PhD in Aquatic Biosciences // No. 36 - 2020

DNA hydroxymethylation and improved growth of Nile tilapia (*Oreochromis niloticus*) during domestication

Ioannis Konstantinidis

FACULTY OF BIOSCIENCES AND AQUACULTURE

NORD University

www.nord.no

DNA hydroxymethylation and improved growth of Nile tilapia (*Oreochromis niloticus*) during domestication

Ioannis Konstantinidis

A thesis for the degree of
Philosophiae Doctor (PhD)

PhD in Aquatic Biosciences no. 36 (2020)
Faculty of Biosciences and Aquaculture

PhD in Aquatic Biosciences no. 36 (2020)

Ioannis Konstantinidis
DNA hydroxymethylation and improved growth of Nile tilapia (*Oreochromis niloticus*) during domestication

ISBN: 978-82-93165-36-1

Print: Trykkeriet NORD

Nord University
N-8049 Bodø
Tel: +47 75 51 72 00
www.nord.no

Preface

This dissertation is submitted in fulfilment of the requirements for the degree of Philosophiae Doctor (PhD) at the Faculty of Biosciences and Aquaculture (FBA), Nord University, Bodø, Norway. The research article and manuscripts included in this dissertation represent original research that was carried out over a period of 4 years from 04.07.2016 to 20.09.2020. This work was part of the EPIFISH project funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreements nos 683210 and 812986), the Research Council of Norway under the Toppforsk programme (grant agreement no 250548/F20) and Nord University (Norway).

The project team consisted of the following members:

Ioannis Konstantinidis, MSc: PhD fellow

Jorge Fernandes, Professor, FBA, Nord University: primary supervisor

Pål Sætrom, Professor, Faculty of Medicine and Health Sciences and Faculty of Information Technology and Electrical Engineering, NTNU: co-supervisor

Ioannis Konstantinidis,

Bodø, 20th September 2020

I dedicate this thesis to my beloved parents,

who supported me in the path I chose to follow

The following Greek phrase is a favourite of my Father's and a constant reminder:

"τά ἀγαθά κόποις κτῶνται"

"good things come after hard work"

Aristotle

Acknowledgements

First and foremost, I would like to thank my supervisor, Jorge, for all the support and guidance throughout this endeavor. I am deeply grateful for entrusting me to carry out part of his project. This PhD started as a once-in-a-lifetime experience; a Portuguese and a Greek, starting their trip to Egypt with one goal; bring wild Nile tilapia to Norway! The past few years have been an emotional roller coaster but Jorge was always there to celebrate and commiserate with me. He was not only a supervisor but a good friend and leader. I am truly grateful.

I would like to thank also my co-supervisor, Pål Sætrom, for providing his valuable knowledge on the analysis and interpretation of epigenetic datasets. His contribution was key for the completion of this thesis and I am truly thankful to him for investing his limited time in me.

From the very start of the EPIFISH project, I spent countless hours with Tom Podgorniak. We have shared every frustration and celebrated every success. I already miss our "corridor meetings" and our collective efforts to resolve every issue around us and beyond. He will always be a good friend and a great colleague, and I hope our research paths will cross again in the near future.

I would like to thank all the members of the EPIFISH team, my colleagues and friends, both past and present, for helping, sharing and contributing their invaluable knowledge through countless meetings and brainstorming sessions. Also, I can never give enough thanks to our administration and technicians on campus, and especially those at the research station in Bukta. They have had an unparallel commitment to keeping our fish healthy, as well as anticipating potential problems throughout the years. I am also truly thankful to all our collaborators, brilliant researchers and scientists who contributed for the completion of this thesis and shared their knowledge with me.

Last but not least, I would like to thank my Greek and British family for their love and support, and my wife, Megan, who never stopped believing in me during this journey.

Table of Contents

Preface i
Acknowledgements iv
List of figures vii
List of papers viii
List of abbreviations ix
Annex xi
Abstract 1
1. Introduction 3
1.1 Animal domestication and its relevance for human society 3
1.2 The history of fish domestication 5
1.3 Nile tilapia biology and culture 10
1.4 Breeding and artificial selection of fish domesticates 11
1.5 Molecular basis of muscle growth 15
1.6 Molecular changes in early domestication of fish 16
1.7 Epigenetics 17
1.7.1 DNA methylation 19
1.7.2 DNA hydroxymethylation 21
1.7.3 Histone modifications 23
1.7.4 Non-coding RNAs 25
1.8 The role of epigenetics in animal domestication 26
1.8.1 Nutritional epigenetics 26
1.8.2 Epigenetic mechanisms of growth 27
1.8.3 Epigenetic mechanisms in immunity 29
1.8.4 Transgenerational epigenetic inheritance 30
1.9 Methods for 5hmC profiling 31
2. Objectives 35
3. General Discussion 36
3.1 Importance of DNA 5-hydroxymethylation during the early stages of fish domestication 37

3.2 A link between DNA hydroxymethylation of myogenic genes and somatic growth..38
3.3 Epigenetic regulation of immune-related genes during fish domestication 41
3.4 Impact of DNA hydroxymethylation in adaptation and genome evolution 43
3.5 Epigenetic modifications as biomarkers.....45
4. Conclusions.....48
5. Future Perspectives.....49
6. References.....51

List of figures

Figure 1. Fragment of a polychrome tomb-painting representing the pool in Nebamun's estate garden.

Figure 2. Front view of the remains of the Grotto of Tiberius (42 BC – 37 AD) at Sperlonga, Italy.

Figure 3. Human population growth from 10,000 BC to 2015 AD.

Figure 4. Nile tilapia global aquaculture production in million tonnes.

Figure 5. Wild Nile tilapia (*Oreochromis niloticus*) male.

Figure 6. Venn diagram representing how genetic, epigenetic and environmental factors are interconnected and contribute to the resulted phenotype.

Figure 7. Active demethylation cycle of cytosine nucleosides.

Figure 8. Schematic flow representing the RRHP library preparation protocol.

Figure 9. Graphical summary of the dissertation

List of papers

Paper I **Konstantinidis, I**., Sætrom, P., Mjelle, R., Nedoluzhko, A.V., Robledo, D. and Fernandes, J.M.O. (2020). Major gene expression changes and epigenetic remodelling in Nile tilapia muscle after just one generation of domestication. *Epigenetics* 15(10): 1052-1067. doi:10.1080/15592294.2020.1748914

Paper II **Konstantinidis, I**., Sætrom, P., Brieuc, M.S., Baaksrud, H.T., Jakobsen, K.S., Liedtke, H., Pohlmann, C., Tsoulia, T. and Fernandes, J.M.O. DNA hydroxymethylation differences underlie early phenotypic divergence of somatic growth during fish domestication. Manuscript

Paper III **Konstantinidis, I**., Sætrom, P., Anastasiadi, D., Nedoluzhko, A.V., Mjelle, R., Podgorniak, T., Piferrer, F. and Fernandes, J.M.O. Epigenetic mapping of the somatotropic axis in Nile tilapia reveals differential DNA hydroxymethylation involved in growth. Manuscript

List of abbreviations

5hmC	-	5-hydroxymethylcytosine
QTLs	-	Quantitative trait loci
SNPs	-	Single nucleotide polymorphisms
GH	-	Growth hormone
5mC	-	5-methylcytosine
Dnmts	-	DNA methyltransferases
TET	-	Ten-eleven translocation
5fC	-	5-formylcytosine
5caC	-	5-carboxylcytosine
BER	-	Base excision repair
TDG	-	Thymine DNA glycosylase
ncRNAs	-	non-coding RNAs
snoRNAs	-	small nucleolar RNAs
siRNAs	-	short interfering RNAs
miRNAs	-	micro RNAs
piRNAs	-	piwi-interacting RNAs
lncRNAs	-	long non-coding RNAs
nt	-	nucleotides
ssRNAs	-	single-stranded RNAs
dsRNAs	-	double-stranded RNAs
mRNA	-	messenger RNA
xiRNA	-	X chromosome non-coding RNA

MyHC	-	Myosin heavy chain
RSSs	-	Recombination signal sequences
TAB-Seq	-	Tet-assisted bisulfite sequencing
RRHP	-	Reduced representation 5-hydroxymethylcytosine profiling
LC-MS/MS	-	Liquid chromatography with tandem mass spectrometry
RNA-Seq	-	RNA sequencing
DhmCs	-	Differentially hydroxymethylated cytosines
fgfs	-	fibroblast growth factors
fgfrs	-	fibroblast growth factor receptors
frs	-	fibroblast receptor substrates
gas	-	growth arrest specific proteins
gdfs	-	growth differentiation factors
igfbps	-	insulin-like growth factor binding proteins
igfrs	-	insulin-like growth factor receptors
MAS	-	Marker-assisted selection
IPN	-	Infectious pancreatic necrosis
GWAS	-	Genome-wide association studies
CRISPR	-	Clustered regularly interspaced palindromic repeats

Annex

Supplemental tables are available upon request. Please contact the author at ioannis.constant@outlook.com

Abstract

The worldwide demand for fish consumption is highly dependent on aquaculture production because commercial fishing reached its maximum exploitation since the 90's. However, continued expansion of the aquaculture industry in a sustainable manner is dependent on several factors, including domestication of new species and establishment of selective breeding programmes. Domestication of new fish species is a rather complex and long process but the deployment of new molecular tools could improve and accelerate it through the holistic characterization of fish genomes. This thesis is based on the hypothesis that epigenetic mechanisms underlie genome-wide adaptation under captivity, since genetic mutations and allele shifts alone cannot explain the rapid transcriptomic changes of fish undergoing domestication. In a series of experiments, we investigated the role of DNA hydroxymethylation during the early stages of fish domestication, and its potential involvement in regulating somatic growth.

At first, we discovered that the DNA hydroxymethylome in fast muscle changes rapidly within a single generation of domestication. As a result of our 5-hydroxymethylcytosine (5hmC) profiling at single nucleotide resolution, we were able to identify that 5hmCs are largely enriched within gene bodies, which supports the notion that they are functionally relevant epigenetic modifications. The annotation of differentially hydroxymethylated genes between wild and first-generation of fish under captivity revealed that the changes occurred primarily within genes involved in immunity, growth and neuronal activity. By comparing gene expression profiles in muscle between wild and first generation of Nile tilapia in captivity, we showed that immune-related genes were upregulated in the wild fish, while genes involved in metabolism and muscle-specific functions were downregulated. These findings revealed that the first generation of fish undergoing domestication is strongly influenced by the environmental conditions under captivity, namely the lack of pathogens and the optimal conditions of water temperature, oxygen, pH and diet. Thus, we provided for the first time a link between environmentally-mediated DNA hydroxymethylation and gene regulation in fish undergoing domestication.

To further explore the connection between DNA hydroxymethylation and somatic growth, we compared the liver hydroxymethylomes of slow- and fast-growing full siblings and examined the 5hmC abundance in three major tissues that compose the somatotropic axis. Interestingly, we identified several differentially hydroxymethylated genes between slow- and fast-growers. These genes were involved in signaling pathways related to cell growth, survival and proliferation such as the PI3K-Akt, the Ras- and Rho- protein signal transduction pathways. By comparing the DNA hydroxymethylomes among the muscle, liver and pituitary, we identified several differentially hydroxymethylated growth factors, receptors and enhancers with major implications in growth, metabolism and skeletal muscle development. Taken together, this thesis provides for the first time a direct link between DNA hydroxymethylation and fish domestication and associates epigenetic marks at single nucleotide resolution to somatic growth using cutting-edge molecular tools. When validated, these epigenetic markers can potentially improve current breeding strategies in aquaculture by providing a holistic approach for broodstock selection.

1 Introduction

1.1 Animal domestication and its relevance for human society

Domestication is a word of Latin origin meaning “belonging to the house”. Nowadays, it can be defined as the process by which plants and animals are bred in captivity and their phenotype gradually diversifies, compared to their wild progenitors, in ways that are beneficial to humans (Diamond, 2002). The environmental conditions and management practices result in the enhancement of specific traits of interest that usually take various forms such as color and shape in terms of aesthetics or improved growth, sexual maturation, disease and stress resistance in terms of large-scale food production (Balon, 2004; Diamond, 2002; Gjerdem, 2012). The domestication of wild animals dates back thousands of years in human history. However, it is rather unlikely that the early domestication events were a result of conscious selective breeding. Compared to methodical selection (systematic breeding for the enhancement of specific traits), and natural selection (predominance of the fittest), the unconscious selection, as described in Darwin’s early work, might explain the beginning of preservation of valued phenotypes that were selected unintentionally by humans (Darwin, 1875). Nowadays, unconscious selection is recognized as part of unintentional or weak artificial selection (Driscoll et al., 2009). In that regard, both conscious and unconscious selection, result in permanent genetic modifications, and should not be confused with taming, which is a process where wild animals become more tolerant to human presence. Consequently, wild animals at the early stages of domestication have been previously characterized as “exploited captives” (Balon, 2004).

Within the timeline of human evolution, modern humans (*Homo sapiens*) were initially hunters/gatherers (15-30 thousand years ago). Their lifestyle was primarily nomadic with a few exceptions of adaptive hunting strategies and temporary settlements (Vigne, 2011). After the peak of the last glacial maximum (21-15 thousand years ago), hostile climatic conditions did not allow for food production. However, a gradual improvement of the climatic conditions during the early Holocene (11.7 thousand years ago) led humans to take advantage of the favorable environmental conditions and intensify their efforts at improving their lifestyles. The beginning of domestication of plants and animals at approximately 12-

15 thousand years ago (Diamond, 2002; Larson et al., 2012; Larson et al., 2014; Vigne, 2011) was a prerequisite to the creation of civilizations, as it transformed nomadic groups of people into societies (Cauvin, 2000). This demographic transition was achieved by making primary food resources accessible and in close proximity throughout the seasons without the need to hunt animals and gather plants across vast geographical areas. The beginning of the agricultural era expanded the capabilities of our species to an enormous extent, bypassing barriers that were impenetrable for our hunter/gatherer ancestors. Thus, humans were able to create food surpluses, increase birth rates and support larger population sizes which enabled producers to displace gatherers (Bocquet-Appel, 2011; Diamond, 1997). Because this transition did not occur for all civilizations across the world at the same time and intensity, the genetic evolution of humans followed the expansion of the primary homelands of agriculture. Therefore, the process of domestication of wild plants and animals has brought by far the most significant changes in the human history and genome evolution (Diamond, 2002).

The first wild animal to be domesticated by humans was the wolf (*Canis lupus*), which is the predecessor of several modern dog breeds (*Canis lupus familiaris*) (Freedman and Wayne, 2017). However, previous studies have pointed out that the wild ancestor of dogs is possibly extinct because dogs are genetically distant to the modern wolf (Frantz et al., 2016; Freedman et al., 2014). Even though dog domestication has been extensively studied for several decades, there are still contradictions regarding its origins (Clutton-Brock, 1995; Olsen and Olsen, 1977). Due to the coexistence of wolves and humans across large geographical areas, it is most likely that dogs have diverged from multiple domestication events (Clutton-Brock, 1987). Within the next few thousand years (8-12 thousand years ago), humans had domesticated a large variety of plants (rice, wheat, barley, corn, millet, lentils, sweet potatoes and chickpeas) and animals (sheep, pigs, cattle and horses), which were necessary for the establishment and development of major civilizations across the world. Although domestication of plants and animals is strongly connected to humans as a source of food, pet keeping has been linked even to our hunter-gatherer ancestors (Larson and Fuller, 2014). Flowers were domesticated for their medicinal properties, as offerings to the gods as well as for displaying purposes (ornamentally) (Goody, 1993).

1.2 The history of fish domestication

In contrast to terrestrial plants and animals, fish domestication appeared much later, approximately 2-3.5 thousand years ago (Harache, 2002; Larson and Fuller, 2014). Evidence of Nile tilapia in captivity is depicted on a fragment from the tomb of Nebamun (ca 1350 BC) in ancient Egypt (Fig. 1) (Middleton and Uprichard, 2008). However, it is difficult to decipher whether fish in ponds were seen exclusively as ornaments, as a food source or both. Records of carp farming date back to the Zhou dynasty in ancient China (1066-256 BC) with one of the earliest monographs on aquaculture being the "*Treatise on Pisciculture*" by Fan Li (5th Century BC) (Jia et al., 2018).

At approximately the same time in Europe, aquaculture began in ancient Rome. The Romans adopted the Egyptian and Assyrian vivarium, which was a pond-like system built in wealthy villas and great estates (Nash, 2000). Owning a fishpond was a symbol of power, wealth and social status. Fish were initially transferred to these ponds to serve as ornaments but reports suggest that they were also kept for consumption (Nash, 2000). Breeding was likely a process that occurred naturally in these ponds but overall fishing was considered to be an occupation of the lower class in ancient Rome. From the 1st century BC to the 1st century AD, aquaculture developed rapidly as depicted in the remains of highly sophisticated and specialized structures, the "piscinae" (Fig. 2), which were found across the shores of Italy, such as the Grotto of Tiberius at Sperlonga (Higginbotham, 1997). In early writings of 40 BC – 100 AD, as they were reviewed by Higginbotham J.A., the main cultured fish species in ancient Rome were grey mullets (*Mugil cephalus*), seabream (*Sparus aurata*), seabass (*Dicentrarchus labrax*), sole (*Solea solea*) and flounders (*Paralichthys dentatus*). Four hundred years later (ca. 1500 years ago), evidence for intensive aquaculture was present in ancient Hawaii with fishponds stretching 1.5-2 km from the shore towards the mountains (Costa-Pierce, 1987). Different approaches had been developed including fresh, brackish and seawater fish ponds. Freshwater ponds were mostly used for the cultivation of taro, green algae and some euryhaline and freshwater fish species such as mullets (*Mugil cephalus*, *Neomyxus leuciscus*) and silver perch (*Bidyanus bidyanus*), however, additional freshwater ponds were used exclusively for fish and prawns. Similarly, the brackish-water ponds were

Figure 1. Fragment of a polychrome tomb-painting representing the pool in Nebamun's estate garden: date-palms, sycomores and mandrakes hedge the pool which teems with fish and fowl. This is one of the earliest indications of fish kept in captivity (1350 BC) either for ornamental purposes or food consumption. This image is released by the British Museum (© The Trustees of the British Museum) for the purposes of this doctoral thesis under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license.

located in coastal zones and they were connected to both freshwater streams and the sea through ditches. These ponds were very productive since they could host a large number of fish species.

Lastly, the highest achievement of the ancient Hawaiian engineering were the seawater ponds which were built from the shore towards the sea. The isolation of these ponds was achieved with a seawall constructed by coral and lava rock. The seawater ponds were automatically stocked from the sea with the use of a grate. The grate was part of the seawall and it was constructed by timbers separated by 0.5-2 cm from each other; thus, only small fish could enter or exit the pond. The fish were fed cut grass, mussels, clams, seaweeds and

Figure 2. Front view of the remains of the Grotto of Tiberius (42 BC – 37 AD) at Sperlonga, Italy. Grottos were highly sophisticated structures belonging to wealthy Romans and used as fish ponds "piscinae". Photo provided by Troels Myrup Kristensen, Professor in classical archaeology at the School of Culture and Society, Aarchus University.

taro leaves (Titcomb, 1952; Wilder, 1923). Due to the fattening within the ponds, the fish would grow and stay in the pond until their harvest (Costa-Pierce, 1987).

During these early times, fish were mostly exploited captives rather than domesticates since they were in captivity but not selectively bred. With the exception of early indications in China, selection of different body shapes as well as scaleless phenotypes took place from the late 12th to mid-14th century AD while selection for the enhancement of specific traits such as growth, disease resistance and sexual maturity have been only reported within the 20th century (Balon, 1995; Embody and Hyford, 1925). The development of breeding practices and optimization of traits were necessary to support our population growth, especially from the 20th century. This is also depicted by the human population growth curve in Figure 3 (Roser et al., 2013) that shows a growth rate of 0.04% from 10,000 BC to 1700

AD, reaching 1 billion in 1800 and 1.65 billion in 1900. Within the following century, our population growth changed fundamentally, reaching 7.35 billion in 2015. In response to the ever-increasing population and world-wide demand for fish consumption, humans overexploited natural stocks, with fishery catches reaching their maximum exploitation three decades ago (FAO, 2020b). In that regard, the boundaries of fish domestication and aquaculture had to be expanded across many continents and species with the first family-based breeding program of Atlantic salmon (*Salmo salar*) being established in 1975 (Gjerdem, 2010). Since then, the production of fish species such as Nile tilapia (*Oreochromis niloticus*) increased considerably from 127,000 tonnes in 1988 to 2 million in 2008 and 4.2 million tonnes in 2016 (FAO, 2020a) (Fig. 4). In terms of trait enhancement efficiency, reports for Atlantic salmon (*Salmo* salar) have showed 13-15% growth gain per generation within 5 generations, for Rohu carp (*Labeo rohita*) a substantial growth of 30% per generation within 4 generations and for the GIFT strain of Nile tilapia an average of 17% growth within 5 generations (Gjerdem, 2012). Among the most important fish species that are currently being cultivated in large quantities around the world are the grass carp (*Ctenopharyngodon idellus*), silver carp (*Hypophthalmichthys molitrix*), Nile tilapia (*Oreochromis niloticus*) and common carp (*Cyprinus carpio*). In terms of value, Atlantic salmon (*Salmo salar*) is by far the most economically important fish species with a market value of approximately USD 27.5 billion in 2018 (FAO, 2020b). In a continuous effort to support the worldwide demand for animal protein, aquaculture is facing several challenges including the expansion and development of its costal and deep-sea infrastructure, fish species diversification, and climate change as well as the further optimization of breeding strategies using molecular tools.

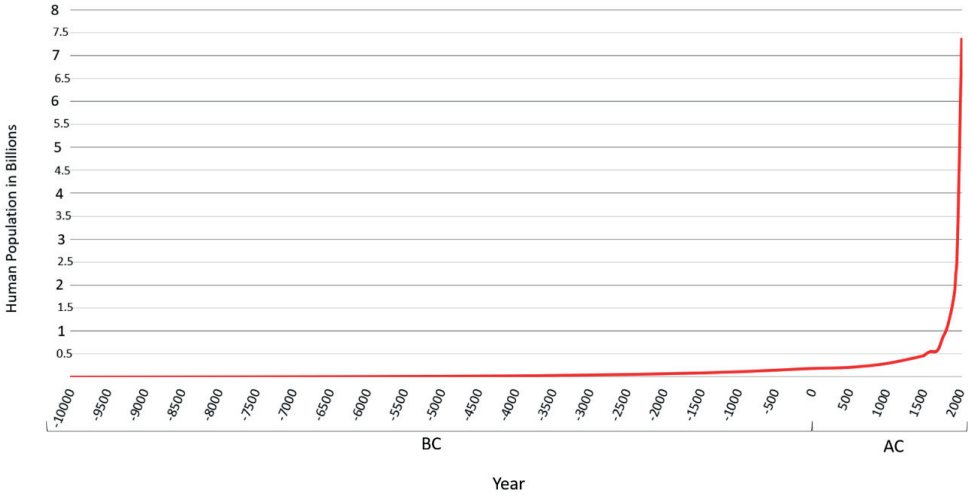


Figure 3. Human population growth from 10,000 BC to 2015 based on spline interpolation until 1950. Population estimates retrieved from the History Database of the Global Environment (HYDE) and the United Nations, while the annual data were downloaded from OurWorldinData.org.

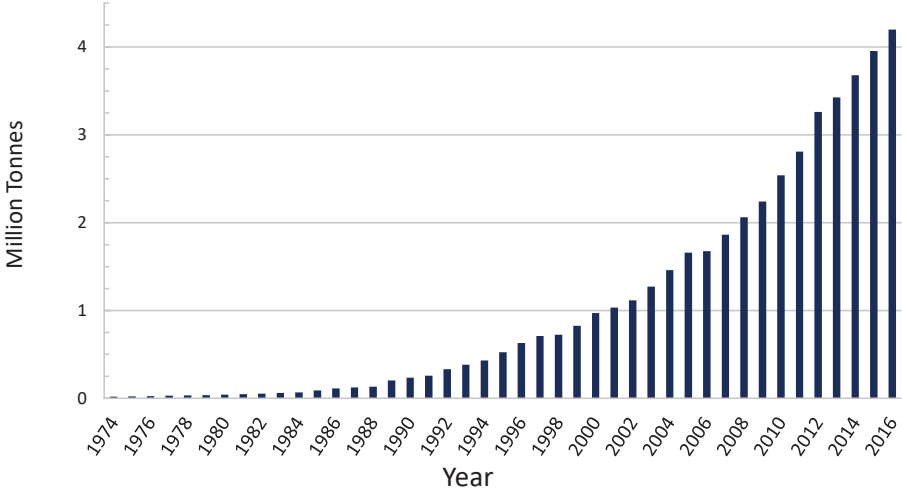


Figure 4. Nile tilapia global aquaculture production in million tonnes from 1974 to 2016 according to the Food and Agriculture Organization (FAO) FishStat.

1.3 Nile tilapia biology and culture

Nile tilapia is a freshwater fish species, native to the Nile river basin, the tropical and subtropical Africa and middle east (Yarkon River, Israel) and adapted to shallow ponds and river banks (Trewavas, 1983). Their body shape can be characterized as oval, deep and laterally compressed. Their skin is covered with cycloid scales and has a distinct dark green colour with vertical yellow stripes. The caudal fin is generally dark coloured with dense vertical yellow stripes. The dorsal and anal fins are also dark with vertical yellow stripes along their soft rays. During the reproductive season, the ventral side including the lower jaw and lower gill arches down to the anus, as well as the pectoral and caudal fins have a characteristic red colour (Fig 5). Reproduction usually occurs when the temperature is between 20 and 35 °C (Chervinski, 1982). Nile tilapia is capable of withstanding temperatures that range from 11 to 42 °C while the preferred range is 27 to 32 °C.

Reduced appetite and growth, as well as increased stress and vulnerability to diseases can be observed when exposed to temperatures out of their preferred ranges. Temperatures below 16 and 12 °C, significantly reduce their development (Wakefield, 2014) and induce severe mortalities (Chervinski, 1982), respectively. Even though it is an omnivorous fish species, Nile tilapia primarily feeds on phytoplankton, aquatic plants and periphyton mats by surface grazing (FAO, 2020a). Their structural adaptations for plant consumption such as specialized jaw, pharyngeal teeth, and long intestines, places them very close to the basis of the food chain (Pullin and Lowe-McConnell, 1982). Their wide range of feeding and environmental adaptability (i.e. omnivorous and opportunistic species, tolerance to low concentration of dissolved oxygen, wide range of alkalinity or acidity, fresh, brackish and salt water, high concentration of ammonia) makes their culture easier compared to other finfish and sustainable for protein production (da Silva Maciel et al., 2014).

Figure 5. Wild Nile tilapia (*Oreochromis niloticus*) male caught in Nile river during the reproductive season. Photo by Jorge M.O. Fernandes, and Ioannis Konstantinidis, August 2016.

1.4 Breeding and artificial selection of fish domesticates

Numerous studies in commercially important salmonids (mainly *Salmo salar, Oncorhynchus mykiss, Oncorhynchus tshawytscha* and *Oncorhynchus kisutch*) have reported breeding practices for the improvement of traits such as growth rate, disease resistance, number of eggs, age at sexual maturity and fillet quality (Kim et al., 2004; Neira et al., 2006; Rye et al., 2010). Overall, technological developments in genetics and bioinformatics have resulted in a big step towards the detection of genomic regions that are affected by artificial selection, selective breeding and the process of domestication. The main idea that prevails is the concept of genetic hitchhiking, which suggests that any type of trait selection leaves footprints like signatures upon the genome (Vasemägi et al., 2012). The identification of loci that are associated with traits of interest are well described by two complementary approaches (Lopez et al., 2014). The first, referred to as “top-down” starts from a known phenotype and uses molecular techniques to identify gene regions, such as quantitative trait loci (QTLs). The second, referred to as “bottom-up” starts from a known genetic profile or pathway and with the use of statistical packages, focuses on the genetic differentiation and linkage disequilibrium. The primary goal of both approaches is the application of such signatures and their association with economic valuable traits in the aquaculture.

Evidence of bad broodstock management and breeding practices in the aquaculture industry can be detected through the genetic structure of the populations using microsatellite DNA markers. Additionally, genetic variability is an important attribute of domestication, since the populations that exhibit high levels of variation are more likely to have higher levels of additive genetic variance for productive traits (Aung et al., 2010). However, many microsatellite markers are located in non-coding regions, therefore they cannot be associated directly with economic valuable traits. In a study using five microsatellites on striped catfish (*Pangasianodon hypophthalmus*), no genetic differentiation was reported between wild and hatchery-reared populations (Ha et al., 2009). These results can be explained by the low number of markers used, the short exposure to domestication (2 generations), and the disability of the markers to identify possible loci, responsible for trait selection. However, a signature of selection during domestication could be a result of subtle shifts in allele frequencies of QTLs instead of strong selection on a few particular loci. In that case, detection of such signatures might be challenging with standard outlier statistics (Makinen et al., 2015). For example, the bottom-up approach has been used to identify selection signatures in Atlantic salmon populations prior and after their domestication (Gutierrez et al., 2016). Both farmed and wild fish were genotyped using single nucleotide polymorphisms (SNPs) and the genetic statistics and structure of the populations as well as the Hardy-Weinberg equilibrium, genetic differentiation and inbreeding coefficients were calculated. The most significant markers were linked to biological pathways associated with cellular process, metabolic process and biological regulation but there was no significant over-representation. Even though the number of outlier markers was low (44 markers, approx. 1% of the total markers used in this study), the authors suggest that some of them are associated with molecular functions that could be related to artificial selection of desirable traits during domestication. Finally, they suggest *pgrc1* gene as a candidate gene for further research, since it is expressed in hormone metabolism and plays an important role determining grilsing (early maturation) in Atlantic salmon. According to the authors, footprints of selection are traceable across the genome after about 15 generations. In a similar study with the domesticated Norwegian Mowi strain (Bicskei et al., 2016), it was shown that trait-related gene pathways were differentially expressed on eyed embryos of

wild and domesticated origin within 10 generations of domestication, while in another study (Roberge et al., 2006) the authors demonstrated that five to seven generations under domestication can impact heritable changes in gene transcription.

Technological advancements in sequencing and new molecular approaches have greatly contributed towards the identification of genetic improvement and selection of desirable traits. For example, QTL mapping was used for the selection of important traits with notable results in growth improvement of rainbow trout (*Oncorhynchus mykiss*) (Wringe et al., 2010) and Asian sea bass (*Lates calcarifer*) (Wang et al., 2006), as well as disease resistance in Atlantic salmon (Moen et al., 2009) and Japanese flounder (*Paralichthys olivaceus*) (Fuji et al., 2006). However, when traits were characterized by low heritability rates and difficulty in their determination and measurement, marker-assisted selection (MAS) was used by identifying adjacent DNA markers linked to QTLs. On the other hand, MAS can only be used for strong associations between a genomic location and a targeted trait but, most commercially important traits are highly polygenic and affected by a large number of genes. In that regard, genotyping by sequencing and genomic selection allowed the synchronous identification of thousands of markers across the genomes of potential breeders. This was achieved through genome-wide association studies (GWAS) which examined genome-wide sets of genetic variants and their association with desirable traits. Within a short period of time, the utilization of the full genetic variance, the increased efficiency of genetic predictors and the reduction of inbreeding values, contributed to the overall increased genetic gain (Zenger et al., 2018).

Direct comparisons of wild and selectively bred fish populations reveal interesting molecular changes that occur during fish domestication. For example, domesticated coho salmon (*Oncorhynchus kisutch*) and growth hormone (GH) transgenic wild strains show strong positive correlation between the same modified gene pathways (Devlin et al., 2009). Wild individuals under GH treatment had 17.6 times growth enhancement compared to the wild control but introducing the same treatment to the domesticated strain did not cause any further improvement. Furthermore, the levels of GH in muscle were 3 times higher in the domesticated compared to the wild non transgenic strain, suggesting a possible overexpression mechanism of growth hormone. The most up-regulated gene were the *gh*

and *igf-1* (insulin-like growth factor I), which contribute to the muscle development and maximum growth, respectively, as well as the *tald* gene which is responsible for lipid biosynthesis. Based on these findings, the use of growth hormone in selective breeding programs with domesticated strains does not ensure increased growth rates. On the contrary, it might have negative impact by further down-regulating important genes such as *tcap*, *isci*, *myosin* as well as other receptors and enzymes (*actb*, *thrb*, *igf-2*, *igf-ir* and cathepsins) with major disadvantages for the well-being of fishes. As a response to increased *igf-1*, proteases such as cathepsins are highly reduced (Nomura and Katunuma, 2005). Deficiency of the iron-sulfur cluster assembly enzyme expressed by the *iscu* gene can lead to myopathy and widespread muscle weaknesses (Kollberg et al., 2009). Myosin and actin play a major role in muscle structure and are the main contributors for muscle contraction and cargo movement along actin filaments of the cytoskeleton. The energy for these movements is generated during a complex mechanochemical reaction cycle which involves ATP hydrolysis (von der Ecken et al., 2016). These indications suggest that a significant downregulation of those genes can possibly induce negative effects for the health and growth of fish undergoing domestication.

Captive-reared individuals are important for the restoration of threatened fish populations (Lorenzen et al., 2012). However, hatchery-reared fish that are being re-introduced into the wild appear to have several morphological, physiological or behavioural differences compared to their wild conspecifics. Thus, maintaining the genetic variation of a population in captivity is not enough. Advanced measures and methods are necessary to minimize the exposure to artificial and social environments, particularly of generations that are meant to be reintroduced in the wild (Fleming et al., 1997). It is well known, that these behavioural and physiological changes are linked to genome-based modifications over generations such as the reduced anti-predator responses (Berejikian, 1995; Fleming et al., 1997; Houde et al., 2010; Huntingford and Adams, 2005) aggressiveness (Huntingford and Adams, 2005), reduced responsiveness to stress (Solberg et al., 2012; Solberg et al., 2013), deficiency in the immune response and reduced adaptive potential to pathogens exposition (Glover et al., 2006a; Glover et al., 2006b; Sauvage et al., 2010). The ability for adaptation to captivity is certainly variable across fish species; however, our current methodologies for the discovery

of signatures that result from both unintentional and intentional selection is limited. Additionally, recent studies have shown that changes on the genome do not occur exclusively in long-term exposure to artificial environments. For example, behavioural adaptation to captivity due to unintentional selection for reproductive success can occur in a single generation for rainbow trout (*Oncorhynchus mykiss*) and has detrimental effects in the wild (Christie et al., 2012). Additionally, several studies have indicated that both hatchery programs that intend to enhance threatened wild populations (Araki et al., 2007; Araki and Schmid, 2010; Stringwell et al., 2014) as well as sea-cage escapees (McGinnity et al., 2003), have a detrimental effect on the fitness of wild populations.

1.5 Molecular basis of muscle growth

One of the most important factors in the culture of animals for food consumption is reaching rapidly the marketable size. Over the last decades, there have been numerous approaches and efforts towards the optimization of muscle development and growth in both marine and freshwater fish species (i.e. *Salmo salar, Oncorhynchus mykiss*, *Oreochromis niloticus*, *Cyprinus carpio*, *Dicentrarchus labrax*, *Sparus aurata*). From the early stages of embryonic development, the formation of muscle (myogenesis) and the different phases of its partial and temporal activation determine muscle development and ultimately muscle mass and overall growth. Stem cells are programmed to form the myogenic lineage by forming myogenic precursor cells. These cells exit irreversibly the cell cycle and form the early myoblasts which through fusion turn into multi-nucleated fibres, the mature myotubes (Walsh and Perlman, 1997). Genes that transcribe myogenic regulatory factors such as *myoD* and *Myf5* are responsible for the establishment of the muscle lineage, while *myogenin* and *mrf4* are responsible for the termination of cell differentiation, maintenance of the muscle and cellular response to growth factors (Rescan, 2001). In zebrafish (*Danio rerio*) the formation of the initial functional embryonic myotome is defined as the primary myogenic wave (Rossi and Messina, 2014). During the secondary phase of myogenesis, muscle fibres differentiate in multiple zones along the body and they are characterized by stratified growth and hyperplasia. Myogenic cell recruitment is marked by the expression of *myoD*

and *myogenin* in these growth zones in trout (Steinbacher et al., 2007) while the strong increase of the *nfixa* mRNA in zebrafish denotes similarities with the embryonic development in mammals and the transcriptional transition from embryonic to fetal muscle (Messina et al., 2010; Pistocchi et al., 2013). During the second phase of myogenesis. *Pax7*$^{+}$ myogenic progenitor cells which originate from the dermomyotome possibly detach and migrate deeper into the fast muscle creating mosaic hyperplasia (Gros et al., 2005; Pistocchi et al., 2013; Steinbacher et al., 2007). Both stratified and mosaic hyperplasia are therefore responsible for the increase of the total number of muscle fibres (Weatherley et al., 1988). Myotubes are formed on the surface of pre-existing muscle fibres and continue to increase in numbers until the fish reaches approximately 50% of its total length (Fernandes et al., 2005; Johnston et al., 2011). Although both hyperplasia and hypertrophy are important for muscle growth, the efficiency of new fibre recruitment throughout the early stages of development introduce growth variations with critical consequences in growth potential during adulthood (Lima et al., 2017; Valente et al., 2013).

1.6 Molecular changes in early domestication of fish

Captive rearing causes both genetic and phenotypic changes in fish. This occurs primarily due to the human-based conditions under captivity (i.e. crowding, nutrition, predator- and pathogen-free environment) and the heritability of related traits that alter wild phenotypes after crossings. Transcriptional responses to domestication for brook charr (*Salvelinus fontinalis*) reveal that three generations of selective breeding are enough for the emergence of significant changes in gene regulation (Sauvage et al., 2010). In total, 156 genes were found to be differentially expressed between control and selected groups of Brook charr that were bred for sexual maturation and growth performance. These genes are involved in biological processes related to nucleic acid, protein and coenzyme metabolism as well as immunity and defence. Gene regulation appears to be highly adaptive during the earliest stages of domestication. For example, Christie et al. (2016) have reported that within a single generation of domestication, steelhead trout (*Oncorhynchus mykiss)* undergoes a rapid transcriptional adaptation that is likely driven due to crowding. In total, there were 723

differentially expressed genes between wild and first-generation hatchery fish, which were enriched for biological processes related to blood coagulation, wound healing, immune response and metabolism.

The driving force of genome evolution and adaptation is greatly influenced by changes within cis- and trans-regulatory elements that affect transcription rate, stability and the activity of transcriptionally relevant factors (Wittkopp et al., 2004). Consequently, the identification of changes within such regions demonstrate that the early stages of a species domestication can potentially be associated with specific, strong and direct molecular changes within cis-regulatory elements such as promoters, enhancers, silencers and operators.

Our understanding of the genomic plasticity during the early stages of fish domestication is scarce, and the identification of genetic modifications in response to captivity is proven to be only a single piece of a large molecular puzzle. Deciphering the mechanisms that can rapidly affect the regulation of genes and their expression without causing any physical changes to the DNA sequence will be crucial to improve our knowledge and understanding in fish domestication, with major implications in wild stock management and aquaculture.

1.7 Epigenetics

The word “epigenetics” derives from the Greek prefix epi- that translates to over, above or upon, and -genetics which means “origin” in ancient Greek. Therefore, the term epigenetics describes molecular modifications and features that are “on top of” the traditional genomic sequence that constitutes the basis of the genetic code and inheritance. The ability of epigenetic modifications to regulate transcriptional factors, alter gene expression and determine cell fate is a fundamental property of eukaryotic genomes during the early stages of development. Since this definition was firstly introduced by Waddington (1942), a plethora of studies have adopted a more specific definition as proposed by Holliday (2006), whereby the epigenome promotes alterations in gene expression that persist through cell division without any changes in the DNA sequence. The epigenome consists of all chemical modifications to the genome, including DNA methylation (Holliday and Pugh, 1975;

Hotchkiss, 1948), DNA hydroxymethylation (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009), histone modifications (Kaufman and Rando, 2010), small non-coding RNAs (Buhler, 2009) and protein-protein interactions (Harvey et al., 2018). Each mechanism plays a unique role in gene regulation (i.e. activation, repression and silencing), but they also have stage-specific and multi-layered functionality to fine tune gene expression. For example, while DNA methylation, DNA hydroxymethylation and histone modifications can act directly upon the genetic sequence and alter transcription, small-non coding RNAs are capable of mRNA degradation affecting the translation of existing transcripts within a cell.

Environmental cues usually prompt specific cell responses that allow adaptation to new conditions. Because the rate of epimutations is considerably higher than that of genetic mutations (Schmitz et al., 2011), it is rather likely that populations respond initially to environmental cues through epigenetic variability (Klironomos et al., 2013). The epigenetic adaptation of an organism to new environmental conditions occurs through mitotically stable epigenetic modifications. Epigenetic modifications within germline cells are the basis of an unprecedented variability in the offspring even when the parents are not adapted to new conditions (Jablonka, 2013). Indeed, Rodriguez-Barreto et al. (2019) have shown that the sperm of Atlantic salmon reared in captivity had distinct DNA methylation profile and that the phenotypic variability in hybrid fish (crosses of hatchery reared and wild individuals) could be explained by differential DNA methylation patterns.

Taking into consideration the existing knowledge of epigenetic systems, their ability to alter a phenotype through the regulation of multiple genomic and cellular elements in response to environmental cues and their persistence from one generation to the other, it would therefore be logical to assume that epigenetics play a key role in adaptation, evolution and speciation.

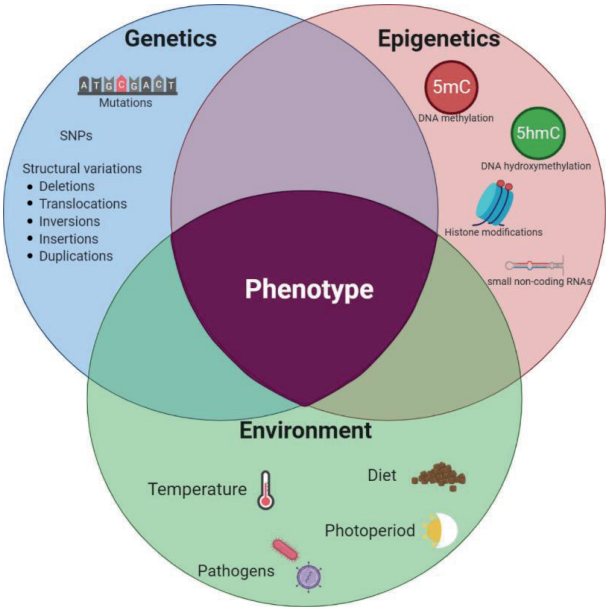


Figure 6. Venn diagram representing how genetic, epigenetic and environmental factors are interconnected and contribute to the resulted phenotype. Created with BioRender.com

1.7.1 DNA methylation

DNA methylation is an epigenetic mechanism by which a single methyl- group is added to the 5th carbon of a cytosine resulting in 5-methylcytosine (5mC) (Hotchkiss, 1948). In mammals, DNA methylation patterns reset following fertilization and the establishment of new modifications upon the DNA sequence arise from the earliest stages of development (Kobayashi et al., 2013; Smith et al., 2012). On the contrary, Jiang et al. (2013) showed that the paternal DNA methylome is inherited in zebrafish early embryos whereas the maternal is gradually discarded through passive demethylation. Although this difference is observed during the early stages of embryonic development, DNA methylation patterns seem to share functionality between mammals and fish at later developmental stages. DNA methylation is an active modification with profound effects in genome stability and involved in several biological processes such as gene transcription, silencing, transposable elements, genetic imprinting and X chromosome inactivation. The enzymes that facilitate the establishment and maintenance of 5mCs across the genome are DNA methyltransferases. In mammals,

three active DNA methyltransferases (*dnmts*) have been identified. DNMT3A and DNMT3B are responsible for establishing de novo DNA methylation (Okano et al., 1998), while DNMT1 serves the role for maintenance (Hermann et al., 2004). A third member of the DNMT3 family, the regulatory protein DNMT3L, shares a similar function with DNMT3A and DNMT3B, however, it lacks catalytic activity (Bourc'his et al., 2001; Hata et al., 2002). Finally, a rather specific methyltransferase that is responsible for the methylation of tRNAAsp is the DNMT2 (tRNA aspartic acid methyltransferase) which is classified as tRNA methyltransferase rather than DNA methyltransferase (Goll et al., 2006). In the absence of these molecules, passive DNA demethylation occurs where 5mCs turn into unmodified cytosines after consecutive cell divisions and genome duplications (Kagiwada et al., 2013). In zebrafish, there are six *dnmt3* paralogues responsible for de novo DNA methylation. In particular, *dnmt6* and *dnmt8* are duplicated *dnmt3a* genes, while *dnmt3*, *dnmt4*, *dnmt5* and *dnmt7* are all *dnmt3b* paralogues (Campos et al., 2012). An important mechanism that can rapidly change DNA methylation states of particular regions across the genome is the active demethylation (Fig. 7) that requires specific methylcytosine dioxygenases, the ten-eleven translocation (TET) enzymes. They actively demethylate 5mCs to 5-hydroxymethylcytosines (5hmCs) by catalyzing the oxidation of 5-methyl groups (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009) while further oxidation results in rather unstable epigenetic modifications, the 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (Ito et al., 2011). The removal of the two latter modifications takes effect by base excision repair (BER) and is carried out by a DNA repair enzyme, the thymine DNA glycosylase (TDG) (Cortellino et al., 2011). Interestingly, TET activity and in particular TET1 and TET2 are more active on 5mC rather than 5hmC substrates (Hu et al., 2015). This supports the notion that 5hmC is a rather stable epigenetic modification that forms gradually after the process of DNA replication (Bachman et al., 2014; Hahn et al., 2014).

