)	20.48	2	PF, PM, M
Q7ZVR1	WD repeat-containing protein 75	95.3	32	13(4)	31.19	9	PM, M
Q7ZW33	U3 small nucleolar RNA-associated protein 15 homolog	58.0	18	11(1)			PM, F, M
Q7ZW63	Kinetochore protein Nuf2 (Cell division cycle-associated protein 1)	53.9	20	7(1)			PF, PM, M
Q7ZWC4	Protein pelota homolog (EC 3.1.-.-)	43.5	20	9(1)			PF
Q7ZWJ4	60S ribosomal protein L18a	20.8	97	259(26)	97.34	288	PF, PM, F, M
Q7ZWJ7	60S ribosomal protein L34	13.4	20	6(1)			PF, PM, F
Q7ZZX9	GTP-binding nuclear protein Ran (GTPase Ran) (Ras-related nuclear protein)	24.6	18	2(1)			M

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q800H9	Heparan-sulfate 6-O-sulfotransferase 2 (HS 6-OST-2) (HS6ST-2) (EC 2.8.2.-)	55.3	57	39(7)	52.54	79	PF, PM, F, M
Q802G6	Methionine-R-sulfoxide reductase B1-A (MsrB) (MsrB1-A) (EC 1.8.4.-) (Selenoprotein X-A) (SePR) (SelX-A)	0.0	36	2(2)	35.80	5	PF, PM
Q802R8	Structural maintenance of chromosomes protein 6 (SMC protein 6) (SMC-6)	125.9	20	17(2)			PF
Q802R9	Structural maintenance of chromosomes protein 5 (SMC protein 5) (SMC-5)	128.8	22	12(1)	22.17	1	PF
Q803G5	Sorting and assembly machinery component 50 homolog A	52.1	19	6(2)			PF
Q803I2	Endoplasmic reticulum-Golgi intermediate compartment protein 3	43.7	22	4(1)	21.91	1	PM
Q803I8	E3 ubiquitin-protein ligase synoviolin (EC 6.3.2.-) (Synovial apoptosis inhibitor 1)	69.3	19	10(1)			PF, F
Q803Q2	Nuclear distribution protein nudE-like 1-B	38.6	20	6(1)			F
Q803U7	Exonuclease 1 (EC 3.1.-.-) (Exonuclease I)	89.7	23	11(2)			PM, F, M
Q803Z2	Protein YIPF3 (YIP1 family member 3)	38.0	128	27(17)	128.35	190	PF, PM, F, M
Q804S5	E3 ubiquitin-protein ligase mib1 (EC 6.3.2.-) (Protein mind bomb)	115.0	19	2(1)			M
Q804T6	Dixin-A (Coiled-coil protein DIX1-A) (Coiled-coil-DIX1-A) (DIX domain-containing protein 1-A)	50.6	62	11(8)	60.92	75	PF, PM, F, M
Q804W2	Parvalbumin-7 (Parvalbumin alpha) (Parvalbumin-4a)	12.0	20	4(1)	20.40	2	PF, PM, F, M
Q8AWZ2	Homeobox protein Hox-A3a	44.6	18	5(1)			F
Q8JFG3	Tumor necrosis factor (TNF-alpha)	28.2	21	7(2)			PF, PM, F, M
Q8JFQ6	Keratin, type I cytoskeletal 13 (Cytokeratin-13) (CK-13) (Keratin-13) (K13)	51.9	149	12(7)			PF, PM, F, M

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q8JFW4	Probable inactive tRNA-specific adenosine deaminase-like protein 3 (tRNA-specific adenosine-34 deaminase subunit adat3)	38.6	21	4(1)	20.88	1	PM
Q8JG70	Alpha-2Da adrenergic receptor (Alpha-2Da adrenoceptor) (Alpha(2Da)AR) (Alpha-2Da adrenoreceptor)	47.0	26	7(1)	26.47	5	PM, F
Q8JH70	Fructose-bisphosphate aldolase C-B (EC 4.1.2.13) (Brain-type aldolase-B)	39.7	110	9(7)	121.56	50	PF, PM, F, M
Q8UVC1	Inversin	113.9	47	12(4)	47.38	25	PF, PM, F, M
Q8UVK2	Transcription elongation factor SPT6 (Protein pandora)	199.3	22	13(1)	21.66	3	PF
Q90304	CD166 antigen homolog (Activated leukocyte cell adhesion molecule) (DM-GRASP homolog) (Neurolin) (CD antigen CD166)	61.0	23	16(2)	21.67	3	PF
Q90327	Mitogen-activated protein kinase 8A (MAP kinase 8A) (MAPK 8A) (EC 2.7.11.24) (Stress-activated protein kinase JNKa) (c-Jun N-terminal kinase A)	49.0	50	13(7)			PF, PM, F, M
Q90339	Myosin heavy chain, fast skeletal muscle	222.3	18	10(1)			PM, M
Q90473	Heat shock cognate 71 kDa protein (Heat shock 70 kDa protein 8)	71.2	34	4(2)	33.38	7	PM, F, M
Q90486	Hemoglobin subunit beta-1 (Beta-1-globin) (Beta-A1-globin) (Hemoglobin beta-1 chain)	16.6	661	130(73)	624.81	950	PF, PM, F, M
Q90487	Hemoglobin subunit alpha (Alpha-globin aa1) (Hemoglobin alpha chain)	15.5	5749	341(267)	5670.23	2368	PF, PM, F, M
Q90WG6	N-acylneuraminate cytidylyltransferase (EC 2.7.7.43) (CMP-N-acetylneuraminic acid synthase) (CMP-NeuNAc synthase)	48.7	20	26(2)			F, M
Q90X49	Coiled-coil domain-containing protein 80	100.3	18	7(1)			PM, F
Q90Y50	Coxsackievirus and adenovirus receptor homolog (CAR)	41.3	22	6(1)	21.83	2	PF, M

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q90YI3	Methionine--tRNA ligase, mitochondrial (EC 6.1.1.10) (Methionyl-tRNA synthetase 2) (Mitochondrial methionyl-tRNA synthetase) (MtMetRS)	68.2	61	16(11)	53.80	74	PF, PM, F, M
Q90YJ2	Neuroglobin	17.8	20	3(2)			F, M
Q90YS2	40S ribosomal protein S3 (EC 4.2.99.18)	27.0	18	3(1)			F
Q90YV0	60S ribosomal protein L18	21.5	44	10(3)	43.96	13	PF, PM, F, M
Q90Z00	Fibroblast growth factor receptor 1-A (FGFR-1-A) (bFGF-R-1-A) (EC 2.7.10.1) (Basic fibroblast growth factor receptor 1-A)	92.0	19	3(1)			PF, F
Q90ZE2	Zinc finger protein 703 (NocA-like zinc finger protein 1) (NocA-related zinc finger protein 1)	62.2	25	9(2)	24.83	16	PF, PM, F, M
Q90ZY6	Serine/threonine-protein kinase tousled-like 1-B (EC 2.7.11.1) (PKU-beta) (Tousled-like kinase 1-B)	86.1	18	3(1)			F, M
Q91191	Peroxiredoxin (EC 1.11.1.15) (Natural killer enhancement factor-like protein) (RBT-NKEF) (Thioredoxin peroxidase) (Thioredoxin-dependent peroxide reductase)	22.2	298	8(8)	297.50	70	PF, PM, F, M
Q91233	Heat shock 70 kDa protein (HSP70)	71.3	27	7(2)	26.95	3	PF, M
Q91428	Transcription factor GATA-3 (GATA-binding factor 3)	48.1	26	9(2)	25.36	4	PF, PM, F, M
Q91487	60S ribosomal protein L13a (Transplantation antigen P198 homolog) (Fragment)	21.7	35	3(1)	35.41	1	PM, F
Q92079	Serotransferrin (Fragment)	69.5	637	27(23)	634.68	240	PF, PM, F, M
Q92148	Cytochrome P450 1A1 (EC 1.14.14.1) (CYPIA1)	58.4	40	32(3)	40.13	32	PF, PM, F, M
Q92155	Vimentin	52.5	102	22(11)	96.69	48	PF, PM, F, M

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q98893	Vitellogenin-2 (Vitellogenin II) (VTG II) [Cleaved into: Lipovitellin-1 (LV1); Phosvitin (PV); Lipovitellin-2 (LV2)]	187.3	19	6(1)			PM, F, M
Q98SN9	Protein transport protein Sec61 subunit alpha isoform A	52.6	18	11(1)			PF
Q98UF7	Pecanex-like protein 1 (Pecanex homolog)	193.4	19	10(1)			PM, F
Q9DE49	Platelet-derived growth factor receptor alpha (PDGF-R-alpha) (PDGFR-alpha) (EC 2.7.10.1) (Alpha platelet-derived growth factor receptor) (Alpha-type platelet-derived growth factor receptor)	119.2	30	9(2)	29.46	2	PM
Q9DGD9	Mitogen-activated protein kinase 8 (MAP kinase 8) (MAPK 8) (EC 2.7.11.24) (Stress-activated protein kinase JNK1) (c-Jun N-terminal kinase 1)	44.5	41	10(5)	46.19	59	PM, M
Q9I8N6	Macrophage colony-stimulating factor 1 receptor (CSF-1 receptor) (CSF-1-R) (CSF-1R) (M-CSF-R) (EC 2.7.10.1) (Proto-oncogene c-Fms homolog)	111.5	25	8(1)	25.17	3	PF, PM, F, M
Q9I8V2	Mothers against decapentaplegic homolog 1 (MAD homolog 1) (Mothers against DPP homolog 1) (SMAD family member 1) (SMAD 1) (Smad1)	53.0	26	10(1)			PF, PM, F
Q9I9K7	T-box transcription factor TBX20 (T-box protein 20) (H15-related T-box transcription factor hrT)	49.4	20	3(2)			PF, M
Q9I9N5	Tumor necrosis factor receptor type 1-associated DEATH domain protein (TNFR1-associated DEATH domain protein) (TNFRSF1A-associated via death domain protein)	33.5	18	2(1)			F
Q9IAK1	Estrogen receptor beta (ER-beta) (Nuclear receptor subfamily 3 group A member 2)	65.1	22	2(1)	22.45	2	PF, PM, M

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q9IAT6	Delta-like protein C (DeltaC) (delC)	75.9	18	7(1)			PF
Q9IB84	Proteasome subunit beta type-1-A	26.0	30	2(1)			M
Q9MIY0	NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3) (NADH dehydrogenase subunit 5)	68.5	21	8(2)			F, M
Q9PTN2	Vitamin D3 receptor A (VDR-A) (1,25-dihydroxyvitamin D3 receptor A) (Nuclear receptor subfamily 1 group I member 1-A)	51.7	47	5(2)	46.29	5	PF, PM, F, M
Q9PTS2	Uroporphyrinogen decarboxylase (UPD) (URO-D) (EC 4.1.1.37)	42.0	21	18(1)	20.52	3	PF, F
Q9PTY0	ATP synthase subunit beta, mitochondrial (EC 3.6.3.14)	55.3	119	14(7)	118.97	11	PF, PM, M
Q9PUQ1	ETS translocation variant 4 (Polyomavirus enhancer activator 3 homolog) (Protein PEA3)	56.0	52	45(11)	49.66	74	PF, PM, F, M
Q9PVM2	Hemoglobin subunit beta-A (Beta-A-globin) (Hemoglobin beta-A chain)	16.5	2430	211(165)	2373.47	1570	PF, PM, F, M
Q9PVM3	Hemoglobin subunit alpha-B (Alpha-B-globin) (Hemoglobin alpha-B chain)	15.9	173	224(39)	171.11	614	PF, PM, F, M
Q9PVM4	Hemoglobin subunit alpha-A (Alpha-A-globin) (Hemoglobin alpha-A chain)	15.6	1562	124(109)	1510.55	1149	PF, PM, F, M
Q9PVU6	Hemoglobin embryonic subunit alpha (Alpha-globin, embryonic) (Hemoglobin alpha-chain, embryonic)	15.6	18	3(1)			PF
Q9PVW7	Complement component C8 beta chain (Complement component 8 subunit beta)	67.5	56	17(10)	38.74	10	PF, PM, F, M
Q9PWL6	Homeobox protein Hox-A9a (Hoxx9)	28.2	39	4(3)	37.71	19	PM, F, M
Q9W686	Semaphorin-3ab (Semaphorin-1B) (Semaphorin-Z1B) (Sema Z1B)	88.8	54	6(3)	52.08	9	PF, PM, F, M
Q9W6C7	Neurogenic differentiation factor 6-B (NeuroD6-B) (Protein atonal homolog 2-B)	36.5	25	15(3)			F, M

Table S5.2. List of proteins. (continued).

