Intrinsic and environmental influences on DNA methylation and gene expression in fish

Dafni Anastasiadi

Ph.D. thesis 2016

Submitted in partial fulfillment of the requirements for the Ph.D. degree from the Universitat Pompeu Fabra (UPF). This work has been carried out at the Group of Biology of Reproduction (GBR), at the Department of Renewable Marine Resources of the Institute of Marine Sciences (ICM-CSIC).

Thesis supervisor:

Dr. Francesc Piferrer

Professor d'Investigació

INSTITUT DE CIENCIES DEL MAR (CSIC)

upf. Universitat Pompeu Fabra Barcelona

To my family

Acknowledgements

This PhD thesis is the result of the intellectual, technical and psychological support of many people during these years. Firstly, I would like to give my most sincere and profound acknowledgements to my supervisor, Dr. Francesc Piferrer, for everything. For his incredible support, encouragement and trust from even before I started, constantly during these years and at the end. For his advice and guidance from designing experiments to producing figures; from critically studying every single piece of information to extracting the message and transmitting it. I always admired you for your scientific curiosity, enthusiasm and willing to constantly learn. I have immeasurably enjoyed doing this thesis under your supervision and most of everything, I have enjoyed thinking how to bring about your crazy ideas and our exploratory, scientific and intellectual conversations. Millions of thanks from the deepest part of my heart!

This thesis was overseen by a PhD committee that I would like to thank for useful advices: Drs. Roderic Guigó and Luciano di Croce (CRG) and Joan Cerdà (IRTA). A great deal of support and help came from the ex and current Group of Biology of Reproduction members. My enormous thanks to Dr. Noelia Díaz for always being there to answer all my questions, for teaching me fish and lab stuff, for endless conversations about everything, for trips, congresses, samplings, everything that we have shared and I have learnt from you. To Sílvia Joly for teaching me lab protocols, maintaining the lab order, helping with samplings and for performing qPCRs and a great part of the MBS protocol. To Dr. Laia Ribas for helping with my insertion to the group and for all our nice conversations. To my office mates Alejandro Valdivieso and Susana Pla and to the students I have taught and been taught from: Ivet, Paula and Núria. Also to Elvira Martínez and Gemma Fuster for fish care.

During this PhD thesis I performed three short stays and I learnt techniques on which many of the experiments were based. I would like to thank Dr. Paloma Morán, for receiving me at the University of Vigo and, together with Pilar Alvariño, teaching me all the MSAP experimental procedure; and Dr. Andrés Pérez-Figueroa, for the *msap* package and all the analyses. Also, Dr. Christoph Bock from the Center for Molecular Medicine in Vienna and his awesome

team: Dr. Matthias Farlik, Paul Datlinger and Johanna Klughammer, for teaching me everything regarding RRBS and more about fancy cutting-edge methods from the biomedical world. A lot of thanks to Dr. Marco Álvarez Santana, from the Universidad Andrés Bello in Chile, for receiving me in his group for 3 months, for his efforts to carry out Chromatin Immunoprecipitation with my samples and for the moral support. Also to all members of his group, Dr. Marcelo Salazar, J.P. Sepúlveda, Nacho Ramos and the ex-member Nicolas Simonet. Moreover, many thanks to Drs. Marc Vandeputte and Béatrice Chatain, to Alain Vergnet, François Allal, Marie-Odile Vidal and François Ruelle and all the french students and technicians from Ifremer in Palavas for raising the fish of Chapter 3 and for the astonishing industrial chain 2-days-2000 sea bass sampling. Furthemore, I would like to thank Dr. Esteban Ballestar (IDIBELL) for guidance on the paper of Chapter 4.

This thesis would not have been possible without the contribution of many people who I don't even know. I would like to greatly thank all the bioinformatics bloggers, anonymous and not, everyone sharing scripts and R/perl/python/unix codes from the simplest single line command to conversations, advices and debugging in forums and for all open-source packages, data, scripts and programs. Thanks for every analysis I thought for which a code was available and thanks for sharing. Among them, the non-anonymous non-blogger, my personal debugger Federico Quattrocchi.