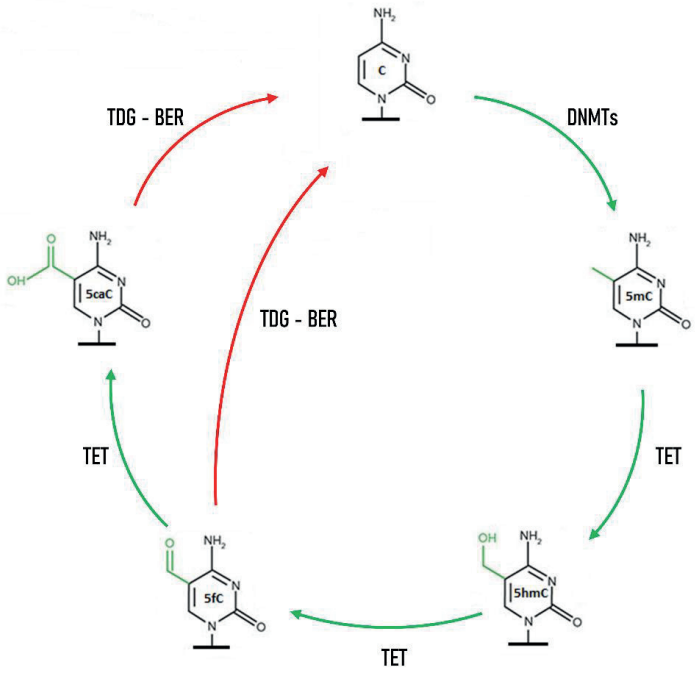


Figure 7. Active demethylation cycle of cytosine nucleoside that involves establishment and maintenance of 5mC by the DNMT enzymes, oxidation of 5mC to 5hmC, 5hmC to 5fC, and 5fC to 5caC by TET enzymes, and final base excision repair by the DNA repair enzyme TDG.

1.7.2 DNA hydroxymethylation

Since its discovery in 2009, DNA 5-hydroxymethylcytosine has been extensively studied and linked to several biological processes such as gene expression (Gao et al., 2019; Greco et al., 2016; Wu et al., 2011a), tissue-specificity (Globisch et al., 2010; Nestor et al., 2012; Ponnaluri et al., 2017) and embryonic development (Ruzov et al., 2011; Sun et al., 2014; Tan et al., 2013; Wu et al., 2011a). It has also been associated with non-coding RNAs (Pan et al., 2016; Zhang et al., 2017), aging (Szulwach et al., 2011) and diseases such as cancer (Bhattacharyya et al., 2017; Wu et al., 2019) and Alzheimer's (Bernstein et al., 2016; Lardenoije et al., 2019). Five years after its discovery, it was recognized as a predominantly stable epigenetic modification that forms after DNA replication as opposed to DNA methylation that occurs

during replication (Bachman et al., 2014). The primary regulator of 5hmC abundance are the TET enzymes as well as their binding sites across genetic elements and structures. Alteration of TET dioxygenase levels within HEK293 cells has been shown to affect DNA methylation dynamics (Grosser et al., 2015). In general, *TET1* overexpression induces overall 5mC oxidation without specific biases across genomic features (i.e. promoters, gene bodies and CpG islands), however, in the presence of toxic substances such as inorganic arsenic, DNA hydroxymethylation reprogramming follows a TET-mediated regulation of CTCF binding sites which are located within the proximal and distal promoter regions of *TET1* and *TET2* (Rea et al., 2018).

An important characteristic of 5hmC, is its enrichment within gene bodies of actively transcribed genes and its positive correlation with gene expression, as opposed to 5mC that has a primarily negative correlation with expression levels (Mellen et al., 2012). In mammals, 5hmCs can be found in all tissues and cell types but their abundance is highly variable in comparison to 5mCs that are relatively consistent among tissues (Globisch et al., 2010; Song et al., 2011). Similar patterns are also observed among vertebrates, with the highest levels of 5hmCs being detected in the nervous system and the brain, intermediate levels in the liver, the muscle and the kidney, and low levels in the thymus gland and testis (Almeida et al., 2012; Kamstra et al., 2015; Wen and Tang, 2014). DNA hydroxymethylation has been characterized by several studies as an important epigenetic modification that plays a key role in gene activation and repression. An early report by Wu et al. (2011a), investigated the impact of 5hmCs within mouse embryonic stem cells. Surprisingly, they found that 5hmCs are probably involved in the establishment and maintenance of chromatin structure for both actively transcribed genes and polycomb-repressed developmental regulators. In later studies, the localization and abundance of 5hmCs within the genome were associated with gene transcription in normal and hypertrophic mouse cardiomyocytes. Interestingly, 5hmCs within distal regulatory regions were positively correlated with gene activation, while within gene bodies were primarily act as pre-activating marks for chromatin remodelling (Greco et al., 2016). Furthermore, DNA hydroxymethylation was also found to be enriched within genomic regions responsible for alternative splicing (Gao et al., 2019). Considering that alternative splicing is largely used by eukaryotes for the production of tissue- and cell-

specific protein isoforms (Yeo et al., 2004), and that 5hmC levels differ significantly between tissues (Globisch et al., 2010), it is rather likely that 5hmC may play key role in tissue-specific alternative pre-mRNA splicing and consequently contribute significantly in gene transcription and function.

Although DNA demethylation is involved in several biological processes, it remains unknown how this molecular system operates in regulating gene expression. For example, 5hmC modifications are generally associated with gene activation, however interactions of methyl-CpG-binding proteins with 5hmC residues are responsible for the recruitment of chromatin remodeling enzymes and gene repression (Yildirim et al., 2011). In any case, these findings reveal that DNA hydroxymethylation can drastically affect the function of an organism at multiple levels but more research is necessary to unfold its true potential. Considering the impact of the environment upon the epigenome and the dynamic nature of DNA hydroxymethylation within cells, tissues and organisms, it is likely that changes of 5hmC profiles during animal domestication contribute to phenotypic adaptation and evolution under captivity.

1.7.3 Histone modifications

Even though the molecular pathways and gene expression changes are often well-studied, little is known about the underlying mechanisms that integrate these cues at the chromatin level. Chromatin consists of nucleosome units and each unit contains an octamer of histones and 147 base pairs of DNA that are active and susceptible to structural changes (Luger et al., 1997). H2A, H2B, H3 and H4 are the four pairs of core histones that form an octamer. These histones can be epigenetically modified with various molecules, including acetylation, methylation, phosphorylation, ubiquitination, deamination and ADP ribosylation among others (Bannister and Kouzarides, 2011). Histone modifications occur in different domains, namely, the N-terminal tail and core domains (Mersfelder and Parthun, 2006). The N-terminal tails flank the core domains and are relatively unstructured with multiple sites that are susceptible to post-translational modifications. Due to their ability to extend around the nucleosome, modified tail domains are capable of influencing the chromatin through

changes in electrostatic charge or changes in the interaction with non-histone proteins (Grewal and Moazed, 2003; Iizuka and Smith, 2003; Jenuwein and Allis, 2001; Martin and Zhang, 2005). Similarly to N-terminal tails, the nucleosome face domain which is one of the core domains, can potentially regulate non-histone protein interactions with the nucleosome or directly influence the structure by nucleosome-nucleosome interactions. One of the most well-characterized modifications of the nucleosome face is the histone H3 lysine 79 methylation that is highly conserved across eukaryotes and belongs to a genetic region that is crucial for transcriptional silencing (Park et al., 2002). Apart from the nucleosome face domain, several histone modifications have been identified and mapped to the nucleosome lateral surface that is either in close proximity or binds directly to the DNA that is wrapped around the histone octamers. Despite the lack of specific binding sites and reading frames, modifications on the nucleosome lateral surface have direct impact on transcription *in vitro*. These findings pinpoint the significance of those modifications and places them as drivers of gene expression rather than post-translational by-products (Lawrence et al., 2016). The last and most basic, but equally important for the structural integrity of the nucleosome interface, are the histone-histone modifications that allow the formation of the H3/H4 tetramer, the H2A/H2B dimer and the tetramer-dimer interaction that comprise the histone octamer. For example, the acetylation of H4 lysine 91, is found within the interaction region of H4 and H2B histones and plays an important role of tetramer-dimer stabilization (Lawrence et al., 2016; Santisteban et al., 1997).

In the context of animal domestication, Fan et al. (2012) have identified sulforaphane supplementation to act as histone deacetylase inhibitor in porcine satellite cells. This inhibition is responsible for the repression of myostatin with major implications in growth potential. Furthermore, in changing environments (i.e. wild animals in captivity), histone variants play major roles in mediating both short- and long-term responses (Talbert and Henikoff, 2014). The same histone modifications that contribute to adaptation can also be passed to the next generations through their mitotic and meiotic persistence, suggesting that such marks are involved in transgenerational epigenetic inheritance in both mammals (Brykczynska et al., 2010) and zebrafish (Wu et al., 2011b).

1.7.4 Non-coding RNAs

The central dogma of molecular biology in its generalized form is the transfer of the genetic information from DNA to transcribed RNA and from RNA to translated protein. However, not all RNAs are translated. These RNAs are called non-coding RNAs (ncRNAs) and are subdivided into housekeeping and regulatory ncRNAs (Wei et al., 2017). Among the regulatory ncRNAs, two clusters are distinct based on their size (Ponting et al., 2009; Zaratiegui et al., 2007). The short-chain ncRNA cluster includes the small nuclear RNAs (snRNAs) which are involved in splicing, the small nucleolar RNAs (snoRNAs), which are responsible for nucleotide modifications on ribosomal and other RNAs (Collins et al., 2011) as well as the short interfering RNAs (siRNAs), microRNAs (miRNAs) and Piwi-interacting RNAs (piRNAs) that since their discovery have opened a new frontier in understanding how gene expression is regulated (Carmell et al., 2002; Hamilton and Baulcombe, 1999; Lee et al., 1993; Lin and Spradling, 1997; Reinhart et al., 2000). The second cluster consists of long ncRNAs (lncRNAs), which have a typical length of above 200 nucleotides (nt) and are also associated with the regulation of gene expression.

As mentioned above, an epigenetic mechanism has the ability to affect gene expression without any changes in the DNA sequence itself. Therefore, ncRNAs that retain the ability to introduce translational changes are considered epigenetic regulators. For example, miRNAs have a length of approximately 21-25 nt, they are single-stranded RNAs (ssRNAs) that are able to fold into double-stranded RNAs (dsRNAs) and after processing and binding to complexes such as the RNA-induced silencing complex, can target specific messenger RNA (mRNA) regions and regulate their translation by silencing or degradation (Lim et al., 2005; Zhang and Su, 2009). For example, in Senegalese sole (*Solea senegalensis*) high expression of certain miRNAs such as miR-17a, miR-181-5p and miR-206-3p promote myogenesis (Campos et al., 2014). The miRNA miR-206 was also found to modulate *igf-1* expression in Nile tilapia, with major implications in growth (Yan et al., 2013). Furthermore, siRNAs are involved in the post-transcriptional gene silencing and transcriptional gene silencing pathways (Golden et al., 2008; Verdel et al., 2009), while piRNAs interact with Piwi family proteins that are linked to transposon control and silencing (Brennecke et al., 2008; Malone and Hannon, 2009). Because epigenetics are inextricably linked to early development and

cell fate, it is well known that X-chromosome inactivation in mammals is controlled by both long and short-chain ncRNAs which are produced by the transcription of two genes, the *Xist* and *Tsix*. XiRNAs which are slightly longer (24-42 nt) than the rest of the short-chain ncRNAs, are involved in X-chromosome inactivation by controlling the methylation of the inactive X chromosome as well as their parent long ncRNAs that are critical for the recruitment of different chromatin remodeling complexes (Autuoro et al., 2014; Gontan et al., 2011).

1.8 The role of epigenetics in animal domestication

Animal domestication is a long process and it is linked to significant changes of various molecular pathways across several generations. However, substantial changes in gene expression appear from the very first generation, when animals become exposed to new environmental conditions (Christie et al., 2016; Konstantinidis et al., 2020). Such rapid changes cannot be explained solely by genetic mutations and allele shifts, therefore we hypothesized that epigenetic mechanisms may play key role in animal domestication. Although epigenetic modifications such as DNA methylation, histone modifications, and non-coding RNAs have been well-known for several decades, the study and association of epigenetics to different phenotypes was hindered by the lack of genomic resources. Major leaps in sequencing technologies have increased the interest for the research of epigenetic modifications within various fields such as nutrition, growth, immunity and disease. In fish and more importantly in aquaculture, traits that can potentially be epigenetically regulated, such as disease resistance and growth performance, are of great interest.

1.8.1 Nutritional epigenetics

One of the most significant changes in wild animals that undergo domestication is their nutrition. Diet has been associated with strong epigenetic changes that can be mitotically stable and persist across generations through the epigenetic modification of the germline. Studies in mammals (Faulk and Dolinoy, 2011; Feinberg et al., 2010; Pons et al., 2011; Tobi et al., 2011; Wiedmeier et al., 2011), other vertebrates including teleost species (Adam et

al., 2019; Dhanasiri et al., 2020; Skjaerven et al., 2018) and invertebrates (Kaminen-Ahola et al., 2011; Niculescu and Lupu, 2011) have shown that nutrition is a major environmentally-induced epigenetic modifier in several tissues, and therefore, could play key role in animal domestication and adaptation. In particular, parental one-carbon micronutrient deficiency and high dietary arachidonic acid in zebrafish, were both associated with liver DNA methylation changes and differential gene transcription of the progeny that was fed control diet (Adam et al., 2019; Skjaerven et al., 2018). Additionally, polyunsaturated fatty acids are capable of altering the epigenome (Burdge and Lillycrop, 2014) and are involved in eicosanoid synthesis, a critical element for fish development and growth (Sargent et al., 1995).

1.8.2 Epigenetic mechanisms of growth

The identification of genetic markers using high-density single nucleotide polymorphism (SNP) chips has improved previous breeding practices and favored the selection of beneficial traits in aquaculture (Houston et al., 2014). However, traits are not exclusively regulated by genetic variations and allele frequencies. In particular, growth is a multi-variable trait, highly species- and context-dependent and affected by the expression of a large number of genes involved in several pathways and expressed in various tissues, as well as by environmental cues.

One of the most crucial environmental factors is temperature. For example, in several teleost fishes such as zebrafish (*Danio rerio*) (Schnurr et al., 2014; Scott and Johnston, 2012), Senegalese sole (Campos et al., 2013b), Nile tilapia (Sun et al., 2016a) and Atlantic salmon (*Salmo salar*) (Finstad and Jonsson, 2012). differences in embryonic incubation temperature affect sex determination, growth and metabolism. In fish, several studies have identified a phenomenon called “thermal imprinting”, i.e. a long-lasting effect on muscle structure (hypertrophy and hyperplasia) and growth rate that is triggered by the ambient temperature during embryonic development (Alami-Durante et al., 2007; Fernandes et al., 2007; Galloway et al., 1999; Johnston et al., 2000). Even though the underlying mechanisms of thermal imprinting were largely unknown, recent studies have linked several epigenetic

mechanisms to the observed thermal regulation. In particular, DNA methylation of crucial myogenin and myosin heavy chain (MyHC) genes as well as metabolic enzymes and several microRNAs related to the mTOR and MAPK pathways are affected by the embryonic incubation temperature (Burgerhout et al., 2017; Campos et al., 2013a; Johnston et al., 2009; Schnurr et al., 2014). MicroRNAs have also been implicated in growth by modulating the expression of important growth-related genes. For example, in Nile tilapia miR-206 was found to repress the expression of insulin-like growth factor 1 (IGF-1) while miR-206 antagomir reduced the expression of mir-206 and increased the expression of IGF-1 (Yan et al., 2013). Furthermore, in Senegalese sole (*Solea senegalensis*) several miRNAs where found to be linked to growth (miR-17a, miR-181-5p and miR-206), lipid metabolism (miR-122) and myogenesis (miR-181-3p) (Campos et al., 2014). These findings reflect the importance of epigenetic mechanisms towards the regulation of major genes and pathways involved in growth with long-lasting effects.

Total body mass and growth potential are often associated with sexual dimorphism in most studied teleost species. Notably, tilapia aquaculture is dominated by all-male populations due to the higher growth rates of males as opposed to females (Lorenzen, 2000; Turner and Robinson, 2000). Despite the identification of sequence-specific markers linked to sex determination in Nile tilapia (Caceres et al., 2019; Conte et al., 2017; Eshel et al., 2011; Li et al., 2015), gonadal differentiation is highly malleable and temperature has been recognized as a strong aspect in sex determination and ultimately total body mass and growth (Baroiller and D'Cotta, 2001; Devlin and Nagahama, 2002).

Other environmental factors such as pathogens, parasites, salinity, crowding, dissolved oxygen and ammonia affect growth performance. However, the molecular responses to these cues and the involvement of epigenetic mechanisms are not well-studied.

1.8.3 Epigenetic mechanisms in immunity

The immune system is a fundamental biological network of cells and proteins that is designed to distinguish self from non-self and defend against environmental challenges and foreign bodies. Immunity relies both in the innate and adaptive immune systems, their heterogeneity, their immune cell repertoire and their highly diverse antigen receptor rearrangements (Busslinger and Tarakhovsky, 2014). Epigenetic profiles among immune cells differ in cell- and lineage-specific manner, and they are related to enhanced diversification and cell adaptation as a response to environmental stimuli. In particular, the recombination substrates of the antigen receptor genes, namely, the variable (V), diversity (D) and joining (J) gene segments as well as the recombination signal sequences (RSSs) (Stanhope-Baker et al., 1996; Yancopoulos and Alt, 1985) are under stringent epigenetic control (Hewitt et al., 2010). Epigenetic mechanisms are also involved in chromatin conformation and looping in hematopoietic cells including monocyte differentiation (Phanstiel et al., 2017), B cell activation (Bunting et al., 2016; Kieffer-Kwon et al., 2017), T cell development (Hu et al., 2018) and differentiation (Mumbach et al., 2017).

In mice, the presence of TET enzymes in T and B cell differentiation and function is critical. Hydroxymethylation of gene bodies is positively correlated with gene expression and enhancer activity (Huang et al., 2014; Tsagaratou et al., 2014; Wu et al., 2011a). *Tet2* is responsible for global B cell genome demethylation following differentiation (Kulis et al., 2015) while mutations in *Tet2* and *Tet3* genes that lead to the loss of TET function are often associated with cell transformation and arrested cell differentiation in the hematopoietic system (Ko et al., 2015; Lio et al., 2016; Orlanski et al., 2016). Recent studies using multiple fish species have identified both genome-wide and gene-specific epigenetic modifications as a response to environmental cues and particularly to pathogens. For example, histone modifications can result in increased cytokine production (Galindo-Villegas et al., 2012), DNA methylation patterns of genomic regions that are associated with disease susceptibility (Shang et al., 2015; Shang et al., 2016; Shang et al., 2017) and several miRNAs that target and block the translation of viral proteins (Yan et al., 2015) as well as regulate specific host immune responses under bacterial infections (Andreassen and Hoyheim, 2017).

1.8.4 Transgenerational epigenetic inheritance

Recent studies have shown that epigenetic modifications have a significant impact in cell fate and reprogramming during the early stages of development (Perez and Lehner, 2019). Epigenetic modifications and particularly DNA hydroxymethylation, which is the focus of this thesis, play a major role during early embryonic reprogramming. For example, from the early stages of mammalian zygotic formation to the later stages of embryonic development and prior to zygotic genome activation, cells accumulate 5hmC residues along with a reduction of 5mC (Wossidlo et al., 2011). In the absence of appropriate DNA methyltransferases, DNA demethylation could be explained by passive loss of 5mC during DNA replication. However, the accumulation of 5hmC suggests an active mechanism that both methylates and oxidizes during cell division. Therefore, the presence of parental DNA methyltransferases and ten-eleven translocation enzymes in the egg and sperm potentially play a significant role in epigenetic reprogramming, gene regulation and normal development. Furthermore, epigenetic inheritance can be a major factor for adaptation in new environments since it is closely associated with phenotypic plasticity (Burggren, 2016) and transcriptionally relevant epimutations occur much faster than genetic mutations (van der Graaf et al., 2015). Repetitive and transposable elements are important features of most eukaryotic genomes. Because of their ability to duplicate and change their position within the genome (transposition), they have been characterized as major mechanisms of genome evolution. Although transposons can give rise to novel coding sequences that promote survival and offer advantageous characteristics to their hosts, they can also create dysfunctional DNA combinations that result in abnormal phenotypes. Epigenetic mechanisms are largely involved in the regulation of these elements. A previous report has shown that DNA demethylation of transposable elements can lead to rapid changes in cell behaviour, arrest of cell proliferation and cell death (Coluccio et al., 2018). Understanding the epigenetic regulation of such elements that facilitate genomic and regulatory evolution, will require an in-depth mapping distinguishing between DNA methylation and DNA hydroxymethylation modifications.

1.9 Methods for 5hmC profiling

After the discovery of the TET enzymes in 2009 (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009) several research papers investigated the abundance and mapping of 5hmCs in various experimental settings (Ficz et al., 2011; Jin et al., 2011; Munzel et al., 2010; Pastor et al., 2011; Song et al., 2011; Williams et al., 2011; Wu et al., 2011a; Xu et al., 2011). However, the resolution of those approaches was poor and the conclusions drawn were broad. The first quantitative sequencing of 5hmCs at single-nucleotide resolution was introduced with the use of the oxidative bisulfite sequencing (oxBS-Seq) (Booth et al., 2012) and Tet-assisted bisulfite sequencing (TAB-Seq) (Yu et al., 2012) methods. Research was necessary for the optimization of 5hmC profiling, since there were several disadvantages of these two methods, namely harsh treatment of the DNA with bisulfite, laborious and time consuming, need for parallel subtractive sequencing, high sequencing depth and high starting concentration of DNA. In that regard, a novel approach called Reduced Representation 5-Hydroxymethylcytosine Profiling (RRHP) (Petterson et al., 2014) enabled faster workflows and avoided bisulfite treatments due to the enzymatic 5hmC-specific fragmentation of the DNA. Additionally, it only required 100 ng of DNA as input which made it suitable and applicable in many more research fields where DNA samples are scarce. By exploiting the properties of the beta-glucosyltransferase, which inhibits digestion of 5hmC junctions between ligated adapters and genomic libraries, selective sequencing of exclusive 5hmC adapterized fragments was possible (Figure 8). After sequencing, the second cytosine of each fragment corresponds to a 5hmC which is then mapped to the reference genome. A direct comparison of 5hmC levels among several biological samples can be obtained by comparing the total number of fragments per 5hmC site. RRHP uses the restriction enzyme MspI which recognizes CCGG sequences across the genome. The combination of multiple restriction enzymes (e.g., Taqα1) and the use of sequence specific adapters can increase the CpG coverage and total number of 5hmC sites across the genome. Additionally, RRHP can be used for the detection of 5hmCs in non-CpG context with the use of AluI, HaeIII, HpyCH4V and CviQI restriction enzymes (Sun et al., 2016b). Furthermore, methylation-sensitive restriction enzymes such as the HpaII and parallel library preparation with MspI, provides both 5mC and 5hmC profiles from the same biological sample. Finally, by avoiding harsh

chemicals such as bisulfite-treatment that deaminates unmethylated cytosines to uracils, allows the use of RRHP sequencing reads to be used for the discovery of SNPs across the genome.

For the preparation of RRHP libraries, all samples have to be prepared simultaneously. The initial genomic DNA concentration of every sample has to be quantified using a method that provides high accuracy and sensitivity, such as fluorimetry or quantitative eloctrophoresis. Randomization of samples (blinding) is highly recommended because it diminishes group-related biases. Even though the same concentration of DNA is being used across all libraries, samples with higher levels of 5hmCs will produce a higher number of adapterized fragments. That is the case because 5hmCs at the junction of each fragment are protected against the second MspI digestion, as shown in Figure 8, and samples with higher levels of 5hmC contain more of these fragments. Finally, the protocol uses limited amplification (10 - 12 PCR cycles) to avoid increased bias due to the amplification of shorter fragments and the sequencing is performed on the basis of equal volume as opposed to equal molarity, since the latter diminishes the 5hmC differences among samples.

Currently, commercially available RRHP library preparation kits are established only for small number of samples. In total, 6 single indices are provided which are based on the TruSeq-Illumina sequencing technology. Taking into account the sequencing depth required per library (20-30 M reads) and the number of samples per experiment, the sequencing capabilities and output of Illumina sequencing platforms such as the HiSeq, are necessary. However, HiSeq has faced issues with index-hoping in single indexed libraries and it was necessary to modify both the P5 and P7 adapters by introducing dual-indexing. Also, RRHP libraries are designed to have a specific pattern at their 5’ end (CCGG), which creates a low-complexity issue across the entire flow cell for the first 4 sequencing cycles in all Illumina platforms. To compensate for the low-complexity, PhiX control is added at high concentrations (approximately 30-35%), with significant reduction in sequencing output. Finally, a bioinformatic pipeline for the analysis of RRHP libraries is not available yet. Thus, advanced bioinformatic knowledge in python and R programming language, as well as Linux command line, are a prerequisite for the analysis of RRHP data sets. The final output of RRHP

datasets are count-based, which makes their statistical analysis similar to that of RNA-Seq datasets.

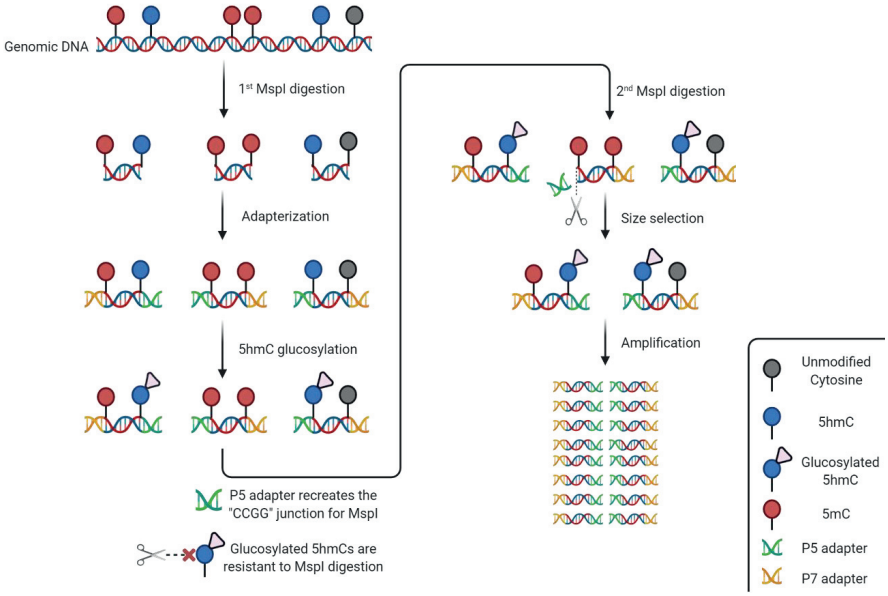


Figure 8. Schematic flow representing the RRHP library preparation protocol. Created with BioRender.com.

Therefore, sample sizes and sequencing depth were decided based on the literature and the experimental conditions. For example, Ching et al. (2014) have reported that a sample size of 5 individuals (n=5) and a sequencing depth of 20 million reads are able to distinguish a 10-fold transcriptional difference between mRNA and lincRNAs for an RNASeq count-based dataset. Furthermore, according to Petterson et al. (2014), RRHP can detect low 5hmC abundances (down to 20%) with a false discovery rate of less than 5% ($q<0.05$) and a sequencing depth of only 20-30 million reads. This was shown in human samples that contain approximately 4.6 million CCGG sites. Throughout all the experiments of this thesis, we have targeted an average sequencing depth of approximately 40 million reads, even if the Nile tilapia genome contains only 1.6 million CCGG sites. This multiplies the 5hmC detection rate in our datasets by 6-8 times compared to the above report. For our second experiment, we also took into consideration the weight and length of the large and small full-siblings since we aimed towards the association of 5hmC changes in response to the

phenotype. Based on the formula below, we calculated the sample size (n) for comparisons between two groups when the endpoint is quantitative data (5hmC counts).

$$n = 2 \cdot (SD \cdot (Z_0 + Z_\emptyset)/d)^2$$

where “SD=214” is the standard deviation among all individuals, “Z_0=1.96“ from Z table and a=0.05, “$Z_\emptyset$=0.842” from Z table (power at 80%) and “d=422” the difference between mean weights. This formula resulted to a sample size of n=4.04, which is also in agreement with the literature. We kept the sample size to 5-6 replicates in order to minimize the number of animals used in the experiments, in agreement with the guidelines from the Norwegian Animal Research Authority Mattilsynet.

Another approach for fast and sensitive 5hmC quantification of genomic DNA is the liquid chromatography mass spectrometry (LC-MS/MS) (Le et al., 2011). Even though LC-MS/MS is the gold standard in quantitation methodology, this method can measure only global 5hmC levels among biological samples in contrast to RRHP, which quantifies at single nucleotide resolution. The identification of key regions and genes that are captured by RRHP can easily be obtained by annotating sequenced reads to the reference genome. However, to investigate whether these modifications are functionally-relevant, further experiments are required. While the use of knowledge bases and chromatin state predictions are usually available only for mammals, associations of epigenetic profiles with gene expression levels provide a more holistic view on the potential role of these modifications. RNA sequencing (RNA-Seq) is a technique that uses DNA-based next-generation sequencing technologies in order to construct data sets that contain the transcriptomic profiles of the genome. In particular, researchers are able to quantify precisely transcript levels, and map interesting regions that are involved in splicing events such as exon-intron junctions and transcription start sites (Tsuchihara et al., 2009). The combination of 5hmC profiles, RNA-Seq and SNPs from a single biological sample increases the robustness of data integration and allows for the assessment of their functional importance.

2. Objectives

Current aquaculture breeding programs rely on genetic and phenotypic screening of broodstock for improvement of relevant traits. However, in a continuous pursuit to close the gap between supply and demand for fish consumption, species diversification and selective breeding are essential. Because the domestication of new fish species is greatly affected by new environmental stimuli, we hypothesized that epigenetic mechanisms are involved in domestication and are likely responsible for the earliest phenotypic changes that occur in captivity. Since growth is a trait of great importance for the aquaculture, the overall goal of this thesis was to explore the complex network of epigenetic and gene expression changes that occur during the first stages of fish domestication, with focus on DNA hydroxymethylation and myogenesis.

Specific objectives include:

1. The comparison of DNA hydroxymethylation and gene expression profiles between wild Nile tilapia and their offspring reared in captivity (Paper I).
2. The investigation of DNA hydroxymethylation differences associated with distinct growth phenotypes in Nile tilapia (Paper II).
3. The characterization of hydroxymethylomes within the somatotropic axis (Paper III).

3. General Discussion

Aquaculture is a food production industry, that has expanded dramatically and continuously for almost four decades. The primary characteristic that changed farming practices and production output was its transformation from a casual- to a research-based industry. This contributed greatly towards the improvement of breeding strategies and selection for traits that are of great importance such as immunity, growth, sexual maturation and flesh quality. Nevertheless, aquaculture is still facing major challenges including climate change (Bell et al., 2013; Rosa et al., 2012), ocean warming (Lyman et al., 2010) and acidification (Clements and Chopin, 2016). One of the most promising options to overcome such barriers is the diversification of the industry along with further development of the existing breeding programmes. Currently, aquaculture production and profitability are heavily based upon the cultivation of a few freshwater fish species, including the common carp (*Cyprinus carpio*), Nile tilapia (*Oreochromis niloticus*), rainbow trout (*Oncorhynchus mykiss*) and the striped catfish (*Pangasianodon hypophthalmus*) as well as marine species such as the Atlantic salmon (*Salmo salar*), the European seabass (*Dicentrarchus labrax*) and seabream (*Sparus aurata*) and the Atlantic bluefin tuna (*Thunnus thynnus*). To achieve a significant increase in global fish production and support the worldwide demand for fish, the integration of new fish species is essential. However, without the identification of the key mechanisms that promote adaptation to farming conditions and drive genome evolution, the domestication of new fish species will be a process that requires several trials and potentially decades.

From the complex experimental design to the establishment of a small-scale breeding programme from wild Nile tilapia and the interpretation of the DNA 5-hydroxymethylcytosine profiles, we believe that so far this thesis constitutes the most complete characterization of the DNA hydroxymethylome in fish undergoing domestication.

3.1 Importance of DNA 5-hydroxymethylation during the early stages of fish domestication

The hydroxymethylome of Nile tilapia was greatly affected by the environmental conditions under captivity (Papers I and II). In particular, the comparison of DNA hydroxymethylation profiles between wild and first generation of wild-born fish reared in captivity revealed significant differences within genes involved in immune and growth-related processes (Paper I). These differences can easily be explained as early domestication effects. Overall, wild fish were exposed to various biotic factors (such as viruses, parasites and other pathogens) that potentially increase the expression of immune-related genes as opposed to their captive offspring reared in a comparatively aseptic environment. Furthermore, the energy status of wild fish compared to their offspring in captivity was expected to be largely different. In general, organisms have to use their finite amount of energy to grow and protect themselves by mobilizing their immune system. The interplay between nutrition (energy source), growth and immunity is highly dependent of an organism's diet and environmental conditions (e.g., presence of pathogens). The diet of wild Nile tilapia is primarily plant-based although sometimes it contains small vertebrates, such as worms, snails and crustaceans. However, in our recirculating aquaculture system, Nile tilapia were fed highly formulated diets that promote growth and improve fish health. Taking into consideration all the above, it is expected that domestication and all its associated black box of nutritional and environmental changes, will most certainly affect the epigenome and consequently the regulation of genes and phenotype. In similar studies, DNA methylation profiles of European seabass (Anastasiadi and Piferrer, 2019), Pacific salmon (Le Luyer et al., 2017) and Atlantic salmon (Rodriguez-Barreto et al., 2019) were associated with phenotypic divergence between wild and first generation of hatchery-reared individuals with modifications that persisted in adulthood as well as the next generation through the germline. Phenotypic divergence in terms of total weight and length among full siblings was also observed during our breeding program. By investigating both their genetic and hydroxymethylome profiles, we found that the differences between large- and small-sized individuals were primarily attributed to differentially hydroxymethylated genes involved in growth and growth-related pathways such as the PI3K-Akt, MAPK, Rap1 signaling and focal

adhesion pathways, rather than to genetic mutations and single nucleotide polymorphisms (Paper II). Most of these genes were found to be hyper-hydroxymethylated in the large group of fish, which is in agreement with the overall positive correlation between 5hmC enrichment and gene expression in mammalian cells (Ficz et al., 2011; Pastor et al., 2011; Wu et al., 2011a). Functionally relevant 5-hydroxymethylcytosines were also observed when different tissues were compared among full siblings. DNA hydroxymethylation appeared to be significantly different among tissues and particularly enriched within the liver when compared to muscle and pituitary (Paper III). Regardless of the experimental setup, DNA hydroxymethylation was always identified as a functionally relevant mechanism that could explain the observed phenotypic characteristics.

3.2 A link between DNA hydroxymethylation of myogenic genes and somatic growth

The first associations between DNA hydroxymethylation and growth were unveiled by comparing the fast muscle of wild Nile tilapia from Egypt against the first generation of wild-born individuals reared in captivity (Paper I). Differentially hydroxymethylated cytosines (DhmCs) were identified within growth-related differentially expressed genes between groups, namely *myo5b*, *creb5* and *bmp1*. Bone morphogenetic proteins are growth factors involved in the formation of bone, cartilage and their regeneration. They are critical for cell-cell adhesion and cell-extracellular matrix interactions and play an orchestrating role in tissue architecture (Kanjilal and Cottrell, 2019; Reddi and Reddi, 2009; Wagner et al., 2010). *Bmp1* was found to be both hyper-hydroxymethylated and upregulated in wild individuals compared to the first-generation of Nile tilapia undergoing domestication. Considering that several genes involved in metabolism and muscle-specific functions were overall downregulated in the wild fish, the overexpression of *bmp1* is likely related to immune and wound healing processes rather than growth-related functions (Muir et al., 2016). This assumption is based upon the presence of parasites across several organs of the wild individuals causing overall inflammation and potentially skin lesions. Hyperhydroxymethylation and overexpression of both *myo5b* and *creb5* in the wild group of

fish could potentially be associated with the function of PI3K-Akt pathway, which is crucial in cell cycle and proliferation. Nascimento et al. (2013) showed that myosin Vb contains a lysine cluster which is conserved among mammals and it is involved in the binding of phosphatase and tensin homolog protein. This interaction can potentially play important role in the downstream regulation of PI3K-Akt signaling pathway because *myo5b* targets and activates *pten* which in turn supresses the phosphorylation of AKT and PDK1. Although, we did not identify significant differences in the expression of downstream kinases such as the serine/threonine-protein kinase MTOR, our study revealed for the first time that *myo5b* is epigenetically-regulated in skeletal muscle during fish domestication with potential implications in somatic growth.

Surprisingly, the PI3K-Akt pathway was linked with changes in DNA hydroxymethylation at multiple levels including tissue-specific profiles (Paper III). By comparing second-generation large- and small-sized full sibs, we identified several differentially hydroxymethylated genes such as *akt1*, *atrn*, *dab2ip*, *magi2*, *prkcq* and *ptk2* (Paper II). AKT serine/threonine kinase 1 (*akt1*) is a critical signaling molecule involved in growth factor receptor-mediated signaling. Overexpression of *akt1* has been linked to cell growth (Franke, 2008) while its downregulation is responsible for the reduction of satellite cell proliferation in the skeletal muscle and the reduction of muscle hypertrophy (Moriya and Miyazaki, 2018). *Atrn* codes for the protein attractin and is indirectly involved within the PI3K-Akt pathway. In particular, it interacts with the transducer COP9 signalosome subunit 5 (*cops5*) which in turn targets *p53* to degradation (Bech-Otschir et al., 2001; Wang et al., 2008). Furthermore, the disabled homolog 2-interacting protein is coded by the *dab2ip* gene and is able to supress the PI3K-Akt pathway by activating the ASK1-JNK that leads to cell apoptosis. Its knockdown in the glandular epithelium is associated with hyperplasia and apoptotic defects (Xie et al., 2009). Similarly to the function of *myo5b* within the PI3K-Akt from paper I, the membrane-associated guanylate kinase, WW and PDZ domain-containing protein 2 (*magi2*) cooperates with *pten* to supress the PI3K-Akt signaling pathway (Matsuda and Kitagishi, 2013). Additionally, protein kinase c theta type (*prkcq*) is responsible for the phosphorylation of the IRS1 at a serine 1101 (Ser^{1101}) in mice and inhibits insulin signaling and reduction of IRS1 signaling to the PI3K-Akt cascade. Lastly, focal adhesion kinase 1 (*ptk2*) controls fibroblast

survival through the simultaneous phosphorylation of p85 subunit of PI3K and the Ser^{473} of Akt (Xia et al., 2004). Interestingly, all of the aforementioned genes excluding *akt1*, where also differentially hydroxymethylated between tissues and especially between liver and muscle (Paper III). These findings reveal that DNA hydroxymethylation is highly involved in the regulation of growth-related signaling pathways. This notion was further supported by performing 5hmC mapping of the somatotropic axis (Paper III). By comparing the hydroxymethylome profiles of the muscle, liver and pituitary we discovered several differentially hydroxymethylated genes among these tissues that comprise the somatotropic axis. Among these genes, we identified several fibroblast growth factors (*fgfs*), receptors (*fgfrs*) and receptor substrates (*frs*). Genes that code for FGF signaling proteins and their receptors are involved in a broad spectrum of biological processes including cellular proliferation, survival and normal development. The Ras/MAP, PI3K-Akt and PLCγ signaling pathways are the predominant pathways affected by their activity (Yun et al., 2010). In addition, we identified several growth differentiation factors (*gdfs*) which are involved in central nervous system development, bone ossification and skeletal morphogenesis (Buxton et al., 2001; Cunningham et al., 1995; Hino et al., 2004). Interestingly, we also identified differentially hydroxymethylated insulin-like growth factor binding proteins (*igfbps*) and receptors (*igfrs*) as well as the growth hormone releasing hormone (*ghrh*) and the growth hormone secretagogue receptor type 1 (*ghsr*). Collectively, they regulate IGF signaling and their interactions with insulin are responsible for the modulation of carbohydrate metabolism. Finally, the myocyte enhancer factors *mef2c* and *mef2d*, the myogenic factor 5 (*myf5*), the multiple epidermal growth factor-like domain protein 10 (*megf10*) and the myosin heave chain fast skeletal muscle (*myh*), also found to be differentially hydroxymethylated suggesting that DNA hydroxymethylation is linked to muscle-specific functions. Across all experiments (Paper I, II and III) we used the RRHP data for the discovery of SNPs. It is a shortcoming that our sample size was too small to provide enough statistical power for a comprehensive detection of SNPs across the genome. For example, to ensure the discovery of both common and rare single nucleotide polymorphisms of poorly characterized primate species, Trask et al. (2011) have concluded that at least 10 individuals are required. Furthermore, traits such as growth are highly polygenic and affected by a vast

number of mechanisms and genomic loci. Additionally, RRHP is a reduced representation genome-wide profiling method which captures polymorphisms upon or close to CCGG sites as opposed to genome-wide association studies (GWAS) that are able to capture a much wider proportion of the genome and detect many more polymorphisms. Taking all the above into consideration, the use of RRHP datasets for the discovery of SNPs seem to be lacking power to convincingly display an association between polymorphisms and phenotypic divergence. However, CCGG sites occur throughout the whole genome and particularly within genes providing a positive representation of transcriptionally relevant genomic regions. Also, the large phenotypic differences in terms of weight and length among full siblings (Paper II), could have been due to strong genetic effects such as fixed SNPs rather than random and synchronous mutations across multiple individuals. Using the RRHP, we did not detect fixed SNPs between fast- and slow-growers. These findings, pinpoint towards a rapid adaptive plasticity due to environmental factors which potentially can be epigenetically inherited in a few individuals due to variational dynamics among gametes.