UniProt n°	Protein name	MW [kDa]	Highest mascot Score	Matches	MaxQUANT Highest Score	MaxQUANT Σ# Peptides	Studied group
Q9W747	Zinc finger protein draculin	48.0	18	9(1)			PF
Q9W7E7	Mothers against decapentaplegic homolog 5 (MAD homolog 5) (Mothers against DPP homolog 5) (Protein somitabun) (SMAD family member 5) (SMAD 5) (Smad5)	52.5	23	11(2)	26.31	6	PF, M
Q9W7R3	Teneurin-4 (Ten-4) (Protein Odd Oz/ten-m homolog 4) (Tenascin-M4) (Ten-m4) (Teneurin transmembrane protein 4)	317.8	19	19(1)			PM, F, M
Q9YGM4	CaM kinase-like vesicle-associated protein	47.7	19	6(2)			PF, PM, F, M
Q9YGM8	Caveolin-1	21.1	19	14(1)			PF, F, M
Q9YGN1	Protein SAND	59.0	22	11(1)	21.65	3	PM
Q9YH26	Sodium/potassium-transporting ATPase subunit alpha-1 (Na(+)/K(+) ATPase alpha-1 subunit) (EC 3.6.3.9) (Sodium pump subunit alpha-1)	112.6	18	16(1)			PF, PM
Q9YH92	Claudin-7-B (Claudin-like protein ZF4A22)	22.9	33	11(2)	30.80	6	PF, PM, F, M
Q3B7P7	Ubiquitin A-52 residue ribosomal protein fusion product 1	8.6			187.4798329	80	PF, PM
Q6DC64	Methyltransferase-like protein 16 (EC 2.1.1.-) (Methyltransferase 10 domain-containing protein)	53.6			23.9678866	2	F

CHAPTER 6

6. Chapter 6

SUMMARY OF FINDINGS

SUMMARY OF FINDINGS

Chapter 2 - ddRAD sequencing in *Arapaima gigas*: analysis of genetic diversity and structure in populations from Amazon and Araguaia-Tocantins River Basins

- The overall extent of polymorphism in *A. gigas* was surprisingly low with only 2.3 % of identified RAD-tags (135 bases long) containing SNPs;
- A panel of 393 informative SNPs was identified and screened across five populations;
- Mean pairwise relatedness among individuals was significant, suggesting a degree of inbreeding in all populations;
- Higher genetic diversity indices were found in populations from Amazon and Solimões, intermediate levels in Tocantins and Captive, and very low levels in Araguaia;
- Populations of Amazon and Solimões were found to be genetically differentiated from Araguaia, with Tocantins comprising individuals from both identified stocks.

Chapter 3 - Effects of GnRHa implants and size pairing on plasma and cephalic secretion sex steroids in *Arapaima gigas* (Schinz, 1822)

- Despite the lack of clear spawning induction, the potency of mGnRHa on sex steroid production was shown regardless of size pairing groups (large, small and mixed);
- Compared to control (non-implanted), implanted broodstock showed a significant increase in plasma 11-KT (large and small males) and T (large and mixed females) post GnRHa implantation;

- Significant correlations between blood plasma and cephalic secretion levels of 11-KT in males and T in females were observed, suggesting pheromone release through the cephalic canals of *A. gigas*.

Chapter 4 –Development of tools to monitor female reproductive function and hormonal manipulation in selected mature females

Part A - Endoscopy application in broodstock management of *Arapaima gigas* (Schinz, 1822)

- Preliminary tests showed a non-surgical endoscopic (gonoductoscopy) can be useful to assess ovary development in adult females of *A. gigas*.

Part B – Gender identification and monitoring of female reproductive function using endoscopy and cannulation techniques in the giant *Arapaima gigas*

- Gonoductoscopy procedure was validated under field operational conditions in 12 adult females and for gender identification in 20 juveniles;
- Cannulation assisted through the description of the genital anatomy made ovary biopsy sampling possible to describe oocyte development from primary growth to ovulation, providing a tool to monitor reproductive function in females;
- Females paired with males in earthponds can undergo oocyte maturation, possibly failing to ovulate and spawn, or alternatively release eggs which are not fertilised.

Part C – Effects of multiple GnRHa injections on ovary development and plasma sex steroids in *Arapaima gigas*

- After 12 days post-injections with GnRHa (combination of types 1 and 3), the leading cohort oocyte diameter advanced in 80 % of the selected females through biopsy, though no spawning was apparent;
- Levels of E_2 in females and 11-KT in males remained unchanged post-treatment.

Chapter 5 - Comparative proteome and peptidome analysis of the cephalic fluid secreted by *Arapaima gigas* (Teleostei: Osteoglossidae) during and outside parental care

- Multiple comparisons revealed 28 peptides were significantly different between males (M) and parental males (PM), 126 between females (F) and parental females (PF), 51 between M and F and 9 between PM and PF;
- Identification revealed peptides were produced in the inner ear (*pcdh15b*), eyes (*tetraspanin* and *ppp2r3a*), central nervous system (*otud4*, *ribeye a*, *tjp1b* and *syn1*) among others;
- A total of 422 proteins were also identified and gene ontology analysis revealed 28 secreted extracellular proteins;
- 2 hormones (*prolactin* and *stanniocalcin*) and 12 proteins associated to immunological processes (*serotransferrin*, *α-1-antitrypsin homolog*, *apolipoprotein A-I*, and others) were identified;
- Novel biochemical data on the lateral line fluid was obtained which will enable future hypothesis-driven experiments to better understand the physiological roles of the lateral line in chemical communication.

CHAPTER 7

7. Chapter 7

GENERAL DISCUSSION

7.1. Overview

The general aims of this thesis were to better understand the reproductive physiology of *A. gigas*, develop new tools to identify gender and monitor gonadal development, test hormonal therapies to induce spawning and characterise the cephalic secretion and its potential roles during parental care. As a fundamental milestone in animal breeding, the genomic variability and diversity of the species was initially assessed in captive and wild populations (Chapter 2). When the work was initiated at the start of the PhD, there was a complete lack of knowledge on the reproductive dysfunction(s) of *A. gigas* and tools to assess gonadal stage were not available to appropriately apply hormonal treatments. The initial approach to this problem included testing the potency of mGnRHa slow-release implants to induce ovulation and spawning in paired couples of different sizes. This strategy has proven to be successful in many other fish species. The effects of the hormonal induction were assessed through the profiling of plasma sex steroids, since viable spawning was not induced although egg release could not be detected reliably (Chapter 3). At this stage, it became very clear that methods to confirm gender of paired couples and assess oogenesis in females were lacking. Endoscopy and cannulation were therefore developed and validated as tools to select broodstocks and form couples. Oogenesis in *A. gigas* was described and the reproductive dysfunction in a captive broodstock was investigated (Chapter 4- Parts A and B). With such technical advances, maturing females could now be selected for a second trial which tested a combination of mGnRHa and sGnRHa administered by dual injections to induce ovulation and spawning, and this time effects on oogenesis were monitored (Chapter 4- Part C). Considering current doubts on the potential roles of the cephalic secretion in *A. gigas*, the potential of pheromone release was investigated (Chapter 3) and putative roles of this biological fluid during parental care were studied through a comparative analysis of the peptidome and

proteome in males and females during and outside the reproductive period (Chapter 5). While this doctoral study did not solve the reproductive bottlenecks in *A. gigas*, it significantly advanced knowledge of the species' reproductive biology and developed essential tools to be used in future experimentation. In this section, the main findings are discussed in view of experimental limitations, research gaps and opportunities for future studies.

7.2. Genomic diversity and population structuring

A critical aspect for the development of aquaculture of any new candidate species is the characterisation of the genetic diversity available within the populations used in breeding programs. Tools to estimate genetic diversity and select broodstocks are key to maintain diversity and ensure good egg productivity and quality over time. To date, no study has assessed the genome-wide diversity of *Arapaima*, although a loss of genetic diversity was previously suggested given the severe depletion of natural stocks (Hrbek et al., 2005; Araripe et al., 2013). Recently, multiple species were proposed for *Arapaima*, and aquaculture was suggested to play a negative role by translocating individuals between river basins, thus causing hybridization with possible genetic introgression (Stewart, 2013a; Watson et al., 2016). Using ddRAD sequencing to genotype five populations from two hydrographic basins, this research could not identify genetically distinct species in Arapaima, though it must be acknowledged that only a small number of individuals were studied from few locations. Unfortunately, morphological data especially on characteristics defining the species *A. leptosoma* and *A. agassizii* were not available for the samples analysed in this study. Considering the possibility of multiple species in Arapaima, it is also arguable that some Amazon fish genus (*i.e. Cichla*) can be represented by one main widespread species and several rare and endemic sub-species,

and this could be the case in Arapaima (Castello & Stewart, 2010; Stewart, 2013a). Therefore, more populations and larger samples should be analysed using this new SNP marker panel to elucidate this important basic question.

Results from this study found very low degrees of genomic polymorphism in all studied populations, and analyses of population found significant pairwise relatedness indicating overall inbreeding with loss of genetic diversity in all populations. Current plans for the conservation of *A. gigas* in the Amazon involve community-based and sustainable development reservoirs, which are helping to reduce pressure on natural stocks (Arantes et al., 2010; Petersen et al., 2016). However, the loss of genetic diversity reported here suggests that conservation efforts should take into account the maintenance of genetic diversity and promote the adaptive potential of the species. While reproduction of Arapaima is not controlled in captivity to sustain an industry, pressure on natural stocks will be constant and unfortunately includes the illegal capture of the high-valued offspring from the wild. Advances towards the control of captive reproduction in Arapaima, aimed at in this thesis, could help develop an industry thus reducing pressure on wild stocks. Moreover, conservation plans as similar to done for species like the endangered razorback sucker, *Xyrauchen texanus,* in the Colorado River (USA), genetic markers have proved useful to help maintain genetic diversity through conservation based on captive breeding and restocking (Dowling et al., 2005). In this thesis, control of spawning was unfortunately not achieved, though advances were made towards this goal, and also a powerful marker panel was identified that can in the future aid in the identification and selection of stocks for *A. gigas* breeding purposes.

Aquaculture is playing a role in the translocation of fishes between distinct populations, and the possible impacts of such translocations have been neglected so far. Genetic studies in several species have shown that introduction of few migrants can be

used to “rescue” populations from inbreeding, as reviewed by Tallmon et al. (2004), and this could be beneficial to populations from Araguaia-Tocantins basin where loss of genetic diversity seems to be more concerning. However, it would be interesting to determine, through controlled studies, the mating compatibility between different stocks (*i.e.* Araguaia x Amazon) and possible outbreeding effects in the fitness of offspring (*i.e.* hybrid vigour or loss of fitness in the first generation F_1), as reviewed in Edmands (2007). Further studies should now select a reduced number of highly polymorphic markers to be validated in a larger number of samples using cheaper genotyping techniques. After validation, the marker panel identified herein could potentially be applied to select broodstocks used in captive reproduction and help identify illegally wild-captured offspring. Recently in Brazil, there is a trend to use molecular markers to identify and regulate commercialised fishery products (Carvalho et al., 2017a), though this would require investments to register Arapaima hatcheries in Brazil.

7.3. Hormonal manipulations and tools to assess gender and reproductive condition

To date, experiments that aimed to better understand the reproductive dysfunction of *A. gigas* and induce oocyte maturation, ovulation and spawning have been limited by the lack of tools to identify genders and the unusual gonadal morphology which limited the assessment of fish reproductive condition. So far, attempts to hormonally and environmentally induce spawning in *A. gigas* have not been reported. The isolation of couples in earthen ponds was shown to promote reproduction during the rainy season (Rebouças et al., 2014; Lima et al., 2015b), however results have been unpredictable with low success due partly to the fact that fish gender could not be confirmed. The present thesis studied for the first time the reproductive biology of *A. gigas* using couples,

allowed by the recent development of a method to identify genders through the assay of vitellogenin in the blood (Dugue et al., 2008; Chu-Koo et al., 2009). It must be acknowledged that other possible mating strategies (*e.g.* triplets with two males) should also be tested given the importance of social factors in the control of reproduction, as recently suggested for *A. gigas* (Farias et al., 2015).

During this thesis research, an initial experiment tested the effects of mGnRHa slow-release implants on randomly selected couples with different size pairings, showing the potency of mGnRHa on the BPG axis of *A. gigas* (Chapter 3). At this stage of the research, several drawbacks were encountered such as the lack of tools to stage gonadal development prior to and post treatment, the lack of knowledge on the GnRH system that remains uncharacterised in Arapaima, the possible dopaminergic inhibition and finally, the difficulty to observe spawning directly. While the hormonal induction clearly had an effect on sex steroid levels in the blood, no spawning was observed. Therefore, the work focused on the development of endoscopy and cannulation techniques as novel tools for gender identification and assessment of reproductive function in females. Importantly, these tools, while helping to ascertain genders in the couples formed, also showed for the first time that the majority of females randomly paired with males in earthen ponds underwent oocyte maturation but appeared to arrest development or regress as no ovulation and spawning were detected (Chapter 4-Parts A and B). Cannulation in males remained unresolved and future investigations should describe the position and gauge of the spermiduct opening. The development of methods to cannulate males could enable the collection of milt without sacrificing the individual. The feasibility of cannulation was further tested in selected mature females, aiming to induce ovulation and spawning and test the dual effects mGnRHa and sGnRHa on oocyte development and circulating sex steroid levels (Chapter 4 – Part C). However, in this trial, repeated handling required

for monitoring oocyte development and perform multiple injections led to broodstock mortalities and potentially oocyte regression and no spawning. This demonstrated the difficulties of working with *A. gigas* broodstock which despite their large size can be fragile and difficult to manipulate.