At the Institute of Marine Sciences, I had the luck of sharing precious moments with amazing people. Among them, Ainhoa, Alba, Amalia, Anabel, Ariadna, Balbina, Batis, Carol, Ciscu, Cristina, Daniela, Fernando, François, Giulia, Guiomar, Isabel, Jacopo, John, José Antonio, Laura, Marc, Marta, Mónica, Montse, Morane, Paula, Raquel, Sonia, Toni, Valerio, Xavi. Very special thanks to Noelia for all the lunches, beers, rakés, concerts, movies etc. Thanks for always being there; to Ulla, the eternal sunshine, for the excursions, mountains, "espectáculos", "la coope". My life in Barcelona is excitingly diverse because of you and your non-stop positiveness; to Federico, for his incredible capacity to make me laugh, smoke, drink coffees and beers and to listen to me for at least 20 seconds and for the puzzle discussions; to Chiara, for listening, helping with everything and for sharing this writing experience a step ahead, like an advanced mirror, for always making me feel well

one way or another and for making me eat; to Valeria, for being a true friend and to mi compañero Claudio, for all the times he was there to see me laugh, cry, scream, fall down or get up. Also to the crack beach volley team, Fran, Eli, María, Ana María, Claudio, John, Fede, Xavi…, for all the nice breaks and trophies!

This thesis is the result of years of studying, growing up, and travelling around the world. To the ones in Barcelona: Αννούλα ήσουν και θα είσαι ανεκτίμητη, Veronica y Despina gracias por ser las mejores compis de piso, a los balcánicos y guapos, a Πένυ, Γιάννη, Ξανθή; thank you Kathrin and Steff for our flat moments; Levent and Sertan for our starts here; moltes gràcies a tots els ballarins de Catalunya amb els quals he compartit experiències màgiques. To a bunch of Greek scientists: Alexandros Triantafyllidis, for your inspirational teaching of evolutionary genomics; Costas Tsiggenopoulos and Giorgos Kotoulas, for helping me "catch the sea bass"; Athina Mavridou, for being my mentor. To the ones in Greece that even if I'm away 10 years are still there for me: Όλοι με τον έναν ή τον άλλον τρόπο με έχετε βοηθήσει όλα αυτά τα χρόνια...ενώ σπουδάζαμε: Βάσω, Λιάνα, Λία, Τόλη, Γιώργο, Φοίβη· ενώ χορεύαμε: Γεννηματά, Βακάλη· ενώ πηγαίναμε σχολείο: Αθηνά, Κώστα, Άθα ή από τότε που γεννηθήκαμε: Άσπα, Αργίνη. Pendant les années de mon master en France, je remercie les meilleurs directeurs de thèse: Bruno Guinand et Erick Desmarais, Nolwenn Quéré; merci à Melek, Rim, Lorenzo, Δήμητρα, Αντρέα, Gaelle. For when I found out I got the PhD scholarship in the middle of the desert in South Africa, thank you Lisa and John. Por hacer fantástica mi estancia en Viena: Cecilia, Adrián, also Katerina, Niko, Rebecca.

Last but not least, I would like to thank my family, τη γιαγιά Μιμόζα που θα ήταν περήφανη, τη θεία Λήδα γιατί είναι η καλύτερη θεία του κόσμου και την Ελμίνα γιατί απλά δεν μπορώ να φανταστώ τη ζωή μου χωρίς εσένα. The words are not enough to thank my parents because apart from my existence, I owe them my liberty, independence and unconditional support for all the crazy things I want to do and all the crazy places I want to go. Thank you for your patience, love and support the years I was growing up and all these years that I am away.

Barcelona, 20 of September 2016

Abstract

Epigenetic mechanisms, influenced by intrinsic and environmental factors are crucial for the regulation of gene expression and, ultimately, the phenotype. The European sea bass is used as a model to study these influences on DNA methylation and the phenotype during early development and later in life. We identify loci altered with age, suggestive of the existence of a piscine epigenetic clock. We show that moderate early developmental temperature increases are associated with genome-wide changes in DNA methylation and with parent-specific responses of genes involved in sexual development. Furthermore, we highlight a genome-wide inverse relationship of gene expression with the DNA methylation of the first intron. Lastly, we provide one of the first empirical demonstrations in support of the neural crest cell deficit hypothesis to explain Darwin's domestication syndrome. Together, these results constitute the most integrative analysis of DNA methylation patterns in a fish species under intrinsic and ecologically relevant contexts.

Resum

Els mecanismes epigenètics són crucials per a la regulació de l'expressió gènica. El llobarro s'utilitza com a model per estudiar aquestes influències sobre la metilació de l'ADN i el fenotip durant el desenvolupament primerenc. S'identifiquen loci alterats amb l'edat, suggeridors de l'existència d'un rellotge epigenètic en peixos. Mostrem que augments moderats de la temperatura a l'inici del desenvolupament s'associen amb canvis en la metilació de l'ADN en tot el genoma i amb les respostes de gens implicats en el sistema de desenvolupament sexual. D'altra banda, destaquem una relació inversa de l'expressió gènica amb la metilació de l'ADN del primer intró. Finalment, oferim una de les primeres demostracions en suport de la hipòtesi del dèficit de cèl·lules de la cresta neural per a explicar la síndrome de domesticació de Darwin. En conjunt, aquests resultats constitueixen l'anàlisi més integradora de la metilació de l'ADN en un peix sota contextos intrínsics i ecològicament pertinents.