3.3 Epigenetic regulation of immune-related genes during fish domestication

Disease resistance is one of the most critical traits in intensive aquaculture systems. That is primarily due to disease outbreaks within a crowded environment. For that purpose, genetic evaluations (i.e., QTLs and marker assisted selection) have been integrated for selection of broodstock. A prime example of genetic selection for resistance against a viral infection is the infectious pancreatic necrosis (IPN) in farmed Atlantic salmon that was reduced by 90% within 8 years. However, not all infectious diseases are dependent upon a single or a few genetic markers because most traits are highly polygenic. In that regard, epigenetic modifications can be a promising tool for immunity-related trait selection. For example, Boltana et al. (2018) have shown that behavioural fever (the choice of an ectotherm organism to seek the best thermal condition) impacts both histone modifications and the transcription of genes that improve immune responses through lymphocyte proliferation and proinflammatory cytokine release. Although this behaviour can be observed in the wild

(i.e. vertical migration in the water column above or below the thermocline), it is rather unlikely in captivity and particularly in recirculating aquaculture systems where the temperature is monitored and stable throughout an organism's life. Therefore, the process of domestication does not exclusively provide optimal living conditions but also eliminates mechanisms of adaptation and potentially reduces immunocompetence. The innate immune system is considered to have limited specificity against pathogens but several studies (reviewed by Netea and van der Meer (2017)), have shown that epigenetic programming is involved in trained immunity , i.e. by the presence of adaptive characteristics upon differentiation of monocytes to macrophages in humans (Saeed et al., 2014). The latter study, identified histone modifications such as the acetylation of the 27th lysine residue of the Histone H3 (H3K27ac) as well as the methylation and trimethylation of the 4th lysine residue of the same histone (H3K4m1 and H3K4m3) within promoter and enhancer regions that likely provide memory-function in macrophages. Recent studies in mice have linked ten-eleven translocation enzymes and consequently DNA hydroxymethylation in immune cell differentiation and development (Lio and Rao, 2019; Lio et al., 2016; Tsagaratou et al., 2017) as well as in proinflammatory cytokine production (Liu et al., 2020). Here we have found that within a single generation of domestication several immune-related genes were both differentially hydroxymethylated and expressed. Higher levels of DNA hydroxymethylation were observed in the wild individuals compared to the first generation of wild-born Nile tilapias reared in captivity (Paper I). Since DNA hydroxymethylation within gene bodies is primarily associated with gene activation and increased expression (Tsagaratou et al., 2014), we assume that the wild individuals had a more active immune system in response to various pathogens such as viruses and parasites. The overall downregulation and hypo-hydroxymethylation of immune-related genes persisted in the second generation of Nile tilapia (Papers II and III) possibly due to the pathogen-free conditions of our recirculating aquaculture system.

3.4 Impact of DNA hydroxymethylation in adaptation and genome evolution

Throughout our experiments we have shown that DNA hydroxymethylation is not only a common epigenetic modification but it is also affected by the environment and likely functional (Papers I, II and III). From the early developmental stages environmental and endogenous factors shape the epigenetic landscape and several DNA hydroxymethylation marks persist through adulthood in mammals (Kochmanski et al., 2018; Massart et al., 2014). In that regard, aging is also a major factor linked to specific patterns and changes of the epigenome (Wilson et al., 1987) including the germline. For example, human sperm is characterized by hypermethylation at repetitive elements and hypomethylation of specific CpGs with age (Jenkins et al., 2014), while in aged mice, oocytes appear to have hypomethylated histone levels (Manosalva and Gonzalez, 2010; Shao et al., 2015) and genes that encode chromatin modifiers are greatly down-regulated (Hamatani et al., 2004). Taking into account these findings, it is rather likely that similar changes are present in fish. Indeed, a recent study revealed that the DNA methylation of the sperm in Atlantic salmon differed among wild and captive-reared individuals (Rodriguez-Barreto et al., 2019). Furthermore, these changes were linked to signatures of domestication, supporting the notion that environmentally-induced epigenetic changes are not random but function-specific. In addition to the above, it is also well known that both DNA methylation and hydroxymethylation have impact on nucleotide transversion rates, promoting mutations and single nucleotide polymorphisms. Anastasiadi and Piferrer (2019) showed that epimutations in wild sea bass (*Dicentrarchus labrax*) are positively correlated and overlap with cytosine-to-thymine (C to T) polymorphisms when compared to domesticated counterparts after 25 years of selective breeding. In both humans and mice, 5-hydroxymethylcytosines have a higher C to G transversion rate than 5-methylcytosines and unmodified cytosines (Supek et al., 2014). We cannot draw direct conclusions for all vertebrates including teleosts, since nucleotide transversion rates by epimutations appear to be species-specific. However, it is plausible that changes in 5hmC abundance could play major role in fish domestication, adaptation, and evolution through the transgenerational inheritance of epigenetic marks and epimutations through the germline. Throughout our

experiments, we identified 5hmCs within the Nile tilapia mitochondrial DNA (mtDNA), which contains 38 CCGG sites. Although 5hmC levels were significantly different among these 38 CCGG sites, they were surprisingly stable across all three experiments. Thus, 5hmC levels within mtDNA were generally stable compared to the 5hmC levels within the nuclear DNA. This stability is potentially a key factor for reduced evolutionary effects and nucleotide transversion rates within the mitochondria. Taking this notion further, modern breeding practices in aquaculture are based upon the phenotypic and genetic screening of the broodstock. However, considering the impact of epimutations and ageing upon the germline, epigenetic profiling might reveal transgenerational epigenetic effects that play a critical role during the early stages of development with long-term consequences on the offspring. In paper I, we identified genome-wide changes of the hydroxymethylome in fast muscle between wild and first-generation of wild-born fish reared under captivity. Such rapid changes potentially occur across all tissues including the ovaries and testis, thus having potential intergenerational effects. Because the environmental factors are largely stable under captivity (especially in the case of recirculating aquaculture systems as opposed to open systems such as lakes, rivers and the sea), the detection of initial differences between wild and first generation in captivity could potentially pinpoint not only towards inter- but also trans-generational effects through the stable and environmentally-induced accumulation of 5hmCs in functionally relevant locations across the genome.

To better understand the impact of domestication on the genome, epigenome and consequently the phenotype during the early stages of captivity, a combination of GWAS, DNA methylation and DNA hydroxymethylation profiles are required among families within and across generations. GWAS will provide sufficient data for the discovery of SNPs and additive genetic variance among and within groups while epigenetic profiling will provide evidence of transgenerational epigenetic inheritance. Additionally, possible overlaps between stable epigenetic modifications and single nucleotide polymorphisms in subsequent generations will verify the impact and importance of epigenetic modifications in animal evolution under captivity. Finally, using this holistic approach it would be possible to identify candidate genomic loci related to trait enhancement (i.e. quantitative trait loci) and stable epigenetic markers.

3.5 Epigenetic modifications as biomarkers (epimarkers)

Epigenetic processes are an inextricable part of the molecular machinery in both eukaryotes and prokaryotes. This thesis supports the notion that epimutations are likely the first molecular mechanism involved in response to environmental stimuli and regulate gene transcription. However, epigenetic markers are rarely taken into consideration for the prediction of a phenotype with the exception of cancers in the field of biomedicine. The lessons from cancer evolution through the mapping of epimutations brought to light the significance of epigenetic heterogeneity among cells and their contribution in discriminating genetically distinct cells without integrating genotypic and transcriptomic profiles (Gaiti et al., 2019). Although abnormal signals of epigenetic biomarkers can be used to detect, diagnose and predict the course of disease, little is known whether they can also be utilized for the prediction of specific phenotypes and traits that are useful in the aquaculture sector.

Recently, the epigenetic characterization of the sperm of low and high fertility Holstein bulls, revealed that DNA methylation differences can likely explain transcriptional reprogramming during the early embryonic stages with implications in successful delivery of transcription factors and reproductive efficiency (Kropp et al., 2017). Furthermore, in a study comparing the DNA methylome of wild boars and domesticated pig breeds, Li et al. (2016) showed that DNA methylation differences within the imprinting gene cluster H19-IGF2 could play an important role in muscle development, which is associated with pig domestication and breed divergence.

To better understand how epigenetic differences promote phenotypic divergence and potentially contribute towards the enhancement of beneficial traits such as growth, disease, and stress resistance, we have to take into account several factors that determine these markers. Animal domestication is greatly based on trait heritability; therefore, epigenetic biomarkers should be inherited from one generation to the next. However, Jiang et al. (2013) and (Potok et al., 2013) showed that DNA methylation reprogramming in early zebrafish development is entirely different compared to mammals. In particular, they showed that the paternal methylome was stably inherited to the embryo whereas the maternal methylome was remodelled by both passive demethylation and *de novo* methylation prior to zygotic

genome activation. In mammals, a complete reset of both the maternal and paternal methylome occurs right after fertilization through passive demethylation (Mayer et al., 2000; Oswald et al., 2000). Currently, it is not known whether other fish species share the same mechanism of transgenerational epigenetic inheritance with zebrafish. However, recent studies in salmon and trout have shown that epigenetic variation within germ cells can potential contribute to heritable changes that pass on to future generations (Gavery et al., 2018; Rodriguez-Barreto et al., 2019). Furthermore, the identification of epigenetic patterns that are positively associated with traits related to the domestication syndrome (i.e., reduced fitness and lower jaw malformations) are evident in subsequent generations (Anastasiadi and Piferrer, 2019; Le Luyer et al., 2017). These findings suggest that transgenerational epigenetic inheritance can largely contribute towards the enhancement of traits that are important for the further development of the aquaculture. In this thesis, we reported a significantly different DNA hydroxymethylome between wild and first generation of Nile tilapia in captivity (Paper I). Genes related to immunity and particularly the function of the phagosome were hyper-hydroxymethylated in the wild individuals, concomitantly with the overexpression of hundreds of immune genes. Whether these changes in DNA hydroxymethylation can be directed to enhance immunity and disease resistance is a question that remains unanswered.

It is also of high importance for selected epigenetic biomarkers to be stable throughout developmental or age-specific stages. Several studies in fish have shown that early embryonic stages are critical and determine traits such as muscle growth and immune resistance that persist through adulthood. Although the precise mechanisms by which DNA hydroxymethylation marks are retained across generations have not yet been identified, during our experiments we identified over 100,000 significantly hydroxymethylated cytosines, most of which were stable regardless of the experimental setting (Papers I, II and III). This is an indication that DNA hydroxymethylation is serving specific roles in normal development and function, thus carrying valuable molecular information that can be utilized in selective breeding.

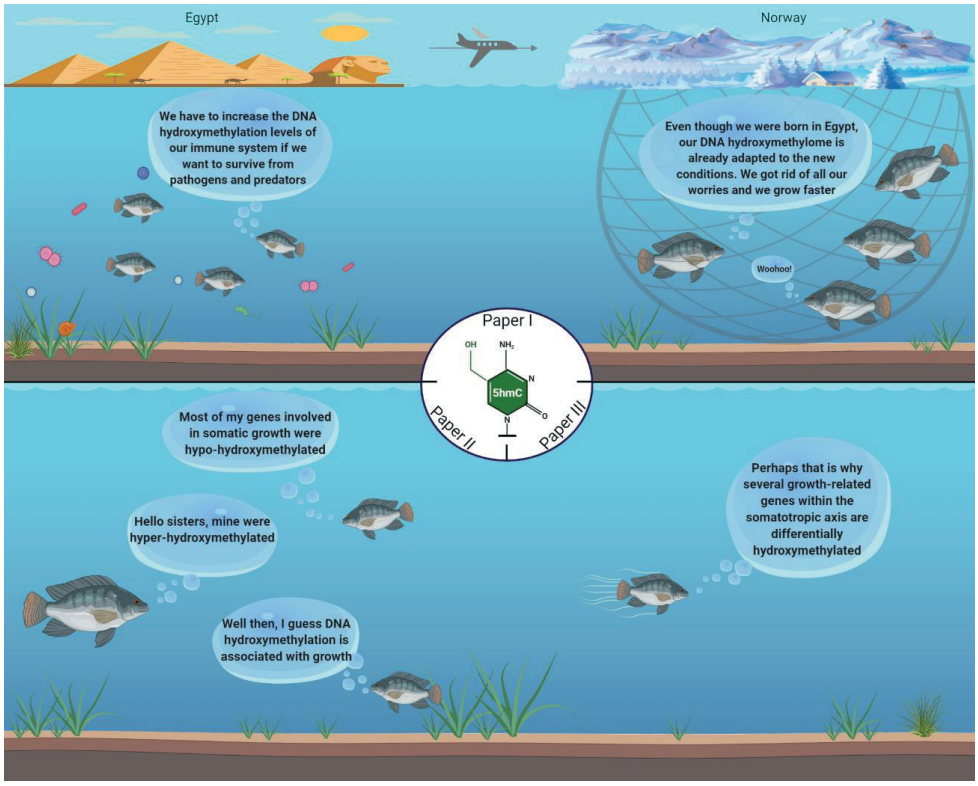


Figure 9. Summary of the main findings of the dissertation. Created with Biorender.com

4. Conclusions

This thesis provides for the first time a complete characterization of the DNA hydroxymethylome at single nucleotide resolution in a teleost species. We showed that DNA hydroxymethylation is present throughout the genome, it occurs primarily within gene bodies and particularly within introns, and it changes rapidly after a single generation of domestication (Paper I). Concomitantly, more than 2 thousand genes were identified as differentially expressed between the wild and the first generation of fish that were wild-born and reared in captivity. These genes were linked to immune responses, metabolism and muscle-specific functions. The first generation of fish undergoing domestication had overall decreased DNA hydroxymethylation levels within immune-related genes compared to their wild progenitors. This persisted in the second generation (Papers II and III), revealing an association between environmental conditions and epigenetic profiles.

Several hydroxymethylated genes involved in major growth-related pathways were identified. In particular, the comparison of liver hydroxymethylation profiles between large- and small-sized full sibs revealed more than 80 differentially hydroxymethylated genes involved in growth, skeletal system development, muscle structure and tissue development, striated muscle tissue development and skeletal system morphogenesis as well as the PI3K-Akt, Ras-MAPK and ERK1, ERK2 signaling pathways.

By profiling the hydroxymethylome of the muscle, liver and pituitary, we showed that DNA hydroxymethylation is particularly enriched within genes that are responsible for the function and regulation of the somatotropic axis. Through this series of experiments, we were able to demonstrate that DNA hydroxymethylation is associated with genes that are involved in the establishment of desirable traits for the aquaculture industry. Taking this innovative perspective further, this thesis will help strengthen the foundation of our knowledge within the field of epigenetics and breeding practices in aquaculture.

Overall, we determined that 5hmC is a promising epigenetic modification that could be involved in the regulation of valuable production traits, such as growth. Although, the associations between DNA hydroxymethylation and growth were evident across all three experiments, further studies are needed to ascertain conclusively if the observed gene

expression changes and phenotypic divergence are attributed to DNA hydromethylation. The findings of this thesis provide a baseline towards the incorporation of epigenetic profiles in the breeding nucleus of the aquaculture industry and set clear goals for further investigations.

5. Future Perspectives

Epigenetics is a timely research field that provides valuable insights towards the understanding of complex biological systems. To incorporate the findings of this thesis into the research needs of the aquaculture industry, experiments that confirm the biological significance of the herein identified epigenetic modifications are necessary. To validate these epigenetic markers *in vivo* and across generations, in line with fish domestication and trait enhancement, we could utilize genome editing tools and directly replace DNA bases (i.e., site-specific CpGs). By comparing controls to epigenetically edited full-siblings, we could test whether epigenetic modifications have impact on the phenotype. Currently, genome editing can be performed using the Clustered Regularly Interspaced Palindromic Repeats – Cas9 system (CRISPR-Cas9) which is an RNA-guided DNA nuclease system, and the prime editing technique which uses a catalytically impaired Cas9 endonuclease fused with a reverse transcriptase and contains a prime editing guide RNA that encodes the desired edit. Recently, a CRISPR-based approach which has some similarities with the prime editing technique, showed that a mutant form of Cas9 which has de-activated endonuclease (dCas9), can be fused with MS2 coat proteins that contain a catalytic domain for Tet1 tethering. Thus, the fused system can precisely detect a genetic sequence and demethylate without any breaks or mutations to the DNA sequence itself (Xu et al., 2016).

Interestingly, because 5hmCs appear to have higher C to G transversion rates compared to 5mCs and unmodified cytosines, it is likely that DNA hydroxymethylation is associated with genetic mutations during fish domestication and selective breeding. In addition, 5hmCs have been found to be both functionally relevant and primarily within gene bodies, thus mutations in functionally relevant regions could potentially drive genome evolution. To test this hypothesis, we will perform an association between DNA hydroxymethylation profiles

and genetic variations across all generations within our lab-scale Nile tilapia breeding program (W and F0-F3 generations).

The genotypic selection of domesticated animals and the optimization of breeding practices have revolutionized aquaculture and improved production of animal protein for an ever-increasing human population. However, epigenetic variation greatly affects the phenotype by regulating gene expression. Epigenetic research in the context of aquaculture, fish domestication and selective breeding is still in its infancy. Therefore, it becomes essential to unravel the potential of the epigenome and utilize it for the selection of superior individuals. The present work provides for the first-time associations between DNA hydroxymethylation and somatic growth, using as model organism one of the most cultivated fish species in the world. Our findings put forward the notion that the integration of epigenetic variation in the breeding nucleus might improve current selection programmes and establish fast-paced adaptation and trait enhancement for new domesticates.

6. References

Adam, A.C., Lie, K.K., Whatmore, P., Jakt, L.M., Moren, M. & Skjaerven, K.H. (2019). Profiling DNA methylation patterns of zebrafish liver associated with parental high dietary arachidonic acid. *PLoS One,* 14**:** e0220934. DOI 10.1371/journal.pone.0220934

Alami-Durante, H., Olive, N. & Rouel, M. (2007). Early thermal history significantly affects the seasonal hyperplastic process occurring in the myotomal white muscle of Dicentrarchus labrax juveniles. *Cell Tissue Res,* 327**:** 553-70. DOI 10.1007/s00441-006-0321-2

Almeida, R.D., Sottile, V., Loose, M., De Sousa, P.A., Johnson, A.D. & Ruzov, A. (2012). Semi-quantitative immunohistochemical detection of 5-hydroxymethyl-cytosine reveals conservation of its tissue distribution between amphibians and mammals. *Epigenetics,* 7**:** 137-40. DOI 10.4161/epi.7.2.18949

Anastasiadi, D. & Piferrer, F. (2019). Epimutations in developmental genes underlie the onset of domestication in farmed European sea bass. *Mol Biol Evol,* 36**:** 2252-2264. DOI 10.1093/molbev/msz153

Andreassen, R. & Hoyheim, B. (2017). miRNAs associated with immune response in teleost fish. *Dev Comp Immunol,* 75**:** 77-85. DOI 10.1016/j.dci.2017.02.023

Araki, H., Cooper, B. & Blouin, M.S. (2007). Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. *Science,* 318**:** 100-3. DOI 10.1126/science.1145621

Araki, H. & Schmid, C. (2010). Is hatchery stocking a help or harm? Evidence, limitations and future directions in ecological and genetic surveys. *Aquaculture,* 308**:** S2-S11. DOI https://doi.org/10.1016/j.aquaculture.2010.05.036

Aung, O., Nguyen, T.T.T., Poompuang, S. & Kamonrat, W. (2010). Microsatellite DNA markers revealed genetic population structure among captive stocks and wild populations of mrigal, *Cirrhinus cirrhosus* in Myanmar. *Aquaculture,* 299**:** 37-43. DOI https://doi.org/10.1016/j.aquaculture.2009.12.010

Autuoro, J.M., Pirnie, S.P. & Carmichael, G.G. (2014). Long noncoding RNAs in imprinting and X chromosome inactivation. *Biomolecules,* 4**:** 76-100. DOI 10.3390/biom4010076

Bachman, M., Uribe-Lewis, S., Yang, X., Williams, M., Murrell, A. & Balasubramanian, S. (2014). 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nat Chem,* 6**:** 1049-55. DOI 10.1038/nchem.2064

Balon, E.K. (1995). Origin and domestication of the wild carp, *Cyprinus carpio*: From roman gourmets to the swimming flowers. *Aquaculture,* 129**:** 3-48. DOI https://doi.org/10.1016/0044-8486(94)00227-F

Balon, E.K. (2004). About the oldest domesticates among fishes. *J FIsh Biol,* 65**:** 1-27. DOI https://doi.org/10.1111/j.0022-1112.2004.00563.x

Bannister, A.J. & Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res,* 21**:** 381-95. DOI 10.1038/cr.2011.22

Baroiller, J.F. & D'cotta, H. (2001). Environment and sex determination in farmed fish. *Comp Biochem Physiol C Toxicol Pharmacol,* 130**:** 399-409. DOI 10.1016/s1532-0456(01)00267-8

Bech-Otschir, D., Kraft, R., Huang, X., Henklein, P., Kapelari, B., Pollmann, C. & Dubiel, W. (2001). COP9 signalosome-specific phosphorylation targets p53 to degradation by the ubiquitin system. *EMBO J,* 20**:** 1630-9. DOI 10.1093/emboj/20.7.1630

Bell, J.D., Ganachaud, A., Gehrke, P.C., Griffiths, S.P., Hobday, A.J., Hoegh-Guldberg, O., Johnson, J.E., Borgne, R., Lehodey, P., Lough, J.M., Matear, R.J., Pickering, T.D., Pratchett, M.S., Gupta, A.S., Senina, I. & Waycott, M. (2013). Mixed responses of tropical Pacific fisheries and aquaculture to climate change. *Nat Clim Chnage,* 3**:** 591-599. DOI https://doi.org/10.1038/nclimate1838

Berejikian, B.A. (1995). The effects of hatchery and wild ancestry and experience on the relative ability of steelhead trour fry (*Oncorhynchus mykiss*) to avoid a benthic predator. *Can J Fish Aquat,* 52**:** 2476-2482. DOI 10.1139/f95-838

Bernstein, A.I., Lin, Y., Street, R.C., Lin, L., Dai, Q., Yu, L., Bao, H., Gearing, M., Lah, J.J., Nelson, P.T., He, C., Levey, A.I., Mulle, J.G., Duan, R. & Jin, P. (2016). 5-Hydroxymethylation-associated epigenetic modifiers of Alzheimer's disease modulate Tau-induced neurotoxicity. *Hum Mol Genet,* 25**:** 2437-2450. DOI 10.1093/hmg/ddw109

Bhattacharyya, S., Pradhan, K., Campbell, N., Mazdo, J., Vasantkumar, A., Maqbool, S., Bhagat, T.D., Gupta, S., Suzuki, M., Yu, Y., Greally, J.M., Steidl, U., Bradner, J., Dawlaty, M., Godley, L., Maitra, A. & Verma, A. (2017). Altered hydroxymethylation is seen at regulatory regions in pancreatic cancer and regulates oncogenic pathways. *Genome Res,* 27**:** 1830-1842. DOI 10.1101/gr.222794.117

Bicskei, B., Taggart, J.B., Glover, K.A. & Bron, J.E. (2016). Comparing the transcriptomes of embryos from domesticated and wild Atlantic salmon (Salmo salar L.) stocks and examining factors that influence heritability of gene expression. *Genet Sel Evol,* 48**:** 20. DOI 10.1186/s12711-016-0200-6

Bocquet-Appel, J.P. (2011). When the world's population took off: the springboard of the Neolithic Demographic Transition. *Science,* 333**:** 560-1. DOI 10.1126/science.1208880

Boltana, S., Aguilar, A., Sanhueza, N., Donoso, A., Mercado, L., Imarai, M. & Mackenzie, S. (2018). Behavioral Fever Drives Epigenetic Modulation of the Immune Response in Fish. *Front Immunol,* 9**:** 1241. DOI 10.3389/fimmu.2018.01241

Booth, M.J., Branco, M.R., Ficz, G., Oxley, D., Krueger, F., Reik, W. & Balasubramanian, S. (2012). Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science,* 336**:** 934-7. DOI 10.1126/science.1220671

Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B. & Bestor, T.H. (2001). Dnmt3L and the establishment of maternal genomic imprints. *Science,* 294**:** 2536-9. DOI 10.1126/science.1065848

Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A. & Hannon, G.J. (2008). An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science,* 322**:** 1387-92. DOI 10.1126/science.1165171

Brykczynska, U., Hisano, M., Erkek, S., Ramos, L., Oakeley, E.J., Roloff, T.C., Beisel, C., Schubeler, D., Stadler, M.B. & Peters, A.H. (2010). Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. *Nat Struct Mol Biol,* 17**:** 679-87. DOI 10.1038/nsmb.1821

Buhler, M. (2009). RNA turnover and chromatin-dependent gene silencing. *Chromosoma,* 118**:** 141-51. DOI 10.1007/s00412-008-0195-z

Bunting, K.L., Soong, T.D., Singh, R., Jiang, Y., Beguelin, W., Poloway, D.W., Swed, B.L., Hatzi, K., Reisacher, W., Teater, M., Elemento, O. & Melnick, A.M. (2016). Multi-tiered reorganization of the genome during B cell affinity maturation anchored by a germinal center-specific locus control region. *Immunity,* 45**:** 497-512. DOI 10.1016/j.immuni.2016.08.012

Burdge, G.C. & Lillycrop, K.A. (2014). Fatty acids and epigenetics. *Curr Opin Clin Nutr Metab Care,* 17**:** 156-61. DOI 10.1097/MCO.0000000000000023

Burgerhout, E., Mommens, M., Johnsen, H., Aunsmo, A., Santi, N. & Andersen, O. (2017). Genetic background and embryonic temperature affect DNA methylation and expression of myogenin and muscle development in Atlantic salmon (*Salmo salar*). *PLoS One,* 12**:** e0179918. DOI 10.1371/journal.pone.0179918

Burggren, W. (2016). Epigenetic inheritance and its role in evolutionary biology: Re-evaluation and new perspectives. *Biology (Basel),* 5. DOI 10.3390/biology5020024

Busslinger, M. & Tarakhovsky, A. (2014). Epigenetic control of immunity. *Cold Spring Harb Perspect Biol,* 6. DOI 10.1101/cshperspect.a019307

Buxton, P., Edwards, C., Archer, C.W. & Francis-West, P. (2001). Growth/differentiation factor-5 (GDF-5) and skeletal development. *J Bone Joint Surg Am,* 83-A Suppl 1**:** S23-30. DOI 10.2106/00004623-200100001-00004

Caceres, G., Lopez, M.E., Cadiz, M.I., Yoshida, G.M., Jedlicki, A., Palma-Vejares, R., Travisany, D., Diaz-Dominguez, D., Maass, A., Lhorente, J.P., Soto, J., Salas, D. & Yanez, J.M. (2019). Fine Mapping Using Whole-Genome Sequencing Confirms Anti-Mullerian Hormone as a Major Gene for Sex Determination in Farmed Nile Tilapia (*Oreochromis niloticus* L.). *G3 (Bethesda),* 9**:** 3213-3223. DOI 10.1534/g3.119.400297

Campos, C., Sundaram, A.Y., Valente, L.M., Conceicao, L.E., Engrola, S. & Fernandes, J.M. (2014). Thermal plasticity of the miRNA transcriptome during Senegalese sole development. *BMC Genomics,* 15**:** 525. DOI 10.1186/1471-2164-15-525

Campos, C., Valente, L.M., Conceicao, L.E., Engrola, S. & Fernandes, J.M. (2013a). Epigenetic regulation of muscle development and growth in senegalese sole larvae. *Commun Agric Appl Biol Sci,* 78**:** 60-1. DOI https://doi.org/10.1016/j.aquaculture.2014.04.035

Campos, C., Valente, L.M., Conceicao, L.E., Engrola, S. & Fernandes, J.M. (2013b). Temperature affects methylation of the myogenin putative promoter, its expression and muscle cellularity in Senegalese sole larvae. *Epigenetics,* 8**:** 389-97. DOI 10.4161/epi.24178

Campos, C., Valente, L.M. & Fernandes, J.M. (2012). Molecular evolution of zebrafish dnmt3 genes and thermal plasticity of their expression during embryonic development. *Gene,* 500**:** 93-100. DOI 10.1016/j.gene.2012.03.041

Carmell, M.A., Xuan, Z., Zhang, M.Q. & Hannon, G.J. (2002). The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev,* 16**:** 2733-42. DOI 10.1101/gad.1026102

Cauvin, J. (2000). *The birth of the Gods and the beginnings of agriculture,* Cambridge UK: Cambridge Univ. Press.

Chervinski, J. (1982). Environmental physiology of tilapias. IN Pullin, R.S.V. & Lowe-Mcconnell, R.H. (Eds.) *The Biology and Culture of Tilapias.* ICLARM Conference Proceedings No. 7: ICLARM.

Ching, T., Huang, S. & Garmire, L.X. (2014). Power analysis and sample size estimation for RNA-Seq differential expression. *RNA,* 20**:** 1684-96. DOI 10.1261/rna.046011.114

Christie, M.R., Marine, M.L., Fox, S.E., French, R.A. & Blouin, M.S. (2016). A single generation of domestication heritably alters the expression of hundreds of genes. *Nat Commun,* 7**:** 10676. DOI 10.1038/ncomms10676

Christie, M.R., Marine, M.L., French, R.A. & Blouin, M.S. (2012). Genetic adaptation to captivity can occur in a single generation. *Proc Natl Acad Sci U S A,* 109**:** 238-42. DOI 10.1073/pnas.1111073109

Clements, J.C. & Chopin, T. (2016). Ocean acidification and marine aquaculture in North America: potential impacts and mitigation strategies. *Rev Aquacult,* 9**:** 326-341. DOI 10.1111/raq.12140

Clutton-Brock, J. (1987). *A natural history of domesticated mammals,* Cambridge: Cambridge Univ. Press.

Clutton-Brock, J. (1995). Origins of the dog: The archaelogical evidence. IN Serpell, J. (Ed.) *The domestic dog, its evolution, behaviour and interactions with people.* Cambridge: Cambridge Univ. Press.

Collins, L.J., Schönfeld, B. & Chen, X.S. (2011). The epigenetics of non-coding RNA. IN Tollefsbol, T. (Ed.) *Handbook of Epigenetics: The New Molecular and Medical Genetics.* San Diego, CA: Academic Press.

Coluccio, A., Ecco, G., Duc, J., Offner, S., Turelli, P. & Trono, D. (2018). Individual retrotransposon integrants are differentially controlled by KZFP/KAP1-dependent histone methylation, DNA methylation and TET-mediated hydroxymethylation in naive embryonic stem cells. *Epigenetics Chromatin,* 11**:** 7. DOI 10.1186/s13072-018-0177-1

Conte, M.A., Gammerdinger, W.J., Bartie, K.L., Penman, D.J. & Kocher, T.D. (2017). A high quality assembly of the Nile Tilapia (*Oreochromis niloticus*) genome reveals the structure of two sex determination regions. *BMC Genomics,* 18**:** 341. DOI 10.1186/s12864-017-3723-5

Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D., Abramowitz, L.K., Bartolomei, M.S., Rambow, F., Bassi, M.R., Bruno, T., Fanciulli, M., Renner, C., Klein-Szanto, A.J., Matsumoto, Y., Kobi, D., Davidson, I., Alberti, C., Larue, L. & Bellacosa, A. (2011). Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell,* 146**:** 67-79. DOI 10.1016/j.cell.2011.06.020

Costa-Pierce, B.A. (1987). Aquaculture in ancient Hawaii. *BioScience,* 37**:** 320-331. DOI 10.2307/1310688

Cunningham, N.S., Jenkins, N.A., Gilbert, D.J., Copeland, N.G., Reddi, A.H. & Lee, S.J. (1995). Growth/differentiation factor-10: a new member of the transforming growth factor-beta superfamily related to bone morphogenetic protein-3. *Growth Factors,* 12**:** 99-109. DOI 10.3109/08977199509028956

Da Silva Maciel, E., Kimie Savay-Da Silva, L., Galvão, J.A. & Oetterer, M. (2014). Tilapia Production: From water to human consumption. IN Wakefield, R. (Ed.) *Tilapia: Biology, Management Practices and Human Consumption.* New York: Nova Science Publishers.

Darwin, C.R. (1875). *The variation of animals and plants under domestication*: London: John Murray.

Devlin, R.H. & Nagahama, Y. (2002). Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. *Aquaculture,* 208**:** 191-366. DOI https://doi.org/10.1016/S0044-8486(02)00057-1

Devlin, R.H., Sakhrani, D., Tymchuk, W.E., Rise, M.L. & Goh, B. (2009). Domestication and growth hormone transgenesis cause similar changes in gene expression in coho salmon (Oncorhynchus kisutch). *Proc Natl Acad Sci U S A,* 106**:** 3047-52. DOI 10.1073/pnas.0809798106

Dhanasiri, A.K.S., Chen, X., Dahle, D., Siriyappagouder, P., Faeste, C.K. & Fernandes, J.M.O. (2020). Dietary inclusion of plant ingredients induces epigenetic changes in the intestine of zebrafish. *Epigenetics,* 15**:** 1035-1051. DOI 10.1080/15592294.2020.1747777

Diamond, J. (1997). *Guns, Germs, and Steel: The Fates of Human Societies,* Norton, New York, .

Diamond, J. (2002). Evolution, consequences and future of plant and animal domestication. *Nature,* 418**:** 700-7. DOI 10.1038/nature01019

Driscoll, C.A., Macdonald, D.W. & O'brien, S.J. (2009). From wild animals to domestic pets, an evolutionary view of domestication. *Proc Natl Acad Sci U S A,* 106 Suppl 1**:** 9971-8. DOI 10.1073/pnas.0901586106

Embody, G.C. & Hyford, C.D. (1925). The advantages of rearing brook trout fingerlings from selected breeders. *T Am Fish Soc,* 55**:** 135-138. DOI https://doi.org/10.1577/1548-8659(1925)55[135:TAORBT]2.0.CO;2

Eshel, O., Shirak, A., Weller, J.I., Slossman, T., Hulata, G., Cnaani, A. & Ron, M. (2011). Fine-mapping of a locus on linkage group 23 for sex determination in Nile tilapia (Oreochromis niloticus). *Anim Genet,* 42**:** 222-4. DOI 10.1111/j.1365-2052.2010.02128.x

Fan, H., Zhang, R., Tesfaye, D., Tholen, E., Looft, C., Holker, M., Schellander, K. & Cinar, M.U. (2012). Sulforaphane causes a major epigenetic repression of myostatin in porcine satellite cells. *Epigenetics,* 7**:** 1379-90. DOI 10.4161/epi.22609

Fao (2020a). Cultured Aquatic Species Information Programme. *Oreochromis niloticus*. Text by Rakocy, J. E. Rome, Updated 18 February 2005: FAO Fisheries and Aquaculture Department.

Fao (2020b). *The state of world fisheries and aquaculture. Sustainability in action,* Rome, Italy: FAO.

Faulk, C. & Dolinoy, D.C. (2011). Timing is everything: the when and how of environmentally induced changes in the epigenome of animals. *Epigenetics,* 6**:** 791-7. DOI 10.4161/epi.6.7.16209

Feinberg, A.P., Irizarry, R.A., Fradin, D., Aryee, M.J., Murakami, P., Aspelund, T., Eiriksdottir, G., Harris, T.B., Launer, L., Gudnason, V. & Fallin, M.D. (2010). Personalized epigenomic signatures that are stable over time and covary with body mass index. *Sci Transl Med,* 2**:** 49ra67. DOI 10.1126/scitranslmed.3001262

Fernandes, J.M., Mackenzie, M.G., Elgar, G., Suzuki, Y., Watabe, S., Kinghorn, J.R. & Johnston, I.A. (2005). A genomic approach to reveal novel genes associated with myotube formation in the model teleost, Takifugu rubripes. *Physiol Genomics,* 22**:** 327-38. DOI 10.1152/physiolgenomics.00087.2005

Fernandes, J.M., Mackenzie, M.G., Kinghorn, J.R. & Johnston, I.A. (2007). FoxK1 splice variants show developmental stage-specific plasticity of expression with temperature in the tiger pufferfish. *J Exp Biol,* 210**:** 3461-72. DOI 10.1242/jeb.009183

Ficz, G., Branco, M.R., Seisenberger, S., Santos, F., Krueger, F., Hore, T.A., Marques, C.J., Andrews, S. & Reik, W. (2011). Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature,* 473**:** 398-402. DOI 10.1038/nature10008

Finstad, A.G. & Jonsson, B. (2012). Effect of incubation temperature on growth performance in Atlantic salmon. *Mar Ecol Prog Ser,* 454**:** 75-82. DOI 10.3354/MEPS09643

Fleming, I.A., Lamberg, A. & Jonsson, B. (1997). Effect of early experience on the reproductive perfomance of Atlantic salmon. *Behavioral Ecology,* 8**:** 470-480. DOI https://doi.org/10.1093/beheco/8.5.470

Franke, T.F. (2008). PI3K/Akt: getting it right matters. *Oncogene,* 27**:** 6473-88. DOI 10.1038/onc.2008.313

Frantz, L.A., Mullin, V.E., Pionnier-Capitan, M., Lebrasseur, O., Ollivier, M., Perri, A., Linderholm, A., Mattiangeli, V., Teasdale, M.D., Dimopoulos, E.A., Tresset, A., Duffraisse, M., Mccormick, F., Bartosiewicz, L., Gal, E., Nyerges, E.A., Sablin, M.V., Brehard, S., Mashkour, M., Balasescu, A., Gillet, B., Hughes, S., Chassaing, O., Hitte, C., Vigne, J.D., Dobney, K., Hanni, C., Bradley, D.G. & Larson, G. (2016). Genomic and archaeological evidence suggest a dual origin of domestic dogs. *Science,* 352**:** 1228-31. DOI 10.1126/science.aaf3161

Freedman, A.H., Gronau, I., Schweizer, R.M., Ortega-Del Vecchyo, D., Han, E., Silva, P.M., Galaverni, M., Fan, Z., Marx, P., Lorente-Galdos, B., Beale, H., Ramirez, O., Hormozdiari, F., Alkan, C., Vila, C., Squire, K., Geffen, E., Kusak, J., Boyko, A.R., Parker, H.G., Lee, C., Tadigotla, V., Wilton, A., Siepel, A., Bustamante, C.D., Harkins, T.T., Nelson, S.F., Ostrander, E.A., Marques-Bonet, T., Wayne, R.K. & Novembre, J. (2014). Genome sequencing highlights the dynamic early history of dogs. *PLoS Genet,* 10**:** e1004016. DOI 10.1371/journal.pgen.1004016

Freedman, A.H. & Wayne, R.K. (2017). Deciphering the Origin of Dogs: From Fossils to Genomes. *Annu Rev Anim Biosci,* 5**:** 281-307. DOI 10.1146/annurev-animal-022114-110937

Fuji, K., Kobayashi, K., Hasegawa, O., Coimbra, M.R.M., Sakamoto, T. & Okamoto, N. (2006). Identification of a single major genetic locus controlling the resistance to lymphocystis disease in Japanese flounder (*Paralichthys olivaceus*). *Aquaculture,* 254**:** 203-210. DOI 10.1016/j.aquaculture.2005.11.024

Gaiti, F., Chaligne, R., Gu, H., Brand, R.M., Kothen-Hill, S., Schulman, R.C., Grigorev, K., Risso, D., Kim, K.T., Pastore, A., Huang, K.Y., Alonso, A., Sheridan, C., Omans, N.D., Biederstedt, E., Clement, K., Wang, L., Felsenfeld, J.A., Bhavsar, E.B., Aryee, M.J., Allan, J.N., Furman, R., Gnirke, A., Wu, C.J., Meissner, A. & Landau, D.A. (2019). Epigenetic evolution and lineage histories of chronic lymphocytic leukaemia. *Nature,* 569**:** 576-580. DOI 10.1038/s41586-019-1198-z

Galindo-Villegas, J., Garcia-Moreno, D., De Oliveira, S., Meseguer, J. & Mulero, V. (2012). Regulation of immunity and disease resistance by commensal microbes and chromatin modifications during zebrafish development. *Proc Natl Acad Sci U S A,* 109**:** E2605-14. DOI 10.1073/pnas.1209920109

Galloway, T.F., Kjorsvik, E. & Kryvi, H. (1999). Muscle growth in yolk-sac larvae of the Atlantic halibut as influenced by temperature in the egg and yolk-sac stage. *J Fish Biol,* 55**:** 26-43. DOI https://doi.org/10.1111/j.1095-8649.1999.tb01044.x

Gao, D., Pinello, N., Nguyen, T.V., Thoeng, A., Nagarajah, R., Holst, J., Rasko, J.E. & Wong, J.J. (2019). DNA methylation/hydroxymethylation regulate gene expression and alternative splicing during terminal granulopoiesis. *Epigenomics,* 11**:** 95-109. DOI 10.2217/epi-2018-0050

Gavery, M.R., Nichols, K.M., Goetz, G.W., Middleton, M.A. & Swanson, P. (2018). Characterization of Genetic and Epigenetic Variation in Sperm and Red Blood Cells from Adult Hatchery and Natural-Origin Steelhead, *Oncorhynchus mykiss*. *G3 (Bethesda),* 8**:** 3723-3736. DOI 10.1534/g3.118.200458

Gjerdem, T. (2010). The first family-based breeding program in aquaculture. *Rev Aquacult,* 2**:** 2-15. DOI 10.1111/j.1753-5131.2010.01011.x

Gjerdem, T. (2012). Genetic improvement for the development of efficient global aquaculture: A personal opinion review. *Aquaculture,* 344-349**:** 12-22. DOI https://doi.org/10.1016/j.aquaculture.2012.03.003

Globisch, D., Munzel, M., Muller, M., Michalakis, S., Wagner, M., Koch, S., Bruckl, T., Biel, M. & Carell, T. (2010). Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS One,* 5**:** e15367. DOI 10.1371/journal.pone.0015367

Glover, K.A., Bergh, Y., Rudra, H. & Skaala, Y. (2006a). Juvenile growth and susceptibility to *Aeromonas salmonicida* subsp. *salmonicida* in Atlantic salmon (*Salmo salar* L.) of farmed, hybrid and wild parentage. *Aquaculture,* 254**:** 72-81. DOI https://doi.org/10.1016/j.aquaculture.2005.10.040

Glover, K.A., Skâr, C., Christie, K.E., Glette, J. & Rudra, H. (2006b). Size-dependent susceptibility to infectious salmon anemia virus (ISAV) in Atlantic salmon (*Salmo salar* L.) of farm, hybrid and wild parentage. *Aquaculture,* 254**:** 82-91. DOI https://doi.org/10.1016/j.aquaculture.2005.10.041

Golden, D.E., Gerbasi, V.R. & Sontheimer, E.J. (2008). An inside job for siRNAs. *Mol Cell,* 31**:** 309-12. DOI 10.1016/j.molcel.2008.07.008

Goll, M.G., Kirpekar, F., Maggert, K.A., Yoder, J.A., Hsieh, C.L., Zhang, X., Golic, K.G., Jacobsen, S.E. & Bestor, T.H. (2006). Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. *Science,* 311**:** 395-8. DOI 10.1126/science.1120976

Gontan, C., Jonkers, I. & Gribnau, J. (2011). Long noncoding RNAs and X chromosome inactivation. *Prog Mol Subcell Biol,* 51**:** 43-64. DOI 10.1007/978-3-642-16502-3_3

Goody, J. (1993). *The culture of flowers,* Cambridge: Cambridge Univ. Press.