Although advances were achieved in this thesis, methods tested to control spawning were not successful and require further investigations. Future experiments with hormonal and environmental manipulations will greatly benefit from the endoscopy and cannulation techniques developed and validated for females. It is expected that such tools will help broodstock management in farms, and future experiments studying the proximate environmental factors controlling gametogenesis in the species. Regarding future experiments aiming to induce ovulation and spawning in *A. gigas*, it would be interesting to conduct ecophysiological studies to evaluate the effects of flooding and water quality parameters on gonadal development. For example in the African arowana *H. niloticus*, a species very similar phylogenetically to *A. gigas*, induced spawning has been achieved by simulating rising water in ponds (Monentcham et al., 2009). Such a study was planned in the present thesis however the technical difficulties to artificially control flooding in large earthen ponds in remote areas in Brazil prevented it. Other factors suggested to promote spawning in *A. gigas* relate to broodstock nutrition, a research topic poorly investigated to date in the species, and rain-related cues (*i.e.* pH, temperature, shading) (Núñez et al., 2011). Likewise, studies have also demonstrated the importance of sound in communication of Arapaima for predation avoidance (Olsen, 2014), but the potential role in cueing reproduction has not been studied yet. The roles of these (and others such as photoperiod or light intensity) should also be tested on *A. gigas* reproduction.

7.4. Cephalic secretion of *A. gigas*

The potential release of pheromonal molecules through the cephalic secretion was investigated by comparing levels of sex steroids in the blood plasma and cephalic secretion (Chapter 3). When comparing both fluids, significant positive correlations were found for 11-KT assayed in males and T in females. Although there is no published study confirming the role of these steroids as pheromones in *A. gigas*, many studies in other fish species showed that 11-KT acts as a pheromone used to attract females (Sorensen & Wisenden, 2015). To confirm the potential pheromonal communication through the cephalic secretion in *A. gigas*, studies are required to evaluate ecophysiological responses through hypothesis-driven experiments. Such studies, while technically difficult to perform in *A. gigas*, could confirm the involvement of the lateral line system in the chemical communication of fish. Likewise, investigations of already known pheromones in the cephalic secretion of other teleosts would support the use of the anterior lateral line system as a source of pheromone release in teleosts. Such investigations are lacking for other species and known routes of pheromone release include the gills, urine, seminal and ovarian fluids (Sorensen & Wisenden, 2015). Given the complex reproductive behaviour and strategy seen in *A. gigas*, a better understanding and characterisation of pheromones in the species could help to develop methods to induce spawning. For example in species like *Clarias gariepinus*, it has been found that ovulation in females can be induced when these are reared in the same water environment as a spawning couple (Resink et al., 1989). In *A. gigas*, earthen ponds for breeding Arapaima have been built using connecting canals believed to help synchronise reproduction between couples while providing the necessary territory for couples to build nests and spawn (Fontenele, 1948). Such a pond system with connecting canals was indeed used in Chapter 4 – Parts B and C, where monitoring of female oogenesis also showed synchronised oogenesis in 11 out of 12 studied females.

Whether such synchronisation among females was controlled through pheromones remains to be investigated. Future studies could use the cephalic secretion to assay candidate molecules (*i.e.* prostaglandins, MIS, sex steroids) to characterise individual phenotypes including gender, reproductive condition but also dominance or other social factors.

Based on the preliminary results obtained above, a study then aimed to characterise the proteome and peptidome of the cephalic secretion to better understand possible functions of this secretion in parental care (Chapter 5). It must be acknowledged that this study analysed low number of samples of parental fish, being limited by the lack of spawning in the species. Therefore, analyses aimed to identify the highest number of proteins and peptides as possible. The low concentrations of protein in all analysed samples suggested this fluid is unlikely to be a source of proteins for the offspring (nutritional role) as suggested in prior studies (Quay, 1972; Noakes, 1979). Immunological and hormonal roles of the secretion for the offspring appeared more likely. While the biochemical composition of the cephalic secretion clearly identified putative immunological molecules that could play a role for developing offspring, future hypothesis-driven experiments are required, and the list of proteins and peptides provided in this study will be a good starting point. Other avenues for research were highlighted by the finding of proteins whose functions are restricted to developing embryos (somitogenesis, mesoderm segmentation, embryonic brain development, among others) and were present in the cephalic fluid of adults. If these proteins are produced by developing embryos and not adult fish, mouthbrooding or a close nest guarding behaviour should be considered very likely where the head of adults remain in constant contact with developing egg mass.

7.5. General conclusions

This thesis addressed key issues on the reproductive biology of *A. gigas*. Initially, population genomics was studied for the first time and identified a SNP marker panel for broodstock identification. Although control of spawning could not be achieved in the timeframe of this thesis, the potency of mGnRHa on the BPG axis was shown for the species and different hormonal treatment approaches were tested and reported. Advances were also made on novel methods for gender identification and for monitoring female reproductive function. These methods can be promptly applied to improve broodstock management and future in vivo experiments to investigate proximate factors controlling reproduction of the species. Also, novel data on the oogenesis was presented describing development from primary growth to ovulation with description of the micropyle morphology in this ancient group of fish. The proteomic and peptidomic study of the cephalic secretion, together with sex steroids, suggested a potential complex role of this lateral line fluid for developing offspring. This opens up many scientific hypotheses that should be tested in future investigations with interesting phylogenetic and evolutionary questions in fish. Overall, this research contributed significantly to advance knowledge on the species and to improve broodstock management on farms to help to realise the full commercial potential of the species in Brazil.

PUBLICATIONS AND CONFERENCES

Publications

Torati, L.S., Varges, A.P.S., Galvão, J.A.S., Mesquita, P.E.C. & Migaud, H. (2016). Endoscopy application in broodstock management of *Arapaima gigas* (Schinz, 1822). Journal of Applied Ichthyology 32, 353-355.

Torati, L.S., Migaud, H, Doherty, M.K., Siwy, J., Mullen, W., Mesquita, P.E.C. & Albalat, A. (in press). Comparative proteome and peptidome analysis of the cephalic fluid secreted by *Arapaima gigas* (Teleostei: Osteoglossidae) during and outside parental care. PlosOne.

Conferences

Torati, L.S., Lima, A.F., Mesquita, P.E.C, Davie, A. & H. Migaud. (2015). Reproductive biology of Arapaima: sex dimorphism, pheromones, parental care and captive reproduction under investigation. Poster presentation. UoS PhD Conference 2015.

Torati, L.S., Lima, A.F., Mesquita, P.E.C, J.B. Taggart, A. Albalat, J. Taylor & H. Migaud. (2015). Advances into the captive reproduction of the Amazon Pirarucu (Osteoglossidae: Arapaima). Oral presentation. Latin American & Caribbean Aquaculture 2015, Fortaleza, Brazil.

Torati, L.S., Taylor, J., Davie, A., Mesquita, P.E.C. & H. Migaud. (2016). Effects of GnRHa slow-release implants in the steroid profile of isolated couples of *Arapaima gigas* (Schinz, 1822). Oral presentation. Aquaculture Europe 2016, Edinburgh, UK.

REFERENCES

Albalat, A., Franke, J., Gonzalez, J., Mischak, H. & Zürbig, P. (2013). Urinary proteomics based on Capillary Electrophoresis coupled to Mass Spectrometry in kidney disease. In: Phillips, T.M. & Kalish, H., eds., Clinical Applications of Capillary Electrophoresis: Methods in Molecular BiologyHumana Press, 203-213.

Albers, J.L., Wildhaber, M.L. & DeLonay, A.J. (2013). Gonadosomatic index and fecundity of lower missouri and middle Mississippi River endangered pallid sturgeon estimated using minimally invasive techniques. *Journal of Applied Ichthyology* **29**, 968-977.

Almeida, I.G., Ianella, P., Faria, M.T., Paiva, S.R. & Caetano, A.R. (2013). Bulked segregant analysis of the pirarucu (*Arapaima gigas*) genome for identification of sex-specific molecular markers. *Genetics and Molecular Research* **12**, 6299-6308.

Almendras, J.M., Duenas, C., Nacario, J., Sherwood, N.M. & Crim, L.W. (1988). Sustained Hormone Release. III. Use of Gonadotropin Releasing Hormone Analogues to Induce Multiple Spawnings in Sea Bass, *Lates calcarifer*. *Aquaculture (Amsterdam, Netherlands)* **74**, 97-111.

Alvan-Aguilar, M.A., Chu-Koo, F., Monge, G.C.B., Panduro, L.A.C. & Ríos, D.A.V. (2016). Analisis de las estadísticas de producción de carne y semilla de paiche *Arapaima gigas* en Loreto y Ucayali (Perú). *Folia Amazonica* **25**, 183-190.

Amaral, J.S. (2009). Gonad steroids and lipid metabolism along the reproductive cycle of *Arapaima gigas* (Schinz, 1822) in natural environments. Unpublished thesis (M.Sc.), University of São Paulo.

Andrews, K.R., Good, J.M., Miller, M.R., Luikart, G. & Hohenlohe, P.A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics* **17**, 81-92.

Arantes, C.C., Castello, L., Stewart, D.J., Cetra, M. & Queiroz, H.L. (2010). Population density, growth and reproduction of arapaima in an Amazonian river-floodplain. *Ecology of Freshwater Fish* **19**, 455-465.

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 marker. *PloS One* **8**, 1-7.

Asai-Coakwell, M., March, L., Dai, X.H., DuVal, M., Lopez, I., French, C.R., Famulski, J., De Baere, E., Francis, P.J., Sundaresan, P., Sauve, Y., Koenekoop, R.K., Berry, F.B., Allison, W.T., Waskiewicz, A.J. & Lehmann, O.J. (2013). Contribution of growth differentiation factor 6-dependent cell survival to early-onset retinal dystrophies. *Human Molecular Genetics* **22**, 1432-1442.

Babin, P.J., Carnevali, O., Lubzens, E. & Schneider, W.J. (2007a). Molecular aspects of oocyte vitellogenesis in fish. In: Babin, P.J., Cerdá, J. & Lubzens, E., eds., The fish oocyte: From basic studies to biotechnological applications, Dordrecht, Netherlands: Springer, 39-76.

Babin, P.J., Cerdà, J. & Lubzens, E. (2007b). The fish oocyte: froom basic studies to biotechnological applications, Dordrecht, Netherlands: Springer.

Babin, P.J., Thisse, C., Durliat, M., Andre, M., Akimenko, M.A. & Thisse, B. (1997). Both apolipoprotein E and A-I genes are present in a nonmammalian vertebrate and are highly expressed during embryonic development. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 8622-8627.

Bailly, D., Agostinho, A.A. & Suzuki, H.I. (2008). Influence of the flood regime on the reproduction of fish species with different reproductive strategies in the Cuiaba river, upper Pantanal, Brazil. *River Research and Applications* **24**, 1218-1229.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A. & Johnson, E.A. (2008). Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PloS One* **3**, e3376.

Balshine, S. & Sloman, K.A. (2011). Parental Care in Fishes. In: Farrell, A., ed., Encyclopedia of Fish Physiology-From Genome to EnvironmentAcademic Press, 670-677.

Barber, I. (2007). Parasites, behaviour and welfare in fish. *Applied Animal Behaviour Science* **104**, 251-264.

Barcellos, L.J., Wassermann, G.F., Scott, A.P., Woehl, V.M., Quevedo, R.M., Ittzes, I., Krieger, M.H. & Lulhier, F. (2001). Steroid profiles in cultured female jundia, the Siluridae *Rhamdia quelen* (Quoy and Gaimard, Pisces Teleostei), during the first reproductive cycle. *General and Comparative Endocrinology* **121**, 325-32.

Bartolini, P., Carvalho, R.F., Sevilhano, T.C., Oliveira, J.E. & Garcez, R. (2015). Molecular cloning, characterization and phylogenetic analysis of pirarucu (*Arapaima gigas*) FSH and LH β-subunits. *Journal of Biotechnology* **208**, S5-S120.

Bartsch, P. & Britz, R. (1997). A single micropyle in the eggs of the most basal living actinopterygian fish, *Polypterus* (Actinopterygii, Polypteriformes). *Journal of Zoology, London* **241**, 589-592.

Bazzoli, N. & Godinho, H.P. (1994). Cortical alveoli in oocytes of freshwater neotropical teleost fish. *Bolletino di zoologia* **61**, 301-308.

Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc B* **57**, 125-133.

Bhatia, V.N., Perlman, D.H., Costello, C.E. & McComb, M.E. (2009). Software tool for researching annotations of proteins (STRAP): open-source protein annotation software with data visualization. *Analytical Chemistry* **81**, 9819-9823.

Bian, L., Su, Y.Q. & Gaffney, P.M. (2016). Development of SNP markers for analysis of population structure in white perch (*Morone americana*) using double digest restriction site-associated DNA sequencing. *Conservation of Genetic Resources* **8**, 403-406.

Bian, X., Zhang, X., Sakurai, Y., Jin, X., Gao, T., Wan, R. & Yamamoto, J. (2014). Envelope surface ultrastructure and specific gravity of artificially fertilized Pacific cod *Gadus macrocephalus* eggs. *Journal of Fish Biology* **84**, 403-421.

Billard, R. (1986). Spermatogenesis and spermatology of some teleost fish species. *Reproduction Nutrition Development* **26**, 877-920.

Blanquer, A. (1990). Phylogeographie intraspecifique d'un poisson marin, le flet *Platichthys flesus* L. (Heterosomata). Polymorphisme des marqueurs nucleaires et mitochondriaux. Unpublished thesis (Ph.D), University of Montpellier.