Greco, C.M., Kunderfranco, P., Rubino, M., Larcher, V., Carullo, P., Anselmo, A., Kurz, K., Carell, T., Angius, A., Latronico, M.V., Papait, R. & Condorelli, G. (2016). DNA hydroxymethylation controls cardiomyocyte gene expression in development and hypertrophy. *Nat Commun,* 7**:** 12418. DOI 10.1038/ncomms12418

Grewal, S.I. & Moazed, D. (2003). Heterochromatin and epigenetic control of gene expression. *Science,* 301**:** 798-802. DOI 10.1126/science.1086887

Gros, J., Manceau, M., Thome, V. & Marcelle, C. (2005). A common somitic origin for embryonic muscle progenitors and satellite cells. *Nature,* 435**:** 954-8. DOI 10.1038/nature03572

Grosser, C., Wagner, N., Grothaus, K. & Horsthemke, B. (2015). Altering TET dioxygenase levels within physiological range affects DNA methylation dynamics of HEK293 cells. *Epigenetics,* 10**:** 819-33. DOI 10.1080/15592294.2015.1073879

Gutierrez, A.P., Yanez, J.M. & Davidson, W.S. (2016). Evidence of recent signatures of selection during domestication in an Atlantic salmon population. *Marine Genomics,* 26**:** 41-50. DOI 10.1016/j.margen.2015.12.007

Ha, H.P., Nguyen, T.T.T., Poompuang, S. & Na-Nakorn, U. (2009). Microsatellites revealed no genetic differentiation between hatchery and contemporary wild populations of striped catfish, *Pangasianodon hypophthalmus* (Sauvage 1878) in Vietnam. *Aquaculture,* 291**:** 154-160. DOI https://doi.org/10.1016/j.aquaculture.2009.03.017

Hahn, M.A., Szabo, P.E. & Pfeifer, G.P. (2014). 5-Hydroxymethylcytosine: a stable or transient DNA modification? *Genomics,* 104**:** 314-23. DOI 10.1016/j.ygeno.2014.08.015

Hamatani, T., Falco, G., Carter, M.G., Akutsu, H., Stagg, C.A., Sharov, A.A., Dudekula, D.B., Vanburen, V. & Ko, M.S. (2004). Age-associated alteration of gene expression patterns in mouse oocytes. *Hum Mol Genet,* 13**:** 2263-78. DOI 10.1093/hmg/ddh241

Hamilton, A.J. & Baulcombe, D.C. (1999). A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science,* 286**:** 950-2. DOI 10.1126/science.286.5441.950

Harache, Y. (2002). Development and diversification issues in aquaculture. A historical and dynamic view of fish culture diversification. *Seafood market studies for the introduction of new aquaculture products.* Zaragosa: Cah Options Mediterr.

Harvey, Z.H., Chen, Y. & Jarosz, D.F. (2018). Protein-based inheritance: Epigenetics beyond the chromosome. *Mol Cell,* 69**:** 195-202. DOI 10.1016/j.molcel.2017.10.030

Hata, K., Okano, M., Lei, H. & Li, E. (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development,* 129**:** 1983-93.

Hermann, A., Goyal, R. & Jeltsch, A. (2004). The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J Biol Chem,* 279**:** 48350-9. DOI 10.1074/jbc.M403427200

Hewitt, S.L., Chaumeil, J. & Skok, J.A. (2010). Chromosome dynamics and the regulation of V(D)J recombination. *Immunol Rev,* 237**:** 43-54. DOI 10.1111/j.1600-065X.2010.00931.x

Higginbotham, J.A. (1997). *Piscinae: Artificial fishponds in Roman Italy,* Chapel Hill and London: University of North Carolina Press.

Hino, J., Kangawa, K., Matsuo, H., Nohno, T. & Nishimatsu, S. (2004). Bone morphogenetic protein-3 family members and their biological functions. *Front Biosci,* 9**:** 1520-9. DOI 10.2741/1355

Holliday, R. (2006). Epigenetics: a historical overview. *Epigenetics,* 1**:** 76-80. DOI 10.4161/epi.1.2.2762

Holliday, R. & Pugh, J.E. (1975). DNA modification mechanisms and gene activity during development. *Science,* 187**:** 226-32. DOI 10.1126/science.187.4173.226

Hotchkiss, R.D. (1948). The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. *J Biol Chem,* 175**:** 315-32.

Houde, A.L.S., Fraser, D.J. & Hutchings, J.A. (2010). Reduced anti-predator responses in multi-generational hybrids of farmed and wild Atlantic salmon (*Salmo salar* L.). *Conserv Genet,* 11**:** 785-794. DOI 10.1007/s10592-009-9892-2

Houston, R.D., Taggart, J.B., Cezard, T., Bekaert, M., Lowe, N.R., Downing, A., Talbot, R., Bishop, S.C., Archibald, A.L., Bron, J.E., Penman, D.J., Davassi, A., Brew, F., Tinch, A.E., Gharbi, K. & Hamilton, A. (2014). Development and validation of a high density SNP genotyping array for Atlantic salmon (*Salmo salar*). *BMC Genomics,* 15**:** 90. DOI 10.1186/1471-2164-15-90

Hu, G., Cui, K., Fang, D., Hirose, S., Wang, X., Wangsa, D., Jin, W., Ried, T., Liu, P., Zhu, J., Rothenberg, E.V. & Zhao, K. (2018). Transformation of accessible chromatin and 3D nucleome underlies lineage commitment of early T cells. *Immunity,* 48**:** 227-242 e8. DOI 10.1016/j.immuni.2018.01.013

Hu, L., Lu, J., Cheng, J., Rao, Q., Li, Z., Hou, H., Lou, Z., Zhang, L., Li, W., Gong, W., Liu, M., Sun, C., Yin, X., Li, J., Tan, X., Wang, P., Wang, Y., Fang, D., Cui, Q., Yang, P., He, C., Jiang, H., Luo, C. & Xu, Y. (2015). Structural insight into substrate preference for TET-mediated oxidation. *Nature,* 527**:** 118-22. DOI 10.1038/nature15713

Huang, Y., Chavez, L., Chang, X., Wang, X., Pastor, W.A., Kang, J., Zepeda-Martinez, J.A., Pape, U.J., Jacobsen, S.E., Peters, B. & Rao, A. (2014). Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. *Proc Natl Acad Sci U S A,* 111**:** 1361-6. DOI 10.1073/pnas.1322921111

Huntingford, F. & Adams, C. (2005). Behavioural syndromes in farmed fish: implications for production and welfare. *Behaviour,* 142**:** 1213-1227. DOI 10.1163/156853905774539382

Iizuka, M. & Smith, M.M. (2003). Functional consequences of histone modifications. *Curr Opin Genet Dev,* 13**:** 154-60. DOI 10.1016/s0959-437x(03)00020-0

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C. & Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science,* 333**:** 1300-3. DOI 10.1126/science.1210597

Jablonka, E. (2013). Epigenetic inheritance and plasticity: The responsive germline. *Prog Biophys Mol Biol,* 111**:** 99-107. DOI 10.1016/j.pbiomolbio.2012.08.014

Jenkins, T.G., Aston, K.I., Pflueger, C., Cairns, B.R. & Carrell, D.T. (2014). Age-associated sperm DNA methylation alterations: possible implications in offspring disease susceptibility. *PLoS Genet,* 10**:** e1004458. DOI 10.1371/journal.pgen.1004458

Jenuwein, T. & Allis, C.D. (2001). Translating the histone code. *Science,* 293**:** 1074-80. DOI 10.1126/science.1063127

Jia, J., Miao, W., Cai, J. & Yuan, X. (2018). Contribution of Chinese aquaculture to the sector, globally, and to overall food security. IN Gui, J.F., Tang, Q., Li, Z., Liu, J. & De Silva, S.S. (Eds.) *Aquaculture in China: Success stories and modern trends.* London, Eng.: Wiley-Blackwell.

Jiang, L., Zhang, J., Wang, J.J., Wang, L., Zhang, L., Li, G., Yang, X., Ma, X., Sun, X., Cai, J., Zhang, J., Huang, X., Yu, M., Wang, X., Liu, F., Wu, C.I., He, C., Zhang, B., Ci, W. & Liu, J. (2013). Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. *Cell,* 153**:** 773-84. DOI 10.1016/j.cell.2013.04.041

Jin, S.G., Wu, X., Li, A.X. & Pfeifer, G.P. (2011). Genomic mapping of 5-hydroxymethylcytosine in the human brain. *Nucleic Acids Res,* 39**:** 5015-24. DOI 10.1093/nar/gkr120

Johnston, I.A., Bower, N.I. & Macqueen, D.J. (2011). Growth and the regulation of myotomal muscle mass in teleost fish. *J Exp Biol,* 214**:** 1617-28. DOI 10.1242/jeb.038620

Johnston, I.A., Lee, H.T., Macqueen, D.J., Paranthaman, K., Kawashima, C., Anwar, A., Kinghorn, J.R. & Dalmay, T. (2009). Embryonic temperature affects muscle fibre recruitment in adult zebrafish: genome-wide changes in gene and microRNA expression associated with the transition from hyperplastic to hypertrophic growth phenotypes. *J Exp Biol,* 212**:** 1781-93. DOI 10.1242/jeb.029918

Johnston, I.A., Mclay, H.A., Abercromby, M. & Robins, D. (2000). Early thermal experience has different effects on growth and muscle fibre recruitment in spring- and autumn-running Atlantic salmon populations. *J Exp Biol,* 203**:** 2553-64.

Kagiwada, S., Kurimoto, K., Hirota, T., Yamaji, M. & Saitou, M. (2013). Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. *EMBO J,* 32**:** 340-53. DOI 10.1038/emboj.2012.331

Kaminen-Ahola, N.J., Ahola, A.I. & Whitelaw, E. (2011). Epigenetic inheritance: Both mitotic and meiotic. IN Niculescu, M.D. & Haggarty, P. (Eds.) *Nutrition in Epigenetics.* John Wiley & Sons.

Kamstra, J.H., Loken, M., Alestrom, P. & Legler, J. (2015). Dynamics of DNA hydroxymethylation in zebrafish. *Zebrafish,* 12**:** 230-7. DOI 10.1089/zeb.2014.1033

Kanjilal, D. & Cottrell, J.A. (2019). Bone morphogenetic proteins (BMPs) and bone regeneration. *Methods Mol Biol,* 1891**:** 235-245. DOI 10.1007/978-1-4939-8904-1_17

Kaufman, P.D. & Rando, O.J. (2010). Chromatin as a potential carrier of heritable information. *Curr Opin Cell Biol,* 22**:** 284-90. DOI 10.1016/j.ceb.2010.02.002

Kieffer-Kwon, K.R., Nimura, K., Rao, S.S.P., Xu, J., Jung, S., Pekowska, A., Dose, M., Stevens, E., Mathe, E., Dong, P., Huang, S.C., Ricci, M.A., Baranello, L., Zheng, Y., Tomassoni Ardori, F., Resch, W., Stavreva, D., Nelson, S., Mcandrew, M., Casellas, A., Finn, E., Gregory, C., St Hilaire, B.G., Johnson, S.M., Dubois, W., Cosma, M.P., Batchelor, E., Levens, D., Phair, R.D., Misteli, T., Tessarollo, L., Hager, G., Lakadamyali, M., Liu, Z., Floer, M., Shroff, H., Aiden, E.L. & Casellas, R. (2017). Myc regulates chromatin decompaction and nuclear architecture during B cell activation. *Mol Cell,* 67**:** 566-578 e10. DOI 10.1016/j.molcel.2017.07.013

Kim, J., Withler, R., Ritland, C. & Cheng, K. (2004). Genetic variation within and between domesticated Chinook salmon, *Oncorhynchus tshawytscha* strains and their progenitor populations. IN Gharrett, R., Gustafson, J., Nielsen, J., Seeb, L., Seeb, W., Smoker, W.W., Thorgaard, G.H. & Wilmot, R.L. (Eds.) *Genetics of Subpolar Fish and Invertebrates.* Dordrecht: Springer.

Klironomos, F.D., Berg, J. & Collins, S. (2013). How epigenetic mutations can affect genetic evolution: model and mechanism. *Bioessays,* 35**:** 571-8. DOI 10.1002/bies.201200169

Ko, M., An, J., Pastor, W.A., Koralov, S.B., Rajewsky, K. & Rao, A. (2015). TET proteins and 5-methylcytosine oxidation in hematological cancers. *Immunol Rev,* 263**:** 6-21. DOI 10.1111/imr.12239

Kobayashi, H., Sakurai, T., Miura, F., Imai, M., Mochiduki, K., Yanagisawa, E., Sakashita, A., Wakai, T., Suzuki, Y., Ito, T., Matsui, Y. & Kono, T. (2013). High-resolution DNA methylome analysis of primordial germ cells identifies gender-specific reprogramming in mice. *Genome Res,* 23**:** 616-27. DOI 10.1101/gr.148023.112

Kochmanski, J.J., Marchlewicz, E.H., Cavalcante, R.G., Perera, B.P.U., Sartor, M.A. & Dolinoy, D.C. (2018). Longitudinal effects of developmental bisphenol A exposure on epigenome-wide DNA hydroxymethylation at imprinted loci in mouse blood. *Environ Health Perspect,* 126**:** 077006. DOI 10.1289/EHP3441

Kollberg, G., Tulinius, M., Melberg, A., Darin, N., Andersen, O., Holmgren, D., Oldfors, A. & Holme, E. (2009). Clinical manifestation and a new ISCU mutation in iron-sulphur cluster deficiency myopathy. *Brain,* 132**:** 2170-9. DOI 10.1093/brain/awp152

Konstantinidis, I., Saetrom, P., Mjelle, R., Nedoluzhko, A.V., Robledo, D. & Fernandes, J.M.O. (2020). Major gene expression changes and epigenetic remodelling in Nile tilapia muscle after just one generation of domestication. *Epigenetics,* 15**:** 1052-1067. DOI 10.1080/15592294.2020.1748914

Kriaucionis, S. & Heintz, N. (2009). The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science,* 324**:** 929-30. DOI 10.1126/science.1169786

Kropp, J., Carrillo, J.A., Namous, H., Daniels, A., Salih, S.M., Song, J. & Khatib, H. (2017). Male fertility status is associated with DNA methylation signatures in sperm and transcriptomic profiles of bovine preimplantation embryos. *BMC Genomics,* 18**:** 280. DOI 10.1186/s12864-017-3673-y

Kulis, M., Merkel, A., Heath, S., Queiros, A.C., Schuyler, R.P., Castellano, G., Beekman, R., Raineri, E., Esteve, A., Clot, G., Verdaguer-Dot, N., Duran-Ferrer, M., Russinol, N., Vilarrasa-Blasi, R., Ecker, S., Pancaldi, V., Rico, D., Agueda, L., Blanc, J., Richardson, D., Clarke, L., Datta, A., Pascual, M., Agirre, X., Prosper, F., Alignani, D., Paiva, B., Caron, G., Fest, T., Muench, M.O., Fomin, M.E., Lee, S.T., Wiemels, J.L., Valencia, A., Gut, M., Flicek, P., Stunnenberg, H.G., Siebert, R., Kuppers, R., Gut, I.G., Campo, E. & Martin-Subero, J.I. (2015). Whole-genome fingerprint of the DNA methylome during human B cell differentiation. *Nat Genet,* 47**:** 746-56. DOI 10.1038/ng.3291

Lardenoije, R., Roubroeks, J.a.Y., Pishva, E., Leber, M., Wagner, H., Iatrou, A., Smith, A.R., Smith, R.G., Eijssen, L.M.T., Kleineidam, L., Kawalia, A., Hoffmann, P., Luck, T., Riedel-Heller, S., Jessen, F., Maier, W., Wagner, M., Hurlemann, R., Kenis, G., Ali, M., Del Sol, A., Mastroeni, D., Delvaux, E., Coleman, P.D., Mill, J., Rutten, B.P.F., Lunnon, K., Ramirez, A. & Van Den Hove, D.L.A. (2019). Alzheimer's disease-associated (hydroxy)methylomic changes in the brain and blood. *Clin Epigenetics,* 11**:** 164. DOI 10.1186/s13148-019-0755-5

Larson, G. & Fuller, D.Q. (2014). The evolution of animal domestication. *Annu Rev Ecol Evol Syst,* 45**:** 115-136. DOI 10.1146/annurev-ecolsys-110512-135813

Larson, G., Karlsson, E.K., Perri, A., Webster, M.T., Ho, S.Y., Peters, J., Stahl, P.W., Piper, P.J., Lingaas, F., Fredholm, M., Comstock, K.E., Modiano, J.F., Schelling, C., Agoulnik, A.I., Leegwater, P.A., Dobney, K., Vigne, J.D., Vila, C., Andersson, L. & Lindblad-Toh, K. (2012). Rethinking dog domestication by integrating genetics, archeology, and biogeography. *Proc Natl Acad Sci U S A,* 109**:** 8878-83. DOI 10.1073/pnas.1203005109

Larson, G., Piperno, D.R., Allaby, R.G., Purugganan, M.D., Andersson, L., Arroyo-Kalin, M., Barton, L., Climer Vigueira, C., Denham, T., Dobney, K., Doust, A.N., Gepts, P., Gilbert, M.T., Gremillion, K.J., Lucas, L., Lukens, L., Marshall, F.B., Olsen, K.M., Pires, J.C., Richerson, P.J., Rubio De Casas, R., Sanjur, O.I., Thomas, M.G. & Fuller, D.Q. (2014). Current perspectives and the future of domestication studies. *Proc Natl Acad Sci U S A,* 111**:** 6139-46. DOI 10.1073/pnas.1323964111

Lawrence, M., Daujat, S. & Schneider, R. (2016). Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet,* 32**:** 42-56. DOI 10.1016/j.tig.2015.10.007

Le Luyer, J., Laporte, M., Beacham, T.D., Kaukinen, K.H., Withler, R.E., Leong, J.S., Rondeau, E.B., Koop, B.F. & Bernatchez, L. (2017). Parallel epigenetic modifications induced by hatchery rearing in a Pacific salmon. *Proc Natl Acad Sci U S A,* 114**:** 12964-12969. DOI 10.1073/pnas.1711229114

Le, T., Kim, K.P., Fan, G. & Faull, K.F. (2011). A sensitive mass spectrometry method for simultaneous quantification of DNA methylation and hydroxymethylation levels in biological samples. *Anal Biochem,* 412**:** 203-9. DOI 10.1016/j.ab.2011.01.026

Lee, R.C., Feinbaum, R.L. & Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell,* 75**:** 843-54. DOI 10.1016/0092-8674(93)90529-y

Li, C., Wang, X., Cai, H., Fu, Y., Luan, Y., Wang, W., Xiang, H. & Li, C. (2016). Molecular microevolution and epigenetic patterns of the long non-coding gene H19 show its potential function in pig domestication and breed divergence. *BMC Evol Biol,* 16**:** 87. DOI 10.1186/s12862-016-0657-5

Li, M., Sun, Y., Zhao, J., Shi, H., Zeng, S., Ye, K., Jiang, D., Zhou, L., Sun, L., Tao, W., Nagahama, Y., Kocher, T.D. & Wang, D. (2015). A Tandem Duplicate of Anti-Mullerian Hormone with a Missense SNP on the Y Chromosome Is Essential for Male Sex Determination in Nile Tilapia, *Oreochromis niloticus*. *PLoS Genet,* 11**:** e1005678. DOI 10.1371/journal.pgen.1005678

Lim, L.P., Lau, N.C., Garrett-Engele, P., Grimson, A., Schelter, J.M., Castle, J., Bartel, D.P., Linsley, P.S. & Johnson, J.M. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature,* 433**:** 769-73. DOI 10.1038/nature03315

Lima, E.C., Povh, J.A., Otonel, R.A., Leonhardt, J.H., Alfieri, A.A., Headley, S.A., Souza, F.P., Poveda-Parra, A.R., Furlan-Murari, P.J. & Lopera-Barrero, N.M. (2017). Morphology and muscle gene expression in GIFT and Supreme Nile tilapia varieties reared in two cultivation systems. *Genet Mol Res,* 16. DOI 10.4238/gmr16019407

Lin, H. & Spradling, A.C. (1997). A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. *Development,* 124**:** 2463-76.

Lio, C.J. & Rao, A. (2019). TET enzymes and 5hmC in adaptive and innate immune systems. *Front Immunol,* 10**:** 210. DOI 10.3389/fimmu.2019.00210

Lio, C.W., Zhang, J., Gonzalez-Avalos, E., Hogan, P.G., Chang, X. & Rao, A. (2016). Tet2 and Tet3 cooperate with B-lineage transcription factors to regulate DNA modification and chromatin accessibility. *Elife,* 5. DOI 10.7554/eLife.18290

Liu, X., Wang, X., Liu, N., Zhu, K., Zhang, S., Duan, X., Huang, Y., Jin, Z., Jaypaul, H., Wu, Y. & Chen, H. (2020). TET2 is involved in DNA hydroxymethylation, cell proliferation and inflammatory response in keratinocytes. *Mol Med Rep,* 21**:** 1941-1949. DOI 10.3892/mmr.2020.10989

Lopez, M.E., Neira, R. & Yanez, J.M. (2014). Applications in the search for genomic selection signatures in fish. *Front Genet,* 5**:** 458. DOI 10.3389/fgene.2014.00458

Lorenzen, K. (2000). Population dynamics and management. IN Beveridge, M.C.M. & Mcandrew, B.J. (Eds.) *Tilapia: Biology and Exploitation.* The Netherlands: Kluwer Academic Publishers.

Lorenzen, K., Beveridge, M.C. & Mangel, M. (2012). Cultured fish: integrative biology and management of domestication and interactions with wild fish. *Biological reviews of the Cambridge Philosophical Society,* 87**:** 639-60. DOI 10.1111/j.1469-185X.2011.00215.x

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. & Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature,* 389**:** 251-60. DOI 10.1038/38444

Lyman, J.M., Good, S.A., Gouretski, V.V., Ishii, M., Johnson, G.C., Palmer, M.D., Smith, D.M. & Willis, J.K. (2010). Robust warming of the global upper ocean. *Nature,* 465**:** 334-7. DOI 10.1038/nature09043

Makinen, H., Vasemagi, A., Mcginnity, P., Cross, T.F. & Primmer, C.R. (2015). Population genomic analyses of early-phase Atlantic Salmon (*Salmo salar*) domestication/captive breeding. *Evol Appl,* 8**:** 93-107. DOI 10.1111/eva.12230

Malone, C.D. & Hannon, G.J. (2009). Small RNAs as guardians of the genome. *Cell,* 136**:** 656-68. DOI 10.1016/j.cell.2009.01.045

Manosalva, I. & Gonzalez, A. (2010). Aging changes the chromatin configuration and histone methylation of mouse oocytes at germinal vesicle stage. *Theriogenology,* 74**:** 1539-47. DOI 10.1016/j.theriogenology.2010.06.024

Martin, C. & Zhang, Y. (2005). The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol,* 6**:** 838-49. DOI 10.1038/nrm1761

Massart, R., Suderman, M., Provencal, N., Yi, C., Bennett, A.J., Suomi, S. & Szyf, M. (2014). Hydroxymethylation and DNA methylation profiles in the prefrontal cortex of the non-human primate rhesus macaque and the impact of maternal deprivation on hydroxymethylation. *Neuroscience,* 268**:** 139-48. DOI 10.1016/j.neuroscience.2014.03.021

Matsuda, S. & Kitagishi, Y. (2013). MAGI scaffolding molecules involved in cancer cell signaling. *J Carcinog Mutagen,* S7-005. DOI 10.4172/2157-2518.S7-005

Mayer, W., Niveleau, A., Walter, J., Fundele, R. & Haaf, T. (2000). Demethylation of the zygotic paternal genome. *Nature,* 403**:** 501-2. DOI 10.1038/35000656

Mcginnity, P., Prodohl, P., Ferguson, A., Hynes, R., Maoileidigh, N.O., Baker, N., Cotter, D., O'hea, B., Cooke, D., Rogan, G., Taggart, J. & Cross, T. (2003). Fitness reduction and potential extinction of wild populations of Atlantic salmon, Salmo salar, as a result of interactions with escaped farm salmon. *Proc Biol Sci,* 270**:** 2443-50. DOI 10.1098/rspb.2003.2520

Mellen, M., Ayata, P., Dewell, S., Kriaucionis, S. & Heintz, N. (2012). MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell,* 151**:** 1417-30. DOI 10.1016/j.cell.2012.11.022

Mersfelder, E.L. & Parthun, M.R. (2006). The tale beyond the tail: histone core domain modifications and the regulation of chromatin structure. *Nucleic Acids Res,* 34**:** 2653-62. DOI 10.1093/nar/gkl338

Messina, G., Biressi, S., Monteverde, S., Magli, A., Cassano, M., Perani, L., Roncaglia, E., Tagliafico, E., Starnes, L., Campbell, C.E., Grossi, M., Goldhamer, D.J., Gronostajski, R.M. & Cossu, G. (2010). Nfix regulates fetal-specific transcription in developing skeletal muscle. *Cell,* 140**:** 554-66. DOI 10.1016/j.cell.2010.01.027

Middleton, A. & Uprichard, K. (Eds.) (2008). *The Nebamun wall paintings: conservation, scientific analysis and display at the British Museum,* London: Archetype Publications.

Moen, T., Baranski, M., Sonesson, A.K. & Kjoglum, S. (2009). Confirmation and fine-mapping of a major QTL for resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar*): population-level associations between markers and trait. *BMC Genomics,* 10**:** 368. DOI 10.1186/1471-2164-10-368

Moriya, N. & Miyazaki, M. (2018). Akt1 deficiency diminishes skeletal muscle hypertrophy by reducing satellite cell proliferation. *Am J Physiol Regul Integr Comp Physiol,* 314**:** R741-R751. DOI 10.1152/ajpregu.00336.2017

Muir, A.M., Massoudi, D., Nguyen, N., Keene, D.R., Lee, S.J., Birk, D.E., Davidson, J.M., Marinkovich, M.P. & Greenspan, D.S. (2016). BMP1-like proteinases are essential to the structure and wound healing of skin. *Matrix Biol,* 56**:** 114-131. DOI 10.1016/j.matbio.2016.06.004

Mumbach, M.R., Satpathy, A.T., Boyle, E.A., Dai, C., Gowen, B.G., Cho, S.W., Nguyen, M.L., Rubin, A.J., Granja, J.M., Kazane, K.R., Wei, Y., Nguyen, T., Greenside, P.G., Corces, M.R., Tycko, J., Simeonov, D.R., Suliman, N., Li, R., Xu, J., Flynn, R.A., Kundaje, A., Khavari, P.A., Marson, A., Corn, J.E., Quertermous, T., Greenleaf, W.J. & Chang, H.Y. (2017). Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. *Nat Genet,* 49**:** 1602-1612. DOI 10.1038/ng.3963

Munzel, M., Globisch, D., Bruckl, T., Wagner, M., Welzmiller, V., Michalakis, S., Muller, M., Biel, M. & Carell, T. (2010). Quantification of the sixth DNA base hydroxymethylcytosine in the brain. *Angew Chem Int Ed Engl,* 49**:** 5375-7. DOI 10.1002/anie.201002033

Nascimento, A.F., Trindade, D.M., Tonoli, C.C., De Giuseppe, P.O., Assis, L.H., Honorato, R.V., De Oliveira, P.S., Mahajan, P., Burgess-Brown, N.A., Von Delft, F., Larson, R.E. & Murakami, M.T. (2013). Structural insights into functional overlapping and differentiation among myosin V motors. *J Biol Chem,* 288**:** 34131-45. DOI 10.1074/jbc.M113.507202

Nash, C. (2000). Aquatic Animals. IN Kiple, K. & Ornelas, K. (Eds.) *The Cambridge World History of Food.* Cambridge: Cambridge Univ. Press.

Neira, R., Diaz, N.F., Gall, G.a.E., Gallardo, J.A., Lhorente, J.P. & Manterola, R. (2006). Genetic improvement in Coho salmon (*Oncorhynchus kisutch*). I: selection response and inbreeding depression on harvest weight. *Aquaculture,* 257**:** 9-17. DOI https://doi.org/10.1016/j.aquaculture.2006.03.002

Nestor, C.E., Ottaviano, R., Reddington, J., Sproul, D., Reinhardt, D., Dunican, D., Katz, E., Dixon, J.M., Harrison, D.J. & Meehan, R.R. (2012). Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Res,* 22**:** 467-77. DOI 10.1101/gr.126417.111

Netea, M.G. & Van Der Meer, J.W. (2017). Trained immunity: An ancient way of remembering. *Cell Host Microbe,* 21**:** 297-300. DOI 10.1016/j.chom.2017.02.003

Niculescu, M.D. & Lupu, D.S. (2011). Nutritional influence on epigenetics and effects on longevity. *Curr Opin Clin Nutr Metab Care,* 14**:** 35-40. DOI 10.1097/MCO.0b013e328340ff7c

Nomura, T. & Katunuma, N. (2005). Involvement of cathepsins in the invasion, metastasis and proliferation of cancer cells. *J Med Investig,* 52**:** 1-9. DOI 10.2152/jmi.52.1

Okano, M., Xie, S. & Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat Genet,* 19**:** 219-20. DOI 10.1038/890

Olsen, S.J. & Olsen, J.W. (1977). The chinese wolf, ancestor of new world dogs. *Science,* 197**:** 533-5. DOI 10.1126/science.197.4303.533

Orlanski, S., Labi, V., Reizel, Y., Spiro, A., Lichtenstein, M., Levin-Klein, R., Koralov, S.B., Skversky, Y., Rajewsky, K., Cedar, H. & Bergman, Y. (2016). Tissue-specific DNA demethylation is required

for proper B-cell differentiation and function. *Proc Natl Acad Sci U S A,* 113**:** 5018-23. DOI 10.1073/pnas.1604365113

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W. & Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. *Curr Biol,* 10**:** 475-8. DOI 10.1016/s0960-9822(00)00448-6

Pan, Z., Zhang, M., Ma, T., Xue, Z.Y., Li, G.F., Hao, L.Y., Zhu, L.J., Li, Y.Q., Ding, H.L. & Cao, J.L. (2016). Hydroxymethylation of microRNA-365-3p regulates nociceptive behaviors via Kcnh2. *J Neurosci,* 36**:** 2769-81. DOI 10.1523/JNEUROSCI.3474-15.2016

Park, J.H., Cosgrove, M.S., Youngman, E., Wolberger, C. & Boeke, J.D. (2002). A core nucleosome surface crucial for transcriptional silencing. *Nat Genet,* 32**:** 273-9. DOI 10.1038/ng982

Pastor, W.A., Pape, U.J., Huang, Y., Henderson, H.R., Lister, R., Ko, M., Mcloughlin, E.M., Brudno, Y., Mahapatra, S., Kapranov, P., Tahiliani, M., Daley, G.Q., Liu, X.S., Ecker, J.R., Milos, P.M., Agarwal, S. & Rao, A. (2011). Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature,* 473**:** 394-7. DOI 10.1038/nature10102

Perez, M.F. & Lehner, B. (2019). Intergenerational and transgenerational epigenetic inheritance in animals. *Nat Cell Biol,* 21**:** 143-151. DOI 10.1038/s41556-018-0242-9

Petterson, A., Chung, T.H., Tan, D., Sun, X. & Jia, X.Y. (2014). RRHP: a tag-based approach for 5-hydroxymethylcytosine mapping at single-site resolution. *Genome Biol,* 15**:** 456. DOI 10.1186/s13059-014-0456-5

Phanstiel, D.H., Van Bortle, K., Spacek, D., Hess, G.T., Shamim, M.S., Machol, I., Love, M.I., Aiden, E.L., Bassik, M.C. & Snyder, M.P. (2017). Static and dynamic DNA loops form AP-1-bound activation hubs during macrophage development. *Mol Cell,* 67**:** 1037-1048 e6. DOI 10.1016/j.molcel.2017.08.006

Pistocchi, A., Gaudenzi, G., Foglia, E., Monteverde, S., Moreno-Fortuny, A., Pianca, A., Cossu, G., Cotelli, F. & Messina, G. (2013). Conserved and divergent functions of Nfix in skeletal muscle development during vertebrate evolution. *Development,* 140**:** 1528-36. DOI 10.1242/dev.076315

Ponnaluri, V.K., Ehrlich, K.C., Zhang, G., Lacey, M., Johnston, D., Pradhan, S. & Ehrlich, M. (2017). Association of 5-hydroxymethylation and 5-methylation of DNA cytosine with tissue-specific gene expression. *Epigenetics,* 12**:** 123-138. DOI 10.1080/15592294.2016.1265713

Pons, D., Trompet, S., De Craen, A.J., Thijssen, P.E., Quax, P.H., De Vries, M.R., Wierda, R.J., Van Den Elsen, P.J., Monraats, P.S., Ewing, M.M., Heijmans, B.T., Slagboom, P.E., Zwinderman, A.H., Doevendans, P.A., Tio, R.A., De Winter, R.J., De Maat, M.P., Iakoubova, O.A., Sattar, N., Shepherd, J., Westendorp, R.G., Jukema, J.W., Group, P.S., Group, W.S. & Group, G.S. (2011). Genetic variation in PCAF, a key mediator in epigenetics, is associated with reduced vascular morbidity and mortality: evidence for a new concept from three independent prospective studies. *Heart,* 97**:** 143-50. DOI 10.1136/hrt.2010.199927

Ponting, C.P., Oliver, P.L. & Reik, W. (2009). Evolution and functions of long noncoding RNAs. *Cell,* 136**:** 629-41. DOI 10.1016/j.cell.2009.02.006

Potok, M.E., Nix, D.A., Parnell, T.J. & Cairns, B.R. (2013). Reprogramming the maternal zebrafish genome after fertilization to match the paternal methylation pattern. *Cell,* 153**:** 759-72. DOI 10.1016/j.cell.2013.04.030

Pullin, R.S.V. & Lowe-Mcconnell, R.H. (1982). *The biology and culture of tilapias,* Manila, Philippines: International Centre for Living Aquatic Resource Management.

Rea, M., Gripshover, T. & Fondufe-Mittendorf, Y. (2018). Selective inhibition of CTCF binding by iAs directs TET-mediated reprogramming of 5-hydroxymethylation patterns in iAs-transformed cells. *Toxicol Appl Pharmacol,* 338**:** 124-133. DOI 10.1016/j.taap.2017.11.015

Reddi, A.H. & Reddi, A. (2009). Bone morphogenetic proteins (BMPs): from morphogens to metabologens. *Cytokine Growth Factor Rev,* 20**:** 341-2. DOI 10.1016/j.cytogfr.2009.10.015

Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R. & Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature,* 403**:** 901-6. DOI 10.1038/35002607

Rescan, P.Y. (2001). Regulation and functions of myogenic regulatory factors in lower vertebrates. *Comp Biochem Physiol B Biochem Mol Biol,* 130**:** 1-12. DOI 10.1016/s1096-4959(01)00412-2

Roberge, C., Einum, S., Guderley, H. & Bernatchez, L. (2006). Rapid parallel evolutionary changes of gene transcription profiles in farmed Atlantic salmon. *Molecular Ecology,* 15**:** 9-20. DOI 10.1111/j.1365-294X.2005.02807.x

Rodriguez-Barreto, D., Garcia De Leaniz, C., Verspoor, E., Sobolewska, H., Coulson, M. & Consuegra, S. (2019). DNA methylation changes in the sperm of captive-reared fish: A route to epigenetic introgression in wild populations. *Mol Biol Evol,* 36**:** 2205-2211. DOI 10.1093/molbev/msz135

Rosa, R., Marques, A. & Nunes, M.L. (2012). Impact of climate change in Mediterranean aquaculture. *Rev Aquacult,* 4**:** 163-177. DOI https://doi.org/10.1111/j.1753-5131.2012.01071.x

Roser, M., Ritchie, H. & Ortiz-Ospina, E. (2013). World population growth. OurWorldInData.org.

Rossi, G. & Messina, G. (2014). Comparative myogenesis in teleosts and mammals. *Cell Mol Life Sci,* 71**:** 3081-99. DOI 10.1007/s00018-014-1604-5

Ruzov, A., Tsenkina, Y., Serio, A., Dudnakova, T., Fletcher, J., Bai, Y., Chebotareva, T., Pells, S., Hannoun, Z., Sullivan, G., Chandran, S., Hay, D.C., Bradley, M., Wilmut, I. & De Sousa, P. (2011). Lineage-specific distribution of high levels of genomic 5-hydroxymethylcytosine in mammalian development. *Cell Res,* 21**:** 1332-42. DOI 10.1038/cr.2011.113

Rye, M., Gjerde, B. & Gjerdem, T. (2010). Genetic improvement programs for aquaculture species in developed countries. *Proceedings of the 9th World Congress on Genetics Applied to Livestock Production.* Leipzig, Germany.

Saeed, S., Quintin, J., Kerstens, H.H., Rao, N.A., Aghajanirefah, A., Matarese, F., Cheng, S.C., Ratter, J., Berentsen, K., Van Der Ent, M.A., Sharifi, N., Janssen-Megens, E.M., Ter Huurne, M., Mandoli, A., Van Schaik, T., Ng, A., Burden, F., Downes, K., Frontini, M., Kumar, V., Giamarellos-Bourboulis, E.J., Ouwehand, W.H., Van Der Meer, J.W., Joosten, L.A., Wijmenga, C., Martens, J.H., Xavier, R.J., Logie, C., Netea, M.G. & Stunnenberg, H.G. (2014). Epigenetic programming of monocyte-to-macrophage differentiation and trained innate immunity. *Science,* 345**:** 1251086. DOI 10.1126/science.1251086

Santisteban, M.S., Arents, G., Moudrianakis, E.N. & Smith, M.M. (1997). Histone octamer function in vivo: mutations in the dimer-tetramer interfaces disrupt both gene activation and repression. *EMBO J,* 16**:** 2493-506. DOI 10.1093/emboj/16.9.2493

Sargent, J.R., Bell, M.V., Bell, J.G., Henderson, R.J. & Tocher, D.R. (1995). Evolution and roles of (n-3) polyunsaturated fatty acids in marine organisms. IN Cevic, G. & Paltauf, F. (Eds.) *Phospholipids: Characterisation, Metabolism and Novel Biological Applications.* Champaign, IL, USA: American Oil Chemistry Society Press.

Sauvage, C., Derome, N., Normandeau, E., St-Cyr, J., Audet, C. & Bernatchez, L. (2010). Fast transcriptional responses to domestication in the brook charr *Salvelinus fontinalis*. *Genetics,* 185**:** 105-12. DOI 10.1534/genetics.110.115071

Schmitz, R.J., Schultz, M.D., Lewsey, M.G., O'malley, R.C., Urich, M.A., Libiger, O., Schork, N.J. & Ecker, J.R. (2011). Transgenerational epigenetic instability is a source of novel methylation variants. *Science,* 334**:** 369-73. DOI 10.1126/science.1212959

Schnurr, M.E., Yin, Y. & Scott, G.R. (2014). Temperature during embryonic development has persistent effects on metabolic enzymes in the muscle of zebrafish. *J Exp Biol,* 217**:** 1370-80. DOI 10.1242/jeb.094037

Scott, G.R. & Johnston, I.A. (2012). Temperature during embryonic development has persistent effects on thermal acclimation capacity in zebrafish. *Proc Natl Acad Sci U S A,* 109**:** 14247-52. DOI 10.1073/pnas.1205012109

Shang, X., Su, J., Wan, Q. & Su, J. (2015). CpA/CpG methylation of CiMDA5 possesses tight association with the resistance against GCRV and negatively regulates mRNA expression in grass carp, *Ctenopharyngodon idella*. *Dev Comp Immunol,* 48**:** 86-94. DOI 10.1016/j.dci.2014.09.007

Shang, X., Wan, Q., Su, J. & Su, J. (2016). DNA methylation of CiRIG-I gene notably relates to the resistance against GCRV and negatively-regulates mRNA expression in grass carp, *Ctenopharyngodon idella*. *Immunobiology,* 221**:** 23-30. DOI 10.1016/j.imbio.2015.08.006

Shang, X., Yang, C., Wan, Q., Rao, Y. & Su, J. (2017). The destiny of the resistance/susceptibility against GCRV is controlled by epigenetic mechanisms in CIK cells. *Sci Rep,* 7**:** 4551. DOI 10.1038/s41598-017-03990-5

Shao, G.B., Wang, J., Zhang, L.P., Wu, C.Y., Jin, J., Sang, J.R., Lu, H.Y., Gong, A.H., Du, F.Y. & Peng, W.X. (2015). Aging alters histone H3 lysine 4 methylation in mouse germinal vesicle stage oocytes. *Reprod Fertil Dev,* 27**:** 419-26. DOI 10.1071/RD13293

Skjaerven, K.H., Jakt, L.M., Fernandes, J.M.O., Dahl, J.A., Adam, A.C., Klughammer, J., Bock, C. & Espe, M. (2018). Parental micronutrient deficiency distorts liver DNA methylation and expression of lipid genes associated with a fatty-liver-like phenotype in offspring. *Sci Rep,* 8**:** 3055. DOI 10.1038/s41598-018-21211-5

Smith, Z.D., Chan, M.M., Mikkelsen, T.S., Gu, H., Gnirke, A., Regev, A. & Meissner, A. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature,* 484**:** 339-44. DOI 10.1038/nature10960

Solberg, M.F., Kvamme, B.O., Nilsen, F. & Glover, K.A. (2012). Effects of environmental stress on mRNA expression levels of seven genes related to oxidative stress and growth in Atlantic salmon *Salmo salar* L. of farmed, hybrid and wild origin. *BMC Res Notes,* 5**:** 672. DOI 10.1186/1756-0500-5-672

Solberg, M.F., Skaala, O., Nilsen, F. & Glover, K.A. (2013). Does domestication cause changes in growth reaction norms? A study of farmed, wild and hybrid Atlantic salmon families exposed to environmental stress. *PLoS One,* 8**:** e54469. DOI 10.1371/journal.pone.0054469

Song, C.X., Szulwach, K.E., Fu, Y., Dai, Q., Yi, C., Li, X., Li, Y., Chen, C.H., Zhang, W., Jian, X., Wang, J., Zhang, L., Looney, T.J., Zhang, B., Godley, L.A., Hicks, L.M., Lahn, B.T., Jin, P. & He, C. (2011). Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat Biotechnol,* 29**:** 68-72. DOI 10.1038/nbt.1732

Stanhope-Baker, P., Hudson, K.M., Shaffer, A.L., Constantinescu, A. & Schlissel, M.S. (1996). Cell type-specific chromatin structure determines the targeting of V(D)J recombinase activity in vitro. *Cell,* 85**:** 887-97. DOI 10.1016/s0092-8674(00)81272-6

Steinbacher, P., Haslett, J.R., Obermayer, A., Marschallinger, J., Bauer, H.C., Sanger, A.M. & Stoiber, W. (2007). MyoD and Myogenin expression during myogenic phases in brown trout: a precocious onset of mosaic hyperplasia is a prerequisite for fast somatic growth. *Dev Dyn,* 236**:** 1106-14. DOI 10.1002/dvdy.21103

Stringwell, R., Lock, A., Stutchbury, C.J., Baggett, E., Taylor, J., Gough, P.J. & Garcia De Leaniz, C. (2014). Maladaptation and phenotypic mismatch in hatchery-reared Atlantic salmon Salmo salar released in the wild. *J Fish Biol,* 85**:** 1927-45. DOI 10.1111/jfb.12543

Sun, L.X., Wang, Y.Y., Zhao, Y., Wang, H., Li, N. & Ji, X.S. (2016a). Global DNA methylation changes in Nile tilapia gonads during high temperature-induced masculinization. *PLoS One,* 11**:** e0158483. DOI 10.1371/journal.pone.0158483

Sun, W., Guan, M. & Li, X. (2014). 5-hydroxymethylcytosine-mediated DNA demethylation in stem cells and development. *Stem Cells Dev,* 23**:** 923-30. DOI 10.1089/scd.2013.0428

Sun, X., Chung, T.H., Tan, D. & Kim, A. (2016b). Practical guidelines and consideration of using RRHP for 5hmC detection. *Epigenomics,* 8**:** 225-35. DOI 10.2217/epi.15.105

Supek, F., Lehner, B., Hajkova, P. & Warnecke, T. (2014). Hydroxymethylated cytosines are associated with elevated C to G transversion rates. *PLoS Genet,* 10**:** e1004585. DOI 10.1371/journal.pgen.1004585

Szulwach, K.E., Li, X., Li, Y., Song, C.X., Wu, H., Dai, Q., Irier, H., Upadhyay, A.K., Gearing, M., Levey, A.I., Vasanthakumar, A., Godley, L.A., Chang, Q., Cheng, X., He, C. & Jin, P. (2011). 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nat Neurosci,* 14**:** 1607-16. DOI 10.1038/nn.2959

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L. & Rao, A. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science,* 324**:** 930-5. DOI 10.1126/science.1170116

Talbert, P.B. & Henikoff, S. (2014). Environmental responses mediated by histone variants. *Trends Cell Biol,* 24**:** 642-50. DOI 10.1016/j.tcb.2014.07.006

Tan, L., Xiong, L., Xu, W., Wu, F., Huang, N., Xu, Y., Kong, L., Zheng, L., Schwartz, L., Shi, Y. & Shi, Y.G. (2013). Genome-wide comparison of DNA hydroxymethylation in mouse embryonic stem cells and neural progenitor cells by a new comparative hMeDIP-seq method. *Nucleic Acids Res,* 41**:** e84. DOI 10.1093/nar/gkt091

Titcomb, M. (1952). *Native use of fish in Hawaii,* Wellington, New Zealand: Polynesian Society Press.

Tobi, E.W., Heijmans, B.T., Kremer, D., Putter, H., Delemarre-Van De Waal, H.A., Finken, M.J., Wit, J.M. & Slagboom, P.E. (2011). DNA methylation of IGF2, GNASAS, INSIGF and LEP and being born small for gestational age. *Epigenetics,* 6**:** 171-6. DOI 10.4161/epi.6.2.13516

Trask, J.A., Malhi, R.S., Kanthaswamy, S., Johnson, J., Garnica, W.T., Malladi, V.S. & Smith, D.G. (2011). The effect of SNP discovery method and sample size on estimation of population genetic data for Chinese and Indian rhesus macaques (Macaca mulatta). *Primates,* 52**:** 129-38. DOI 10.1007/s10329-010-0232-4

Trewavas, E. (1983). *Tilapiine fishes of the genera Sarotherodon, Oreochromis and Danakilia,* London: British Museum (Natural History).

Tsagaratou, A., Aijo, T., Lio, C.W., Yue, X., Huang, Y., Jacobsen, S.E., Lahdesmaki, H. & Rao, A. (2014). Dissecting the dynamic changes of 5-hydroxymethylcytosine in T-cell development and differentiation. *Proc Natl Acad Sci U S A,* 111**:** E3306-15. DOI 10.1073/pnas.1412327111

Tsagaratou, A., Lio, C.J., Yue, X. & Rao, A. (2017). TET methylcytosine oxidases in T cell and B cell development and function. *Front Immunol,* 8**:** 220. DOI 10.3389/fimmu.2017.00220

Tsuchihara, K., Suzuki, Y., Wakaguri, H., Irie, T., Tanimoto, K., Hashimoto, S., Matsushima, K., Mizushima-Sugano, J., Yamashita, R., Nakai, K., Bentley, D., Esumi, H. & Sugano, S. (2009). Massive transcriptional start site analysis of human genes in hypoxia cells. *Nucleic Acids Res,* 37**:** 2249-63. DOI 10.1093/nar/gkp066

Turner, G.T. & Robinson, R.L. (2000). Reproductive biology, mating systems and parental care. IN Beveridge, M.C.M. & Mcandrew, B.J. (Eds.) *Tilapias: Biology and Exploitation.* The Netherlands: Kluwer Academin Publishers.

Valente, L.M.P., Moutou, K.A., Conceição, L.E.C., Engrola, S., Fernandes, J.M.O. & Johnston, I.A. (2013). What determines growth potential and juvenile quality of farmed fish species? *Rev Aquacult,* 5**:** S168-S193. DOI https://doi.org/10.1111/raq.12020

Van Der Graaf, A., Wardenaar, R., Neumann, D.A., Taudt, A., Shaw, R.G., Jansen, R.C., Schmitz, R.J., Colome-Tatche, M. & Johannes, F. (2015). Rate, spectrum, and evolutionary dynamics of spontaneous epimutations. *Proc Natl Acad Sci U S A,* 112**:** 6676-81. DOI 10.1073/pnas.1424254112

Vasemägi, A., Nilsson, J., Mcginnity, P., Cross, T., O'reilly, P., Glebe, B., Peng, B., Berg, P.R. & Primmer, C.R. (2012). Screen for footprints of selection during domestication/captive breeding of Atlantic salmon. *Comp Funct Genom,* 2012**:** 628204. DOI 10.1155/2012/628204

Verdel, A., Vavasseur, A., Le Gorrec, M. & Touat-Todeschini, L. (2009). Common themes in siRNA-mediated epigenetic silencing pathways. *Int J Dev Biol,* 53**:** 245-57. DOI 10.1387/ijdb.082691av

Vigne, J.D. (2011). The origins of animal domestication and husbandry: a major change in the history of humanity and the biosphere. *C R Biol,* 334**:** 171-81. DOI 10.1016/j.crvi.2010.12.009

Von Der Ecken, J., Heissler, S.M., Pathan-Chhatbar, S., Manstein, D.J. & Raunser, S. (2016). Cryo-EM structure of a human cytoplasmic actomyosin complex at near-atomic resolution. *Nature,* 534**:** 724-8. DOI 10.1038/nature18295

Waddington, C.H. (1942). Canalization of development and the inheritance of acquired characters. *Nature,* 150**:** 563-565. DOI https://doi.org/10.1038/150563a0

Wagner, D.O., Sieber, C., Bhushan, R., Borgermann, J.H., Graf, D. & Knaus, P. (2010). BMPs: from bone to body morphogenetic proteins. *Sci Signal,* 3**:** mr1. DOI 10.1126/scisignal.3107mr1

Wakefield, R. (2014). *Tilapia: Biology, Management Practices and Human Consumption,* New York: Nova Science Publishers.

Walsh, K. & Perlman, H. (1997). Cell cycle exit upon myogenic differentiation. *Curr Opin Genet Dev,* 7**:** 597-602. DOI 10.1016/s0959-437x(97)80005-6

Wang, C.M., Lo, L.C., Zhu, Z.Y. & Yue, G.H. (2006). A genome scan for quantitative trait loci affecting growth-related traits in an F1 family of Asian seabass (*Lates calcarifer*). *BMC Genomics,* 7**:** 274. DOI 10.1186/1471-2164-7-274

Wang, J., Yuan, Y., Zhou, Y., Guo, L., Zhang, L., Kuai, X., Deng, B., Pan, Z., Li, D. & He, F. (2008). Protein interaction data set highlighted with human Ras-MAPK/PI3K signaling pathways. *J Proteome Res,* 7**:** 3879-89. DOI 10.1021/pr8001645

Weatherley, P.M., Clark, H.S. & Lobo, A.F. (1988). Recruitment and maximal diameter of axial muscle fibres in teleosts and their relationship to somatic growth and ultimate size. *Fish Biol,* 33**:** 851-859. DOI https://doi.org/10.1111/j.1095-8649.1988.tb05532.x

Wei, J.W., Huang, K., Yang, C. & Kang, C.S. (2017). Non-coding RNAs as regulators in epigenetics (Review). *Oncol Rep,* 37**:** 3-9. DOI 10.3892/or.2016.5236

Wen, L. & Tang, F. (2014). Genomic distribution and possible functions of DNA hydroxymethylation in the brain. *Genomics,* 104**:** 341-6. DOI 10.1016/j.ygeno.2014.08.020

Wiedmeier, J.E., Joss-Moore, L.A., Lane, R.H. & Neu, J. (2011). Early postnatal nutrition and programming of the preterm neonate. *Nutr Rev,* 69**:** 76-82. DOI 10.1111/j.1753-4887.2010.00370.x

Wilder, G.P. (1923). *The taro, or kalo, in Hawaii,* Honolulu: University of Hawaii Extension Service.

Williams, K., Christensen, J., Pedersen, M.T., Johansen, J.V., Cloos, P.A., Rappsilber, J. & Helin, K. (2011). TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature,* 473**:** 343-8. DOI 10.1038/nature10066

Wilson, V.L., Smith, R.A., Ma, S. & Cutler, R.G. (1987). Genomic 5-methyldeoxycytidine decreases with age. *J Biol Chem,* 262**:** 9948-51.

Wittkopp, P.J., Haerum, B.K. & Clark, A.G. (2004). Evolutionary changes in cis and trans gene regulation. *Nature,* 430**:** 85-8. DOI 10.1038/nature02698

Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C.J., Zakhartchenko, V., Boiani, M., Arand, J., Nakano, T., Reik, W. & Walter, J. (2011). 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat Commun,* 2**:** 241. DOI 10.1038/ncomms1240

Wringe, B.F., Devlin, R.H., Ferguson, M.M., Moghadam, H.K., Sakhrani, D. & Danzmann, R.G. (2010). Growth-related quantitative trait loci in domestic and wild rainbow trout (*Oncorhynchus mykiss*). *BMC Genet,* 11**:** 63. DOI 10.1186/1471-2156-11-63

Wu, H., D'alessio, A.C., Ito, S., Wang, Z., Cui, K., Zhao, K., Sun, Y.E. & Zhang, Y. (2011a). Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev,* 25**:** 679-84. DOI 10.1101/gad.2036011

Wu, S.F., Zhang, H. & Cairns, B.R. (2011b). Genes for embryo development are packaged in blocks of multivalent chromatin in zebrafish sperm. *Genome Res,* 21**:** 578-89. DOI 10.1101/gr.113167.110

Wu, Y., Chen, X., Zhao, Y., Wang, Y., Li, Y. & Xiang, C. (2019). Genome-wide DNA methylation and hydroxymethylation analysis reveal human menstrual blood-derived stem cells inhibit

hepatocellular carcinoma growth through oncogenic pathway suppression via regulating 5-hmC in enhancer elements. *Stem Cell Res Ther,* 10**:** 151. DOI 10.1186/s13287-019-1243-8

Xia, H., Nho, R.S., Kahm, J., Kleidon, J. & Henke, C.A. (2004). Focal adhesion kinase is upstream of phosphatidylinositol 3-kinase/Akt in regulating fibroblast survival in response to contraction of type I collagen matrices via a beta 1 integrin viability signaling pathway. *J Biol Chem,* 279**:** 33024-34. DOI 10.1074/jbc.M313265200

Xie, D., Gore, C., Zhou, J., Pong, R.C., Zhang, H., Yu, L., Vessella, R.L., Min, W. & Hsieh, J.T. (2009). DAB2IP coordinates both PI3K-Akt and ASK1 pathways for cell survival and apoptosis. *Proc Natl Acad Sci U S A,* 106**:** 19878-83. DOI 10.1073/pnas.0908458106

Xu, X., Tao, Y., Gao, X., Zhang, L., Li, X., Zou, W., Ruan, K., Wang, F., Xu, G.L. & Hu, R. (2016). A CRISPR-based approach for targeted DNA demethylation. *Cell Discov,* 2**:** 16009. DOI 10.1038/celldisc.2016.9

Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A.J., Zheng, L., Zhang, H., Huang, S., Min, J., Nicholson, T., Chen, T., Xu, G., Shi, Y., Zhang, K. & Shi, Y.G. (2011). Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol Cell,* 42**:** 451-64. DOI 10.1016/j.molcel.2011.04.005

Yan, B., Zhu, C.D., Guo, J.T., Zhao, L.H. & Zhao, J.L. (2013). miR-206 regulates the growth of the teleost tilapia (*Oreochromis niloticus*) through the modulation of IGF-1 gene expression. *J Exp Biol,* 216**:** 1265-9. DOI 10.1242/jeb.079590

Yan, Y., Guo, C., Ni, S., Wei, J., Li, P., Wei, S., Cui, H. & Qin, Q. (2015). Singapore grouper iridovirus (SGIV) encoded SGIV-miR-13 attenuates viral infection via modulating major capsid protein expression. *Virus Res,* 205**:** 45-53. DOI 10.1016/j.virusres.2015.05.010

Yancopoulos, G.D. & Alt, F.W. (1985). Developmentally controlled and tissue-specific expression of unrearranged VH gene segments. *Cell,* 40**:** 271-81. DOI 10.1016/0092-8674(85)90141-2

Yeo, G., Holste, D., Kreiman, G. & Burge, C.B. (2004). Variation in alternative splicing across human tissues. *Genome Biol,* 5**:** R74. DOI 10.1186/gb-2004-5-10-r74

Yildirim, O., Li, R., Hung, J.H., Chen, P.B., Dong, X., Ee, L.S., Weng, Z., Rando, O.J. & Fazzio, T.G. (2011). Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell,* 147**:** 1498-510. DOI 10.1016/j.cell.2011.11.054

Yu, M., Hon, G.C., Szulwach, K.E., Song, C.X., Zhang, L., Kim, A., Li, X., Dai, Q., Shen, Y., Park, B., Min, J.H., Jin, P., Ren, B. & He, C. (2012). Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell,* 149**:** 1368-80. DOI 10.1016/j.cell.2012.04.027

Yun, Y.R., Won, J.E., Jeon, E., Lee, S., Kang, W., Jo, H., Jang, J.H., Shin, U.S. & Kim, H.W. (2010). Fibroblast growth factors: biology, function, and application for tissue regeneration. *J Tissue Eng,* 2010**:** 218142. DOI 10.4061/2010/218142

Zaratiegui, M., Irvine, D.V. & Martienssen, R.A. (2007). Noncoding RNAs and gene silencing. *Cell,* 128**:** 763-76. DOI 10.1016/j.cell.2007.02.016

Zenger, K.R., Khatkar, M.S., Jones, D.B., Khalilisamani, N., Jerry, D.R. & Raadsma, H.W. (2018). Genomic selection in aquaculture: Application, limitations and opportunities with special reference to marine shrimp and pearl oysters. *Front Genet,* 9**:** 693. DOI 10.3389/fgene.2018.00693

Zhang, R. & Su, B. (2009). Small but influential: the role of microRNAs on gene regulatory network and 3'UTR evolution. *J Genet Genomics,* 36**:** 1-6. DOI 10.1016/S1673-8527(09)60001-1

Zhang, S., Liu, H., Liu, Y., Zhang, J., Li, H., Liu, W., Cao, G., Xv, P., Zhang, J., Lv, C. & Song, X. (2017). miR-30a as potential therapeutics by targeting TET1 through regulation of Drp-1 promoter hydroxymethylation in idiopathic pulmonary fibrosis. *Int J Mol Sci,* 18. DOI 10.3390/ijms18030633

Paper I

EPIGENETICS
2020, VOL. 15, NO. 10, 1052–1067
https://doi.org/10.1080/15592294.2020.1748914

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

Major gene expression changes and epigenetic remodelling in Nile tilapia muscle after just one generation of domestication

Ioannis Konstantinidis[a], Pål Sætrom[a,b,c,d,e], Robin Mjelle[a,b], Artem V. Nedoluzhko[a], Diego Robledo[f], and Jorge M. O. Fernandes[a]

[a]Faculty of Biosciences and Aquaculture, Nord University, Bodø, Norway; [b]Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway; [c]Department of Computer Science, Norwegian University of Science and Technology, Trondheim, Norway; [d]Bioinformatics Core facility-BioCore, Norwegian University of Science and Technology, Trondheim, Norway; [e]K.G. Jebsen Center for Genetic Epidemiology, Norwegian University of Science and Technology, Trondheim, Norway; [f]The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Midlothian, UK

ABSTRACT

The historically recent domestication of fishes has been essential to meet the protein demands of a growing human population. Selection for traits of interest during domestication is a complex process whose epigenetic basis is poorly understood. Cytosine hydroxymethylation is increasingly recognized as an important DNA modification involved in epigenetic regulation. In the present study, we investigated if hydroxymethylation plays a role in fish domestication and demonstrated for the first time at a genome-wide level and single nucleotide resolution that the muscle hydroxymethylome changes after a single generation of Nile tilapia (*Oreochromis niloticus*, Linnaeus) domestication. The overall decrease in hydroxymethylcytosine levels was accompanied by the downregulation of 2015 genes in fish reared in captivity compared to their wild progeni- tors. In contrast, several myogenic and metabolic genes that can affect growth potential were upregulated. There were 126 differentially hydroxymethylated cytosines between groups, which were not due to genetic variation; they were associated with genes involved in immune-, growth- and neuronal-related pathways. Taken together, our data unveil a new role for DNA hydroxy-methylation in epigenetic regulation of fish domestication with impact in aquaculture and implications in artificial selection, environmental adaptation and genome evolution.

ARTICLE HISTORY
Received 28 October 2019
Revised 23 February 2020
Accepted 25 March 2020

KEYWORDS
Domestication; epigenetics; DNA hydroxymethylation; muscle growth; *Oreochromis niloticus*

Introduction

For thousands of years, domestication of plants and terrestrial mammals has shaped human evolution. Our civilizations, societies, global demography and diet have been inextricably linked with this process [1]. Historical evidence of fish domestication is recorded much later than that of terrestrial animals. Its first documented instance was when freshwater fishes were attracted to rice paddies and managed alongside them [2]. Domestication is a rather long and complex process, whereby wild animals are not simply kept in captivity, but bred for enhanced traits that are beneficial to humans. Phenotypic, behavioural as well as genetic changes are known to be linked to mutations and shifts in allele frequencies that occur gradually after consecutive generations of selection [3–6]. Nevertheless, phenotypic changes are often detected within a single generation of domestication.

For example, selection for growth can result in an average increase in body weight of 13% per generation in aquatic species of commercial value [7]. Selected lines of Atlantic salmon (*Salmo salar*) demonstrate an increase of 4.6% in feed efficiency ratio per generation compared to their wild counterparts, [8] and weight gain in Nile tilapia can reach up to 14% per generation [9]. Epigenetic mechanisms that promote phenotypic remodelling without involving changes in the nucleo- tide sequence are expected to be involved in domes-tication. 5-hydroxymethylcytosine (5hmC) is an epigenetic DNA modification produced through the oxidation of 5-methylcytosine (5mC) by ten-eleven translocation enzymes [10]. Recent evidence pin- points the importance of 5hmC in epigenetics and its association with active histone marks, histone accessibility and gene regulation [11–13]. Its particular enrichment within gene-bodies; mainly within

CONTACT Jorge M. O. Fernandes jorge.m.fernandes@nord.no Nord University Bodø 8049, Norway
Supplemental data for this article can be accessed here.

promoters and intragenic regions, as well as its linkage to gene-rich genomic regions suggest that 5hmC is a transcriptionally relevant epigenetic modification [14,15]. Depending on the location, 5hmC enriched regions can be functionally relevant by maintaining an active or inactive chromatin state [14]. Compared to 5mC marks, hydroxymethylation seems to be involved in highly complex mechanisms by activating secondary molecules such as non-coding RNAs. In particular, increased levels of 5hmC at the promoter of microRNA-365-3p, resulted in the regulation of nociceptive behaviours in mice through a potassium channel (Kcnh2), which also pinpoints the accuracy and effectiveness of such modifications to rapidly transfer and regulate environmental signals through the peripheral and central nervous systems [13]. Furthermore, in a state of hypertrophied cardiac mus- cle cells in mice, 5hmCs are shifting from gene bodies to intergenic regions. Cytosine hydroxymethylation is a pre-activating mark of chromatin remodelling, a process that is crucial for chromatin accessibility [11]. To date, the role of 5hmC in teleosts and their domestication remains unknown.

Here we reveal that the muscle hydroxymethy-lome changes when the offspring of wild Nile tilapia are reared in captivity for just one genera- tion of domestication. We observed an overall downregulation of immune genes, possibly linked to pathogen exposure and immune cell infiltration in muscle of wild individuals, concomitantly with an upregulation of metabolic and myogenic genes. By focusing on differentially expressed (DE) genes that were associated with differentially hydroxy-methylated cytosines (DhmCs), we demonstrate that 5hmC may play a key role in the regulation of cell cycle, metabolism and muscle growth.

Results

The nuclear but not the mitochondrial muscle hydroxymethylome differs between wild Nile tilapia and their offspring reared in captivity

Due to the nature of the reduced representation hydro- xymethylation profiling (RRHP) libraries, samples with higher levels of hydroxymethylation produced higher number of hydroxymethylated fragments. Therefore, all libraries were sequenced with equal volumes for proportional sequencing. On average, 69 and 29 M raw reads were obtained per library from the W and D groups, respectively. Approximately 54% of the reads were uniquely mapped to the reference gen- ome, while 14% were multiple mapped (Fig. S1). In total, we covered 1.2 out of 1.8 million possible CCGG sites from both strands across all chromosomes and scaffolds. After removing 5hmC sites having less than 26 counts across 7 or more samples (see Methods), 298,289 CCGG sites were used for further analyses.

Hydroxymethylation was present throughout the whole genome without forming any particular clus- ters in specific chromosomes (Figure 1(a)). The overall 5hmC pattern across the genome showed increased hydroxymethylation levels in the W group (uniquely mapped CCGG sites; one-tailed Student's t-test; $p < 0.01$). Liquid chromatography/ mass spectrometry (LC-MS/MS) confirmed that the global 5hmC levels were significantly higher in the W compared to the D group (Figure 1(b)).

The significant differences in library sizes between the two groups did not reflect similar effects on 5hmC variation, which was higher within the D group (ANOVA; $D_{var} = 1.54 \times 10^{14}$, $W_{var} = 1.51 \times 10^{14}$, $F = 9$, $F_{crit} = 5$, $p = 0.01$ with $a = 0.05$; Table S1), and may be due to relaxed natural selection and pheno- typic divergence in captivity conditions [16]. However, the overall variation of individual CCGG sites after filtering was higher in the W group and this variance was mainly asso- ciated with large differences between 5hmC sites. The opposite effect was observed in the D group, where the variance was attributed mainly to within-site variation (Table S2). Based on the annotation of all CCGG sites as well as the sites that were hydroxymethylated in our data set, we observed that 5hmCs were relatively depleted within intergenic regions (19%) and enriched in exons (31%) and promo- ters (33%) (Figure 2).

A total of 13 of the 38 possible mitochondrial CCGG sites were substantially hydroxymethylated (Figure 1(c)). Four of these sites were detected within the *cox1* (cytochrome c oxidase subunit I) gene, which is involved in a large number of processes including metal ion binding, aerobic respiration, and oxidative phosphorylation. Notably, the highest 5hmC levels per CCGG site were observed within the 12 S and 16 S ribosomal RNA genes.

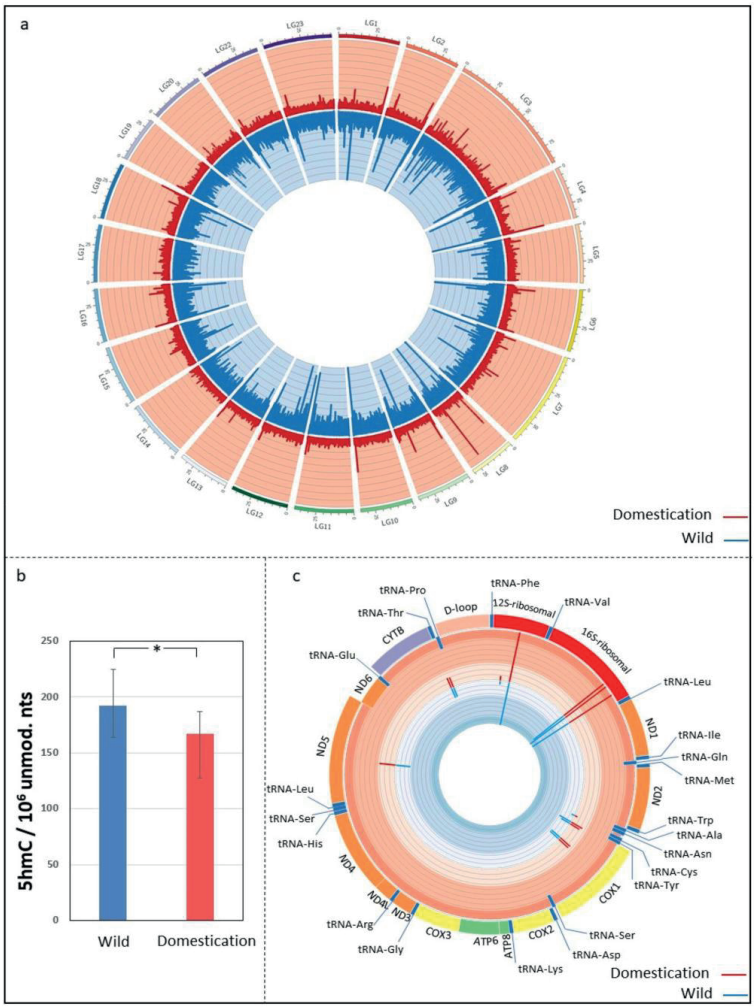


Figure 1. Hydroxymethylation differences between wild Nile tilapia and their offspring reared in captivity. (a), Circular representation of the Nile tilapia nuclear genome (linkage groups LG1-LG23, no scaffolds included for visualization purposes). Only sites with substantial levels of 5hmC were included (298289 CCGG sites). (b), Histogram with the average number of hydroxymethylated cytosines,5-hm(dC), per 1 million nucleotides, in fast muscle of wild fish compared to their offspring undergoing domestication. Significance was evaluated using the one-tailed Student's t-test (n = 6, p = 0.04). Standard errors of the means are shown by black bars. (c), Circular representation of the Nile tilapia mitochondrial genome with its gene features in both L and H strands. (a,c), Blue and red peaks (oriented in and out, respectively) represent the hydroxymethylation levels in the groups of wild fish and their offspring undergoing domestication, respectively. The height of each peak is the average count for each group per location (n = 6). The 5hmC counts per genomic location and individual, as well as the standard deviation for the W and D groups in panels 1a and 1 c can be found in Supplementary Table S9.

Differentially hydroxymethylated genes are involved in immunity, as well as in various intra- and extracellular processes

We identified 126 differentially hydroxymethylated cytosines (DhmCs) between the two groups (p adj. val. <0.05; Figure 3). All sites were hyper-hydroxymethylated in the wild group compared to the first generation of fish reared in captivity. DhmCs were mostly enriched within linkage group 3 but no DhmCs were found in chromosomes LG5,

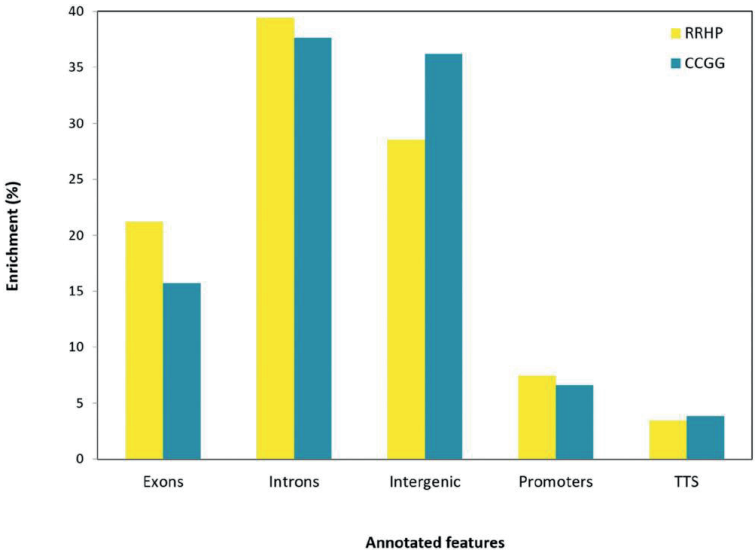


Figure 2. Histogram of gene features (Promoters defined from −1kb to +100bp; TTS defined from −100bp to +1kb). Yellow and turquoise represent the enrichment of sites with substantial levels of 5hmC (RRHP − 295720 sites) and total CCGGs (whole genome − 1791,720 sites) across the Nile tilapia genome, respectively.

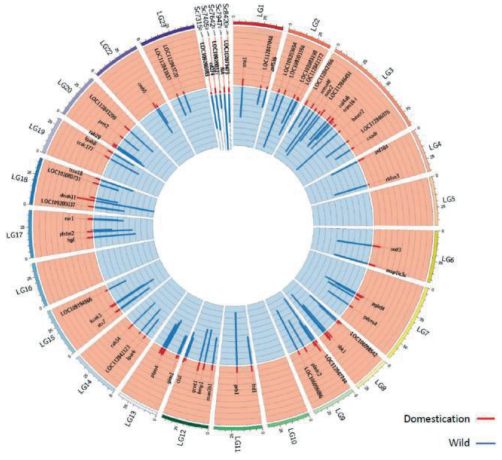


Figure 3. Circular representation of the Nile tilapia nuclear genome (linkage groups LG1-LG23 and scaffolds) showing 126 DhmCs. For visualization purposes the numbers in the scaffold symbols 'Sc' represent the last four digits in the accession number starting with 'NW_02032' in the NCBI database (e.g. 'Sc7315' represents scaffold 'NW_020327315'). Gene symbols refer to the closest gene to the reported CCGG site. See Figure 1(a,c) for explanations of peaks. The 5hmC counts per genomic location and individual, as well as the standard devia- tion for the W and D groups in Figure 3 can be found in Supplementary Table S10.

LG10, LG16 and the mitochondrial genome. Most of these DhmCs were located in gene bodies (79) rather than in intergenic regions (47) (Fig. S2). DhmCs located in intergenic regions were assigned to genes based on their proximal distance to the transcription start site (TSS) of the corresponding gene. In total, 61 and 34 unique genes were linked to intragenic and intergenic DhmC marks, respectively. Proportionally to the genomic and filtered CCGG sites, DhmCs were enriched within transcription termination sites (geno- mic-TTS; p = 0.055, filtered-TTS; p = 0.041, binomial test). The genes with the highest number of DhmCs were *myo5b* (myosin Vb) and *kmt2e* (inactive histone- lysine N-methyltransferase 2E) with 4 DhmCs, as well as the *lpar6* (lysophosphatidic acid receptor 6), *map1lc3 c* (microtubule-associated protein 1 light chain 3 C) and *extensin* with 3 DhmCs (Table S3). Genes that were differentially hydroxymethylated are involved mainly in immune response (*fcgr2b, stx7, cd274, havcr2, ctsl* and *sh2d1a*), intracellular processes (*myo5b, rab29, rab34, map1lc3 c, elovl4* and *gcnt1*), and core extracellular processes (*bmp1, col4a6, ctsl, hgf, muc2, plxnc1* and *plxdc2*) (Table S4). These genes are known to play crucial roles in the modula- tion of response to growth factors and cell growth

inhibition, and they are associated with the canonical pathway of the muscle that is affiliated to structural or functional extracellular matrix proteins (p = 0.007). By using the zebrafish annotated features, we identi- fied genes that are involved in axon (GO:0061564; Log(p) = −2.67), neuron (GO:0048666; Log(p) = −2.37) and neuron projection development (GO:0031175; Log(p) = −2.16) as well as the phago- some pathway (Log(p) = −2.26) (Table S5).

A single generation of domestication is associated with substantial changes in the muscle transcriptome

The average size of the rRNA depleted libraries was 19.2 M raw reads for the W samples and 19.3 M raw reads for the D samples (Fig. S3). Of these, 13.1 M and 12.8 M reads were uniquely mapped to the Nile tilapia for the W and D samples, respectively. Of the mapped reads, 3.6 M and 3.8 M reads were mapped to multiple locations in the genome for the W and D samples, respectively.

We found 2,121 genes (q < 0.05, |log FC|> 1.5) that were differentially expressed between the W and D generations. A substantial number of genes (2015) were downregulated in the D group, while only 106 genes were upregulated (q < 0.05, | logFC|> 1.5; Figure 4). A large proportion of the downregulated genes were enriched for immune functions, such as lymphocyte activation (log-(p) = −24.11), cytokine production (log(p) = −21.6), myeloid leukocyte activation (log(p) = −20.49), chemotaxis (log(p) = −20.26) and leukocyte migra- tion (log(p) = −18.86), as well as the cytokine- mediated (log(p) = −15.68) and immune response signalling pathways (log(p) = −16.98) (Fig. S4; Table S6). Upregulated genes were involved in metabolic processes such as monocarboxylic acids (q = 3.43 × 10-5), fatty acids (q = 15 × 10-2) and glucose catabolic process to pyruvate (q = 3.3 × 10-2), as well as the glycine, serine and threonine metabolism KEGG pathway (q = 0.016). Interestingly, significantly enriched genes within biological functions such as striated muscle tissue development (q = 3.4 × 10-2) and muscle system process (q = 1.9 × 10-3) were also upregulated in the D group, which may indicate a higher growth potential of fish in captivity (Fig. S5, Table S7).

Hydroxymethylation and gene expression are positively associated

A total of 11 DhmCs were found within 7 DE genes (Table 1; q < 0.05, logFC> 1.5). These 7 genes were hyper-hydroxymethylated and upregu- lated in the wild group, indicating a positive asso- ciation between hydroxymethylation and gene expression. The gene that displayed the highest fold change in terms of gene expression was *sh2d1a* (SH2 domain containing protein 1A), which is involved in cellular defence and regula- tion of immune responses.

The remaining DE genes containing DhmCs play roles in cell survival (*creb5*) and broad neu- ronal-related effects such as axon guidance and development, neuron projection guidance and development, regulation of axonogenesis and gen- eration of neurons (*dok2, plxnc1, dnah11* and *myo5b*). A gene essential for the structural main- tenance of the extracellular matrix and wound healing of the skin in mammals (*bmp1*) also con- tained DhmCs and was DE between the two groups. Notably, DhmCs within DE genes were mostly located within introns (8/11), which might affect alternative splicing events or cryptic promoters. Finally, distance-based association between DhmCs and DE genes without the pre- viously mentioned thresholds revealed one addi- tional gene (*fzd1*; p < 0.05, logFC = 1.3) which is involved in a broad range of processes within the Wnt signalling pathway and was differentially hydroxymethylated at a distance of approximately 23 kbp upstream of its transcription start site.

No evidence of population structure between W and D groups

Considering the unknown genetic background of wild Nile tilapia males and their contribution to the D group genotypes, we investigated whether genetic mutations within the restriction sites (CCGG) could potentially explain the observed epigenetic reprogramming during the first genera- tion of domestication. In total, 2,871 single nucleo- tide polymorphisms (SNPs) were identified among all samples from the same sequencing data that was used to identify 5hmCs.

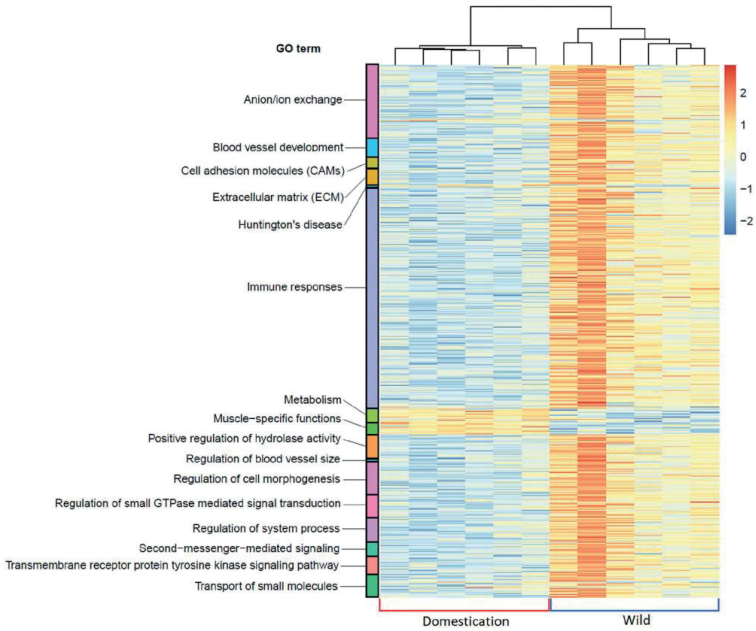


Figure 4. Heatmap showing the differentially expressed genes assigned to their gene ontology (GO) terms between the two clustered groups. Red and blue brackets at the bottom of the heatmap correspond to the 'Domestication' and 'Wild' groups, respectively. The gradient from blue to red colour represents down- to upregulation, respectively. Each GO term is colour-coded as shown at the left side of this heatmap. All related statistics, including significance of each enriched GO term, are shown in tables S5 and S6.

Pairwise SNP dis- tances confirmed that each D-offspring corre- sponds to each W-parent, considering that the smallest distances were observed between each W-D pair (Fig. S6). The fixation index between the W and D groups was negative (Fst = −0.05), which confirms our hypothesis that there was no genetic structure between the two Nile tilapia populations, as shown in Figure 5. The within- group nucleotide diversity estimates were similar for both populations.

Discussion

Major leaps in growth rate during the first genera- tion of fish domestication cannot be explained by genetic adaptation alone, such as mutations and selective sweeps that lead to specific shifts in allele frequencies. As shown before, genetic adaptation to captivity is a relatively slow, species-specific process and can only be detected after several generations [3,4,6,17].

Nevertheless, within a single generation, hatchery-reared fish are selected for traits that are advantageous in captiv- ity (e.g., growth rate, egg size, fecundity, physiolo- gical processes associated with smoltification and relaxed anti-predator behaviour) but can be mala- daptive in the wild, resulting in reduced reproduc- tive fitness [18,19]. Recent evidence has shown that gene expression patterns can be altered within the first generation of captivity rearing and that these changes can be transferred to the next gen- eration [20]. Indeed, wild growth-hormone trans- genic coho salmon (*Oncorhynchus kisutch*) and rainbow trout (*Oncorhynchus mykiss*), appear to have similar modifications in genetic pathways as their counterparts undergoing domestication [21,22]. However, very limited research has been made to uncover the mechanisms underlying these changes. In the present study, 2,121 DE genes were identified between wild Nile tilapia and their first- generation offspring.

Table 1. List of genes that were both differentially hydroxymethylated and differentially expressed between wild Nile tilapia (W) and their offspring undergoing domestication (D). P adj.: FDR adjusted p value (q value), log FC: logarithmic fold-change, where negative values represent upregulation in D and positive values represent upregulation in W .

Gene symbol	*creb5*	*myo5b*	*myo5b*	*myo5b*	*myo5b*	*dok2*	*plxnc1*	*dok2*	*fzd1*	*sh2d1a*	*dnah11*	*bmp1*
Protein	cyclic AMP-responsive element-binding protein 5	myosin VB	myosin VB	myosin VB	myosin VB	docking protein 2	plexin-C1	docking protein 2	frizzled-1	SH2 domain containing 1A	dynein heavy chain 11 axonemal	bone morphogenetic protein 1
Gene Expression (q value)	0.004	0.009	0.009	0.009	0.009	0.013	0.021	0.021	0.023	0.031	0.036	0.048
(log FC) WvsD	2.3	1.991	1.991	1.991	1.991	2.162	2.022	2.022	1.276	2.688	1.703	1.608
DhmC (distance_to_TSS)	3431	4816	4818	3742	3977	−2695	7015	7017	22676	8999	7153	4250
DhmC (Annotation)	intron 3 of 10	intron 1 of 38	intron 1 of 38	intron 1 of 38	intron 1 of 38	TTS	intron 3 of 6	intron 3 of 6	Intergenic	exon 38 of 40	TTS	intron 1 of 19
DhmC (log FC)W vs D	1.457	6.242	6.233	5.655	4.628	6.421	6.534	5.660	6.414	6.279	1.075	6.481
DhmC (P adj.)	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.018	0.008	0.002	0.030	0.009
Strand	-	-	+	-	+	-	+	-	-	+	+	+
Position	25495940	36567964	36567962	36569038	36568803	34318460	37756187	37756189	2292991	31466924	12201282	13634977
Chromosome	NC_031985.2	NC_031972.2	NC_031972.2	NC_031972.2	NC_031972.2	NC_031981.2	NC_031977.2	NC_031977.2	NC_031976.2	NC_031966.2	NC_031982.2	NC_031977.2

A rapid adaptation to captivity conditions was evident at the molecular level, with most genes being downregulated in the group undergoing domestication. In a recent study [20], 723 genes were reported to be DE between wild-born and first generation hatchery-reared offspring Steelhead trout (*Oncorhynchus mykiss*). Unlike our findings in Nile tilapia, most of these genes were upregulated in hatchery-reared fish and only the muscle-related troponin I gene (*tnni1*) was found to be upregulated in both steel- head trout and Nile tilapia reared in captivity. This might be an indication of either species- or tissue- specificity (whole fry vs fast muscle) or even dif- ferences in environmental factors (e.g., rearing conditions, temperature, dissolved oxygen and pH) that contribute to the expression of these genes.

A total of 316 genes directly associated with the immune response were upregulated in wild Nile tila- pia compared to their counterparts undergoing domestication. These results demonstrate the challen- ging and variable environmental factors affecting wild populations, namely, high water temperatures during the summer months (up to 35°C at river banks and shallows), exposure to viruses, bacteria and parasites, as well as predators and fishing pressure compared to the predator- and pathogen-free conditions of our recirculating aquaculture system. For example, one of the major physiological differences between the W and D group was the obvious presence of endo-parasites in the W group (supplementary video 1), which could justify the upregulation of several genes such as interleukins and their receptors (*il21 r, il12rb2, il12rb1, il16, il2rb, il7 r, il18rap, il1rap, il17 f, il1r2, il4l1, il1b, il31ra*) and other genes previously reported to be upregulated as a response to parasitic load in vertebrates [23] (e.g., *irf4, irf1, thbs1b, socs3, junb, prf1, ccr9, itga5* and *slc3a2*). Based on the tissue enrichment analysis and manual inspection of expres- sion patterns of these genes in humans (p adj. val. ≈ 10^{-7}; Fig. S7), it is plausible that wild fish had a higher degree of immune cell infiltration in muscle tissue than their offspring reared in captivity, even though this could not be confirmed by histological analysis due to the unavailability of necessary chemicals and equipment in the field.

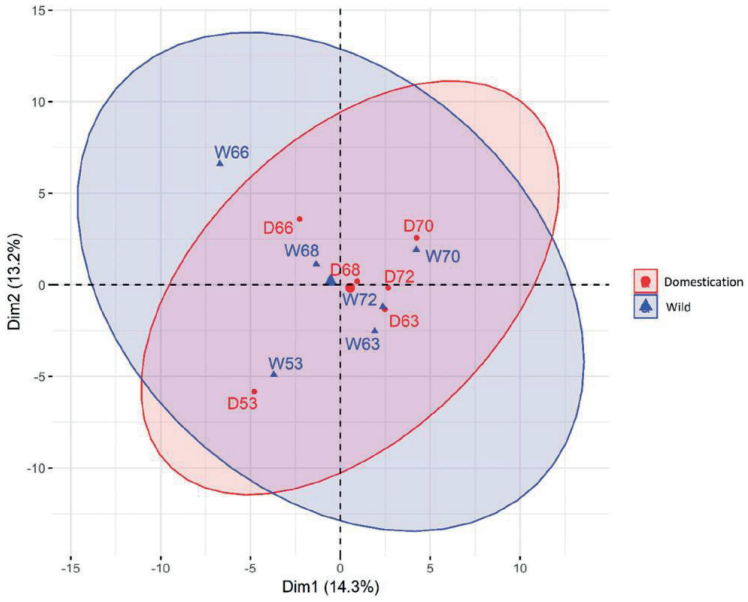


Figure 5. Principal component analysis (PCA) showing the genetic diversity of the 6 wild and 6 first-generation offspring based on their SNPs. Blue triangles and red circles represent the wild fish (W) and their counterparts in the first generation of domestication (D), respectively.