Bocanegra, F.A., Tello, J.S., Chavez, C.V., Rodriguez, L.C., Kohler, C.C., Kohler, S.T. & Camargo, W.M. (2004). Pond culture of *Arapaima gigas* in the Peruvian Amazon. *World Aquaculture* **34**, 45-46.

Bonga, S.E.W. & Pang, P.K.T. (1991). Control of Calcium Regulating Hormones in the Vertebrates: Parathyroid Hormone, Calcitonin, Prolactin, and Stanniocalcin. *International Review of Cytology* **128**, 139-213.

Borella, M.I., Venturieri, R. & Mancera, J.M. (2009). Immunocytochemical identification of adenohypophyseal cells in the pirarucu (*Arapaima gigas*), an Amazonian basal teleost. *Fish Physiology and Biochemistry* **35**, 3-16.

Brauner, C.J. & Val, A.L. (1996). The interaction between O_2 and CO_2 exchange in the obligate air breather, *Arapaima gigas*, and the facultative air breather, *Lipossarcus pardalis* In: Val, A.L., Almeida-Val, V.M.F. & Randall, D.J., eds., Physiology and Biochemistry of the Fishes of the Amazon, Manaus: INPA, 101-110.

Brown-Peterson, N.J., Wyanski, D.M., Saborido-Rey, F., Macewicz, B.J. & Lowerre-Barbieri, S.K. (2011). A standardized terminology for describing reproductive development in fishes. *Marine and Coastal Fisheries: Dynamics, Management, and Ecosystem Science* **3**, 52-70.

Brown, J.K., Taggart, J.B., Bekaert, M., Wehner, S., Palaiokostas, C., Setiawan, A.N., Symonds, J.E. & Penman, D.J. (2016). Mapping the sex determination locus in the hāpuku (*Polyprion oxygeneios*) using ddRAD sequencing. *BMC Genomics* **17**, 1-12.

Buckley, J., Maunder, R.J., Foey, A., Pearce, J., Val, A.L. & Sloman, K.A. (2010). Biparental mucus feeding: a unique example of parental care in an Amazonian cichlid. *Journal of Experimental Biology* **213**, 3787-3795.

Carlisle, S.L., Marxer-Miller, S., K., Canario, A.V.M., Oliveira, R.F., Carneiro, L. & Grober, M.S. (2000). Effects of 11-ketotestosterone on genital papilla morphology in the sex changing fish *Lythrypnus dalli*. *Journal of Fish Biology* **57**, 445-456.

Carosfeld, J., Godinho, H.P., Zaniboni Filho, E. & Harvey, B.J. (2003). Cryopreservation of sperm in Brazilian migratory fish conservation. *Journal of Fish Biology* **63**, 472-489.

Carreiro, C.R.P. (2012). Inovações tecnológicas na sexagem, manejo reprodutivo e crescimento do pirarucu, *Arapaima gigas* (Schinz,1822), (Actinopterygii, Arapaimidae) cultivado no centro de pesquisas em aquicultura Rodolpho Von Ihering (cpa) do DNOCS, Pentecoste, Estado do Ceará. Unpublished thesis (PhD), Federal University of Ceará.

Carreiro, C.R.P., Furtado-Neto, M.A.A., Mesquita, P.E.C. & Bezerra, T.A. (2011). Sex determination in the Giant fish of Amazon Basin, *Arapaima gigas* (Osteoglossiformes, Arapaimatidae), using laparoscopy. *Acta Amazonica* **41**, 415-420.

Carvalho, D.C., Guedes, D., Trindade, M.G., Coelho, R.M.S. & Araujo, P.H.L. (2017a). Nationwide Brazilian governmental forensic programme reveals seafood mislabelling trends and rates using DNA barcoding. *Fisheries Research* **191**, 30-35.

Carvalho, F., Power, M., Forsberg, B.R., Castello, L., Martins, E.G. & Freitas, C.E.C. (2017b). Trophic Ecology of *Arapaima* sp. in a ria lake-river-floodplain transition zone of the Amazon. *Ecology of Freshwater Fish* **early view**.

Carvalho, F.R., Casatti, L., Manzotti, A.R. & Ravazzi, D.C.W. (2015). First record of *Arapaima gigas* (Schinz, 1822) (Teleostei: Osteoglossomorpha), the “pirarucu”, in the upper Paraná River basin, Southeast Brazil. *CheckList* **11**, 1729.

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

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

Castello, L., Arantes, C.C., Mcgrath, D.G., Stewart, D.J. & Sousa, F.S. (2015). Understanding fishing-induced extinctions in the Amazon. *Aquatic Conservation: Marine and Freshwater Ecosystems* **25**, 587-598.

Castello, L. & Stewart, D.J. (2010). Assessing CITES non-detriment findings procedures for *Arapaima* in Brazil. *J Appl Ichthyol* **26**, 49-56.

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

Catchen, J., Hohenlohe, P., Bassham, S., Amores, A. & Cresko, W. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology Notes* **22**, 3124-3140.

Cavero, B.A.S., Fonseca, F.A.L.d., Pereira-Filho, M., Ituassú, D.R., Bordinhon, A.M., Roubach, R. & Ono, E.A. (2004). Tolerância de juvenis de pirarucu ao aumento da concentração de amônia em ambiente confinado. *Pesquisa Agropecuária Brasileira, Brasília* **39**, 513-516.

Cavole, L.M., Arantes, C.C. & Castello, L. (2015). How illegal are tropical small-scale fisheries? An estimate for arapaima in the Amazon. *Fisheries Research* **168**, 1-5.

Cerdá, J., Fabra, M. & Raldúa, D. (2007). Physiological and molecular basis of fish oocyte hydration. In: Babin, P.J., Cerdá, J. & Lubzens, E., eds., The fish oocyte: From basic studies to biotechnological applications, Dordrecht, Netherlands: Springer, 349-396.

Chang, J.P. & Pemberton, J.G. (2017). Comparative aspects of GnRH-Stimulated signal transduction in the vertebrate pituitary – Contributions from teleost model systems. *Molecular and Cellular Endocrinology* **in press**.

Chi, L., Li, X., Liu, Q. & Liu, Y. (2017). Photoperiod regulate gonad development via kisspeptin/kissr in hypothalamus and saccus vasculosus of Atlantic salmon (*Salmo salar*). *PloS One* **12**, e0169569.

Chong, K., Joshi, S., Jin, L.T. & Chong Shu-Chien, A.C. (2005). Proteomics profiling of epidermal mucus secretion of a cichlid (*Symphysodon aequifasciata*) demonstrating parental care behavior. *Proteomics* **5**, 2251–2258.

Chu-Koo, F., Dugue, R., Alvan Aguilar, M., Casanova Daza, A., Alcantara Bocanegra, F., Chavez Veintemilla, C., Duponchelle, F., Renno, J.F., Tello, S. & Nunez, J. (2009). Gender determination in the Paiche or Pirarucu (*Arapaima gigas*) using plasma vitellogenin, 17β-estradiol, and 11-ketotestosterone levels. *Fish Physiology and Biochemistry* **35**, 125-36.

Colombo, G., Grandi, G. & Rossi, R. (1984). Gonad differentiation and body growth in *Anguilla anguilla* L. *Journal of Fish Biology* **24**, 215-228.

Coombs, S., Bleckmann, H., Fay, R.R. & Popper, A.N. (2014). The lateral line system, London: Springer.

Cordero, H., Brinchmann, M.F., Cuesta, A., Meseguer, J. & Esteban, M.A. (2015). Skin mucus proteome map of European sea bass (*Dicentrarchus labrax*). *Proteomics* **15**, 4007-4020.

Cordero, H., Morcillo, P., Cuesta, A., Brinchmann, M.F. & Esteban, M.A. (2016). Differential proteome profile of skin mucus of gilthead seabream (*Sparus aurata*) after probiotic intake and/or overcrowding stress. *Journal of Proteomics* **132** 41-50.

Cowan, M., Azpeleta, C. & López-Olmeda, J.F. (2017). Rhythms in the endocrine system of fish: a review. *Journal of Comparative Physiology B* 1-33.

Damme, P.A.V., Méndez, C.C., Carolsfeld, J. & Olden, J.D. (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**, 375-383.

Denovan-Wright, E.M., Ramsey, N.B., McCormick, C.J., Lazier, C.B. & Wright, J.M. (1996). Nucleotide sequence of transferrin cDNAs and tissue-specific expression of the transferrin gene in Atlantic cod (*Gadus morhua*). *Comparative Biochemistry and Physiology B* **113**, 269-273.

Derby, C.D. & Sorensen, P.W. (2008). Neural processing, perception, and behavioral responses to natural chemical stimuli by fish and crustaceans. *Journal of Chemical Ecology* **34**, 898-914.

Divers, S.J. (2010). Endoscopy equipment and instrumentation for use in exotic animal medicine. *Veterinary Clinics of North America: Exotic Animal Practice* **13**, 171-185.

Divers, S.J., Boone, S.S., Hoover, J.J., Boysen, K.A., Killgore, K.J., Murphy, C.E., George, S.G. & Camus, A.C. (2009). Field endoscopy for identifying gender, reproductive stage and gonadal anomalies in free-ranging sturgeon (*Scaphirhynchus*) from the lower Mississippi River. *Journal of Applied Ichthyology* **25**, 68-74.

Døving, K.B. & Lastein, S. (2009). The Alarm Reaction in Fishes—Odorants, Modulations of Responses, Neural Pathways. *Annals of the New York Academy of Sciences* **1170**, 413-423.

Dowling, T.E., Marsh, P.C., Kelsen, A.T. & Tibbets, C.A. (2005). Genetic monitoring of wild and repatriated populations of endangered razorback sucker (*Xyrauchen texanus*, Catostomidae, Teleostei) in Lake Mohave, Arizona-Nevada. *Molecular Ecology* **14**, 123-135.

Dufour, S., Sebert, M.E., Weltzien, F.A., Rousseau, K. & Pasqualini, C. (2010). Neuroendocrine control by dopamine of teleost reproduction. *Journal of Fish Biology* **76**, 129-160.

Dugue, R., Chu Koo, F., Alcantara, F., Duponchelle, F., Renno, J.F. & Nunez, J. (2008). Purification and assay of *Arapaima gigas* vitellogenin: potential use for sex determination. *Cybium* **32**, 111-111.

Duston, J. & Bromage, N. (1987). Constant photoperiod regimes and the entrainment of the annual cycle of reproduction in the female rainbow trout (*Salmo gairdneri*). *General and Comparative Endocrinology* **65**, 373-384.

Edmands, S. (2007). Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. *Molecular Ecology* **16**, 463-475.

Eigenmann, C.H. & Allen, W.R. (1942). Fishes of the Western South America, Lexington, Kentucky: The University of Kentucky.

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

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

Fänge, R., Larsson, A. & Lidman, U. (1972). Fluids and jellies of the acusticolateralis system in relation to body fluids in *Coryphaenoides rupestris* and other fishes. *Marine Biology* **17**, 180-185.

FAO (2016). Cultured Aquatic Species Information Programme: *Arapaima gigas* (Schinz, 1822) [online], Food and Agriculture Organization of the United Nations, available: http://www.fao.org/fishery/culturedspecies/Arapaima_gigas/en [accessed 26 Dec 2016].

Faria, M.T., Carvalho, R.F., Sevilhano, T.C., Oliveira, N.A., Silva, C.F., Oliveira, J.E., Soares, C.R., Garcez, R., Santo, P.R. & Bartolini, P. (2013). Isolation of the pituitary gonadotrophic α-subunit hormone of the giant amazonian fish: pirarucu (*Arapaima gigas*). *Fish Physiology and Biochemistry* **39**, 683-693.

Farias, I.P., Leão, A., Crossa, M., Almeida, Y.S., Honczaryk, A., Verba, J.T. & Hrbek, T. (2015). Evidence of polygamy in the socially monogamous Amazonian fish *Arapaima gigas* (Schinz, 1822) (Osteoglossiformes, Arapaimidae). *Neotrop Ichthyol* **13**, 195-204.

Farrel, A.P. & Randall, D.J. (1978). Air-breathing mechanics in two Amazonian teleosts, *Arapaima gigas* and *Hoplerythrinus unitaeniatus*. *Canadian Journal of Zoology* **56**, 939-945.

Fernald, R.D. & White, R.B. (1999). Gonadotropin-Releasing Hormone Genes: Phylogeny, Structure, and Functions. *Frontiers in Neuroendocrinology* **20**, 224-240.

Fernandes, M.N., Cruz, A.L., Costa, O.T. & Perry, S.F. (2012). Morphometric partitioning of the respiratory surface area and diffusion capacity of the gills and

swim bladder in juvenile Amazonian air-breathing fish, *Arapaima gigas*. *Micron* **43**, 961-970.

Fernandez-Palacios, H., Schuchardt, D., Roo, J., Hernández-Cruz, C.M. & Izquierdo, M. (2015). Multiple GnRHa injections to induce successful spawning of wild caught greater amberjack (*Seriola dumerili*) matured in captivity. *Aquaculture Research* **46**, 1748-1759.

Fischer, A.H., Jacobson, K.A., Rose, J. & Zeller, R. (2008). Hematoxylin and eosin staining of tissue and cell sections. *Cold Spring Harbor Protocols* **3**, 1-2.

Fogaça, F.H.S., Oliveira, E.G., Carvalho, S.E.Q. & Santos, J.F.S. (2011). Yield and composition of pirarucu fillet in different weight classes. *Acta Scientiarum Biological Sciences, Maringá* **33**, 95-99.