It is noteworth that T and B cell activation, differentiation, and function is mediated by ten- eleven translocation enzymes and the DNA oxida- tion and demethylation pathways [24]. We found that *tet3* was upregulated in the W group ($p < 0.05$, logFC = 1.17), which suggests an overall demethylation process within both muscle and immune related cells, with major implications in gene transcription. These findings demonstrate the complexity of epigenetic mechanisms and their roles in gene regulation, cell defence and growth. The overall muscle composition changes with age and the different cell types present in muscle have distinct DNA methylation patterns [25]. As the exact age of the wild Nile tilapia could not be determined by otolith analysis, one cannot exclude the possibility that the observed upregu- lation of *tet3* in the W group was at least partly due to age differences. Nevertheless, the local fishermen that assisted us in Egypt know that these fish were approximately 4–5 months old due to the fishing pressure imposed by the gill nets and traps used. The changes in gene expres- sion and 5hmC levels identified between wild fish and their progeny undergoing domestication are most likely due to environmental differences, since the latter were also 5 months old at the time of sampling. Stress-related epigenetic changes may also play an important role in immunocompetence and acute stress can increase the transcription of immune-related genes [26]. Even though the authors have taken precautions towards reducing any bias in the experimental design, the collection of fertilized eggs from the mouth of the W females without anaesthesia represented an additional acute stress that could potentially have influenced the epigenome.

Crowded conditions can potentially result in skin abrasions and hypoxia [27] and rearing den- sity plays an important role towards the adaptation to captivity [20]. The rearing density in our recir- culating aquaculture system was higher than in the wild but we found few DE genes related to crowd- ing conditions. Specifically, we identified only two genes, *grn* (granulin) and *phb2* (prohibitin-2) that were upregulated in fish reared in captivity and that were related to wound healing and response to hypoxia, respectively.

Studies on commercial fish [28–31] and model organisms [32], have shown that captivity rearing usually offers the appropriate conditions for increased growth rates. Phenotypic selection has led to the development of the genetically improved farmed tilapia strain, which gradually gained up to 15% in body mass per generation within the first six generations of domestication [33]. Considering the impact of environmental factors on the epigenome and the overall health status during the transition from the wild to an enclosed rearing environment, it is plausible that the first signs of growth-gain are strongly associated with a lower pathogen load in captivity. This notion is further supported by the fact that the immune system is closely associated with the muscle development, since chronic inflammation can lead to muscle damage and fibrosis [34]. It is likely that other mechanisms directly affect muscle growth during the first generation of domestication. We found two genes (*myo5b* and *creb5*) that are both DE and contain DhmCs between wild and first-generation captive fish. These genes are involved in the PI3K-Akt signalling pathway, which is one of the most important intracellular pathways in muscle that regulate cell cycle, metabolism, proliferation and growth. Myosin Vb targets and activates directly the phosphatase and tensin homolog (PTEN) [35], which in turn dephosphorylates PIP_3 (phosphatidy- linositol 3,4,5 − trisphosphate) and prevents cells from growing. Both *myo5b* (LogFC = 1.991; q = 0.009) and *pten* (LogFC = 1.303; q = 0.043) were upregulated in wild individuals, suggesting a possible epigenetic mechanism that regulates mus- cle hypertrophy. Additionally, *myo5b* is directly involved in insulin-dependent glucose transporter 4 (GLUT4) translocation in muscle cells and uses *rab8a* and *rab13* GTPase-activating proteins that require Akt signalling and activation [36,37]. Apart from its involvement within the PI3K-Akt signalling pathway, *myo5b* plays various roles in neuronal pro- cesses. It recycles the endosome delivery into hippo- campal neuron spines in dendrites, targets and activates PTEN, which in turn regulates the soma size, and promotes axonal development in conjunc- tion with the Ras-related protein *rab10* [38]. It would not be surprising if the epigenetic regulation of neu-ronal development was associated with muscle development, considering that neurons are an inte- gral part of the skeletal muscle. *Creb5* codes for a protein that is a substrate for the Akt/PKB and is involved in cell survival. Among other pathways, *creb5* is also involved in insulin resistance and glu- coneogenesis in the muscle, which occurs during periods of starvation and low-carbohydrate diets [39]. *Myo5b* and *creb5* seem to be good candidates for further investigation into their epigenetic regula- tion during domestication. Overall, genes that were upregulated in the D group were mostly related to metabolic processes (43/106). We also identified genes involved in myosin binding and actomyosin structure organization (*actc1*), regulation of skeletal muscle contraction (*casq2, eno1, mb*) and skeletal muscle development (*srpk3*), response to and regula- tion of glucose import (*rtn2, mef2a*) as well as cal- cium release channels and insulin secretion (*fkbp1b*). Interestingly, *srpk3* (serine/arginine-rich specific kinase 3), is a well conserved gene among mamma- lian species and it is closely related with myocyte- specific enhancers that promote muscle growth in the skeletal muscle tissue [40]. These results coupled with the dietary effects in captivity, such as high- quality diet and regular feeding *ad libitum*, suggest an increase in metabolic capacity that leads to a higher growth rate from the very first generation of fish undergoing domestication. It is plausible that this effect is partly due to nutritional programming. Previous studies in zebrafish (*Danio rerio*) have shown that parental nutritional programming through increased dietary intake of arachidonic acid or deficiency in 1-C nutrients results in DNA methylation changes and epigenetic regulation of gene expression in the liver of the offspring [41,42]. The total number of 5hmCs in our reduced repre- sentation of the Nile tilapia genome was substantially high, covering ubiquitously the nuclear genome as well as identifying 5hmCs within the mitochondrial genome. High levels of 5hmC within the 12 S and 16 S ribosomal RNA might indicate that epigenetics could have an effect in the translational dynamics of mitochondrial mRNAs. Although studies have shown 5hmCs to be present in mammalian mito- chondria [43,44], there are no reports of 5hmCs at single nucleotide resolution in the mitochondria of teleosts. Here we provide the first evidence of 5hmC marks across the mitochondrial genome of Nile tila- pia, which also implies the existence or relocation of CpG demethylation mechanisms within the mito- chondria. Strikingly, the large differences of the

nuclear 5hmC levels between wild and first-generation fish, do not reflect the average 5hmC levels within the mitochondria. This suggests that the two demethylation mechanisms are effectively separated. Similarly to our study, 5hmCs in the muscle of zebrafish occur mainly in bodies of genes, some of which have functions with interesting implications in metabolism, development and other muscle-specific functions [45].

We observed higher hydroxymethylation levels in the W group, which resulted to the identification of 126 DhmCs. A previous study that compared the DNA methylation profiles of hatchery-reared and wild coho salmon (*O. kisutch*), identified 37 differ- entially methylated regions (DMRs) that are likely linked with the reduction in fitness of the hatchery- reared fish [46]. These DMRs were found within 52 unique transcripts, but none of them appeared in our DhmC data. Possible explanations for these differ- ences are species-specificity (smoltification related processes that are not present in Nile tilapia), onto- geny stage (fry vs adults), and 5mC/5hmC-specificity (different epigenetic modifications may control dif- ferent pathways). There were two genes in common between our DE genes and these 52 transcripts, namely ankyrin-1 (*ank1*), which is involved in cytos- keleton organization, and *phb2*, which is involved in a large number of processes, such as cellular response to hypoxia and response to wounding. *Ank1* was upregulated in W fish, whereas *phb2* was upregu- lated in the D group, which indicates a common regulation of these genes among fish species during the early stages of domestication. This finding pos- sibly indicates that similar stressors might affect the wild populations of fish with major impact on their fitness and growth potential. However, our data indicate that most molecular networks regulated during the early stages of domestication appear to be species-specific.

According to a recent study, DNA methylation of the first intron has a negative effect on gene expres- sion levels across several species and tissues [47]. In our case, the first intron of *myo5b* is highly hydro- xymethylated in the W group (4 DhmCs, logFC = 5.69, q val = 0.0085) compared to the D group. In addition, we observe a positive correla- tion with its transcript levels (logFC = 1.99, q. val = 0.009). Additionally, a single DhmC was found within the first intron of *bmp1* (logFC = 6.48, q val = 0.008) and it was also upregulated in the W group (logFC = 1.6, q val = 0.048). These results indicate the importance of distinguishing 5mC from 5hmC calls. Unlike our RRHP approach, other com- monly used methods such as reduced representation bisulphite sequencing (RRBS) are unable to do so, which limits the conclusions from those studies.

While genetic variation alone did not explain the large differences in 5hmC patterns between the W and D groups, it is worth mentioning that the SNP coverage was limited due to the nature of the RRHP libraries. A broader genome-wide approach could potentially identify additional SNPs between the two groups. Furthermore, hydroxymethylation levels at the beginning of the embryonic development in fish are very low [45,48], therefore epigenetic reprogram- ming is likely occurring as a result of environmental conditions rather than of maternal effects.

Conclusion

The process of domestication affects animals in ways that still remain unknown. Both genetic and epigenetic changes can alter the phenotype within a single generation. Here, we demonstrated for the first time that DNA hydroxymethylation in muscle differs significantly between wild fish and their progeny reared in captivity. We further identified DhmCs mostly within gene bodies, which might indicate their functional role in epigenetic regula- tion. These DhmCs were associated mainly with immune-, neuronal- and growth-related processes. Concurrently, thousands of genes were differen- tially expressed between groups, with upregulation of immune related genes in the wild and the upre- gulation of metabolic and muscle-specific genes in the first generation of fish undergoing domestica- tion. Notably, genes that were found to be both differentially hydroxymethylated and differentially expressed had a consistently positive correlation. Another remarkable finding was the hyper- hydroxymethylation of the first intron of *myo5b* and its connection with the upregulated *pten* gene in the wild group. *Pten* is involved in the PI3K-Akt signalling pathway and muscle atrophy. Further investigation into the PI3K-Akt pathway could reveal one plausible way for rapid regulation of growth and muscle hypertrophy in fish during the first stages of their domestication. Lastly, we

found no evidence for genetic differentiation between the two groups, showing that the observed hydroxymethylation changes within the first generation in captivity have an epigenetic rather than genetic basis. Understanding the epi- genetic mechanisms, namely hydroxymethylation, involved in early domestication is essential for aquaculture diversification (farming of new spe- cies), sustainable production, and wild fish stock management practices.

Materials and methods

Ethics approval

This study was approved by the Nord University (Bodø, Norway) ethics board and all procedures involving animals were performed according to the instructions of the Norwegian Animal Research Authority (FOTS ID 1042).

Experimental design

A total of 6 wild (W) mouthbrooding Nile tilapia females were captured using traps and gill nets along the river Nile in Luxor, Egypt within a 500 m radius from the following coordinates (latitude: 25.6655° N; longitude: 32.6186° E). After collecting approximately 180 fertilized eggs from each female by gently holding their head in containers filled with river water, the fish were placed in 25 L tanks, transferred from the boats to the sampling location and sacrificed by immer- sion in a clove oil (Sigma Aldrich, USA) emulsion containing 12 ml of a clove oil pre-mix in 96% ethanol (1:10 v/v) in 10 L of water [49].

A 0.5 cm thick cross-section was taken at 0.7 standard length, just anterior to the urogenital open- ing. Fast (white) skeletal muscle was carefully dis- sected from the left dorsal quadrant of the section and preserved in DNA/RNA Shield (Zymo Research, USA) at 4°C. Sixteen days post-sampling, all tissues were stored at −20°C. Each batch of eggs was disin- fected with hydrogen peroxide (1 g H_2O_2/L for
10 minutes) and placed in egg rockers (Cobalt Aquatics, USA) installed in a 60 L tank with UV-treated water, containing 5‰ NaCl (w/v). Approximately, 85% of the eggs hatched within 4 days at 28°C. At 9 days post-fertilization, the larvae were placed in fish transport bags filled with UV- treated and 100% oxygen saturated water, and were air freighted to Nord University's research station (Bodø, Norway). Transportation of the larvae lasted approximately 18 hours and their survival rate exceeded 95%. They were reared for 5 months at a maximum density of 27 fish/m^3 in a freshwater recirculating aquaculture system (pH = 7.6, DO = 100%, temperature = 28°C and photoperiod LD 13:11) and were fed *ad libitum* with 0.15–0.8 mm Amber Neptun pellets (Skretting, Norway). The fish reared in captivity represent the first generation of fish (D) undergoing domestication. Each D group was kept separate in different tanks based on their ancestry. Fast muscle from 6 D females was sampled and stored using the same procedure as above. W and D samples were genetically paired (i.e., each D individual corresponded to a W progenitor) and at approximately the same age and ontogeny stage (i.e., puberty). All sampling conditions at the research facilities were simulated to match the sampling of the wild fish in Egypt. To compensate for the starva- tion period of wild mouthbrooding females, D individuals were starved for 48 hours prior to sampling. All fish were sampled at approximately the same time of the day (7–10 am). Both W and D fish were removed from their environment (Nile river for W and rearing 1500 L tanks for D) and kept for a maximum period of 30 minutes in a 25 L tank with anaesthetic before sampling. All sampling pro- cedures were performed in an air-conditioned room at 25°C and oxygen was provided by an air pump. Additional information on the animals, such as total and standard lengths, weight and general observa- tions at the time of sampling can be found in supplementary material (Table S8). Aliquots of the same fast muscle sample from each individual were used for DNA and RNA isolation.

Global 5hmC quantification by LC-MS/MS

Genomic DNA from fast muscle was isolated from all 12 individuals (6 D and 6 W) using the Quick-DNA miniprep Plus kit (Zymo Research). The concentra- tion of DNA was determined using the Invitrogen Qubit 3.0 fluorometer (ThermoFisher Scientific, USA). Quality and integrity of the DNA was assessed using the NanoDrop 1000 Spectrophotometer (ThermoFisher Scientific) and the Tapestation 2200

(Agilent Technologies, USA), with absorbance ratios 260/280 > 1.93, 260/230 > 1.89 and minimum DNA integrity number (DIN) of 8.4, respectively. For the global quantification of 5hmC, the internal standard d3-5-hydroxymethyl-2'-deoxycytidine (Toronto Research Chemicals, Canada) was added to the DNA samples. Subsequently, they were hydrolysed completely to deoxynucleosides by 20 U benzonase (Santa Cruz Biotech, USA), 0.2 U nuclease P1, and 0.1 U alkaline phosphatase (Sigma Aldrich) in 10 mM ammonium acetate pH 6.0 and 1 mM magnesium chloride at 40°C for 1 hour. Three volumes of acet- onitrile were added to the mix before their centrifuga- tion (16,000 g, 30 min, 4°C). The supernatants were dried and dissolved in 50 µl water for LC-MS/MS analysis of modified and unmodified deoxyribonu- cleosides. Chromatographic separation was per- formed using an Agilent 1290 Infinity II UHPLC system with an ZORBAX RRHD Eclipse Plus C18 150 × 2.1 mm ID (1.8 µm) column protected with an ZORBAX RRHD Eclipse Plus C18 5 × 2.1 mm ID (1.8 µm) guard column (Agilent Technologies). The mobile phase consisted of water and methanol (0.1% formic acid was added to both phases). For the initial run of 5-hm(dC), the flow rate was set at 0.15 ml/min with 5% methanol for 30 seconds, which was adjusted to a gradient of 5-15% methanol for the next 4 min- utes. The flow rate was increased to 0.25 ml/min in a methanol gradient of 15-90% for 3 minutes, and finally re-equilibration was performed at 5% metha- nol for 4 minutes. A portion of each sample was diluted for the analysis of unmodified deoxynucleo- sides. Unmodified deoxynucleosides were chromato- graphed isocratically with 20% methanol. Mass spectrometric detection was performed using an Agilent 6495 Triple Quadrupole system operating in positive electrospray ionization mode, monitoring the mass transitions 261.1/145.1 (d3-5-hm(dC)), 258.1/142.1 (5-hm(dC)), 252.1/136.1 (dA), 228.1/112.1 (dC), 268.1/152.1 (dG), and 243.1/127.1 (dT).

RRHP library preparation and sequencing

The preparation of the reduced representation hydroxymethylation profiling (RRHP) libraries was carried out according to the manufacturer's protocol (Zymo Research). Fast muscle stored in DNA/RNA Shield (Zymo Research) was homoge- nized using ZR bashing beads in lysis tubes (Zymo Research) on a Precellys 24 homogenizer (Bertin Instruments, France) at 5000 rpm for 2 cycles of 20 seconds. DNA was extracted using the Quick- DNA miniprep kit (Zymo Research) with a 25 µg binding capacity. After digestion with the restric- tion enzyme MspI at 37°C for 8 hours, the DNA was washed using the Zymo-Spin IC columns pro- vided in the RRHP 5-hmC library prep kit. Specific adapters that re-create the MspI cutting site were introduced. After library extension, all fragments were glucosylated at 37°C for 2 hours and washed as recommended. During this step, all 5hmCs at the starting CCGG junction of each fragment were glucosylated and protected from the second MspI digestion. At the end of this procedure, only fragments with hydroxymethy-lated CCGG junctions remain adapterized. Double indices were introduced after size selection according to Illumina's instructions [50] for a 6-plex sequencing design. To minimize bias dur- ing library preparation, all samples were handled simultaneously, as suggested in the literature [51,52]. In total, 6 D, 6 W and 2 positive control libraries were sequenced. The positive controls were 5hmC independent-libraries and contained all possible CCGG sites across the genome. Thus, we were able to identify biases in fragment ampli-fication within our count-based matrices. All sam- ples were sequenced at the Norwegian Sequencing Centre with equal volumes in three lanes on a HiSeq4000 (Illumina, USA) to obtain 150 bp paired-end reads. The positive controls were sequenced at 1/3 volume in all three lanes to account for possible lane effects. Due to the low complexity of the libraries (CCGG) at the begin- ning of each read, 30% PhiX control v3 (Illumina) was added to each lane.

Bioinformatic analyses

Adapters and N bases from the 3' end of each read were trimmed using trim_galore v0.4.4 (Babraham Bioinformatics) [53]. Clean reads were aligned to the reference genome O_niloticus_UMD_NMBU [54] with Bowtie v0.12.8 [55] using the parameters -best, -v 1, -n 1 and -m 3. These options compromised alignment rates for robustness and reliability, as they minimized multiple aligned reads. From each SAM alignment file, only reads starting with CCGG from

both strands were extracted and their coordinates (chromosome and position) were used as input into R [56]. A count matrix was created for each sample and afterwards all samples were concatenated into one count matrix. With the aim of avoiding false discovery of 5hmCs, mainly due to within-group var- iation, we applied double filtering and stringent para- meters. At first, a filter was applied in order to remove 5hmCs with low number of counts. All 5hmCs were removed from the matrix when 7 out of 12 samples had below 1 count. This filter not only excluded hydroxymethylated cytosines with low counts but it increased the overall median value and ensured that 5hmC on/off switches remained in the data set. After calculating the median, we also removed 5hmC sites when 7 or more samples had counts lower than the median value of the matrix. In order to perform a comparison of the two groups, our design handled the six different families as blocks while the W and D samples were considered as two separate groups. We then used the R package limma to perform a duplicate correlation based on their counts and the RRHP bioinformatics was performed similarly to gene expression (RNA-seq) rather than methylation (RRBS) data analysis. The Benjamini-Hochberg pro- cedure was used to calculate adjusted p-values for each 5hmC and single cytosines with an adjusted p value ≤ 0.05 were counted as DhmCs between the W and D groups. The annotation of all significant positions was performed using the software HOMER, annotatePeaks.pl [57] and the NCBI *Oreochromis niloticus* annotation release 104. The association of differentially hydroxymethylated genes with their cor- responding gene ontology terms was performed with Metascape [58]. Two separate gene ontology analyses for Nile tilapia genes were performed based on known functions for their annotated (i) human (*Homo sapiens*) and (ii) zebrafish (*Danio rerio*) orthologs.

RNA library preparation, sequencing and analysis

Fast muscle in DNA/RNA Shield was homogenized as described above. Total RNA from the same 12 indivi- duals was isolated using the Quick-RNA miniprep plus kit (Zymo Research) according to the manufac- turer's protocol and recommendations for samples stored in DNA/RNA Shield. Ribosomal RNA was removed using the Ribo-Zero Gold rRNA Removal kit (Illumina). RNA concentration and integrity were assessed using the RNA High-sensitivity tape on the Tapestation 2200 (Agilent Technologies). RNA-seq libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB, USA), according to the manufacturer's instructions. A unique barcode was assigned to each sample and all libraries were pooled at equimolar concentrations. They were then sequenced at the Norwegian Sequencing Centre on a single HiSeq 4000 lane (Illumina), yielding 480 million 150bp paired-end reads. The raw sequencing files were trimmed using trimmomatic v0.33 [59] with the following para- meters: *PE -phred33, LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36*. The trimmed reads were aligned to the Nile tilapia genome (NCBI assembly GCA_001858045.3) using STAR [60] with default settings and paired-end mode. The mapped reads were used as input to *htseq-count* [61] to calcu- late count numbers for each gene using the gff3 gen- ome annotation from the above mentioned genome assembly. Differentially expressed genes were deter- mined using the edgeR package in R. The differentially expression analysis required that each gene should be expressed with at least 1 count per million in more than 25% of the samples. The p-values for the differ- entially expressed genes were adjusted for multiple testing using the Benjamini-Hochberg correction [62]. Based on the gene ontology and annotated human orthologs, a high number of immune- related genes in the muscle of wild fish was observed. To better understand the role of these genes in the muscle we manually inspected tissue-specific patterns in human tissues. The analysis was performed using the GeneAtlas U133A gcrma data- set [63] and a plot for tissue-specific gene expression
was created using the TissueEnrich tool [64].

Genetic diversity using SNPs from the RRHP sequencing

The RRHP method allows extraction of SNP information across the genome using the same sequencing data [52]. After quality control and adapter trimming, the resulting fastq files were sorted in directories as recommended by the CFSAN SNP Pipeline [65]. The output SNP distance matrix was used to plot distance- based radars among samples. The resulting vcf

files were then used as input for the R package PopGenome with default parameters to calcu- late the fixation index (Fst) as well as the nucleotide diversity within and between the two groups. A principal component analysis (PCA) was performed using the R package ada- genet and plotted using the R package factoextra.

Acknowledgments

The LC-MS/MS analysis was performed by the Proteomics and Modomics Experimental Core facility (PROMEC), Norwegian University of Science and Technology (NTNU). PROMEC is funded by the Faculty of Medicine and Health Sciences at NTNU and the Central Norway Regional Health Authority.

We are thankful to Hilde Ribe, Øivind Torslett, Steinar Johnsen and Kaspar Klaudiussen for their assistance and unpar- alleled commitment to the husbandry and welfare of the fish.

Authors' contributions

IK carried out the sampling, wet lab and bioinformatic ana- lysis, interpreted the results and wrote the article. PS con- tributed significantly to the bioinformatic analysis of the RRHP data set, the interpretation of the results and revised the article. RM contributed significantly to the bioinformatic analysis of the RNA-Seq data set and revised the article. AVN contributed significantly in the wet lab and revised the article. DR contributed significantly to the bioinformatic analysis of the SNP data set. JMOF conceived the study, carried out the sampling, provided reagents and consumables, and contrib- uted significantly to the interpretation of the results and the article revision.

Code availability

The scripts used for transcriptomic and hydroxymethylation analyses are available from the corresponding author upon request.

Disclosure statement

The authors declare that they have no competing interests.

Funding

This work was funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreements nos 683210 and 812986) and by the Research Council of Norway under the Toppforsk programme (grant agreement no 250548/F20).

ORCID

Ioannis Konstantinidis http://orcid.org/0000-0003-2560- 5447
Artem V. Nedoluzhko http://orcid.org/0000-0001-7040- 0892
Diego Robledo http://orcid.org/0000-0002-9616-5912
Jorge M. O. Fernandes http://orcid.org/0000-0002-2550- 1640

Data availability statement

The RRHP dataset generated and analysed during the current study is available in the SRA (NCBI) repository, under the accession number PRJNA559277. The RNA-seq dataset gen- erated and analysed during the current study is available in the GEO (NCBI) repository, under the accession number GSE135811.

References

[1] Diamond J. Evolution, consequences and future of plant and animal domestication. Nature. 2002;418:700–707.

[2] Larson G, Fuller DQ. The evolution of animal domestication. Annu Rev Ecol Evol Syst. 2014;45:115–136.

[3] Makinen H, Vasemagi A, McGinnity P, et al. Population genomic analyses of early-phase Atlantic Salmon (Salmo salar) domestication/captive breeding. Evol Appl. 2015;8:93–107.

[4] Lopez ME, Neira R, Yanez JM. Applications in the search for genomic selection signatures in fish. Front Genet. 2014;5:458.

[5] Roberge C, Einum S, Guderley H, et al. Rapid parallel evolutionary changes of gene transcription profiles in farmed Atlantic salmon. Mol Ecol. 2006;15:9–20.

[6] Sauvage C, Derome N, Normandeau E, et al. Fast transcriptional responses to domestication in the brook charr *Salvelinus fontinalis*. Genetics. 2010;185:105–112.

[7] Gjerdem T, Robinson N. Advances by selective breeding for aquatic species: A review. Agri Sci. 2014;5:1152–1158.

[8] Thodesen J, Grisdale-Helland B, Helland SJ, et al. Feed intake, growth and feed utilization of offspring from wild and selected Atlantic salmon (*Salmo salar*). Aquaculture. 1999;180:237–246.

[9] Bentsen HB, Gjerde B, Eknath AE, et al. Genetic improvement of farmed tilapias: response to five gen- erations of selection for increased body weight at har- vest in *Oreochromis niloticus* and the further impact of the project. Aquaculture. 2017;468:206–217.

[10] Tahiliani M, Koh KP, Shen Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science. 2009;324:930–935.

[11] Greco CM, Kunderfranco P, Rubino M, et al. DNA hydroxymethylation controls cardiomyocyte gene expression in development and hypertrophy. Nat Commun. 2016;7:12418.

[12] Sun F, Abreu-Rodriguez I, Ye S, et al. TET1 is an important transcriptional activator of TNFalpha expression in macrophages. PLoS One. 2019;14:e0218551.

[13] Pan Z, Zhang M, Ma T, et al. Hydroxymethylation of microRNA-365-3p regulates nociceptive behaviors via Kcnh2. J Neurosci. 2016;36:2769–2781.

[14] Wu H, D'Alessio AC, Ito S, et al. Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. Genes Dev. 2011;25:679–684.

[15] Song CX, Szulwach KE, Fu Y, et al. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. Nat Biotechnol. 2011;29:68–72.

[16] Runemark A, Brydegaard M, Svensson EI. Does relaxed predation drive phenotypic divergence among insular populations? J Evol Biol. 2014;27:1676–1690.

DOI:10.1111/jeb.12421

[17] Vasemägi A, Nilsson J, McGinnity P, et al. Screen for footprints of selection during domesticatin/captive breeding of Atlantic salmon. Comp Funct Genom. 2012;2012:628204.

[18] Christie MR, Marine ML, French RA, et al. Genetic adaptation to captivity can occur in a single generation. Proc Natl Acad Sci U S A. 2012;109:238–242.

[19] Araki H, Cooper B, Blouin MS. Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. Science. 2007;318:100–103.

[20] Christie MR, Marine ML, Fox SE, et al. A single generation of domestication heritably alters the expression of hundreds of genes. Nat Commun. 2016;7:10676.

[21] Devlin RH, Sakhrani D, White S, et al. Effects of domestication and growth hormone transgenesis on mRNA profiles in rainbow trout (Oncorhynchus mykiss). J Anim Sci. 2013;91:5247–5258.

[22] Devlin RH, Sakhrani D, Tymchuk WE, et al. Domestication and growth hormone transgenesis cause similar changes in gene expression in coho sal- mon (Oncorhynchus kisutch). Proc Natl Acad Sci U S A. 2009;106:3047–3052.

[23] Haase D, Rieger JK, Witten A, et al. Specific gene expression responses to parasite genotypes reveal redundancy of innate immunity in vertebrates. PLoS One. 2014;9:e108001.

[24] Lio CJ, Rao A. TET Enzymes and 5hmC in adaptive and innate immune systems. Front Immunol. 2019;10:210.

[25] Tsumagari K, Baribault C, Terragni J, et al. Early de novo DNA methylation and prolonged demethylation in the muscle lineage. Epigenetics. 2013;8:317–332.

[26] Uren Webster TM, Rodriguez-Barreto D, Martin SAM, et al. Contrasting effects of acute and chronic stress on the transcriptome, epigenome, and immune response of Atlantic salmon. Epigenetics. 2018;13:1191–1207.

[27] Sveen LR, Timmerhaus G, Krasnov A, et al. High fish density delays wound healing in Atlantic salmon (Salmo salar). Sci Rep. 2018;8:16907.

[28] Tymchuk WE, Beckman B, Devlin RH. Altered expres- sion of growth hormone/insulin-like growth factor I axis hormones in domesticated fish. Endocrinology. 2009;150:1809–1816.

[29] Hassin S, Monbrison D, Hanin Y, et al. Domestication of the white grouper, *Epinephelus aeneus* 1. Growth and reproduction. Aquaculture. 1997;156:305–316.

[30] Johnsson JI, Petersson E, Jönsson E, et al. Domestication and growth hormone alter antipredator behaviour and growth patterns in juvenile brown trout, *Salmo trutta*. Can J Fish Aquat Sci. 1996;53:1546–1554.

[31] Fleming IA, Agustsson T, Finstad B, et al. Effects of domestication on growth physiology and endocrinol- ogy of Atlantic salmon (*Salmo salar*). Can J Fish Aquat Sci. 2002;59:1323–1330.

[32] Robison D, Rowland W. A potential model system for studying the genetics of domestication: behavioral var- iation among wild and domesticated strains of zebra danio (*Danio rerio*). Can J Fish Aquat Sci. 2005;62:2046–2054.

[33] Ponzoni RW, Nguyen NH, Khaw HL, et al. Genetic improvement of Nile tilapia (*Oreochromis niloticus*) with special reference to the work conducted by the WorldFish Center with the GIFT strain. Rev Aquacult. 2011;3:27–41.

[34] Tidball JG. Regulation of muscle growth and regenera- tion by the immune system. Nat Rev Immunol. 2017;17:165–178.

[35] Nascimento AF, Trindade DM, Tonoli CC, et al. Structural insights into functional overlapping and dif- ferentiation among myosin V motors. J Biol Chem. 2013;288:34131–34145.

[36] Sun Y, Bilan PJ, Liu Z, et al. Rab8A and Rab13 are activated by insulin and regulate GLUT4 translocation in muscle cells. Proc Natl Acad Sci U S A. 2010;107:19909–19914.

[37] Ishikura S, Klip A. Muscle cells engage Rab8A and myo- sin Vb in insulin-dependent GLUT4 translocation. Am J Physiol Cell Physiol. 2008;295:C1016–25.

[38] Hammer JA 3rd, Wagner W. Functions of class V myosins in neurons. J Biol Chem. 2013;288:28428–28434.

[39] Altarejos JY, Montminy M. CREB and the CRTC co-activators: sensors for hormonal and metabolic signals. Nat Rev Mol Cell Biol. 2011;12:141–151.

[40] Xu Y, Yu W, Xiong Y, et al. Molecular characterization and expression patterns of serine/arginine-rich specific kinase 3 (SPRK3) in porcine skeletal muscle. Mol Biol Rep. 2011;38:2903–2909.

[41] Adam AC, Lie KK, Whatmore P, et al. Profiling DNA methylation patterns of zebrafish liver associated with parental high dietary arachidonic acid. PLoS One. 2019;14:e0220934.

[42] Skjaerven KH, Jakt LM, Fernandes JMO, et al. Parental micronutrient deficiency distorts liver DNA methyla- tion and expression of lipid genes associated with a fatty-liver-like phenotype in offspring. Sci Rep. 2018;8:3055.

[43] Shock LS, Thakkar PV, Peterson EJ, et al. DNA methyl- transferase 1, cytosine methylation, and cytosine hydroxymethylation in mammalian mitochondria. Proc Natl Acad Sci U S A. 2011;108:3630–3635.

[44] Bellizzi D, D'Aquila P, Scafone T, et al. The control region of mitochondrial DNA shows an unusual CpG and non-CpG methylation pattern. DNA Res. 2013;20:537–547.

[45] Potok ME, Nix DA, Parnell TJ, et al. Reprogramming the maternal zebrafish genome after fertilization to match the paternal methylation pattern. Cell. 2013;153:759–772.

[46] Le Luyer J, Laporte M, Beacham TD, et al. Parallel epigenetic modifications induced by hatchery rearing in a Pacific salmon. Proc Natl Acad Sci U S A. 2017;114:12964–12969.

[47] Anastasiadi D, Esteve-Codina A, Piferrer F. Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species. Epigenetics Chromatin. 2018;11:37.

[48] Jiang L, Zhang J, Wang JJ, et al. Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. Cell. 2013;153:773–784.

[49] Simões LN, Lombardi DC, Gomide ATM, et al. Efficacy of clove oil as anesthetic in handling and transportation of Nile tilapia, *Oreochromis niloticus* (Actinopterygii: cichli- dae) juveniles. Zoologia. 2011;28:285–290.

[50] Illumina. Effects of index misassignment on multiplex- ing and downstream analysis. 2017.

[51] Sun X, Chung TH, Tan D, et al. Practical guidelines and consideration of using RRHP for 5hmC detectionPractical guidelines and consideration of using RRHP for 5hmC detection. Epigenomics. 2016;8:225–235.

[52] Petterson A, Chung TH, Tan D, et al. RRHP: a tag-based approach for 5-hydroxymethylcytosine mapping at single-site resolution. Genome Biol. 2014;15:456.

[53] Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal. 2011;17:10–12.

[54] Conte MA, Gammerdinger WJ, Bartie KL, et al. A high quality assembly of the Nile Tilapia (Oreochromis nilo- ticus) genome reveals the structure of two sex determi- nation regions.

BMC Genomics. 2017;18:341.

[55] Langmead B. Aligning short sequencing reads with Bowtie. Curr Protoc Bioinformatics. 2010;32:11.7.1– 11.7.14.

[56] Team RC. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria; 2017.

[57] Heinz S, Benner C, Spann N, et al. Simple combina- tions of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell. 2010;38:576–589.

[58] Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 2019;10:1523.

[59] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–2120.

[60] Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21.

[61] Anders S, Pyl PT, Huber W. HTSeq–a Python frame- work to work with high-throughput sequencing data. Bioinformatics. 2015;31:166–169.

[62] Benjamini Y, Hochberg Y. Controlling the false dis- covery rate: A practical and powerful approach to multiple testing. J R Stat Soc Series B. 1995;57:289–300.

[63] Su A Dataset: GeneAtlas U133A, gcrma. 2019.

[64] Jain A, Tuteja G. TissueEnrich: tissue-specific gene enrichment analysis. Bioinformatics. 2019;1966–1967. DOI:10.1093/bioinformatics/bty890

[65] Davis S, Pettengill JB, Luo Y, et al. CFSAN SNP Pipeline: an automated method for constructing SNP matrices from next-generation sequence data. PeerJ Comput Sci. 2015;1:e20.

Paper III

Epigenetic mapping of the somatotropic axis in Nile tilapia reveals differential DNA hydroxymethylation marks involved in growth

Ioannis Konstantinidis[1], Pål Sætrom[1,2,3,4,5], Artem V. Nedoluzhko[1], Dafni Anastasiadi[6,7], Robin Mjelle[1,2], Tomasz Podgorniak[1], Francesc Piferrer[7] and Jorge M. O. Fernandes[1]*

[1]Faculty of Biosciences and Aquaculture, Nord University, Bodø, Norway

[2]Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway

[3]Department of Computer Science, Norwegian University of Science and Technology, Trondheim, Norway.

[4]Bioinformatics core facility-BioCore, Norwegian University of Science and Technology, Trondheim, Norway

[5]K.G. Jebsen Center for Genetic Epidemiology, Norwegian University of Science and Technology, Trondheim, Norway

[6]The New Zealand Institute for Plant and Food Research, Nelson, New Zealand

[7]Institut de Ciències del Mar, Spanish National Research Council (CSIC), Barcelona, Spain

*Corresponding author: Jorge M. O. Fernandes

Abstract

The somatotropic axis, involving growth hormone production in the pituitary gland and the subsequent responses in liver and muscles, plays a major role in the regulation of metabolic and physiological processes. Its output in terms of growth rate is likely to be epigenetically tuned, since it is highly affected by nutritional and environmental cues. Epigenetic mechanisms, including DNA hydroxymethylation, can regulate transcription of key genes involved in myogenesis. In spite of its importance, a holistic characterization of the hydroxymethylome in the main tissues of the somatotropic axis has not been investigated to date. Here, we used reduced representation 5-hydroxymethylcytosine (5hmC) profiling (RRHP) to map 5hmC at single nucleotide resolution in Nile tilapia tissues that comprise the somatotropic axis, namely, muscle, liver and pituitary.

We show that 5hmC has tissue-specific profiles and differs in abundance between muscle, liver and pituitary. Moreover, 5hmC is highly enriched within gene bodies and present in several genes that regulate growth, such as insulin-like growth factor receptors and binding proteins, fibroblast growth factors, myogenic factors, growth inhibitors and myocyte enhancers. Integration of the fast muscle transcriptome with the corresponding RRHP data set revealed that 5hmC marks on promoters, first introns, and first exons have an overall negative correlation with gene expression. Among the most expressed genes in muscle, the closest 5hmC to the transcription start site was located within the first intron of the troponin T fast skeletal muscle gene and it was hyper-hydroxymethylated in muscle compared to liver, displaying a positive correlation between 5hmC levels and gene expression.

These findings suggest that 5hmCs have a tissue-specific localization within the genome and likely act as a balancing factor in gene activation or repression. Their abundance within gene bodies and particularly introns of several growth-related genes indicates that 5hmC likely modulates myogenesis in Nile tilapia. We propose that cytosine hydroxymethylation may contribute to the phenotypic plasticity of growth through epigenetic regulation of the somatotropic axis.

Keywords: Somatotropic axis, Muscle, Liver, Pituitary, Growth, Epigenetics, DNA hydroxymethylation, RRHP, Teleosts

Introduction

From the production of growth hormone in the pituitary gland to the stimulation and regulation of hundreds of genes in the liver and muscle, the somatotropic axis plays a critical role across several biological pathways involved in somatic growth, carbohydrate and lipid metabolism, energy equilibrium, normal development, reproduction and osmoregulation in fish (Company et al., 2001; Manzon, 2002; Whittington and Wilson, 2013). The interaction of molecules such as the growth hormone, insulin-like growth factors, binding proteins and receptors (GH, GHRs, IGFs, IGFBPs, IGFRs) are greatly affected by both endogenous and exogenous factors, i.e. sex, developmental stage, nutrition, genetic background, temperature, photoperiod, salinity, dissolved oxygen, immunocompetence and stress (Beckman, 2011; Campos et al., 2010; Campos et al., 2013c; Katan et al., 2016). For example, previous research has shown that seasonal differences (Metcalfe et al., 2002), photoperiod, and sexual maturity (Taylor et al., 2008) affect insulin-like growth factor 1 (*igf1*) plasma levels and consequently somatic growth. Similarly to mammals, the liver is the main source of *igf1* production in fish; however, in cases of malnutrition the liver is predominantly unresponsive to GH stimulation (Gray et al., 1992; Perez-Sanchez et al., 1995). Such interactions have been widely demonstrated and both endocrine and paracrine *igf1* have been linked to bone, cartilage and muscle growth in fish (Nordgarden et al., 2006; Valente et al., 2013). Myogenesis, one of the critical processes that determines growth in fish, requires a well-tuned transcriptional synchronization of three main tissue types that compose the somatotropic axis: the pituitary gland, liver and skeletal muscle.

One of the most important hormones secreted by the pituitary gland is the growth hormone (GH). The insulin-like growth factors (IGF-I and -II), their receptors (IGFRS) and binding proteins (IGFBPS) that are present in substantial amounts in liver and fast muscle, are directly associated with GH and collectively regulate somatic growth. Because growth rate

varies considerably with environmental factors (Perez-Sanchez et al., 2018), epigenetic mechanisms have the potential to be major regulators of growth both in early development (e.g., thermal imprinting) (Burgerhout et al., 2017; Campos et al., 2014a; Campos et al., 2013b; Campos et al., 2014b), and adulthood. Indeed, intrauterine growth restriction in rats, induces DNA methylation changes around growth hormone response elements of the IGF-1 gene, resulting in developmental re-programming and persisting long-term effects (Fu et al., 2015). Additionally, epigenetic mechanisms are likely responsible for splicing events leading to the translation of tissue-specific IGF-1 peptides through the usage of alternative promoters and exons (Beresewicz et al., 2010; Schober et al., 2012). We have recently reported that growth-related methylation differences in *igf2bp2* are likely linked to phenotypic differences between small and large Nile tilapia full-sibs (Podgorniak et al., 2019). In teleosts, recent studies have shown that epigenetic modifications regulate the expression of important growth genes during temperature fluctuations such as the DNA methylation of *myogenin* in Atlantic salmon (*Salmo salar*) (Burgerhout et al., 2017) and Senegalese sole (*Solea senegalensis*) (Campos et al., 2013a; Campos et al., 2013b). In our previous study, we have also identified sex-specific DNA methylation marks within the *map3k5*, *akt3*, *gadd45g* and *ppargc1a* genes. Their epigenetic regulation could potentially explain size-related differences among sexes because of their involvement in processes such as cell cycle, proliferation, cellular response to glucose and muscle tissue development (Podgorniak et al., 2019). The epigenetic function of non-coding RNAs, such as microRNAs (miRNAs), has also been shown to play critical roles in somatic growth (Campos et al., 2014a). For example, miR-206 targets the 3'UTR of IGF-1 and directly affects its expression in Nile tilapia (Yan et al., 2013), while miR-1 and miR-133 regulate the sarcomeric actin organization in zebrafish (Mishima et al., 2009).

DNA hydroxymethylation is an epigenetic modification that has received very little attention in relation to growth, with a few exceptions associating it to abnormal growths, cancers and oncogenic pathways such as the PI3K-Akt (Bhattacharyya et al., 2017; Wu et al., 2019). It occurs via ten-eleven-translocation enzymes, which catalyse the oxidation of 5-methyl ($-CH_3$) into 5-hydroxymethyl ($-CH_2OH$) groups in cytosines (5hmCs) (Tahiliani et al., 2009). Additionally, 5hmC has been identified as a stable DNA modification with a pertinent role in

gene regulation and transcription (Bachman et al., 2014; Branco et al., 2011), tissue-specific patterns and developmental stage-dependent profiles (Ponnaluri et al., 2017). In human, 5hmCs are not exclusively correlated with DNA methylcytosine (5mC) content and when found in transcribed gene bodies they have been positively correlated with transcription (Nestor et al., 2012). In spite of 5hmC's significance in regulating gene expression in humans (Nestor et al., 2012; Ponnaluri et al., 2017; Rohde et al., 2015) and mice (Greco et al., 2016; Ma et al., 2018), little is known about its role on growth, particularly in fish.

To investigate whether 5hmC marks were associated with key genes within the somatotropic axis, we compared genome-wide and tissue-specific 5hmC profiles at single nucleotide resolution between liver, muscle and pituitary using Nile tilapia (*Oreochromis niloticus*) as our model organism.

Results

Tissue-specific 5hmC levels across the somatotropic axis

Among the three main tissues that comprise the somatotropic axis, 5hmC levels were highest in liver, followed by muscle and pituitary. The overall 5hmC levels of each tissue follow the sequencing depth of the corresponding reduced representation 5hmC profiling (RRHP) libraries (see Methods). Liver libraries had an average depth of 28.1 M raw reads compared to 22.6 M and 21.7 M raw reads in muscle and pituitary libraries, respectively. After quality control, adapter trimming and alignment of the sequenced reads to the Nile tilapia genome, an average of 14.9, 11.5 and 10.6 M reads were obtained from the libraries prepared from liver, muscle and pituitary, respectively (Additional file 1).

The Nile tilapia genome contains 1,613,446 CCGG sites in both strands across its 22 linkage groups and the mitochondrial genome. RRHP captured 75% (1,211,876 CCGG) of all possible sites. The reduced coverage was mostly attributed to the size selection during library preparation. Compared to the total number of genomic CCGGs, 15.7% was significantly hydroxymethylated based on the median filtering of 5hmC counts (see Methods). The rest 59.3% was dominated by 5hmC sites with very low counts across all tissues. A total of 16,812

5hmCs, which was approximately 1% of the total genomic CCGGs, were identified as differentially hydroxymethylated among the three tissues (Figure 1A).

The highest difference was detected between the muscle and liver (MvL), with 15,757 cytosines being differentially hydroxymethylated, out of which 3,119 were hyper-hydroxymethylated in muscle ($q<0.01$). By comparing the hydroxymethylation profiles between muscle-pituitary (MvP) and liver-pituitary (LvP) we identified 1,728 ($q<0.05$) and 1,467 ($q<0.01$) additional DhmCs, respectively. An overlap of almost 2,000 DhmCs was observed between pairs (MvL-LvP, MvL-MvP and LvP-MvP); however, we identified only four cytosines with significantly different hydroxymethylation levels among all pairwise comparisons (Figure 1B; Additional file 2). Two of these four DhmCs were found within predicted long non-coding RNAs (*LOC102076435* and *LOC112847764*), one DhmC located within the predicted gene gap junction delta-3 protein-like (*LOC100690349*) and one DhmC within the par-3 family cell polarity regulator alpha, b (*pard3ab*).

Tissue-specific DNA hydroxymethylation abundance within genomic features

While DNA hydroxymethylation levels were substantially different among tissues, we also identified differences in DhmC enrichment within the annotated features of the Nile tilapia genome. Compared to liver, the muscle was characterized by an overall hypo-hydroxymethylation across all the annotated features, except the promoter regions (Figure 2A). Striking differences were also observed between pituitary and muscle, with exons being mostly hypo-hydroxymethylated in muscle. (Figure 2B). By comparing liver to pituitary, we observed that the liver genome was hyper-hydroxymethylated across all the annotated features (Figure 2C). The relative enrichment of DhmCs was calculated based on the number of DhmCs per annotated feature compared to the total number of DhmCs in each tissue. Compared to liver and pituitary, DhmCs in muscle were relatively enriched in introns (12%) and promoters (104%), whereas exons were depleted by 58 % compared to liver and pituitary. In contrast, hyper-5hmCs in the liver were relatively depleted within promoters and enriched within exons, compared to muscle and pituitary (Additional file 3). The highest percentage of DhmCs among all tissues was detected within introns (40% on average), while

DhmCs in intragenic regions including promoters represented 70%, 74.6% and 71.8% of the entire dataset in muscle, liver, and pituitary, respectively.

Differentially hydroxymethylated cytosines are found in genes with distinct molecular functions in each tissue

To better understand whether DhmCs were linked to tissue-specific genes and potentially their functions, we performed a gene ontology enrichment analysis based on the genes containing or being closely associated to DhmCs. In total, we identified 2102 hyper-hydroxymethylated genes in muscle compared to liver and pituitary, 4993 genes in liver compared to muscle and pituitary and 1006 genes in the pituitary compared to muscle and liver. Interestingly, we identified 773 genes containing hyper-hydroxymethylcytosines in both the muscle and liver, 40 common genes between muscle and pituitary and 23 genes shared between liver and pituitary (Figure 1C).

In muscle compared to liver, genes containing hypo-5hmCs (lower 5hmC levels in muscle) were associated with molecular functions related to transcription regulator and GTPase activator activity, GTPase binding and Ras GTPase binding. Gene ontologies related to development as well as anatomical structure and system development were among the most enriched biological processes. Furthermore, we identified 12 biological processes linked to metabolism, including regulation of primary, cellular, macromolecule, RNA and nitrogen compound metabolic process. Interestingly, hypo-5hmCs were assigned to gene ontologies such as skeletal system, muscle structure and cartilage development (Additional file 4). On the other hand, hyper-5hmCs (higher 5hmC levels in the muscle) were primarily associated to ionotropic glutamate receptor activity, signaling receptor activity and cell surface receptor signaling pathway, as well as system development and developmental process (Additional file 5). Notably, 773 genes were identified containing both hyper- and hypo-5hmCs that occurred in different intragenic loci. These genes were primarily linked to the regulation of biological and cellular process, as well as the signal transduction and response to stimulus (Additional file 6).

In muscle compared to pituitary, hyper-hydroxymethylated genes (higher 5hmC levels in the muscle) were associated with tissue morphogenesis and the morphogenesis of an epithelium, as well as the developmental process and signaling among other gene ontologies (Additional file 7). Hypo-hydroxymethylated genes (lower 5hmC levels in muscle) were enriched for functions related to transcription regulator activity and the regulation of metabolic and biosynthetic process of macromolecules and RNA (Additional file 8). We identified 40 shared genes between muscle and pituitary containing hyper- and hypo-5hmCs in different loci. However, due to the relatively low number of genes we did not identify any significant enrichment for molecular functions or biological processes.

Compared to pituitary, the liver contained hypo-hydroxymethylated genes (lower 5hmC levels in liver) that were enriched for a cell surface receptor signaling pathway and axon target recognition (Additional file 9). In contrast, hyper-hydroxymethylated genes (higher 5hmC levels in the liver) were enriched for molecular functions related to protein and DNA binding. The most enriched biological processes were the regulation of cellular and biological processes, signaling and signal transduction (Additional file 10). In total, 23 genes containing hyper- and hypo-5hmCs were found within the LvP pairwise comparison. Among them, we identified the gene coding for arginine-glutamic acid dipeptide repeats protein (*rere*), which is involved in cell survival and apoptosis (Waerner et al., 2001), and the protein phosphatase 1 regulatory subunit 3C (*ppp1r3c*) involved in glycogen metabolism as well as the metabolic process of macromolecules (Doherty et al., 1996).

Differentially hydroxymethylated genes involved in somatic growth

Genes that play a critical role in development and somatic growth were identified across all pairwise comparisons. Most genes involved in somatic growth were hypo-hydroxymethylated in muscle compared to liver. These included several fibroblast growth factors (*fgf5*, *fgf6*, *fgf8*, *fgf12*, *fgf14* and *fgf22*), the receptor *fgfr3* and the fibroblast growth factor receptor substrate 2 (*frs2*). Only *fgf10* and *fgfr4* were hyper-hydroxymethylated in muscle. Additionally, growth arrest-specific proteins (*gas2l3* and *gas7*) and growth differentiation factors (*gdf3*, *gdf5*, *gdf6*¸ *gdf6a* and *gdf10*) were also hypo-hydroxymethylated in muscle. Hypo-5hmCs were detected within the last exon of growth

hormone secretagogue receptor type 1 (*ghsr*), the second intron of insulin-like growth factor 2 mRNA binding protein 2 (*igf2bp2*), the first and third introns of insulin-like growth factor binding protein 5 (*igfbp5*) and the 32nd intron of IGF like family receptor 1 (*igflr1*). On the other hand, the promoter of the growth hormone releasing hormone (*ghrh*), the first intron of the insulin-like growth factor binding protein 2-B and the 29th exon of *igflr1* contained a single hyper-5hmC. Surprisingly, several genes involved in somatic growth contained both hyper- and hypo-5hmCs in different intragenic loci. These included the myoD family inhibitor domain-containing protein (*mdfic*), the multiple epidermal growth factor-like domains protein 10 (*megf10*) and the myosin heavy chain fast skeletal muscle (*myh*) (Figure 3A; Additional file 11A). In muscle compared to pituitary, we identified two hyper-5hmCs located 7549 and 7367 bp, from the transcription start site (TSS) of the myostatin (*mstn*) gene, and a single hyper-5hmC at 6945 bp from the TSS of the growth arrest and DNA damage inducible gamma (*gadd45g*) gene. Similar to the comparison between muscle and liver, fibroblast growth factor genes *fgf5*, *fgf12* and *frs2* were hypo-hydroxymethylated in muscle compared to pituitary; however, hyper-5hmCs were detected within the *fgfr2* and *fgfbp3* genes. Additionally, two hyper-5hmCs were identified within the 5th exon and 7th intron of the gene coding for myoD family inhibitor domain-containing protein 2 (*mdfic2*) (Figure 3B; Additional file 11B). Several of the above-mentioned genes were also differentially hydroxymethylated in liver compared to pituitary (Figure 3C; Additional file 11C).

Gene expression profiling in Nile tilapia fast muscle

The average sequencing depth of the RNA-Seq libraries was 26 M reads (Additional file 12). Approximately 1 M reads were excluded from the analysis after quality control (Q >20) and trimming. On average, 2.8 M reads were mapped multiple times and 19.2 M reads were mapped uniquely to the reference Nile tilapia genome. In total, we identified 80,417 transcripts that were ranked based on their normalized counts (Additional file 13). The gene with the highest expression in Nile tilapia fast muscle was the actin alpha skeletal muscle A (*LOC100534413*) which is involved in several biological processes including muscle contraction, myosin binding, skeletal muscle fibre development and thin filament assembly.

Among the genes with the highest expression, we identified several fast skeletal muscle myosin light and heavy chains (*mylpf*, *LOC100712344*, *LOC100707599*, *LOC100698429*), creatine kinases muscle a and b (*ckma* and *ckmb*), ATPase sarcoplasmic/endoplasmic reticulum $Ca2^{+}$ transporting 1 like (*atp2a1l*) and several isoforms of troponin T (*LOC100707421*) and tropomyosin 1 (*tpm1*) (Additional file 14).

Association of muscle-specific gene expression and DNA hydroxymethylation

To investigate the links between DNA hydroxymethylation and gene expression levels in muscle, we performed a principal component analysis using three variables: the 5hmC levels (mean of multiple values per gene), the distance to the transcription start site (TSS) (mean of multiple distances per gene) and gene expression (single value). The first principal component (PC1) was strongly related to variation in both the distance to TSS (ρ=0.63) and gene expression (GE; ρ=0.70), indicating that these variables were strongly positively correlated in our data. On the other hand, 5hmC levels were strongly represented in the second principal component (PC2; ρ=0.84) revealing no correlation with GE (ρ=0.10; Figure 4).

Since the first kilobases up- and downstream of promoter regions have been previously correlated with gene transcription and alternative splicing (Anastasiadi et al., 2018; Bieberstein et al., 2012; Rose, 2018), we also investigated the correlation between gene expression and 5hmC levels in promoters, first introns and first exons. Consistent with our findings above, we identified a negative correlation between 5hmC levels and gene expression in all three comparisons (Figure 5). Among the most expressed genes in fast muscle, we identified 4 differentially hydroxymethylated genes between muscle and liver. Two DhmCs were located 19,650 and 19,551 bp upstream of the myosin heavy chain fast skeletal muscle (*LOC100706261*) promoter, two DhmCs were located 10,176 and 10,040 bp upstream of the nucleoside diphosphate kinase B (*LOC100697135*) promoter, two DhmCs were found within the 16th exon of *atp2a1l*, one DhmC within the 5th exon of myosin light chain 3 skeletal muscle isoform (*LOC100698429*) and one DhmC within the 1st intron of

troponin T fast skeletal muscle isoforms (*LOC100707421*). Surprisingly, all 5hmCs were hypo-hydroxymethylated in muscle compared to liver, except the one that was located closest to the TSS (1318 bp) and within the 1st intron of troponin T fast skeletal muscle isoform (Table 1).

Discussion

This study presents the first DNA hydroxymethylation profiling of the somatotropic axis at single nucleotide resolution in a teleost species. We compared the 5hmC levels among tissues and identified the liver as the tissue with the highest, the muscle with intermediate and the pituitary with the lowest 5hmC levels (Figure 6).

In contrast to these results, global 5hmC levels of the muscle and liver measured by liquid chromatography mass spectrometry (LC/MS) in male zebrafish (*Danio rerio*) revealed that muscle had higher 5hmC levels compared to liver (Kamstra et al., 2015). Similarly, 5hmC quantification by LC/MS in muscle and liver of mice, showed higher 5hmC levels in muscle compared to liver (Globisch et al., 2010). These results can potentially be explained either by species-specific 5hmC patterns or by the methods used. For example, reduced representation of the genome using RRHP provides 5hmC mapping in CCGG context and single nucleotide resolution, however, global 5hmC quantification using LC/MS provides a single value indicative of all CpG and non-CpG sites. The low 5hmC content in pituitary gland in Nile tilapia is in accordance with previous findings in mice (Globisch et al., 2010). Indeed, reports have demonstrated that high 5hmC levels in brain are associated with tissue-specific neuronal functions (Kinde et al., 2015; Santiago et al., 2014), compared to pituitary which is responsible for hormone secretion.

Based on the functional enrichment analysis of hyper-hydroxymethylated genes in muscle compared to liver, we identified several genes involved in ionotropic and metabotropic glutamate receptor activity. These genes were associated with all three types of ionotropic glutamate receptors, namely, the N-methyl-D-aspartate (NMDA), the kainate and the a-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptors while their functions were

mostly associated with fast synaptic transmission and the modulation of synaptic responses (Colombo and Francolini, 2019). A previous study using C2C12 mouse skeletal myoblasts has shown that NMDA receptors play a major role in calcium influx and myoblast fusion, which are fundamental processes for muscle growth (Lee et al., 2004). Furthermore, we also identified several differentially hydroxymethylated metabotropic glutamate receptors (*grm1*, *grm2*, *grm4*, *grm7* and *grm8*) between muscle and liver, which play critical roles in cell survival, metabolism and proliferation through their association with the PI3 kinase and the initiation of the PI3K-Akt-mTOR signaling pathway (Rong et al., 2003). Glutamate receptors were found to be under positive selection across several domesticates (O'Rourke and Boeckx, 2020), including fish (Bertolini et al., 2016). Particularly in sea bass (*Dicentrarchus labrax*), the identification of genetic mutations after many generations in captivity was found to be overlapping with early changes in DNA methylation profiles during domestication, suggesting that epimutations potentially drive genome evolution (Anastasiadi and Piferrer, 2019). Here, we identified significantly hydroxymethylated ionotropic and metabotropic glutamate receptors within the somatotropic axis, which supports the functional importance of these genes in somatic growth.

Finally, we identified several DhmCs within the receptor-type tyrosine-protein phosphatase N2 (*ptprn2*) gene, which is expressed in nervous and endocrine cells and it is associated with lipid metabolism and insulin secretion (Bence, 2013; Cui et al., 1996). Previous studies have shown that *ptprn2* is largely associated with somatic growth, and regulated by epigenetic modifications in various biological processes including hepatocellular carcinomas (Gentilini et al., 2017; Shen et al., 2015), breast cancer (Sengelaub et al., 2016) as well as prenatal growth patterns and birthweight in humans (Chen et al., 2018). Therefore, it is likely that DNA hydroxymethylation may modulate somatic growth indirectly, through the regulation of genes related to neuromuscular and neurohepatic signals.

Growth hormone release depends on three hypothalamic factors, namely, growth hormone releasing hormone (GHRH), somatostatin (SST) and ghrelin multifaceted hormone (GHRL) (Howard et al., 1996; Yin et al., 2014). The latter is stimulated by receptors, such as the growth hormone secretagogue receptor (*ghsr*), which was found to be hyper-hydroxymethylated in the liver, compared to the muscle and pituitary. *Ghsr* encodes for a

tissue-specific hormone receptor that is commonly found across the somatotropic axis and regulates energy homeostasis, fat distribution and the expression of genes involved in lipid metabolism and total body weight (Barazzoni et al., 2005; Pfluger et al., 2008). Notably, among the differentially hydroxymethylated genes, we also identified a hyper-5hmC (higher 5hmC level in muscle compared to liver) within the promoter of the growth hormone *ghrh* as well as several hyper-5hmCs (higher 5hmC levels in muscle and pituitary compared to liver) within predicted genes coding for somatostatin in Nile tilapia (*loc102077970* – somatostatin-1; *loc100698045* – somatostatin-1B; *loc100694069* – somatostatin-2), known also as growth hormone-inhibiting hormone. Additionally, we identified differentially hydroxymethylated genes that play critical roles in growth hormone signaling and are downstream regulators of somatic growth, such as the insulin-like growth factor binding proteins *igfbp2*, *igf2bp2* and *igfbp5* and the receptor *igflr1*. In particular, hyper-5hmCs (higher 5hmC levels in muscle compared to liver) were located within *igfbp2* and *igflr1* while hypo-5hmCs (lower 5hmC levels in muscle compared to liver) were located within igfbp5, *igf2bp2* and *igflr1*. The same genes were also differentially hydroxymethylated between liver and pituitary. A hyper-5hmC (higher 5hmC levels in liver compared to pituitary) was found within *igf2bp2*, and a hypo-5hmC (lower 5hmC levels in liver compared to pituitary) was located within *igflr1*. These differentially hydroxymethylated cytosines suggest a direct link between DNA hydroxymethylation and the function of the somatotropic axis (Figure 6).

Interestingly, several genes were both hyper- and hypo-hydroxymethylated between tissues due to the localization of unique DhmCs across their gene body or their proximal regions. For example, the above mentioned *ptprn2* gene was highly hydroxymethylated in liver (compared to muscle) across the 1st, 3rd and 8th introns as well as the 5th exon, but a hypo-hydroxymethylated cytosine (low 5hmC levels in liver compared to muscle) was also present in the 10th intron. Similar patterns were observed in other genes, including the multiple epidermal growth factor-like domains protein 10 (*mefg10*) and insulin receptor (*insr*), methyltransferases such as *mettl21c* and *mettl4*, adhesion G protein-coupled receptors (*adgrf4*, *adgrl2*, *adgrl3*, *adgrb1*) and several transcription factors (*atf2*, *atf3*, *ebf3*, *grhl1*, *gtf3c1*, *gtf3c3*, *hsf2*, *sox4* and *runx2*). Taking into consideration the enrichment of 5hmCs within introns and their positional dynamics between tissues, these findings could

potentially explain tissue-specific gene expression through the regulation of alternative promoters or splicing junctions. Indeed, recent studies have shown that DNA hydroxymethylation is involved in alternative splicing (Cingolani et al., 2013; Gao et al., 2019) and that positional dynamics of 5hmCs within or in proximal distance of genes result in expression changes (Greco et al., 2016).

Additionally, studies in mammals and mammalian cells have shown that DNA hydroxymethylation is positively correlated to gene expression and is present in gene bodies of actively transcribed genes (Ficz et al., 2011; Pastor et al., 2011; Szulwach et al., 2011; Wu et al., 2011). Here, even though the correlations of gene expression in the muscle and 5hmC levels were generally moderate to weak (ρ=0.47-0.63), we found that 5hmCs display a positive correlation with distance from TSSs; the farther the 5hmCs from a TSS, the higher the expression of genes. Frequently, first introns and exons are positively associated with gene regulation (Anastasiadi et al., 2018; Bieberstein et al., 2012; Rose, 2018). By comparing muscle gene expression and 5hmC levels of promoters, first introns and first exons, we determined that 5hmCs in close proximity to TSSs had an overall negative but weak correlation with gene transcription. However, a close examination of the top 30 most expressed genes in muscle and their hydroxymethylation status, revealed that the only hyper-5hmC in muscle compared to liver was also the closest one to the TSS of *troponin T fast skeletal muscle isoform*. Several hypo-5hmCs (lower 5hmC levels in muscle compared to liver) were found at a distance of 10-19 kbp upstream from the promoter as well as within the open reading frames of highly expressed genes in the muscle but none of them were located within the first intron or exon. These findings suggest that the position of 5hmCs within gene features plays indeed a crucial role in gene repression or activation. The overall negative correlation between 5hmCs and gene expression in our study is likely attributed to the presence of other epigenetic mechanisms that were not profiled here, such as cytosine methylation, histone modifications and non-coding RNAs. Although, high 5hmC levels are usually correlated with gene activation, there is also evidence of 5hmCs being involved in the repression of genes such as the insulin-degrading enzyme (IDE) in depressed individuals (Gross et al., 2017) and polycomb groups of developmental regulators (Wu et al., 2011). In human, there is a high overlap (85.6%) between 5mCs and 5hmCs (Petterson et al., 2014).

Previously, high 5mC levels within the first intron of genes have been strongly associated to low levels of expression across multiple species and tissues (Anastasiadi et al., 2018). Considering the potential overlap between these two cytosine modifications, their overall contrasting roles in gene regulation, their relative abundance (5mCs are estimated to be 14-fold more abundant than 5hmCs in mammalian embryonic stem cells (Tahiliani et al., 2009)), and the negative but weak correlation between substantially hydroxymethylated genes and their expression, we propose that the presence of 5hmCs throughout the genome plays a balancing role in gene regulation.

Conclusions

Somatic growth and normal development are regulated by a complex molecular cascade that is affected by several mechanisms, including epigenetic modifications. Here, we show that DNA hydroxymethylation is present throughout the genome, is characterized by tissue-specific profiles and is highly enriched within gene bodies (Figure 6). Several genes that are both directly or indirectly associated with functions such as growth hormone release, insulin-like growth factor binding, liver metabolism and somatic growth were found to be differentially hydroxymethylated between liver, muscle and pituitary. We identified a weak, but significant, negative correlation between gene expression and 5hmC levels in muscle, which is explained by the fact that DNA hydroxymethylation is not the only mechanism affecting gene activation or repression. However, one of the most expressed genes in the muscle (t*roponin T fast skeletal muscle*) contained a single hyper-5hmC (higher 5hmC levels in the muscle compared to liver) within the first intron, revealing a positive correlation between DNA hydroxymethylation and gene expression. The high overlap between 5hmCs and 5mCs and their overall contrasting roles in gene transcription pinpoints towards a dynamic interrelationship with DNA hydroxymethylation acting as a balancing factor in gene regulation within tissues. The identification of tissue-specific epigenetic modifications within the somatotropic axis adds an additional layer of complexity to our current knowledge about the regulation of growth.

Methods

Ethics statement

All animal handling procedures comply with the EU Directive 2010/63 and were approved by the Nord University ethics committee and the Norwegian Animal Research Authority (FOTS ID 1042).

Experimental design and sampling

Nile tilapia were reared in a recirculating aquaculture system (pH=7.6, oxygen saturation=100%, temperature=28 °C and photoperiod adjusted at 11:13 dark:light). To minimize the effect of environmental factors and sampling bias, the fish were full-sibs reared in common garden in a single tank and randomly sampled. The fish were euthanized by immersion for 3 minutes in 10 L of water containing 15 ml clove oil pre-mix, consisting of pure clove oil (Sigma Aldrich, USA) and 96% ethanol at a 1:9 ratio. Fast (white) muscle, liver and pituitary gland were carefully dissected from five (5) males, snap-frozen in liquid nitrogen and stored at -80 °C. In particular, fast muscle was extracted from a 4 mm cross-section at 0.7 standard length. The samples were taken above the lateral line from the left upper white muscle mass. Liver samples were collected from the left lobe and around the entry point of the portal vein. The pituitary gland was extracted by cutting open the skull, removing the brain from the dorsal side and picking the pituitary out with forceps. Information regarding the measurements (weight, total and standard length) of the fish can be found in Table 2.

DNA extraction and RRHP library preparation

DNA extraction was carried out using the Quick-DNA miniprep plus kit (Zymo Research, USA) according to the manufacturer's instructions. Quantification and integrity control of the extracted DNA was performed using Qubit 3.0 fluorometer and the double-stranded DNA high sensitivity assay kit (ThermoFisher Scientific, USA) and Tapestation 2200 (Agilent Technologies, USA), respectively (Additional file 15A).

Reduced representation of 5-hydroxymethylcytosine profiling (RRHP) was performed according to manufacturer's protocol (Zymo Research). The preparation of RRHP libraries was performed synchronously for all tissues and replicates. The starting DNA concentration was 150 ng diluted in 10µl DNAse and RNAse free water. Sequencing was performed using Nord University's inhouse NextSeq500 (Illumina, USA) and all libraries were pooled with equal volumes. Finally, they were distributed among three High-Output kit v2.5 150 cycle flow cells, yielding approximately 360 million single-end reads.

RRHP bioinformatic pipeline

Raw reads were trimmed for adapters using trim_galore v0.4.4 and quality check was performed using MultiQC. The software Bowtie v0.12.8 was used for the alignment of high quality and trimmed reads to the latest Nile tilapia reference genome (NCBI assembly GCA_001858045.3) (Conte et al., 2017) with the following parameters: -S -v 1 -n 1 -m 3 --strata --best. The chromosome, position and strand information from reads that start with CCGG were extracted in text files for each sample. These files were used as input into R, where a count matrix was created by summing up overlapping chromosomal positions. To minimize the false discovery of 5hmCs the data set was filtered twice, as reported (Konstantinidis et al., 2020). Briefly, the first filter removed reads for which more than two groups of samples (2 out of 3 tissues) had less than 1 count. For the second filter, the median of the entire dataset was calculated and 5hmC sites were removed when at least one tissue among the three (5 or more samples) had counts equal or less than the median (21 counts). For the comparison of 5hmC levels at single nucleotide resolution between the three tissues, we used the R package limma (Ritchie et al., 2015) and the adjusted p-values were calculated using the Benjamini-Hochberg correction. Three contrasts were performed comparing the 5hmC levels of i) the muscle against the liver (MvL), ii) the muscle against the pituitary gland (MvP) and iii) the liver against the pituitary gland (LvP).

Functional enrichment analysis

Differentially hydroxymethylated cytosines located within uncharacterized genes were excluded from further analysis, since they lack information regarding their function. In total, 4,590 DhmCs in MvL, 463 DhmCs in MvP and 367 DhmCs in LvP comparisons were annotated within uncharacterized genes. Functional enrichment analysis was performed using gProfiler (version: e99_eg46_p14_f929183) with default parameters for *Oreochromis niloticus*. Multiple testing correction was performed using the suggested tailor-made algorithm g:SCS and the significance threshold was set at 0.05 (Raudvere et al., 2019). The analysis was separated in 6 stages using hyper- and hypo-hydroxymethylated genes per pairwise tissue comparison as well as 3 additional stages for genes that were both hypo- and hyper-hydroxymethylated in a single tissue per pairwise tissue comparison.

RNA extraction and library preparation

The same fast muscle samples that were used for the RRHP library preparation were homogenised in DNA/RNA Shield (Zymo Research) using ZR bashing beads lysis tubes (Zymo Research). Homogenization was carried out using a Precellys 24 homogenizer (Bertin Instruments, France) with 2 cycles at 5000 rpm and 20 seconds duration. RNA extraction was performed using the Quick-RNA miniprep kit (Zymo Research) following the manufacturer's protocol for samples stored in DNA/RNA Shield. The Ribo-Zero Gold rRNA removal kit (Illumina) was used to ensure the removal of ribosomal RNA. RNA quality was determined using the spectrophotometer NanoDrop ND 1000 (ThermoFisher Scientific), while its concentration and integrity was measured with an RNA High-sensitivity screentape on a Tapestation 2200 (Agilent Technologies) (Additional File 15B). For the library preparation, the NEBNEXT Ultra II Directional RNA library prep kit for Illumina (NEB, USA) was used following the manufacturer's instructions and recommendations. Each sample was tagged with a unique barcode and the libraries were pooled at equimolar concentrations. The pool was sequenced at the Norwegian Sequencing Centre on a HiSeq 4000 lane (Illumina) in paired-end 150 bp mode. The resulting files were trimmed using BBDuk of the BBTools Suite (v. 38.22-0, Joint Genome Institute) which decontaminates using Kmers. Adapters were trimmed as per default (ktrim=r, k=23, mink=11, hdist=1, tpe, tbo) and quality trimmed for

both sides of the read with the Phred threshold set to 20, while reads less than 23 bp after trimming to Q20 were discarded. An index was built with the HISAT2 (v. 2.1.0) program (Kim et al., 2019) using the Nile tilapia reference genome above and trimmed reads were aligned by the core function of HISAT2 with RNA strandness set to FR and reported alignments tailored for downstream transcript assemblers. Alignments were sorted by samtools (v. 1.9) and a summarized report for all samples was generated the MultiQC software (Ewels et al., 2016). Calculation of count numbers for each gene was performed by featureCounts (v 1.6.2) (Liao et al., 2014) of the Subread package (Liao et al., 2013) using the gff3 genome annotation (NCBI, ref_O_niloticus_UMD_NMBU_top_level.gff3). Chimeric fragments were excluded, duplicates were ignored, only primary alignments were counted and fragments were allowed to match more than one meta-feature. Further analysis were performed in R (v. 3.6.1) (Team, 2015) and Rstudio (1.2.1335) (Team, 2020). The DESeq2 package (Love et al., 2014) was used for applying a regularized log transformation to the data in order to minimize differences between small counts and normalize according to the library size in an unbiased manner by prior information.

For each cytosine with 5hmC information, the mean of 5hmC counts across all five biological replicates was used together with the distance of each C from the TSS. Each C corresponded to a single expressed gene, the value of which was estimated based on the mean of gene expression counts across all five biological replicates. In other words, a value of gene expression could be attributed to more than one C. Principal component analysis and visualizations were performed with the FactoMineR (v. 1.42) and factoextra (v. 1.0.5) packages.

Availability of data and materials

The DNA hydroxymethylation dataset of this study is available in the SRA (NCBI) repository, under the accession number PRJNA665628.

The RNA sequencing dataset of this study is available in the GEO (NCBI) repository, under the accession number GSE158910.

References

Anastasiadi, D., Esteve-Codina, A. & Piferrer, F. (2018). Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species. *Epigenetics Chromatin,* 11**:** 37. DOI 10.1186/s13072-018-0205-1

Anastasiadi, D. & Piferrer, F. (2019). Epimutations in developmental genes underlie the onset of domestication in farmed European sea bass. *Mol Biol Evol,* 36**:** 2252-2264. DOI 10.1093/molbev/msz153

Bachman, M., Uribe-Lewis, S., Yang, X., Williams, M., Murrell, A. & Balasubramanian, S. (2014). 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nat Chem,* 6**:** 1049-55. DOI 10.1038/nchem.2064

Barazzoni, R., Bosutti, A., Stebel, M., Cattin, M.R., Roder, E., Visintin, L., Cattin, L., Biolo, G., Zanetti, M. & Guarnieri, G. (2005). Ghrelin regulates mitochondrial-lipid metabolism gene expression and tissue fat distribution in liver and skeletal muscle. *Am J Physiol Endocrinol Metab,* 288**:** E228-35. DOI 10.1152/ajpendo.00115.2004

Beckman, B.R. (2011). Perspectives on concordant and discordant relations between insulin-like growth factor 1 (IGF1) and growth in fishes. *Gen Comp Endocrinol,* 170**:** 233-52. DOI 10.1016/j.ygcen.2010.08.009

Bence, K.K. (2013). *Protein tyrosine phosphatase control of metabolism,* New York: Springer.

Beresewicz, M., Majewska, M., Makarewicz, D., Vayro, S., Zablocka, B. & Gorecki, D.C. (2010). Changes in the expression of insulin-like growth factor 1 variants in the postnatal brain development and in neonatal hypoxia-ischaemia. *Int J Dev Neurosci,* 28**:** 91-7. DOI 10.1016/j.ijdevneu.2009.09.002

Bertolini, F., Geraci, C., Schiavo, G., Sardina, M.T., Chiofalo, V. & Fontanesi, L. (2016). Whole genome semiconductor based sequencing of farmed European sea bass (Dicentrarchus labrax) Mediterranean genetic stocks using a DNA pooling approach. *Mar Genomics,* 28**:** 63-70. DOI 10.1016/j.margen.2016.03.007

Bhattacharyya, S., Pradhan, K., Campbell, N., Mazdo, J., Vasantkumar, A., Maqbool, S., Bhagat, T.D., Gupta, S., Suzuki, M., Yu, Y., Greally, J.M., Steidl, U., Bradner, J., Dawlaty, M., Godley, L., Maitra, A. & Verma, A. (2017). Altered hydroxymethylation is seen at regulatory regions in pancreatic cancer and regulates oncogenic pathways. *Genome Res,* 27**:** 1830-1842. DOI 10.1101/gr.222794.117

Bieberstein, N.I., Carrillo Oesterreich, F., Straube, K. & Neugebauer, K.M. (2012). First exon length controls active chromatin signatures and transcription. *Cell Rep,* 2**:** 62-8. DOI 10.1016/j.celrep.2012.05.019

Branco, M.R., Ficz, G. & Reik, W. (2011). Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat Rev Genet,* 13**:** 7-13. DOI 10.1038/nrg3080

Burgerhout, E., Mommens, M., Johnsen, H., Aunsmo, A., Santi, N. & Andersen, O. (2017). Genetic background and embryonic temperature affect DNA methylation and expression of myogenin and muscle development in Atlantic salmon (*Salmo salar*). *PLoS One,* 12**:** e0179918. DOI 10.1371/journal.pone.0179918

Campos, C., Sundaram, A.Y., Valente, L.M., Conceicao, L.E., Engrola, S. & Fernandes, J.M. (2014a). Thermal plasticity of the miRNA transcriptome during Senegalese sole development. *BMC Genomics,* 15**:** 525. DOI 10.1186/1471-2164-15-525

Campos, C., Valente, L.M., Borges, P., Bizuayehu, T. & Fernandes, J.M. (2010). Dietary lipid levels have a remarkable impact on the expression of growth-related genes in Senegalese sole (Solea senegalensis Kaup). *J Exp Biol,* 213**:** 200-9. DOI 10.1242/jeb.033126

Campos, C., Valente, L.M., Conceicao, L.E., Engrola, S. & Fernandes, J.M. (2013a). Epigenetic regulation of muscle development and growth in senegalese sole larvae. *Commun Agric Appl Biol Sci,* 78**:** 60-1.

Campos, C., Valente, L.M., Conceicao, L.E., Engrola, S. & Fernandes, J.M. (2013b). Temperature affects methylation of the myogenin putative promoter, its expression and muscle cellularity in Senegalese sole larvae. *Epigenetics,* 8**:** 389-97. DOI 10.4161/epi.24178

Campos, C., Valente, L.M., Conceicao, L.E., Engrola, S., Sousa, V., Rocha, E. & Fernandes, J.M. (2013c). Incubation temperature induces changes in muscle cellularity and gene expression in Senegalese sole (Solea senegalensis). *Gene,* 516**:** 209-17. DOI 10.1016/j.gene.2012.12.074

Campos, C., Valente, L.M.P., Conceição, L.E.C., Engola, S. & Fernandes, J.M.O. (2014b). Molecular regulation of muscle development and growth in Senegalese sole larvae exposed to temperature fluctuations. *Aquaculture,* 432**:** 418-425. DOI 10.1016/j.aquaculture.2014.04.035

Chen, P.Y., Chu, A., Liao, W.W., Rubbi, L., Janzen, C., Hsu, F.M., Thamotharan, S., Ganguly, A., Lam, L., Montoya, D., Pellegrini, M. & Devaskar, S.U. (2018). Prenatal growth patterns and birthweight are associated with differential DNA methylation and gene expression of cardiometabolic risk genes in human placentas: A discovery-based approach. *Reprod Sci,* 25**:** 523-539. DOI 10.1177/1933719117716779

Cingolani, P., Cao, X., Khetani, R.S., Chen, C.C., Coon, M., Sammak, A., Bollig-Fischer, A., Land, S., Huang, Y., Hudson, M.E., Garfinkel, M.D., Zhong, S., Robinson, G.E. & Ruden, D.M. (2013). Intronic non-CG DNA hydroxymethylation and alternative mRNA splicing in honey bees. *BMC Genomics,* 14**:** 666. DOI 10.1186/1471-2164-14-666

Colombo, M.N. & Francolini, M. (2019). Glutamate at the vertebrate neuromuscular junction: From modulation to neurotransmission. *Cells,* 8. DOI 10.3390/cells8090996

Company, R., Astola, A., Pendon, C., Valdivia, M.M. & Perez-Sanchez, J. (2001). Somatotropic regulation of fish growth and adiposity: growth hormone (GH) and somatolactin (SL) relationship. *Comp Biochem Physiol C Toxicol Pharmacol,* 130**:** 435-45.

Conte, M.A., Gammerdinger, W.J., Bartie, K.L., Penman, D.J. & Kocher, T.D. (2017). A high quality assembly of the Nile Tilapia (*Oreochromis niloticus*) genome reveals the structure of two sex determination regions. *BMC Genomics,* 18**:** 341. DOI 10.1186/s12864-017-3723-5

Cui, L., Yu, W.P., Deaizpurua, H.J., Schmidli, R.S. & Pallen, C.J. (1996). Cloning and characterization of islet cell antigen-related protein-tyrosine phosphatase (PTP), a novel receptor-like PTP and autoantigen in insulin-dependent diabetes. *J Biol Chem,* 271**:** 24817-23. DOI 10.1074/jbc.271.40.24817

Doherty, M.J., Young, P.R. & Cohen, P.T. (1996). Amino acid sequence of a novel protein phosphatase 1 binding protein (R5) which is related to the liver- and muscle-specific glycogen

binding subunits of protein phosphatase 1. *FEBS Lett,* 399**:** 339-43. DOI 10.1016/s0014-5793(96)01357-9

Ewels, P., Magnusson, M., Lundin, S. & Kaller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics,* 32**:** 3047-8. DOI 10.1093/bioinformatics/btw354

Ficz, G., Branco, M.R., Seisenberger, S., Santos, F., Krueger, F., Hore, T.A., Marques, C.J., Andrews, S. & Reik, W. (2011). Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature,* 473**:** 398-402. DOI 10.1038/nature10008

Fu, Q., Mcknight, R.A., Callaway, C.W., Yu, X., Lane, R.H. & Majnik, A.V. (2015). Intrauterine growth restriction disrupts developmental epigenetics around distal growth hormone response elements on the rat hepatic IGF-1 gene. *FASEB J,* 29**:** 1176-84. DOI 10.1096/fj.14-258442

Gao, D., Pinello, N., Nguyen, T.V., Thoeng, A., Nagarajah, R., Holst, J., Rasko, J.E. & Wong, J.J. (2019). DNA methylation/hydroxymethylation regulate gene expression and alternative splicing during terminal granulopoiesis. *Epigenomics,* 11**:** 95-109. DOI 10.2217/epi-2018-0050

Gentilini, D., Scala, S., Gaudenzi, G., Garagnani, P., Capri, M., Cescon, M., Grazi, G.L., Bacalini, M.G., Pisoni, S., Dicitore, A., Circelli, L., Santagata, S., Izzo, F., Di Blasio, A.M., Persani, L., Franceschi, C. & Vitale, G. (2017). Epigenome-wide association study in hepatocellular carcinoma: Identification of stochastic epigenetic mutations through an innovative statistical approach. *Oncotarget,* 8**:** 41890-41902. DOI 10.18632/oncotarget.17462

Globisch, D., Munzel, M., Muller, M., Michalakis, S., Wagner, M., Koch, S., Bruckl, T., Biel, M. & Carell, T. (2010). Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS One,* 5**:** e15367. DOI 10.1371/journal.pone.0015367

Gray, E.S., Kelley, K.M., Law, S., Tsai, R., Young, G. & Bern, H.A. (1992). Regulation of hepatic growth hormone receptors in coho salmon (Oncorhynchus kisutch). *Gen Comp Endocrinol,* 88**:** 243-52. DOI 10.1016/0016-6480(92)90256-j

Greco, C.M., Kunderfranco, P., Rubino, M., Larcher, V., Carullo, P., Anselmo, A., Kurz, K., Carell, T., Angius, A., Latronico, M.V., Papait, R. & Condorelli, G. (2016). DNA hydroxymethylation controls cardiomyocyte gene expression in development and hypertrophy. *Nat Commun,* 7**:** 12418. DOI 10.1038/ncomms12418

Gross, J.A., Pacis, A., Chen, G.G., Drupals, M., Lutz, P.E., Barreiro, L.B. & Turecki, G. (2017). Gene-body 5-hydroxymethylation is associated with gene expression changes in the prefrontal cortex of depressed individuals. *Transl Psychiatry,* 7**:** e1119. DOI 10.1038/tp.2017.93

Howard, A.D., Feighner, S.D., Cully, D.F., Arena, J.P., Liberator, P.A., Rosenblum, C.I., Hamelin, M., Hreniuk, D.L., Palyha, O.C., Anderson, J., Paress, P.S., Diaz, C., Chou, M., Liu, K.K., Mckee, K.K., Pong, S.S., Chaung, L.Y., Elbrecht, A., Dashkevicz, M., Heavens, R., Rigby, M., Sirinathsinghji, D.J., Dean, D.C., Melillo, D.G., Patchett, A.A., Nargund, R., Griffin, P.R., Demartino, J.A., Gupta, S.K., Schaeffer, J.M., Smith, R.G. & Van Der Ploeg, L.H. (1996). A receptor in pituitary and hypothalamus that functions in growth hormone release. *Science,* 273**:** 974-7. DOI 10.1126/science.273.5277.974

Kamstra, J.H., Loken, M., Alestrom, P. & Legler, J. (2015). Dynamics of DNA hydroxymethylation in zebrafish. *Zebrafish,* 12**:** 230-7. DOI 10.1089/zeb.2014.1033

Katan, T., Nash, G.W., Rise, M.L., Hall, J.A., Fernandes, J.M.O., Boyce, D., Johnsen, C.A. & Gamperl, A.K. (2016). A little goes long way: improved growth in Atlantic cod (Gadus morhua) fed small

amounts of wild zooplankton. *Aquaculture,* 451**:** 271-282. DOI 10.1016/j.aquaculture.2015.09.014

Kim, D., Paggi, J.M., Park, C., Bennett, C. & Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol,* 37**:** 907-915. DOI 10.1038/s41587-019-0201-4

Kinde, B., Gabel, H.W., Gilbert, C.S., Griffith, E.C. & Greenberg, M.E. (2015). Reading the unique DNA methylation landscape of the brain: Non-CpG methylation, hydroxymethylation, and MeCP2. *Proc Natl Acad Sci U S A,* 112**:** 6800-6. DOI 10.1073/pnas.1411269112

Konstantinidis, I., Saetrom, P., Mjelle, R., Nedoluzhko, A.V., Robledo, D. & Fernandes, J.M.O. (2020). Major gene expression changes and epigenetic remodelling in Nile tilapia muscle after just one generation of domestication. *Epigenetics,* 15**:** 1052-1067. DOI 10.1080/15592294.2020.1748914

Lee, K.H., Park, J.Y. & Kim, K. (2004). NMDA receptor-mediated calcium influx plays an essential role in myoblast fusion. *FEBS Lett,* 578**:** 47-52. DOI 10.1016/j.febslet.2004.10.076

Liao, Y., Smyth, G.K. & Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res,* 41**:** e108. DOI 10.1093/nar/gkt214

Liao, Y., Smyth, G.K. & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics,* 30**:** 923-30. DOI 10.1093/bioinformatics/btt656

Love, M.I., Huber, W. & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol,* 15**:** 550. DOI 10.1186/s13059-014-0550-8

Ma, Q., Xu, Z., Lu, H., Xu, Z., Zhou, Y., Yuan, B. & Ci, W. (2018). Distal regulatory elements identified by methylation and hydroxymethylation haplotype blocks from mouse brain. *Epigenetics Chromatin,* 11**:** 75. DOI 10.1186/s13072-018-0248-3

Manzon, L.A. (2002). The role of prolactin in fish osmoregulation: a review. *Gen Comp Endocrinol,* 125**:** 291-310. DOI 10.1006/gcen.2001.7746

Metcalfe, K.M., Bull, C.D. & Mangel, M. (2002). Seasonal variation in catch-up growth reveals state-dependent somatic allocations in salmon. *Evol Ecol Res,* 4**:** 871-881.