Fontenele, O. (1948). Contribuição para o conhecimento da biologia do Pirarucú, "*Arapaima gigas"* (Cuvier), em cativeiro (Actinopterygii, Osteoglossidae). *Revista Brasileira de Biologia* **8**, 445-459.

Fontenele, O. (1953). Habitos de desova do pirarucu, *Arapaima gigas* (Cuvier) (Pisces, Isospondyli, Arapaimidae), e evolução de sua larva, Fortaleza: DNOCS.

Forlano, P.M., Deitcher, D.L., Myers, D.A. & Bass, A.H. (2001). Anatomical distribution and cellular basis for high levels of aromatase activity in the brain of teleost fish: Aromatase enzyme and mRNA expression identify glia as source. *Journal of Neuroscience* **21**, 8943-8955.

Forniés, M.A., Carrillo, M., Mañanós, E., Sorbera, L.A., Zohar, Y. & Zanuy, S. (2003). Relative potency of the forms of GnRH and their analogs on LH release in sea bass. *Journal of Fish Biology* **63**, 73-89.

Froese, R. (2006). Cube law, condition factor and weight–length relationships: history, meta-analysis and recommendations. *Journal of Applied Ichthyology* **22**, 241-253.

García-Castillo, J., Pelegrín, P., Mulero, V. & Meseguer, J. (2002). Molecular cloning and expression analysis of tumor necrosis factor α from a marine fish reveal its constitutive expression and ubiquitous nature. *Immunogenetics* **54**, 200-207.

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

Godinho, H.P., Santos, J.E., Formagio, P.S. & Guimarães-Cruz, R.J. (2005). Gonadal morphology and reproductive traits of the Amazonian fish *Arapaima gigas* (Schinz, 1822). *Acta Zoologica, Stockolm* **86**, 289-294.

Golan, M., Zelinger, E., Zohar, Y. & Levavi-Sivan, B. (2015). Architecture of GnRH-gonadotrope-vasculature reveals a dual mode of gonadotropin regulation in fish. *Neuroendocrinology* **156**, 4163-4173.

Gonzalez, R.J., Brauner, C.J., Wang, Y.X., Richards, J.G., Patrick, M.L., Xi, W., Matey, V. & Val, A.L. (2010). Impact of ontogenetic changes in branchial morphology on gill function in *Arapaima gigas*. *Physiological and Biochemical Zoology* **83**, 322-332.

Good, D.M., Zurbig, P., Argiles, A., Bauer, H.W., Behrens, G., Coon, J.J., Dakna, M., Decramer, S., Delles, C., Dominiczak, A.F., Ehrich, J.H.H., Eitner, F., Fliser, D., Frommberger, M., Ganser, A., Girolami, M.A., Golovko, I., Gwinner, W., Haubitz, M., Herget-Rosenthal, S., Jankowski, J., Jahn, H., Jerums, G., Julian,

B.A., Kellmann, M., Kliem, V., Kolch, W., Krolewski, A.S., Luppi, M., Massy, Z., Melter, M., Neususs, C., Novak, J., Peter, K., Rossing, K., Rupprecht, H., Schanstra, J.P., Schiffer, E., Stolzenburg, J.U., Tarnow, L., Theodorescu, D., Thongboonkerd, V., Vanholder, R., Weissinger, E.M., Mischak, H. & Schmitt-Kopplin, P. (2010). Naturally occurring human urinary peptides for use in diagnosis of chronic kidney disease. *Molecular & Cellular Proteomics* **9**, 2424-2437.

Greenwood, P.H. & Liem, K.F. (1984). Aspiratory respiration in *Arapaima gigas* (Teleostei, Osteoglossomorpha): a reappraisal. *Journal of Zoology*, *London* **203**, 411-425.

Grier, H.J. (2012). Development of the follicle complex and oocyte staging in Red Drum, *Sciaenops ocellatus* Linnaeus, 1776 (Perciformes, Sciaenidae). *Journal of Morphology* **273**, 801-829.

Grier, H.J., Aranzábal, M.C.U. & Patiño, R. (2009). The ovary, folliculogenesis, and oogenesis in teleosts. In: Jamieson, B.G.M., ed., Reproductive biology and phylogeny of fishes (Agnathans and Bony Fishes), Enfield, New Hampshire: Science Publishers, 24-85.

Guzmán, J.M., Cal, R., García-López, A., Chereguini, O., Kight, K., Olmedo, M., Sarasquete, C., Mylonas, C.C., Peleteiro, J.B., Zohar, Y. & Mañanós, E.L. (2011). Effects of in vivo treatment with the dopamine antagonist pimozide and gonadotropin-releasing hormone agonist (GnRHa) on the reproductive axis of Senegalese sole (*Solea senegalensis*). *Comparative Biochemistry and Physiology A* **158**, 235-245.

Guzmán, J.M., Ramos, J., Mylonas, C.C. & Mañanós, E.L. (2009). Spawning performance and plasma levels of GnRHa and sex steroids in cultured female

Senegalese sole (*Solea senegalensis*) treated with different GnRHa-delivery systems. *Aquaculture (Amsterdam, Netherlands)* **291**, 200-209.

Halverson, M. (2013). Manual de boas práticas de reprodução do Pirarucu em cativeiro, Brasília: Sebrae.

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

Hara, T.J. (1994). The diversity of chemical stimulation in fish olfaction and gustation. *Reviews in Fish Biology and Fisheries* **4**, 1-35.

Haslewood, G.A.D. & Tökés, L. (1972). Comparative studies of bile salts, a new type of bile salt from *Arapaima gigas* (Cuvier) (Family Osteoglossidae). *Biochemistry Journal* **126**, 1161-1170.

Hayakawa, E., Landuyt, B., Baggerman, G., Cuyvers, R., Lavigne, R., Luyten, W. & Schoofs, L. (2013). Peptidomic analysis of human reflex tear fluid. *Peptides* **42**, 63-9.

Hernandez-Divers, S.J., Bakal, R.S., Hickson, B.H., Rawlings, C.A., Wilson, H.G., Radlinsky, M., Hernandez-Divers, S.M. & Dover, S.R. (2004). Endoscopic sex determination and gonadal manipulation in Gulf of Mexico Sturgeon (*Acipenser oxyrinchus desotoi*). *Journal of Zoo and Wildlife Medicine* **35**, 459-470.

Hilton, E.J. (2003). Comparative osteology and phylogenetic systematics of fossil and living bony-tongue fishes (Actinopterygii, Teleostei, Osteoglossomorpha). *Zoological Journal of the Linnean Society* **137**, 1-100.

Hiramatsu, N., Todo, T., Sullivan, C.V., Schilling, J., Reading, B.J., Matsubara, T., Ryua, Y.W., Mizuta, H., Luo, W., Nishimiya, O., Wu, M., Mushirobira, Y., Yilmaz, O. & Hara, A. (2015). Ovarian yolk formation in fishes: Molecular mechanisms

underlying formation of lipid droplets and vitellogenin-derived yolk proteins. *General and Comparative Endocrinology* **221**, 9-15.

Hohenlohe, P.A., Catchen, J. & Cresko, W.A. (2012). Population genomic analysis of model and nonmodel organisms using sequenced RAD tags. In: Pompanon, F. & Bonin, A., eds., Data Production and Analysis in Population Genomics: Methods and Protocols, Methods in Molecular Biology, New York: Springer, 235-260.

Holbrook, R.I. (2011). Comment on 'Biparental mucus feeding: a unique example of parental care in an Amazonian cichlid'. *The Journal of Experimental Biology* **214**, 1213-1214.

Holtta, M., Minthon, L., Hansson, O., Holmen-Larsson, J., Pike, I., Ward, M., Kuhn, K., Ruetschi, U., Zetterberg, H., Blennow, K. & Gobom, J. (2015). An integrated workflow for multiplex CSF proteomics and peptidomics-identification of candidate cerebrospinal fluid biomarkers of Alzheimer's disease. *Journal of Proteome Research* **14**, 654-63.

Honczaryk, A. & Inoue, L.A.K.A. (2009). Anesthesia in pirarucu by benzocaine sprays in the gills. *Ciencia Rural, Santa Maria* **40**, 1-4.

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**, 909-917.

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.

Huang, C.J., Lee, M.S., Huang, F.L. & Chang, G.D. (1995). A protease inhibitor of the serpin family is a major protein in carp perimeningeal fluid: II. cDNA cloning,

sequence analysis, and *Escherichia coli* expression. *Journal of Neurochemistry* **64**, 1721-1727.

Huertas, M., Almeida, O.G., Canário, A.V.M. & Hubbard, P.C. (2014). Tilapia male urinary pheromone stimulates female reproductive axis. *General and Comparative Endocrinology* **196**, 106-111.

Hurvitz, A., Jackson, K., Degani, G. & Levavi-Sivan, B. (2007). Use of endoscopy for gender and ovarian stage determinations in Russian sturgeon (*Acipenser gueldenstaedtii*) grown in aquaculture. *Aquaculture (Amsterdam, Netherlands)* **270**, 158-166.

Idahor, K.O. (2014). Microscopic observation of spermatozoa in milt collected with syringe without sacrificing the male African Catfish (*Clarias anguillaris* B. 1911). *International Journal of Fisheries and Aquatic Studies* **2**, 88-9.

Imbiriba, E.P. (2001). Potencial de Criacao de Pirarucu, *Arapaima gigas*, em Cativeiro. *Acta Amazonica* **31**, 299-316.

Iq, K.C. & Shu-Chien, A.C. (2011). Proteomics of Buccal Cavity Mucus in Female Tilapia Fish (*Oreochromis* spp.): A Comparison between Parental and Non-Parental Fish. *PloS One* **6**, e18555.

Isaú, Z.A., Rizzo, E. & Amaral, T.B. (2011). Structural analysis of oocytes, post-fertilization events and embryonic development of the Brazilian endangered teleost *Brycon insignis* (Characiformes). *Zygote* **21**, 85-94.

Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J. & Mann, M. (2005). Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Molecular & Cellular Proteomics* **4**, 1265-1272.

Jeffries, D.L., Copp, G.H., Handley, L.L., Olsén, K.H., Sayer, C.D. & Hänfling, B. (2016). Comparing RADseq and microsatellites to infer complex phylogeographic patterns, an empirical perspective in the Crucian carp, *Carassius carassius*, L. *Molecular Ecology* **25**, 2997-3018.

Johnston, L.D., Brown, G., Gauthier, D., Reece, K., Kator, H. & Veld, P.V. (2008). Apolipoprotein A-I from striped bass (*Morone saxatilis*) demonstrates antibacterial activity in vitro. *Comp Biochem and Physiol B* **151**, 167-175.

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403-1405.

Jones, M.B. & Simons, M.J. (1983). Latitudinal variation in reproductive characteristics of a mud crab, *Helice crassa* (Grapsidae). *Bulletin of Marine Science* **33**, 656-670.

Jurado, J., Fuentes-Almagro, C.A., Guardiola, F.A., Cuesta, A., Esteban, A.M. & Álamo, M.J.P. (2015). Proteomic profile of the skin mucus of farmed gilthead seabream (*Sparus aurata*). *Journal of Proteomics* **120** 21-34.

Kai, W., Nomura, K., Fujiwara, A., Nakamura, Y., Yasuike, M., Ojima, N., Masaoka, T., Ozaki, A., Kazeto, Y., Gen, K., Nagao, J., Tanaka, H., Kobayashi, T. & Ototake, M. (2014). A ddRAD-based genetic map and its integration with the genome assembly of Japanese eel (*Anguilla japonica*) provides insights into genome evolution after the teleost-specific genome duplication. *BMC Genomics* **15**, 1-16.

Karnovsky, M.J. (1965). A formaldehyde–glutaraldehyde fixative of high osmolarity for use in eléctron microscopy. *Journal of Cell Biology* **27**, 137-138A.

Katagiri, T., Hirono, I. & Aoki, T. (1999). Molecular analysis of complement component C8beta and C9 cDNAs of Japanese flounder, *Paralichthys olivaceus*. *Immunogenetics* **50**, 43-48.

Keller-Costa, T., Canário, A.V.M. & Hubbard, P.C. (2015). Chemical communication in cichlids: A mini-review. *General and Comparative Endocrinology* **15**, 64-74.

Kelly, G.M., Lai, C.J. & Moon, R.T. (1993). Expression of wnt10a in the central nervous system of developing zebrafish. *Developmental Biology* **158**, 113-121.

Kembenya, E.M., Marcial, H.S., Outa, N.O., Sakakura, Y. & Hagiwara, A. (2016). Captive Breeding of Threatened African Carp, Labeo victorianus, of Lake Victoria. *Journal of the World Aquaculture Society* **early view**.

Kennedy, J., Gundersen, A.C., Høines, A.S. & Kjesbu, O.S. (2011). Greenland halibut (*Reinhardtius hippoglossoides*) spawn annually but successive cohorts of oocytes develop over 2 years, complicating correct assessment of maturity. *Canadian Journal of Fisheries and Aquatic Sciences* **68**, 201-209.

Kessels, M.Y., Huitema, L.F.A., Boeren, S., Kranenbarg, S., Schulte-Merker, S., Leeuwen, J.L. & Vries, S.C. (2014). Proteomics analysis of the zebrafish skeletal extracellular matrix. *PloS One* **9**, e90568.