Mishima, Y., Abreu-Goodger, C., Staton, A.A., Stahlhut, C., Shou, C., Cheng, C., Gerstein, M., Enright, A.J. & Giraldez, A.J. (2009). Zebrafish miR-1 and miR-133 shape muscle gene expression and regulate sarcomeric actin organization. *Genes Dev,* 23**:** 619-32. DOI 10.1101/gad.1760209

Nestor, C.E., Ottaviano, R., Reddington, J., Sproul, D., Reinhardt, D., Dunican, D., Katz, E., Dixon, J.M., Harrison, D.J. & Meehan, R.R. (2012). Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Res,* 22**:** 467-77. DOI 10.1101/gr.126417.111

Nordgarden, U., Fjelldal, P.G., Hansen, T., Bjornsson, B.T. & Wargelius, A. (2006). Growth hormone and insulin-like growth factor-I act together and independently when regulating growth in vertebral and muscle tissue of atlantic salmon postsmolts. *Gen Comp Endocrinol,* 149**:** 253-60. DOI 10.1016/j.ygcen.2006.06.001

O'rourke, T. & Boeckx, C. (2020). Glutamate receptors in domestication and modern human evolution. *Neurosci Biobehav Rev,* 108**:** 341-357. DOI 10.1016/j.neubiorev.2019.10.004

Pastor, W.A., Pape, U.J., Huang, Y., Henderson, H.R., Lister, R., Ko, M., Mcloughlin, E.M., Brudno, Y., Mahapatra, S., Kapranov, P., Tahiliani, M., Daley, G.Q., Liu, X.S., Ecker, J.R., Milos, P.M., Agarwal, S. & Rao, A. (2011). Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature,* 473**:** 394-7. DOI 10.1038/nature10102

Perez-Sanchez, J., Marti-Palanca, H. & Kaushik, S.J. (1995). Ration size and protein intake affect circulating growth hormone concentration, hepatic growth hormone binding and plasma insulin-like growth factor-I immunoreactivity in a marine teleost, the gilthead sea bream (Sparus aurata). *J Nutr,* 125**:** 546-52. DOI 10.1093/jn/125.3.546

Perez-Sanchez, J., Simo-Mirabet, P., Naya-Catala, F., Martos-Sitcha, J.A., Perera, E., Bermejo-Nogales, A., Benedito-Palos, L. & Calduch-Giner, J.A. (2018). Somatotropic axis regulation unravels the differential effects of nutritional and environmental factors in growth performance of marine farmed fishes. *Front Endocrinol (Lausanne),* 9**:** 687. DOI 10.3389/fendo.2018.00687

Petterson, A., Chung, T.H., Tan, D., Sun, X. & Jia, X.Y. (2014). RRHP: a tag-based approach for 5-hydroxymethylcytosine mapping at single-site resolution. *Genome Biol,* 15**:** 456. DOI 10.1186/s13059-014-0456-5

Pfluger, P.T., Kirchner, H., Gunnel, S., Schrott, B., Perez-Tilve, D., Fu, S., Benoit, S.C., Horvath, T., Joost, H.G., Wortley, K.E., Sleeman, M.W. & Tschop, M.H. (2008). Simultaneous deletion of ghrelin and its receptor increases motor activity and energy expenditure. *Am J Physiol Gastrointest Liver Physiol,* 294**:** G610-8. DOI 10.1152/ajpgi.00321.2007

Podgorniak, T., Brockmann, S., Konstantinidis, I. & Fernandes, J.M.O. (2019). Differences in the fast muscle methylome provide insight into sex-specific epigenetic regulation of growth in Nile tilapia during early stages of domestication. *Epigenetics,* 14**:** 818-836. DOI 10.1080/15592294.2019.1618164

Ponnaluri, V.K., Ehrlich, K.C., Zhang, G., Lacey, M., Johnston, D., Pradhan, S. & Ehrlich, M. (2017). Association of 5-hydroxymethylation and 5-methylation of DNA cytosine with tissue-specific gene expression. *Epigenetics,* 12**:** 123-138. DOI 10.1080/15592294.2016.1265713

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H. & Vilo, J. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res,* 47**:** W191-W198. DOI 10.1093/nar/gkz369

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W. & Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res,* 43**:** e47. DOI 10.1093/nar/gkv007

Rohde, K., Keller, M., Stumvoll, M., Dietrich, A., Bluher, M. & Bottcher, Y. (2015). DNA 5-hydroxymethylation in human adipose tissue differs between subcutaneous and visceral adipose tissue depots. *Epigenomics,* 7**:** 911-20. DOI 10.2217/epi.15.50

Rong, R., Ahn, J.Y., Huang, H., Nagata, E., Kalman, D., Kapp, J.A., Tu, J., Worley, P.F., Snyder, S.H. & Ye, K. (2003). PI3 kinase enhancer-Homer complex couples mGluRI to PI3 kinase, preventing neuronal apoptosis. *Nat Neurosci,* 6**:** 1153-61. DOI 10.1038/nn1134

Rose, A.B. (2018). Introns as gene regulators: A brick on the accelerator. *Front Genet,* 9**:** 672. DOI 10.3389/fgene.2018.00672

Santiago, M., Antunes, C., Guedes, M., Sousa, N. & Marques, C.J. (2014). TET enzymes and DNA hydroxymethylation in neural development and function - how critical are they? *Genomics,* 104**:** 334-40. DOI 10.1016/j.ygeno.2014.08.018

Schober, M.E., Ke, X., Xing, B., Block, B.P., Requena, D.F., Mcknight, R. & Lane, R.H. (2012). Traumatic brain injury increased IGF-1B mRNA and altered IGF-1 exon 5 and promoter region epigenetic characteristics in the rat pup hippocampus. *J Neurotrauma,* 29**:** 2075-85. DOI 10.1089/neu.2011.2276

Sengelaub, C.A., Navrazhina, K., Ross, J.B., Halberg, N. & Tavazoie, S.F. (2016). PTPRN2 and PLCbeta1 promote metastatic breast cancer cell migration through PI(4,5)P2-dependent actin remodeling. *EMBO J,* 35**:** 62-76. DOI 10.15252/embj.201591973

Shen, J., Lefave, C., Sirosh, I., Siegel, A.B., Tycko, B. & Santella, R.M. (2015). Integrative epigenomic and genomic filtering for methylation markers in hepatocellular carcinomas. *BMC Med Genomics,* 8**:** 28. DOI 10.1186/s12920-015-0105-1

Szulwach, K.E., Li, X., Li, Y., Song, C.X., Wu, H., Dai, Q., Irier, H., Upadhyay, A.K., Gearing, M., Levey, A.I., Vasanthakumar, A., Godley, L.A., Chang, Q., Cheng, X., He, C. & Jin, P. (2011). 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nat Neurosci,* 14**:** 1607-16. DOI 10.1038/nn.2959

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L. & Rao, A. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science,* 324**:** 930-5. DOI 10.1126/science.1170116

Taylor, J.F., Porter, M.J., Bromage, N.R. & Migaud, H. (2008). Relationships between environmental changes, maturity, growth rate and plasma insulin-like growth factor-I (IGF-I) in female rainbow trout. *Gen Comp Endocrinol,* 155**:** 257-70. DOI 10.1016/j.ygcen.2007.05.014

Team, R. (2020). RStudio: Integrated development for R. Boston, MA: RStudio, PBC.

Team, R.D.C. (2015). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Valente, L.M.P., Moutou, K.A., Conceição, L., Engrola, S., Fernandes, J.M.O. & Johnston, I.A. (2013). What determines growth potential and juvenile quality of farmed fish species. *Rev Aquacult,* 5**:** 168-193. DOI 10.1111/raq.12020

Waerner, T., Gardellin, P., Pfizenmaier, K., Weith, A. & Kraut, N. (2001). Human RERE is localized to nuclear promyelocytic leukemia oncogenic domains and enhances apoptosis. *Cell Growth Differ,* 12**:** 201-10.

Whittington, C.M. & Wilson, A.B. (2013). The role of prolactin in fish reproduction. *Gen Comp Endocrinol,* 191**:** 123-36. DOI 10.1016/j.ygcen.2013.05.027

Wu, H., D'alessio, A.C., Ito, S., Wang, Z., Cui, K., Zhao, K., Sun, Y.E. & Zhang, Y. (2011). Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev,* 25**:** 679-84. DOI 10.1101/gad.2036011

Wu, Y., Chen, X., Zhao, Y., Wang, Y., Li, Y. & Xiang, C. (2019). Genome-wide DNA methylation and hydroxymethylation analysis reveal human menstrual blood-derived stem cells inhibit hepatocellular carcinoma growth through oncogenic pathway suppression via regulating 5-hmC in enhancer elements. *Stem Cell Res Ther,* 10**:** 151. DOI 10.1186/s13287-019-1243-8

Yan, B., Zhu, C.D., Guo, J.T., Zhao, L.H. & Zhao, J.L. (2013). miR-206 regulates the growth of the teleost tilapia (*Oreochromis niloticus*) through the modulation of IGF-1 gene expression. *J Exp Biol,* 216**:** 1265-9. DOI 10.1242/jeb.079590

Yin, Y., Li, Y. & Zhang, W. (2014). The growth hormone secretagogue receptor: its intracellular signaling and regulation. *Int J Mol Sci,* 15**:** 4837-55. DOI 10.3390/ijms15034837

Tables

Table 1. List of DhmCs between muscle and liver, found within the most expressed genes in muscle. Positive and negative logFC values correspond to hyper- and hypo-hydroxymethylation in muscle compared to liver, respectively.

Chromosome	Position	LogFC	Annotation	Dist. to TSS	Gene Symbol	Description
NC_031972.2	3582578	0.70	Intron (1/11)	1318	*LOC100707421*	troponin T fast skeletal muscle isoforms
NC_031973.2	16443052	-1.51	Exon (16/24)	7879	*LOC100706607*	sarcoplasmic/endoplasmic reticulum calcium ATPase 1
NC_031986.2	22143050	-1.53	Exon (5/6)	2333	*LOC100698429*	myosin light chain 3 skeletal muscle isoform
NC_031969.2	34504991	-1.74	Intergenic	10176	*LOC100697135*	nucleoside diphosphate kinase B
NC_031969.2	34505127	-1.94	Intergenic	10040	*LOC100697135*	nucleoside diphosphate kinase B
NC_031973.2	16442970	-2.01	Exon (16/24)	7797	*LOC100706607*	sarcoplasmic/endoplasmic reticulum calcium ATPase 1
NC_031969.2	23538241	-2.72	Intergenic	19650	*LOC100698429*	myosin heavy chain fast skeletal muscle
NC_031969.2	23538340	-3.18	Intergenic	19551	*LOC100698429*	myosin heavy chain fast skeletal muscle

Table 2. Sampling information in xlsx format, including measurements of weight, total and standard length of all the individuals used for the current study (n=5). Animal and family ID were tracked through passive integrated transponders (PIT).

Sample ID	Weight (g)	Total length (cm)	Standard length (cm)
M1	1001	37.1	29.7
M2	683	33.9	26.8
M3	1298	37.9	29.1
M4	915	35.9	28.0
M5	708	34.2	26.4

Figure Legends

Figure 1. Venn diagrams representing the RRHP dataset. A) Stacked Venn diagram showing the distribution of total CCGG sites captured by RRHP, substantially hydroxymethylated and differentially hydroxymethylated among tissues compare to the total genomic CCGG sites of the Nile tilapia genome. B) Venn diagram depicting total differentially hydroxymethylated sites between and among tissue comparisons (MvL - Muscle vs Liver, MvP – Muscle vs Pituitary, and LvP – Liver vs Pituitary). C) Venn diagram showing the total and shared number of hyper-hydroxymethylated genes among tissues. Genes containing or associated with a DhmC in close proximity based on their distance to TSS were considered differentially hydroxymethylated. Created with BioRender.com

Figure 2. Histograms depicting the number of differentially hydroxymethylated cytosines (DhmCs) between (A) muscle versus liver, (B) muscle versus pituitary gland, and (C) liver versus pituitary.
The number of hyper- and hypo-hydroxymethylated sites are presented separately for promoters, introns, exons, transcription termination sites and intergenic regions.

Figure 3. Heatmaps of growth-related genes containing differentially hydroxymethylated cytosines (DhmCs) among the three tissues. The level of 5hmC enrichment is mapped based on 5hmC counts across samples. The three heatmaps represent the comparisons between (A) muscle and liver, (B) muscle and pituitary, and (C) liver and pituitary. The 5hmC enrichment is represented as a gradient from light blue (low) to red (high).

Figure 4. Principal component analysis using three variables; mean 5hmC levels (orange), mean gene expression (GE; blue) and distance to transcription start site (green) across the first two dimensions that explain 67.5% of the total dataset variation. Colour coding depicts the quality (cos^2) of representation for the three variables, with lower, intermediate and higher quality represented in blue, yellow and orange, respectively.

Figure 5. Line graph depicting the correlation of low (light blue) and high (red) 5hmC levels with gene expression. DNA hydroxymethylation levels are represented on the y-axis in four density quantiles from 0 to 100%, while the x-axis shows mean gene expression. The three graphs from left to right represent the correlations between gene expression and 5hmC levels within first exons (A), first introns (B) and promoters (C).

Figure 6. Graphical summary of the main findings of the study. DNA hydroxymethylation was enriched within gene bodies and promoters (80%) compared to intergenic regions (20%). Tissues of the somatotropic axis were differentially hydroxymethylated. The liver had the highest, muscle intermediate, and pituitary the lowest number of hyper-5hmCs. Several genes involved in somatic growth were found to be differentially hydroxymethylated among all three comparisons (pituitary – muscle, muscle – liver and liver – pituitary). Created with BioRender.com

Figures

Figure 1.

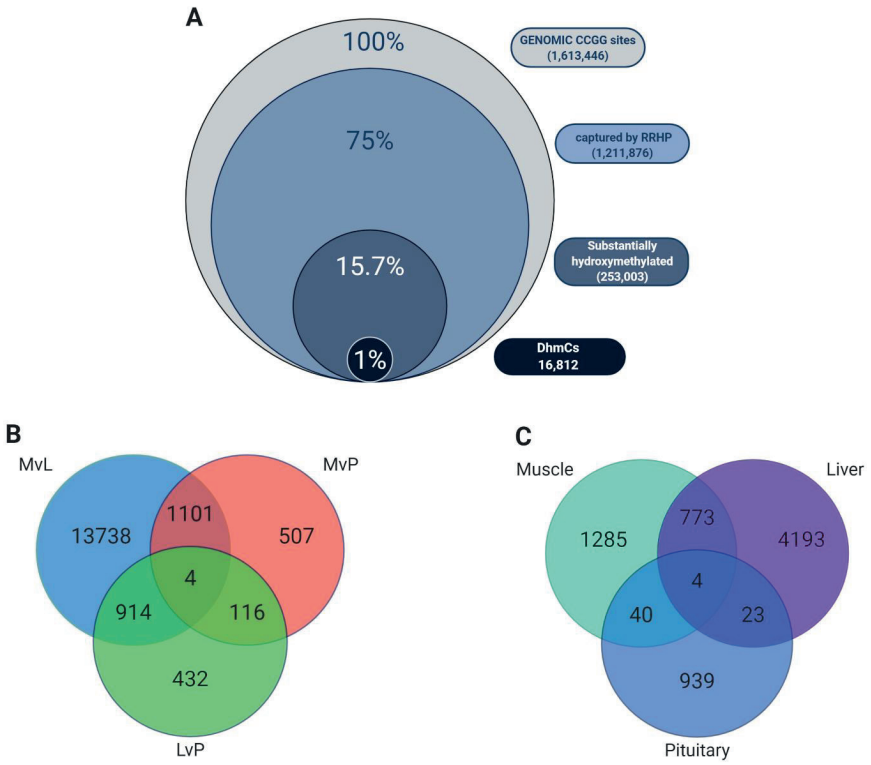


Figure 2.

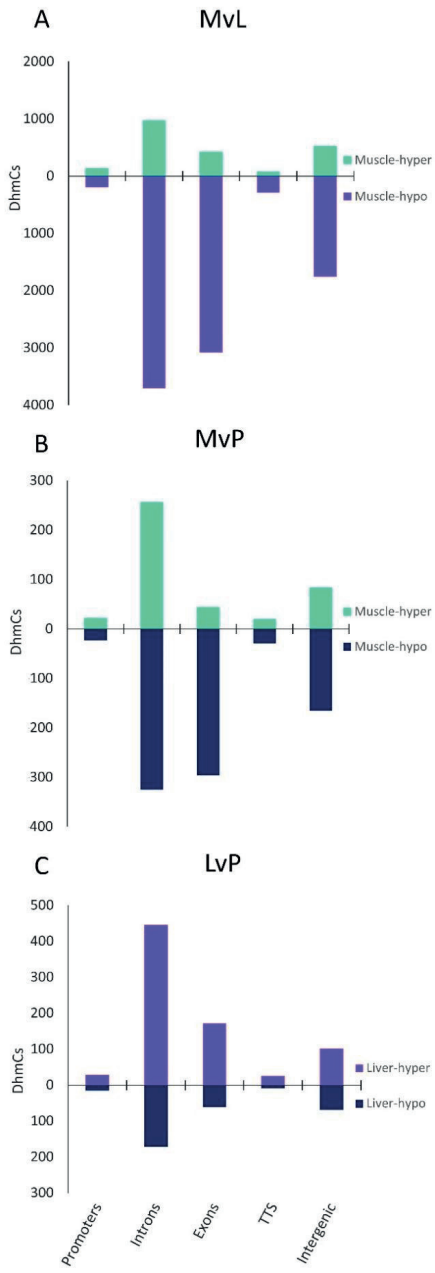

A
MvL
2000
1000
0
1000
2000
3000
4000
DhmCs
Muscle-hyper
Muscle-hypo
B
MvP
300
200
100
0
100
200
300
400
DhmCs
Muscle-hyper
Muscle-hypo
C
LvP
500
400
300
200
100
0
100
200
300
DhmCs
Liver-hyper
Liver-hypo
Promoters
Introns
Exons
TTS
Intergenic

Figure 3.

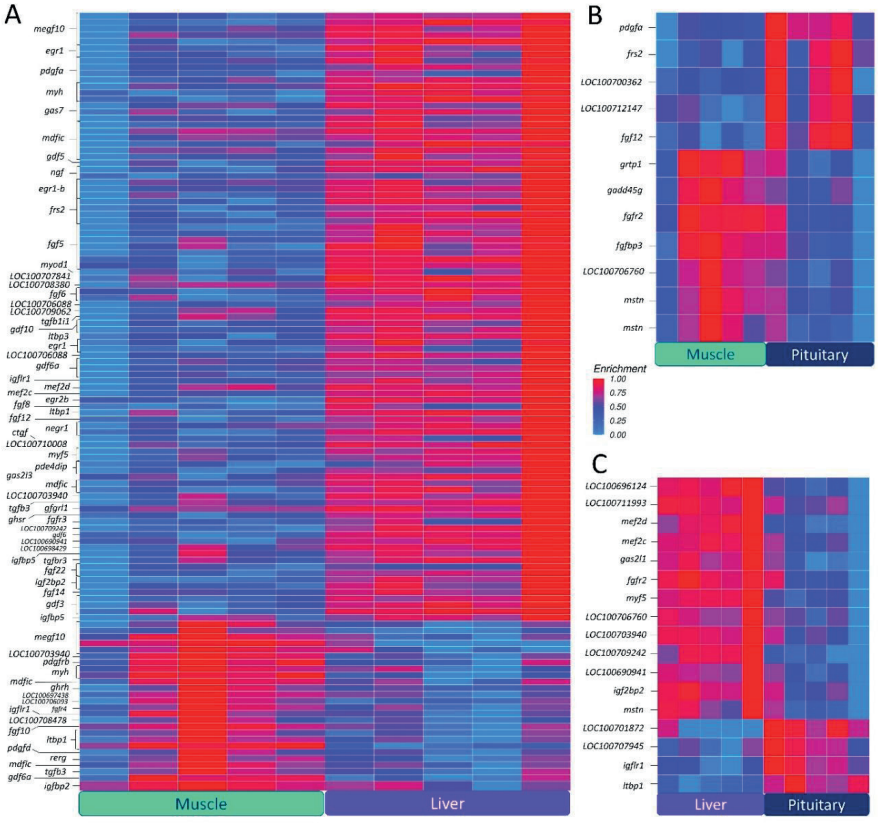

A
megf10
egr1
pdgfa
myh
gas7
mdfic
gdf5
ngf
egr1-b
frs2
fgf5
myod1
LOC100707841
LOC100708380
fgf6
LOC100706088
LOC100709062
tgfb1i1
gdf10
itbp3
egr1
LOC100706088
gdf6a
igfir1
mef2d
mef2c
egr2b
fgf8
itbp1
fgf12
negr1
ctgf
LOC100710008
myf5
pde4dip
gas2l3
mdfic
LOC100703940
tgfb3
gfgrl1
ghsr
fgfr3
LOC100709242
LOC100690941
LOC100698479
igfbp5
tgfbr3
fgf22
igf2bp2
fgf14
gdf3
igfbp5
megf10
LOC100703940
pdgfrb
myh
mdfic
ghrh
LOC100697438
LOC100706093
igfir1
fgf4
LOC100708478
fgf10
itbp1
pdgfd
rerg
mdfic
tgfb3
gdf6a
igfbp2
Muscle
Liver
B
pdgfa
frs2
LOC100700362
LOC100712147
fgf12
grtp1
gadd45g
fgfr2
fgfbp3
LOC100706760
mstn
mstn
Muscle
Pituitary
Enrichment
1.00
0.75
0.50
0.25
0.00
C
LOC100696124
LOC100711993
mef2d
mef2c
gas2l1
fgfr2
myf5
LOC100706760
LOC100703940
LOC100709242
LOC100690941
igf2bp2
mstn
LOC100701872
LOC100707945
igfir1
itbp1
Liver
Pituitary

Figure 4.

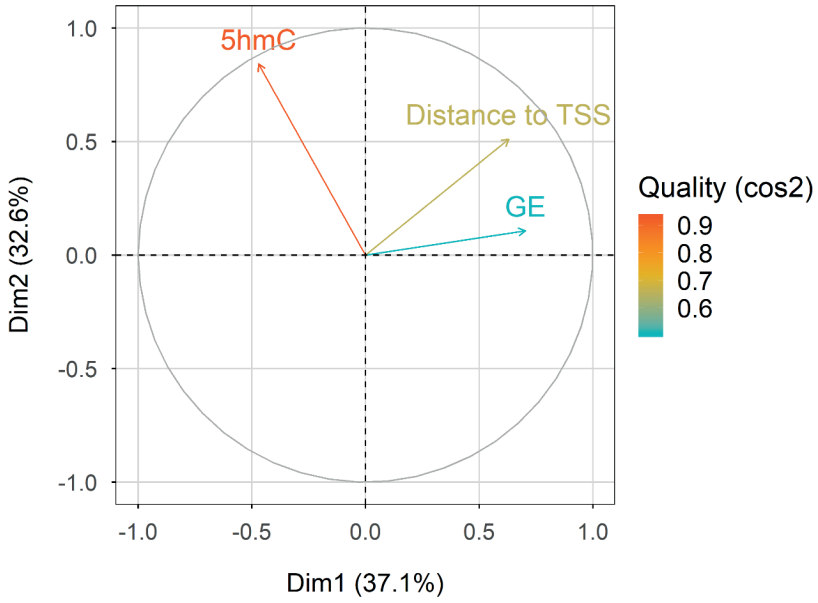

5hmC
Distance to TSS
GE
Quality (cos2)
0.9
0.8
0.7
0.6
1.0
0.5
0.0
-0.5
-1.0
Dim2 (32.6%)
Dim1 (37.1%)

Figure 5.

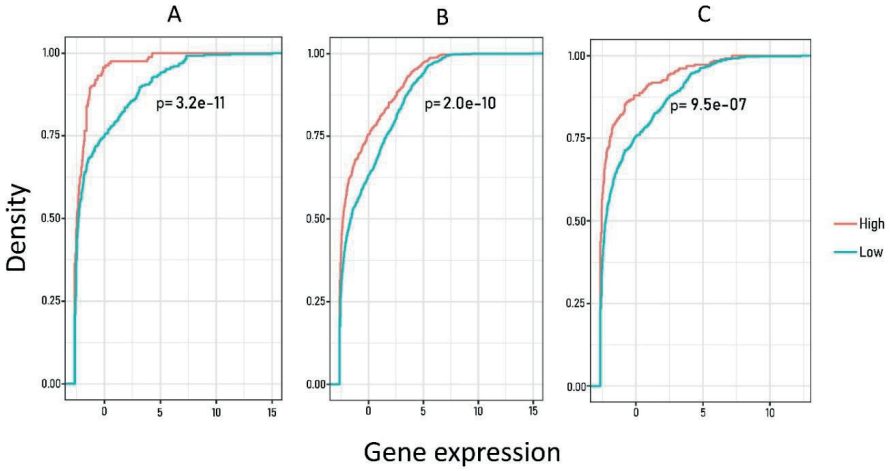
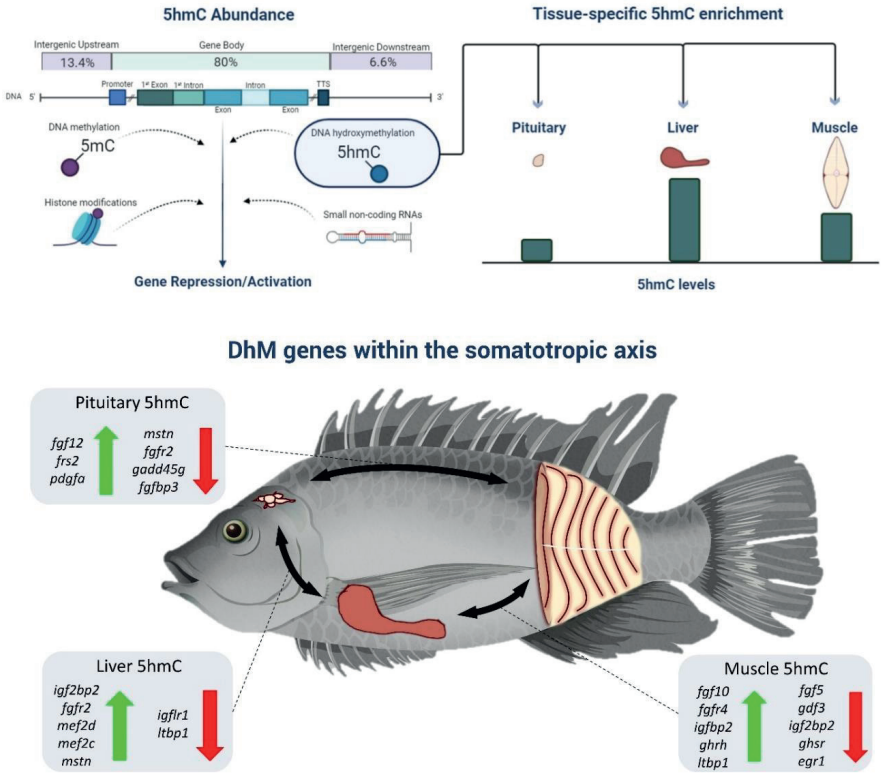


Figure 6.


5hmC Abundance
Intergenic Upstream
13.4%
Gene Body
80%
Intergenic Downstream
6.6%
Promoter
1st Exon
1st Intron
Intron
TTS
Exon
Exon
DNA 5'
3'
DNA methylation
5mC
DNA hydroxymethylation
5hmC
Histone modifications
Small non-coding RNAs
Gene Repression/Activation
Tissue-specific 5hmC enrichment
Pituitary
Liver
Muscle
5hmC levels
DhM genes within the somatotropic axis
Pituitary 5hmC
fgf12
frs2
pdgfa
mstn
fgfr2
gadd45g
fgfbp3
Liver 5hmC
igf2bp2
fgfr2
mef2d
mef2c
mstn
igflr1
ltbp1
Muscle 5hmC
fgf10
fgfr4
igfbp2
ghrh
ltbp1
fgf5
gdf3
igf2bp2
ghsr
egr1

Supplemental Material

Supplemental legends

Additional file 1. Barplot in jpeg format, representing the raw, quality control and trimmed, uniquely aligned, multiple aligned and failed to aligned reads (x-axis) in millions (y-axis) across the three tissues. Muscle, liver and pituitary are represented in green, purple and dark blue, respectively (n=5). Bars and error bars represent means and standard errors, respectively.

Additional file 2. List of differentially hydroxymethylated cytosines (DhmCs) in xlsx format, among all pairwise tissue comparisons. Their location within the Nile tilapia genome (NCBI assembly GCA_001858045.3) is depicted in the first three columns (Chromosome, Position and Strand). Gene symbols and their annotation are based on the output by HOMER. Adjusted p-values for each pairwise tissue comparison: muscle compared to pituitary (MvL), liver compared to pituitary (MvP) and muscle compared to liver (MvL). Finally, the rest columns depict the 5hmC filtered counts and average filtered counts per sample and tissue, respectively.

Additional file 3. Barplot in jpeg format, representing the percentage of DhmC enrichment (y-axis) within the annotated features (TTS, exons, introns, intergenic, promoters; x-axis). Muscle, liver and pituitary are represented in green, purple and dark blue, respectively (n=5).

Additional file 4. List of significantly enriched molecular functions and biological processes in xlsx format, based on all hypo-hydroxymethylated genes in muscle compared to liver (gProfiler, multiple testing correction g:SCS, $q<0.05$).

Additional file 5. List of significantly enriched molecular functions and biological processes in xlsx format, based on all hyper-hydroxymethylated genes in muscle compared to liver (gProfiler, multiple testing correction g:SCS, $q<0.05$).

Additional file 6. List of significantly enriched molecular functions and biological processes in xlsx format, based on 773 genes that contained hyper-hydroxymethylcytosines in both muscle and liver (gProfiler, multiple testing correction g:SCS, q<0.05).

Additional file 7. List of significantly enriched molecular functions and biological processes in xlsx format, based on all hyper-hydroxymethylated genes in muscle compared to pituitary (gProfiler, multiple testing correction g:SCS, q<0.05).

Additional file 8. List of significantly enriched molecular functions and biological processes in xlsx format, based on all hypo-hydroxymethylated genes in muscle compared to pituitary (gProfiler, multiple testing correction g:SCS, q<0.05).

Additional file 9. List of significantly enriched molecular functions and biological processes in xlsx format, based on all hypo-hydroxymethylated genes in liver compared to pituitary (gProfiler, multiple testing correction g:SCS, q<0.05).

Additional file 10. List of significantly enriched molecular functions and biological processes in xlsx format, based on all hyper-hydroxymethylated genes in liver compared to pituitary (gProfiler, multiple testing correction g:SCS, q<0.05).

Additional file 11. List of genes (in xlsx format) associated with somatic growth and found to be both hypo- and hyper-hydroxymethylated in muscle compared to liver (**A**), muscle compared to pituitary (**B**), and liver compared to pituitary (**C**). The information provided in the table contains the precise location of 5hmCs in the Nile tilapia genome (Chromosome, Position and Strand), the log fold change (logFC) of DNA hydroxymethylation levels between the corresponding tissues and their annotation (annotated feature, Distance to TSS, transcript, gene symbol and description) based on software HOMER.

Additional file 12. List of muscle samples (n=5) and their RNA-Seq library characterization in xlsx format. The table provides information for the precise number of reads and fragments during every step of analysis, from raw sequenced to uniquely aligned reads and assigned fragments.

Additional file 13. List of all the RNA transcripts and their expression in xlsx format. The table summarizes the level of expression of each transcript for every fast muscle sample (n=5) based on their normalized counts.

Additional file 14. List of the top 30 most expressed transcripts in fast muscle (n=5) in xlsx format. The table provides information regarding the expression of each transcript based on normalized counts as well as their annotation within the Nile tilapia genome.

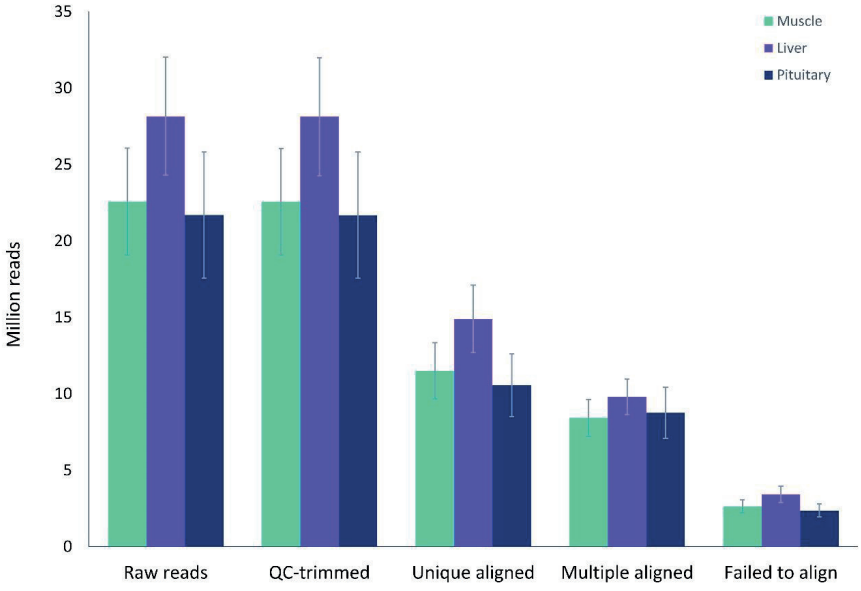
Additional file 15. List of (A) DNA and (B) RNA samples, their integrity number and concentration after DNA and RNA extraction, respectively.

Supplemental tables

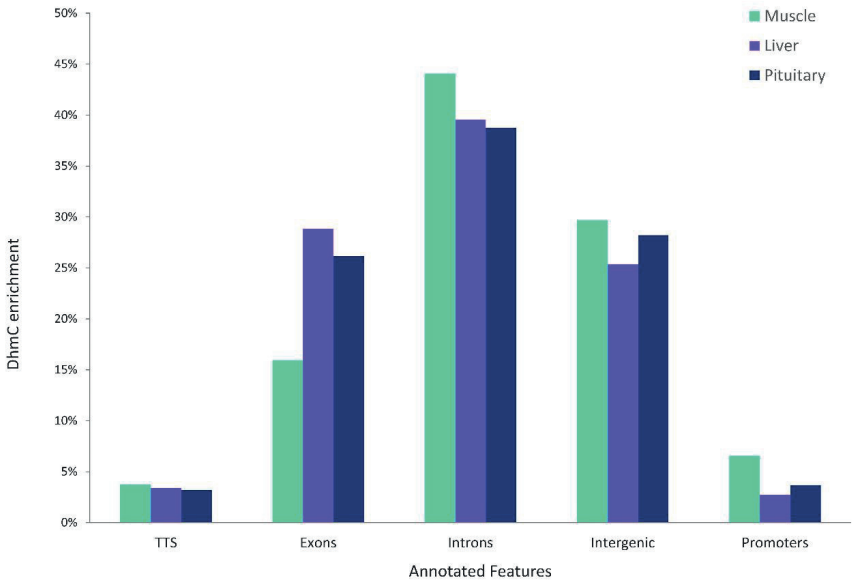
Supplemental tables are available upon request: ioannis.constant@outlook.com

Supplementary Figures

Additional file 1.



Additional file 3.



List of previously published theses for PhD in Aquaculture / PhD in Aquatic Biosciences, Nord University

No. 1 (2011)
PhD in Aquaculture
Chris André Johnsen
Flesh quality and growth of farmed Atlantic salmon (*Salmo salar* L.) in relation to feed, feeding, smolt type and season
ISBN: 978-82-93165-00-2

No. 2 (2012)
PhD in Aquaculture
Jareeporn Ruangsri
Characterization of antimicrobial peptides in Atlantic cod
ISBN: 978-82-93165-01-9

No. 3 (2012)
PhD in Aquaculture
Muhammad Naveed Yousaf
Characterization of the cardiac pacemaker and pathological responses to cardiac diseases in Atlantic salmon (*Salmo salar* L.)
ISBN: 978-82-93165-02-6

No. 4 (2012)
PhD in Aquaculture
Carlos Frederico Ceccon Lanes
Comparative Studies on the quality of eggs and larvae from broodstocks of farmed and wild Atlantic cod
ISBN: 978-82-93165-03-3

No. 5 (2012)
PhD in Aquaculture
Arvind Sundaram
Understanding the specificity of the innate immune response in teleosts: Characterisation and differential expression of teleost-specific Toll-like receptors and microRNAs
ISBN: 978-82-93165-04-0

No. 6 (2012)
PhD in Aquaculture
Teshome Tilahun Bizuayehu
Characterization of microRNA during early ontogeny and sexual development of Atlantic halibut (*Hippoglossus hippoglossus* L.)
ISBN: 978-82-93165-05-7

No. 7 (2013)
PhD in Aquaculture
Binoy Rajan
Proteomic characterization of Atlantic cod skin mucosa – Emphasis on innate immunity and lectins
ISBN: 978-82-93165-06-04

No. 8 (2013)
PhD in Aquaculture
Anusha Krishanthi Shyamali Dhanasiri
Transport related stress in zebrafish: physiological responses and bioremediation
ISBN: 978-82-93165-07-1

No. 9 (2013)
PhD in Aquaculture
Martin Haugmo Iversen
Stress and its impact on animal welfare during commercial production of Atlantic salmon (*Salmo salar* L.)
ISBN: 978-82-93165-08-8

No. 10 (2013)
PhD in Aquatic Biosciences
Alexander Jüterbock
Climate change impact on the seaweed *Fucus serratus*, a key foundational species on North Atlantic rocky shores
ISBN: 978-82-93165-09-5

No. 11 (2014)
PhD in Aquatic Biosciences
Amod Kulkarni
Responses in the gut of black tiger shrimp *Penaeus monodon* to oral vaccine candidates against white spot disease
ISBN: 978-82-93165-10-1

No. 12 (2014)
PhD in Aquatic Biosciences
Carlo C. Lazado
Molecular basis of daily rhythmicity in fast skeletal muscle of Atlantic cod *(Gadus morhua)*
ISBN: 978-82-93165-11-8

No. 13 (2014)
PhD in Aquaculture
Joanna Babiak
Induced masculinization of Atlantic halibut (*Hippoglossus hippoglossus* L.): towards the goal of all-female production
ISBN: 978-82-93165-12-5

No. 14 (2015)
PhD in Aquaculture
Cecilia Campos Vargas
Production of triploid Atlantic cod: A comparative study of muscle growth dynamics and gut morphology
ISBN: 978-82-93165-13-2

No. 15 (2015)
PhD in Aquatic Biosciences
Irina Smolina
Calanus in the North Atlantic: species identification, stress response, and population genetic structure
ISBN: 978-82-93165-14-9

No. 16 (2016)
PhD in Aquatic Biosciences
Lokesh Jeppinamogeru
Microbiota of Atlantic salmon (*Salmo salar L*.), during their early and adult life
ISBN: 978-82-93165-15-6

No. 17 (2017)
PhD in Aquatic Biosciences
Christopher Edward Presslauer
Comparative and functional analysis of microRNAs during zebrafish gonadal development
ISBN: 978-82-93165-16-3

No. 18 (2017)
PhD in Aquatic Biosciences
Marc Jürgen Silberberger
Spatial scales of benthic ecosystems in the sub-Arctic Lofoten-Vesterålen region
ISBN: 978-82-93165-17-0

No. 19 (2017)
PhD in Aquatic Biosciences
Marvin Choquet
Combining ecological and molecular approaches to redefine the baseline knowledge of the genus Calanus in the North Atlantic and the Arctic Oceans
ISBN: 978-82-93165-18-7

No. 20 (2017)
PhD in Aquatic Biosciences
Torvald B. Egeland
Reproduction in Arctic charr – timing and the need for speed
ISBN: 978-82-93165-19-4

No. 21 (2017)
PhD in Aquatic Biosciences
Marina Espinasse
Interannual variability in key zooplankton species in the North-East Atlantic: an analysis based on abundance and phenology
ISBN: 978-82-93165-20-0

No. 22 (2018)
PhD in Aquatic Biosciences
Kanchana Bandara
Diel and seasonal vertical migrations of high-latitude zooplankton: knowledge gaps and a high-resolution bridge
ISBN: 978-82-93165-21-7

No. 23 (2018)
PhD in Aquatic Biosciences
Deepti Manjari Patel
Characterization of skin immune and stress factors of lumpfish, *Cyclopterus lumpus*
ISBN: 978-82-93165-21-7

No. 24 (2018)
PhD in Aquatic Biosciences
Prabhugouda Siriyappagouder
The intestinal mycobiota of zebrafish – community profiling and exploration of the impact of yeast exposure early in life
ISBN: 978-82-93165-23-1

No. 25 (2018)
PhD in Aquatic Biosciences
Tor Erik Jørgensen
Molecular and evolutionary characterization of the Atlantic cod mitochondrial genome
ISBN: 978-82-93165-24-8

No. 26 (2018)
PhD in Aquatic Biosciences
Yangyang Gong
Microalgae as feed ingredients for Atlantic salmon
ISBN: 978-82-93165-25-5

No. 27 (2018)
PhD in Aquatic Biosciences
Ove Nicolaisen
Approaches to optimize marine larvae production
ISBN: 978-82-93165-26-2

No. 28 (2019)
PhD in Aquatic Biosciences
Qirui Zhang
The effect of embryonic incubation temperature on the immune response of larval and adult zebrafish (*Danio rerio*)
ISBN: 978-82-93165-27-9

No. 29 (2019)
PhD in Aquatic Biosciences
Andrea Bozman
The structuring effects of light on the deep-water scyphozoan *Periphylla periphylla*
ISBN: 978-82-93165-28-6

No. 30 (2019)
PhD in Aquatic Biosciences
Helene Rønquist Knutsen
Growth and development of juvenile spotted wolffish (*Anarhichas minor*) fed microalgae incorporated diets
ISBN: 978-82-93165-29-3

No. 31 (2019)
PhD in Aquatic Biosciences
Shruti Gupta
Feed additives elicit changes in the structure of the intestinal bacterial community of Atlantic salmon
ISBN: 978-82-93165-30-9

No. 32 (2019)
PhD in Aquatic Biosciences
Peter Simon Claus Schulze
Phototrophic microalgal cultivation in cold and light-limited environments
ISBN: 978-82-93165-31-6

No. 33 (2019)
PhD in Aquatic Biosciences
Maja Karoline Viddal Hatlebakk
New insights into Calanus glacialis and C. finmarchicus distribution, life histories and physiology in high-latitude seas
ISBN: 978-82-93165-32-3

No. 34 (2019)
PhD in Aquatic Biosciences
Arseny Dubin
Exploration of an anglerfish genome
ISBN: 978-82-93165-33-0

No. 35 (2020)
PhD in Aquatic Biosciences
Florence Chandima Perera Willora Arachchilage
The potential of plant ingredients in diets of juvenile lumpfish (*Cyclopterus lumpus*)
ISBN: 978-82-93165-35-4

DNA hydroxymethylation has been characterized as the sixth DNA base in genomic DNA. Its ability to regulate gene transcription and alter histone accessibility is the driver behind recent developments across several fields of biology. Using one of the most cultivated fish species in the world as our model organism, the Nile tilapia, this thesis represents the first comprehensive characterization of the DNA hydroxymethylome in fish undergoing domestication. Using genome-wide approaches, we identified major changes in DNA hydroxymethylation and gene expression in response to domestication. Genes involved in growth, metabolism, immunity and neuronal-related functions were found to be differentially hydroxymethylated between wild individuals and their progenitors reared in captivity. Furthermore, phenotypic divergence after a single generation in captivity was associated with changes in DNA hydroxymethylation rather than genetic mutations. This supports our notion that epigenetic modifications play a critical role in animal domestication and selective breeding. Moreover, this thesis provides an extensive list of candidate epigenetic biomarkers related to growth. Their validation has the potential to revolutionize current breeding strategies in aquaculture for the selection of superior phenotypes and to contribute towards the diversification of the industry.

NORD University

ISBN: 978-82-93165-36-1
Trykk: Trykkeriet, Nord universitet
www.nord.no