Khong, H.K., Kuah, M.K., Jaya-Ram, A. & Shu-Chien, A.C. (2009). Prolactin receptor mRNA is upregulated in discus fish (*Symphysodon aequifasciata*) skin during parental phase. *Comp Biochem and Physiol B* **153**, 18-28.

Kim, D.K., Cho, E.B., Moon, M.J., Park, S., Hwang, S.I., Kah, O., Sower, S.A., Vaudry, H. & Seong, J.Y. (2011). Revisiting the evolution of gonadotropin-releasing hormones and their receptors in vertebrates: Secrets hidden in genomes. *General and Comparative Endocrinology* **170**, 68-78.

Kobayashi, M., Sorensen, P.W. & Stacey, N. (2002). Hormonal and pheromonal control of spawning behavior in the goldfish. *Fish Physiology and Biochemistry* **26**, 71-84.

Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A. & Mayrose, I. (2015). CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* **15**, 1179-1191.

Kumar, S., Stecher, G. & Tamura, K. (2016). Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**, 1870-1874.

Kynard, B. & Kieffe, M. (2002). Use of a borescope to determine the sex and egg maturity stage of sturgeons and the effect of borescope use on reproductive structures. *Journal of Applied Ichthyology* **18**, 505-508.

Lambris, J.D., Lao, Z., Pang, J. & Alsenz, J. (1993). Third Component of Trout Complement. cDNA Cloning and Conservation of Functional Sites. *Journal of Immunology* **151**, 6123-6134.

Larsson, D.G.J., Mylonas, C.C., Zohar, Y. & Crim, L.W. (1997). Gonadotropin-releasing hormone analogue (GnRH-A) induces multiple ovulations of high-quality eggs in a cold-water, batch-spawning teleost, the yellowtail flounder (*Pleuronectes ferrugineus*). *Canadian Journal of Fisheries and Aquatic Sciences* **54**, 1957-1964.

Lavoue, S. (2016). Was Gondwanan breakup the cause of the intercontinental distribution of Osteoglossiformes? A time-calibrated phylogenetic test combining molecular, morphological, and paleontological evidence. *Molecular Phylogenetics and Evolution* **99** 34-43.

Lawson, L.L., Tuckett, Q.M., Lawson, K.M., Watson, C.A. & Hill, J.E. (2015). Lower lethal temperature for Arapaima *Arapaima gigas*: Potential implications for culture and establishment in Florida. *North American Journal of Aquaculture* **77**, 497-502.

Le Menn, F., Cerdá, J. & Babin, P.J. (2007). Ultrastructural aspects of the ontogeny and differentiation of ray-finned fish ovarian follicles. In: Babin, P.J., Cerdá, J. & Lubzens, E., eds., The fish oocyte: From Basic Studies to Biotechnological Applications, Dordrecht, Netherlands: Springer, 1-37.

Lefevre, S., Bayley, M., Mckenzie, D.J. & Craig, J.F. (2014). EDITORIAL: Air-breathing fishes. *Journal of Fish Biology* **84**, 547-553.

Léger, S. & Brand, M. (2002). Fgf8 and Fgf3 are required for zebrafish ear placode induction, maintenance and inner ear patterning. *Mechanisms of Development* **119**, 91-108.

Lehtonen, T.K., Lindström, K. & Wong, B.B.M. (2015). Body size mediates social and environmental effects on nest building behaviour in a fish with paternal care. *Oecologia* **178**, 699-706.

Levavi-Sivan, B., Bogerd, J., Mañanós, E.L., Gómez, A. & Lareyre, J.J. (2010). Perspectives on fish gonadotropins and their receptors. *General and Comparative Endocrinology* **165**, 412-437.

Lima, A.F., Rodrigues, A.P.O., Varela, E.S., Torati, L.S. & Maciel, P.O. (2015a). Pirarucu culture in the Brazilian Amazon. *The Global Aquaculture Advocate* **26-29**, 54-56.

Lima, A.F., Varela, E.S., Maciel, P.O., Alves, A.L., Rodrigues, A.P.O., Torati, L.S., Matavelli, M. & Bezerra, T.A. (2015b). Manejo de plantel de reprodutores de Pirarucu, Brasília: EMBRAPA.

Lin, Y.S., Wei, C.T., Olevsky, E.A. & Meyers, M.A. (2011). Mechanical properties and the laminate structure of *Arapaima gigas* scales. *Journal of the Mechanical Behavior of Biomedical Materials* **4**, 1145-56.

Lubzens, E., Bobe, J., Young, G. & Sullivan, C.V. (2017). Maternal investment in fish oocytes and eggs: The molecular cargo and its contributions to fertility and early development. *Aquaculture (Amsterdam, Netherlands)* **472**, 107-143.

Lubzens, E., Young, G., Bobe, J. & Cerdà, J. (2010). Oogenesis in teleosts: How fish eggs are formed. *General and Comparative Endocrinology* **165**, 367-389.

Lüling, K.H. (1964). Zur Biologie und Ökologie von *Arapaima gigas* (Pisces, Osteoglossidae). *Zeitschrift fuer Morphologie und Oekologie der Tiere* **54**, 436-530.

Lüling, K.H. (1969). Das Laichverhalten der Vertreter der Familie Osteoglossidae (Versucheiner Ubersicht). *Bonner Zoologische Beiträge* **20**, 228-243.

Lundberg, J.G. & 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.

Lynch, M. & Ritland, K. (1999). Estimation of pairwise relatedness with molecular markers. *Genetics* **152**, 1753-1766.

Maitra, S.K., Chattoraj, A., Mukherjee, S. & Moniruzzaman, M. (2013). Melatonin: A potent candidate in the regulation of fish oocyte growth and maturation. *General and Comparative Endocrinology* **181**, 215-222.

Mañanós, E., Duncan, N. & Mylonas, C.C. (2008). Reproduction and control of ovulation, spermiation and spawning in cultured fish. In: Cabrita, E., Robles, V. & Herráez, M.P., eds., Methods in Reproductive Aquaculture: Marine and Freshwater SpeciesTaylor and Francis Group CRC Press.

Manousaki, T., Tsakogiannis, A., Taggart, J.B., Tsaparis, D., Lagnel, J., Chatziplis, D., Palaiokostas, C., Magoulas, A., Papandroulakis, N., Mylonas, C.C. & Tsigenopoulos, C.S. (2016). Exploring a nonmodel teleost genome through RAD

sequencing—linkage mapping in common Pandora, *Pagellus erythrinus* and comparative genomic analysis. *Genes, Genomes and Genetics* **6**, 509-519.

Manseth, E., Skjervold, P.O. & Flengsrud, R. (2004). Sample displacement chromatography of Atlantic Salmon (*Salmo salar*) thrombin. *Journal of Biochemical and Biophysical Methods* **60**, 39-47.

Mansor, R., Mullen, W., Albalat, A., Zerefos, P., Mischak, H., Barrett, D.C., Biggs, A. & Eckersall, P.D. (2013). A peptidomic approach to biomarker discovery for bovine mastitis. *Journal of Proteomics* **85**, 89-98.

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

McFarland, K.A., Topczewska, J.M., Weidinger, G., Dorsky, R.I. & Appel, B. (2008). Hh and Wnt signaling regulate formation of olig2^{+} zebrafish cerebellum. *Developmental Biology* **318**, 162-171.

Melamed, P., Xue, Y., Poon, J.F.D., Wu, Q., Xie, H., Yeo, J., Foo, T.W.J. & Chua, H.K. (2005). The male seahorse synthesizes and secretes a novel C-type lectin into the brood pouch during early pregnancy. *FEBS Journal* **272** 1221-1235.

Menezes, R.S. (1951). Biological and Economical Notes on the Pirarucu *Arapaima gigas* (Actinopterygii, Arapaimidae), Rio de Janeiro, Brazil Ministério da Agricultura.

Meyers, M.A., Lin, Y.S., Olevsky, E.A. & Chen, P.Y. (2012). Battle in the Amazon: Arapaima versus Piranha. *Advanced Engineering Materials* **14**, B279-B288.

MICES (2017). ALICEWEB-Data on imported fish in Brazil from 01/2012 to 12/2016. [online], available: http://aliceweb.mdic.gov.br/ [accessed March/2017].

Migaud, H., Bell, G., Cabrita, E., McAndrew, B., Davie, A., Bobe, J., Herráez, M.P. & Carrillo, M. (2013). Gamete quality and broodstock management in temperate fish. *Reviews in Aquaculture* **5**, S194-S223.

Migaud, H., Davie, A. & Taylor, J.F. (2010). Current knowledge on the photoneuroendocrine regulation of reproduction in temperate fish species. *Journal of Fish Biology* **76**, 27-68.

Migaud, H., Fontaine, P., Sulistyo, I., Kestemont, P. & Gardeur, J.N. (2002). Induction of out-of-season spawning in Eurasian perch *Perca fluviatilis*: effects of rates of cooling and cooling durations on female gametogenesis and spawning. *Aquaculture (Amsterdam, Netherlands)* **205**, 253-267.

Migaud, H., Wang, N., Gardeur, J.N. & Fontaine, P. (2006). Influence of photoperiod on reproductive performances in Eurasian perch *Perca fluviatilis*. *Aquaculture (Amsterdam, Netherlands)* **252**, 385-393.

Migdalski, E.C. (1957). Contribution to the Life History of the South American Fish *Arapaima gigas*. *Copeia* **1**, 54-56.

Miller-Bertoglio, V.E., Mullins, M.C. & Halpern, M.E. (1997). Differential regulation of chordin expression domains in mutant Zebrafish. *Developmental Biology* **192**, 537-550.

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**, 129-138.

Mommens, M., Lanes, C.F.C. & Babiak, I. (2013). Egg yolk nutritional constituents as indicators of egg quality in Atlantic halibut (*Hippoglossus hippoglossus* L.). *Aquaculture Research* 1-11.

Monentcham, S.E., Kouam, J., Pouomogne, V. & Kestemont, P. (2009). Biology and prospect for aquaculture of African bonytongue, *Heterotis niloticus* (Cuvier, 1829): A review. *Aquaculture (Amsterdam, Netherlands)* **289**, 191-198.

Monteiro, L.B.B., Soares, M.C.F., Catanho, M.T.J. & Honczaryk, A. (2010). Reproductive aspects and sexual steroids hormonal profiles of Pirarucu, *Arapaima gigas* (Schinz,1822), in captivity conditions. *Acta Amazonica* **40**, 435-450.

Moore, J.S., Bourret, V., Dionne, M., Bradbury, I., O' Reilly, P., Kent, M., Chaput, G. & Bernatchez, L. (2014). Conservation genomics of anadromous Atlantic Salmon across its north american range: Outlier loci identify the same patterns of population structure as neutral loci. *Molecular Ecology* **23**, 5680-5697.

Mueller, O. (2006). *Arapaima gigas*, market study, current status of Arapaima global trade and perspectives on the Swiss, French and UK markets, United Nations Conference on Trade and Development.

Mukherjee, D., Majumdera, S., Moulik, S.R., Pal, P., Gupta, S., Guha, P. & Kumar, D. (2017). Membrane receptor cross talk in gonadotropin-, IGF-I-, and insulin-mediated steroidogenesis in fish ovary: An overview. *General and Comparative Endocrinology* **240**, 10-18.

Mylonas, C.C., Bridges, C., Gordin, H., Ríos, A.B., García, A., Gándara, F., Fauvel, C., Suquet, M., Medina, A., Papadaki, M., Heinisch, G., Metrio, G., Corriero, A., Vassallo-Agius, R., Guzmán, J.M., Mañanós, E. & Zohar, Y. (2007). Preparation and administration of gonadotropin-releasing hormone agonist (GnRHa) implants for the artificial control of reproductive maturation in captive-reared Atlantic Bluefin tuna (*Thunnus thynnus thynnus*). *Reviews in Fisheries Science* **15**, 183-210.

Mylonas, C.C., Duncan, N.J. & Asturiano, J.F. (2017). Hormonal manipulations for the enhancement of spermproduction in cultured fish and evaluation of sperm quality. *Aquaculture (Amsterdam, Netherlands)* **472**, 21-44.

Mylonas, C.C., Fostier, A. & Zanuy, S. (2010). Broodstock management and hormonal manipulations of fish reproduction. *General and Comparative Endocrinology* **165**, 516-534.

Mylonas, C.C., Sigelaki, I., Duncan, N.J., Fatira, E., Karkut, P. & Papadaki, M. (2015). Reproduction of hatchery-produced meagre *Argyrosomus regius* in captivity III. Comparison between GnRHa implants and injections on spawning kinetics and egg/larval performance parameters. *Aquaculture (Amsterdam, Netherlands)* **448**, 44-53.

Mylonas, C.C. & Zohar, Y. (2001). Use of GnRHa-delivery systems for the control of reproduction in fish. *Reviews in Fish Biology and Fisheries* **10**, 463-491.

Mylonas, C.C. & Zohar, Y. (2007). Promoting oocyte maturation, ovulation and spawning in farmed fish. In: Babin, P.J., Cerdà, J. & Lubzens, E., eds., The fish oocyte: From basic studies to biotechnological applications, Dordrecht, Netherlands: Springer, 437-474.

Nelson, J.S., Grande, T.C. & Wilson, M.V.H. (2016). Fishes of the world, 5th ed., New Jersey: Wiley.

Neves, A.M.B. (1995). Conhecimento atual sobre o Pirarucu, *Arapaima gigas* (Cuvier 1817). *Boletim do Museu Paraense Emílio Goeldi* **11**, 33-56.

Nissling, A., Florin, A.B., Thorsen, A. & Bergström, U. (2013). Egg production of turbot, *Scophthalmus maximus*, in the Baltic Sea. *Journal of Sea Research* **84**, 77-86.

Noakes, D.L.G. (1979). Parent touching behavior by young fishes: incidence, function and causation. *Environmental Biology of Fishes* **4**, 389-400.

Noce, I.D., Carra, S., Brusegan, C., Critelli, R., Frassine, A., Lorenzo, C., Giordano, A., Bellipanni, G., Villa, E., Cotelli, F., Pistocchi, A. & Schepis, F. (2015). The Coiled-Coil Domain Containing 80 (ccdc80) Gene Regulates gadd45b2 Expression in the Developing Somites of Zebrafish as a New Player of the Hedgehog Pathway. *Journal of Cellular Physiology* **230**, 821-830.

Nunes, E.S.C.L., Franco, R.M., Mársico, E.T., Nogueira, E.B., Neves, M.S. & Silva, F.E.R. (2012). Indigenous bacteria and pathogens in salted and dried Pirarucu (*Arapaima gigas* Shing, 1822) traded in Belém city, Pará. *Revista Brasileira de Ciências Veterinarias* **19**, 98-103.

Núñez, J., Chu-Koo, F., Berland, M., Arévalo, L., Ribeyro, O., Duponchelle, F. & Renno, J. (2011). Reproductive success and fry production of the paiche or pirarucu, *Arapaima gigas* (Schinz), in the region of Iquitos, Perú. *Aquaculture Research* **42**, 815-822.

Núñez, J. & Duponchelle, F. (2009). Towards a universal scale to assess sexual maturation and related life history traits in oviparous teleost fishes. *Fish Physiology and Biochemistry* **35**, 167-80.

Núñez, J., Duponchelle, F., Cotrina-Doria, M., Renno, J.F., Chavez-Veintimilla, C., Rebaza, C., Deza, S., Dávila, C.G., Chu-Koo, F., Tello, S. & Baras, E. (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.

Ogawa, M. (1970). Effects of prolactin on the epidermal mucous cells of the goldfish, *Carassius auratus* L. *Canadian Journal of Zoology* **48**, 501-503.

Ogawa, S. & Parhar, I.S. (2013). Anatomy of the kisspeptin systems in teleosts. *General and Comparative Endocrinology* **181**, 169-174.

Ohta, H., Kagawa, H., Tanaka, H., Okuzawa, K., Iinuma, N. & Hirose, K. (1997). Artificial induction of maturation and fertilization in the Japanese eel, *Anguilla japonica*. *Fish Physiology and Biochemistry* **17**, 163-169.

Okubo, K. & Nagahama, Y. (2008). Structural and functional evolution of gonadotropin-releasing hormone in vertebrates. *Acta Physiologica* **193**, 3-15.

Oliveira, E.G., Pinheiro, A.B., Oliveira, V.Q., Silva Jr., A.R.M., Moraes, M.G., Rocha, I.R.C.B., Sousa, R.R. & Costa, F.H.F. (2012). Effects of stocking density on the performance of juvenile pirarucu (*Arapaima gigas*) in cages. *Aquaculture (Amsterdam, Netherlands)* **370-71**, 96-101.

Oliveira, V., Poleto, S.L. & Venere, P.C. (2005). Feeding of juvenile pirarucu (*Arapaima gigas*, Arapaimidae) in their natural environment, lago Quatro Bocas, Araguaiana-MT, Brazil. *Neotropical Ichthyology* **3**, 312-314.

Oliveros, J.C. (2007). Venny. An interactive tool for comparing lists with Venn's diagrams [online], available: http://bioinfogp.cnb.csic.es/tools/venny/index.html [accessed

Olsen, J.E.B. (2014). Predation Avoidance Mechanisms of Juvenile *Arapaima* spp.: Significance of Synchronized Breathing and Sound Production. Unpublished thesis (Bachelor), SUNY College of Environmental Science and Forestry.

Ortega, H., Mojica, J.I., Alonso, J.C. & Hidalgo, M. (2006). Listado de los peces de la cuenca del río Putumayo en su sector colombo – peruano. *Biota Colombiana* **7**, 95-112.

Palaiokostas, C., Bekaert, M., Davie, A., Cowan, M.E., Oral, M., Taggart, J.B., Karim Gharbi, K., McAndrew, B.J., Penman, D., J. & Migaud, H. (2013). Mapping the sex determination locus in the Atlantic halibut (*Hippoglossus hippoglossus*) using RAD sequencing. *BMC Genomics* **14**, 1-12.

Palaiokostas, C., Bekaert, M., Khan, M.G., Taggart, J.B., Gharb, K., McAndrew, B.J. & Penman, D.J. (2015). A novel sex-determining QTL in Nile tilapia (*Oreochromis niloticus*). *BMC Genomics* **16**, 1-10.

Patiño, R. & Sullivan, C.V. (2002). Ovarian follicle growth, maturation, and ovulation in teleost fish. *Fish Physiology and Biochemistry* **26**, 57-70.

Paulo, J.A. (2014). Practical and Efficient Searching in Proteomics: A Cross Engine Comparison. *Webmedcentral* **4**, 1-15.

Peakall, R. & Smouse, P.E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288-295.

Peakall, R. & Smouse, P.E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* **28**, 2537-2539.

Pereira-Filho, M., Cavero, B.A.S., Roubach, R., Ituassú, D.R., Gandra, A.L. & Crescêncio, R. (2003). Cultivo do Pirarucu (*Arapaima gigas*) em viveiro escavado. *Acta Amazonica* **33**, 715-718.

Petersen, T.A., Brum, S.M., Rossoni, F., Silveira, G.F.V. & Castello, L. (2016). Recovery of *Arapaima* sp. populations by community-based management in floodplains of the Purus River, Amazon. *Journal of Fish Biology* **89**, 241-248.

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S. & Hoekstra, H.E. (2012). Double Digest RADseq: An Inexpensive Method for SNP Discovery and Genotyping in Model and Non-Model Species. *PloS One* **7**, 1-11.

Pontillo, C., Filip, S., Borr, D.M., Mullen, W., Vlahou, A. & Mischak, H. (2015). CE-MS-based proteomics in biomarker discovery and clinical application. *Proteomics: Clinical Applications* **9**, 322-334.

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

Quay, W. (1972). Integument and the environment: glandular composition, function, and evolution. *American Zoologist* **12**, 95-108.

Queiroz, H.L. (2000). Natural history and conservation of Pirarucu, *Arapaima gigas*, at the Amazonian Várzea: Red giants in muddy waters. Unpublished thesis (PhD), University of St Andrews.

R-Core-Team (2014). R: A language and environment for statistical computing.

Rajan, B., Fernandes, J.M.O., Caipang, C.M.A., Kiron, V., Rombout, J.H.W.M. & Brinchmann, M.F. (2011). Proteome reference map of the skin mucus of Atlantic cod (*Gadus morhua*) revealing immune competent molecules. *Fish & Shellfish Immunology* **31**, 224-231.

Ramos, C.A., Fernandes, M.N., da Costa, O.T. & Duncan, W.P. (2013). Implications for osmorespiratory compromise by anatomical remodeling in the gills of *Arapaima gigas*. *The Anatomical Record* **296**, 1664-1675.

Randall, D.J., Farrel, A.P. & Haswell, M.S. (1978). Carbon dioxide excretion in the pirarucu (*Arapaima gigas*), an obligate air-breathing fish. *Canadian Journal of Zoology* **56**, 977-982.

Rasotto, M.B. & Shapiro, D.Y. (1998). Morphology of gonoducts and male genital papilla, in the bluehead wrasse: implications and correlates on the control of gamete release. *Journal of Fish Biology* **52**, 716-725.

Rebouças, P.M., Maciel, R.L., Costa, B.G.B., Galvão, J.A.S. & Barbosa Filho, J.A.D. (2014). Analysis of the welfare of broodstock *Arapaima gigas* (Schinz, 1822) by length-weight relationship, condition factor and fry production. *Bioscience Journal* **30**, 873-881.

Resink, J.W., Van Den Berg, T.W.M., Van Den Hurk, R., Huisman, E.A. & Van Oordt, P.G.W.J. (1989). Induction of gonadotropin release and ovulation by pheromones in the african catfish, *Clarius gariepinus*. *Aquaculture (Amsterdam, Netherlands)* **83**, 167-177.

Rhody, N.R. (2014). Optimisation of common snook *Centropomus undecimalis* broodstock management. Unpublished thesis (PhD), University of Stirling.

Rhody, N.R., Davie, A., Zmora, N., Zohar, Y., Main, K.L. & Migaud, H. (2015). Influence of tidal cycles on the endocrine control of reproductive activity in common snook (*Centropomus undecimalis*). *General and Comparative Endocrinology* **224**, 247-259.

Rhody, N.R., Neidig, C.L., Grier, H.J., Main, K.L. & Migaud, H. (2013). Assessing reproductive condition in captive and wild Common Snook stocks: A comparison between the wet mount technique and histological preparations. *Transactions of the American Fisheries Society* **142**, 979-988.

Rhody, N.R., Puchulutegui, C., Taggart, J.B., Main, K.L. & Migaud, H. (2014). Parental contribution and spawning performance in captive common snook *Centropomus undecimalis* broodstock. *Aquaculture (Amsterdam, Netherlands)* **432**, 144-153.

Rice, W.R. (1989). Analyzing tables of statistical tests. *Evolution* **43**, 223-225.

Rime, H., Guitton, N., Pineau, C., Bonnet, E., Bobe, J. & Jalabert, B. (2004). Post-ovulatory ageing and egg quality: a proteomic analysis of rainbow trout coelomic fluid. *Reproductive Biology and Endocrinology* **2**, 1-10.

Rinchard, J., Kestemont, P., Kühn, E.R. & Fostier, A. (1993). Seasonal changes in plasma levels of steroid hormones in an asynchronous fish the Gudgeon *Gobio gobio* L. (Teleostei: Cyprinidae). *General and Comparative Endocrinology* **92**, 168-178.

Roch, G.J., Busby, E.R. & Sherwood, N.M. (2014). GnRH receptors and peptides: Skating backward. *General and Comparative Endocrinology* **209**, 118-134.

Røed, K.H., Dehli, A.K., Flengsrud, R., Midthjell, L. & Rørvik, K.A. (1995). Immunoassay and partial characterisation of serum transferrin from Atlantic salmon (*Salmo salar* L.). *Fish & Shellfish Immunology* 71-80.

Rojas, H.H.F. (2005). Contribución al conocimiento de la reproducción del pirarucu *Arapaima gigas* (Cuvier, 1817) (Pisces: Arapamidae) en cautiverio. Unpublished thesis (B.Sc.), Universidad de la Amazonia.

Rondeau, E.B., Minkley, D.R., Leong, J.S., Messmer, A.M., Jantzen, J.R., von Schalburg, K.R., Lemon, C., Bird, N.H. & Koop, B.F. (2014). The Genome and Linkage Map of the Northern Pike (*Esox lucius*): Conserved Synteny Revealed between the Salmonid Sister Group and the Neoteleostei. *PloS One* **9**, e102089.

Rosa, R., Rubert, M., Caetano-Filho, M. & Giuliano-Caetano, L. (2009). Conserved Cytogenetic Features in the Amazonian Arapaima, *Arapaima gigas* (Schinz 1822) from Jamari River, Rondônia–Brazil. *The Open Biology Journal* **2**, 91-94.

Rosenfeld, H., Meiri, I. & Elizur, A. (2007). Gonadotrophic regulation of oocyte development. In: Babin, P.J., Cerdà, J. & Lubzens, E., eds., The fish oocyte: From basic studies to biotechnological aplications, Dordrecht, Netherlands: Springer.

Rosenfeld, H., Mylonas, C.C., Bridges, C.R., Heinisch, G., Corriero, A., Vassallo-Aguis, R., Medina, A., Belmonte, A., Garcia, A., De la Gándara, F., Fauvel, C., De Metrio, G., Meiri-Ashkenazi, I., Gordin, H. & Zohar, Y. (2012). GnRHa-mediated stimulation of the reproductive endocrine axis in captive Atlantic bluefin tuna, *Thunnus thynnus*. *General and Comparative Endocrinology* **175**, 55-64.

Rousset, F. (2008). GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**, 103-106.

Saenz-Agudelo, P., Dibattista, J.D., Piatek, M.J., Gaither, M.R., Harrison, H.B., Nanninga, G.B. & Berumen, M.L. (2015). Seascape genetics along environmental gradients in the Arabian Peninsula: insights from ddRAD sequencing of anemonefishes. *Molecular Ecology* **24**, 6241-6255.

Saito, M., Takenouchi, Y., Kunisaki, N. & Kimura, S. (2001). Complete primary structure of rainbow trout type I collagen consisting of α1(I)α2(I)α3(I) heterotrimers. *European Journal of Biochemistry* **268**, 2817-2827.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**, 406-425.

Samarin, A.M., Policar, T. & Lahnsteiner, F. (2015). Fish oocyte ageing and its effect on egg quality. *Reviews in Fisheries Science & Aquaculture* **23**, 302-314.

Sanahuja, I. & Ibarz, A. (2015). Skin mucus proteome of gilthead sea bream: A non-invasive method to screen for welfare indicators. *Fish & Shellfish Immunology* **46**, 426-435.

Sato, R., Saito, T., Ogata, H., Nakanea, N., Namegawaa, K., Sekiguchi, S., Omuraa, K., Kurabuchi, S., Mitamurab, K., Ikegawa, S., Raines, J., Hageyd, L.R., Hofmannd, A.F. & Iida, T. (2016). Novel, major 2α- and 2β-hydroxy bile alcohols and bile acids in the bile of *Arapaima gigas*, a large South American river fish. *Steroids* **107**, 112-120.

Sawaya, P. (1946). Sôbre a biologia de alguns peixes de respiração aérea (*Lepidosiren paradoxa* Fitz, e *Arapaima gigas* Cuv.). *Boletim da Faculdade de Filosofia Ciencias e Letras, Universidade de São Paulo, Zoologia* **11**, 255-286.

Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671-675.

Schreck, C.B., Contreras-Sanchez, W. & Fitzpatrick, M.S. (2001). Effects of stress on fish reproduction, gamete quality, and progeny. *Aquaculture (Amsterdam, Netherlands)* **197**, 3-24.

Schulz, R., Menting, S., Bogerd, J., França, L.R., Vilela, D.A.R. & Godinho, H.P. (2005). Sertoli cell proliferation in the adult testis—evidence from two fish species belonging to different orders. *Biology of Reproduction* **73**, 891-898.

Schulz, R.W., França, L.R., Lareyre, J.J., LeGac, F., Chiarini-Garcia, H., Nobrega, R.H. & Miura, T. (2010). Spermatogenesis in fish. *General and Comparative Endocrinology* **165**, 390-411.

Schulz, R.W. & Nóbrega, R.H. (2011). Regulation of spermatogenesis. In: Farrell, A., ed., Encyclopedia of Fish Physiology-From Genome to EnvironmentAcademic Press, 627-634.

Schütz, M. & Barlow, G., W. (1997). Young of the Midas cichlid get biologically active nonnutrients by eating mucus from the surface of their parents. *Fish Physiology and Biochemistry* **16**, 11-18.

Selvaraj, S., Ohga, H., Nyuji, M., Kitano, K., Nagano, N., Yamaguchi, A. & Matsuyama, M. (2015). Effects of synthetic kisspeptin peptides and GnRH analogue on oocyte growth and circulating sex steroids in prepubertal female chub mackerel (*Scomber japonicus*). *Aquaculture Research* **46**, 1866–1877.

Sherman, V.R., Quan, H., Yang, W., Ritchie, R.O. & Meyers, M.A. (2016). A comparative study of piscine defense: The scales of *Arapaima gigas*, *Latimeria chalumnae* and *Atractosteus spatula*. *Journal of the Mechanical Behavior of Biomedical Materials* **in press**.

Shevchenko, A., Valcu, C.M. & Junqueira, M. (2009). Tools for Exploring the Proteomosphere. *Journal of Proteomics* **72**, 137-144.

Sidi, S., Goutel, C., Peyriéras, N. & Rosa, F.M. (2003). Maternal induction of ventral fate by zebrafish radar. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 3315-3320.

Sink, T.D., Lochmann, R.T. & Fecteau, K.A. (2008). Validation, use, and disadvantages of enzyme-linked immunosorbent assay kits for detection of cortisol in channel catfish, largemouth bass, red pacu, and golden shiners. *Fish Physiology and Biochemistry* **34**, 95-101.

Siqueira-Silva, D.H., Ninhaus-Silveira, A., Silva, A.P.S. & Veríssimo-Silveira, R. (2015). Morphology of the urogenital papilla and its component ducts in *Astyanax altiparanae* Garutti & Britski, 2000 (Characiformes: Characidae). *Neotropical Ichthyology* **13**, 309-316.

Sloman, K.A. & Buckley, J. (2011). Nutritional provision during parental care. In: Farrell, A., ed., Encyclopedia of Fish Physiology-From Genome to Environment, London: Academic Press, 678-683.

Sorensen, P.W. & Stacey, N. (1999). Evolution and Specialization of Fish Hormonal Pheromones. In: al., J.e., ed., Advances in Chemical Signals in Vertebrates, New York: Plenum Publishers, 15-47.

Sorensen, P.W. & Stacey, N.E. (2004). Brief review of fish pheromones and discussion of their possible uses in the control of non-indigenous teleost fishes. *New Zealand Journal of Marine and Freshwater Research* **38**, 399-417.

Sorensen, P.W. & Wisenden, B.D. (2015). Fish pheromones and related cues, Pondicherry, India: Wiley Blackwell.

Stacey, N. (2003). Hormones, pheromones and reproductive behavior. *Fish Physiology and Biochemistry* **28**, 229-235.

Stafford, J.L., Neumann, N.F. & Belosevic, M. (2001). Products of proteolytic cleavage of transferrin induce nitric oxide response of goldfish macrophages. *Developmental and Comparative Immunology* **25**, 101-115.

Stevens, E.D. & Holeton, G.F. (1978). The partitioning of oxygen uptake from air and from water by the large obligate air-breathing teleost pirarucu (*Arapaima gigas*). *Canadian Journal of Zoology* **56**, 974-976.

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

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

Swenson, E.A., Rosenberger, A.E. & Howell, P.J. (2007). Validation of endoscopy for determination of maturity in small salmonids and sex of mature individuals. *Transactions of the American Fisheries Society* **136**, 994-998.

Takács, P., Erős, T., Specziár, A., Sály, P., Vitál, Z., Ferincz, A., Molnár, T., Szabolcsi, Z., Bíró, P. & Csoma, E. (2015). Population Genetic Patterns of Threatened European Mudminnow (Umbra krameri Walbaum, 1792) in a Fragmented Landscape: Implications for Conservation Management. *PloS One* **10**, e0138640.

Tallmon, D.A., Luikart, G. & Waples, R.S. (2004). The alluring simplicity and complex reality of genetic rescue. *TRENDS in Ecology and Evolution* **19**, 489-496.

Taslima, K., Taggart, J.B., Wehner, S., McAndrew, B.J. & Penman, D.J. (2017). Suitability of DNA sampled from Nile tilapia skin mucus swabs as a template for ddRAD-based studies. *Conservation of Genetic Resources* **9**, 39-42.

Torati, L.S., Varges, A.P.S., Galvão, J.A.S., Mesquita, P.E.C. & Migaud, H. (2016). Endoscopy application in broodstock management of *Arapaima gigas* (Schinz, 1822). *J Appl Ichthyol* **32**, 353-355.

Torres-Vázquez, J., Fraser, S.D., Pham, V.N., Childs, S., Gitler, A.D., Berk, J.D., Fishman, M.C. & Weinste, B.M. (2004). Semaphorin-plexin signaling guides patterning of the developing vasculature. *Developmental Cell* **7**, 117-123.

Toyama, R., Kim, M.H., Rebbert, M.L., Gonzales, J., Burgess, H. & Dawid, I.B. (2013). Habenular commissure formation in zebrafish is regulated by the pineal gland specific gene unc119c. *Developmental Dynamics* **242**, 1033-1042.

Trivers, R.L. (1974). Parent-offspring conflict. *American Zoologist* **14**, 249-264.

Tyler, C.R. & Sumpter, J.P. (1996). Oocyte growth and development in teleosts. *Reviews in Fish Biology and Fisheries* **6**, 287-318.

Ubuka, T., Son, Y.L., Tobari, Y. & Tsutsui, K. (2012). Gonadotropin-inhibitory hormone action in the brain and pituitary. *Frontiers in Neuroendocrinology* **3**, 1-13.

Val, A.L. (2000). Organic phosphates in the red blood cells of fish. *Comparative Biochemistry and Physiology A* **125**, 417-435.

Valcu, C.M. & Kempenaers, B. (2014). Proteomics in behavioral ecology. *Behavioral Ecology* **00**, 1-15.

Van Eeden, F.J.M., Granato, M., Schach, U., Brand, M., Furutani-Seiki, M., Haffter, P., Hammerschmidt, M., Heisenberg, C.P., Jiang, Y.J., Kane, D.A., Kelsh, R.N., Mullins, M.C., Odenthal, J., Warga, R.M., Allende, M.L., Weinberg, E.S. & Nüsslein-Volhard, C. (1996). Mutations affecting somite formation and patterning in the zebrafish, *Danio rerio*. *Development* **123**, 153-164.

Viegas, C.S.B., Simes, D.C., Laize, V., Williamson, M.K., Price, P.A. & Cancela, M.L. (2008). Gla-rich Protein (GRP), A new vitamin K-dependent protein identified

from Sturgeon cartilage and highly conserved in vertebrates. *Journal of Biological Chemistry* **283**.

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**, 557-568.

Wagner, G.F., Dimattia, G.E., Davie, J.R., Copp, D.H. & Friesen, H.G. (1992). Molecular cloning and cDNA sequence analysis of coho salmon stanniocalcin. *Molecular and Cellular Endocrinology* **90**, 1-15.

Wallace, R.A. & Selman, K. (1981). Cellular and dynamic aspects of oocyte growth in teleosts. *American Zoologist* **21**, 325-343.

Watson, L.C., Stewart, D.J. & Kretzer, A.M. (2016). Genetic Diversity and Population Structure of the Threatened Giant *Arapaima* in Southwestern Guyana: Implications for Their Conservation. *Copeia* **104**, 864-872.

Watson, L.C., Stewart, D.J. & Teece, M.A. (2013). Trophic ecology of *Arapaima* in Guyana: giant omnivores in Neotropical floodplains. *Neotropical Ichthyology* **11**, 341-349.

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

West, G. (1990). Methods of assessing ovarian development in fishes: A review. *Australian Journal of Marine and Freshwater Research* **41**, 199-222.

Whittington, C.M. & Wilson, A.B. (2013). The role of prolactin in fish reproduction. *General and Comparative Endocrinology* **191** 123-136.

Wiegand, M.D. (1996). Composition, accumulation and utilization of yolk lipids in teleost fish. *Reviews in Fish Biology and Fisheries* **6**, 259-286.

Wildhaber, M.L., Papoulias, D.M., Delonay, A.J., Tillitt, D.E., Bryan, J.L., Annis, M.L. & Allert, J.A. (2005). Gender identification of shovelnose sturgeon using ultrasonic and endoscopic imagery and the application of the method to the pallid sturgeon. *Journal of Fish Biology* **67**, 114-132.

Wosnitza-Mendo, C. (1984). The growth of *Arapaima gigas* after stocking in a Peruvian lake. *Archiv fur Fischereiwissenschaft* **35**, 1-5.

Wyatt, T.D. (2010). Pheromones and signature mixtures: defining species-wide signals and variable cues for identity in both invertebrates and vertebrates. *J Comp Phys A* **196**, 685-700.

Yang, W., Chen, I.H., Gludovatz, B., Zimmermann, E.A., Ritchie, R.O. & Meyers, M.A. (2013). Natural flexible dermal armor. *Advanced Materials* **25**, 31-48.

Yang, W., Sherman, V.R., Gludovatz, B., Mackey, M., Zimmermann, E.A., Chang, E.H., Schaible, E., Qin, Z., Buehler, M.J., Ritchie, R.O. & Meyers, M.A. (2014). Protective role of *Arapaima gigas* fish scales: Structure and mechanical behavior. *Acta Biomaterialia* **10**, 3599-3614.

Yoshida, M., Honda, A., Watanabe, C., Satoh, M. & Yasutake, A. (2014). Neurobehavioral changes in response to alterations in gene expression profiles in the brains of mice exposed to low and high levels of mercury vapor during postnatal development. *Journal of Toxicological Sciences* **39**, 561-570.

Zeng, L.C., Han, Z.G. & Ma, W.J. (2005). Elucidation of subfamily segregation and intramolecular coevolution of the olfactomedin-like proteins by comprehensive phylogenetic analysis and gene expression pattern assessment. *FEBS Letters* **579**, 5443-5453.

Zhang, J.Y. (2006). Phylogeny of Osteoglossomorpha. *Vertebrata Palasiatica* **44**, 43-59.

Zhanjiang, L. (2011). Next generation sequencing and whole genome selection in aquaculture, Iowa: Wiley-Blackwell.

Zhou, X., Ding, Y. & Wang, Y. (2012). Proteomics: present and future in fish, shellfish and seafood. *Reviews in Aquaculture* **4**, 11-20.

Zohar, Y., Muñoz-Cueto, J.A., Elizur, A. & Kah, O. (2010). Neuroendocrinology of reproduction in teleost fish. *General and Comparative Endocrinology* **165**, 438-455.

Zürbig, P., Renfrow, M.B., Schiffer, E., Novak, J., Walden, M., Wittke, S., Just, I., Pelzing, M., Neusü, C., Theodorescu, D., Root, K.E., Ross, M.M. & Mischak, H. (2006). Biomarker discovery by CE-MS enables sequence analysis via MS/MS with platform independent separation. *Electrophoresis* **27**, 2111-2125.

Zürbig, P., Schiffer, E. & Mischak, H. (2009). Capillary electrophoresis coupled to mass spectrometry for proteomic profiling of human urine and biomarker discovery. In: Reinders, J. & Sickmann, A., eds., Proteomics: Methods and Protocols, Totowa, NJ: Humana Press, 105-121.