Prologue

This PhD thesis was carried out at the Group of Biology of Reproduction (GBR), of the Department of Renewable Marine Resources, Institute of Marine Sciences (ICM-CSIC) in Barcelona, under the supervision of Dr. Francesc Piferrer, and under the PhD Programme in Biomedicine of the Department of Experimental and Health Sciences of the Universitat Pompeu Fabra during the years 2011 to 2016. The aim of this thesis was to investigate the influence of intrinsic, such as ageing, sex, genetic variation and tissue specificity, and environmental factors, such as temperature and farming, on the DNA methylation and gene expression patterns and relationships in fish, using the European sea bass as a model. Part of the results of this thesis could be generalized in other fish species and vertebrates.

The thesis is organized in two blocks with three and two chapters each based on the level of technical resolution and coverage.

Block A: Targeted or low resolution DNA methylation patterns

Chapter 1. Age-related DNA methylation changes in somatic and gonadal tissues in the European sea bass –A piscine epigenetic clock?

Chapter 2. Small ocean temperature increases and stage-dependent changes in DNA methylation in fish.

Chapter 3. The genetic make-up and temperature together affect the DNA methylation of gonadal aromatase and other genes important for European sea bass sexual development.

Block B: Genome-wide patterns of DNA methylation

Chapter 4. Methylation of the first intron shows genome-wide negative correlation with gene expression in tissues of different transcriptomic complexity.

Chapter 5. The first steps to domestication in fish support the neural crest cell deficit hypothesis to explain Darwin's domestication syndrome.

Each chapter corresponds to an article either submitted or in preparation for submission:

Anastasiadi, D., Piferrer, F. Age-related DNA methylation changes in somatic and gonadal tissues in the European sea bass –A piscine epigenetic clock? *In preparation.*

Anastasiadi, D., Díaz, N., Piferrer F. Small ocean temperature increases and stage-dependent changes in DNA methylation in fish. *Submitted.*

Anastasiadi, D., Vandeputte, M., Sánchez-Baizán, N., Chatain, B., Piferrer F. The genetic make-up and temperature together affect the DNA methylation of gonadal aromatase and other genes important for European sea bass sexual development. *In preparation.*

Anastasiadi, D., Esteve-Codina, A., Piferrer F. Methylation of the first intron shows genome-wide negative correlation with gene expression in tissues of different transcriptomic complexity. *In preparation.*

Anastasiadi, D., Piferrer, F. The first steps to domestication in fish support the neural crest cell deficit hypothesis to explain Darwin's domestication syndrome. *In preparation.*

Part of the results from Chapter 1 have been published in the following article:

Anastasiadi D., Diaz N., Piferrer F. 2013. Effects of temperature on global DNA methylation during early development in European seabass. In: Advances in Comparative Endocrinology (eds. Isabel Navarro, Joaquim Gutiérrez, Encarnación Capilla), Publicaciones i Edicions de la Universitat de Barcelona, Barcelona, Spain. 7: 103-106.

And I have also participated in the following article:

Ribas, L., M. Genestar, L. Oller, D. Anastasiadi, N. Díaz, F. Piferrer. 2013. La supervivència, el creixement i la proporció de

mascles són alterats degut a l'estrés ocasionat pel confinament en peix zebra. Biologia de la Reproducció, 2013: 15–19.

And in the following article in preparation:

Ribas L., Vanezis K., Imúez M.A., Sarrà L., Ruiz S., Genestar M., Diaz, N., Anastasiadi D. and Piferrer F. Development of an *in vivo* system in zebrafish to investigate the effects of altered DNA methylation during development. *In preparation.*

Results that are part of this PhD thesis have been presented in the following scientific meetings: Sixth International Symposium on Vertebrate Sex Determination. Hawaii (April 2012); Global Questions On Advanced Biology. Barcelona (July 2012); XIII Jornada de Biologia de la Reproduccio. Barcelona (June 2013); 9th Congress of the Asociacion Iberica de Endocrinologia Comparada. Barcelona (July 2013); 7th International Congress of Comparative Endocrinology. Barcelona (July 2013); Genomics in Aquaculture 3rd International Symposium. Bodø, Norway (September 2013); Epiconcept Workshop. Sant Feliu de Guíxols (October 2013); IV BCN-Chromatin and Epigenetics Meeting. Barcelona (March 2014); Epiconcept Workshop. Las Palmas (May 2014); 10th International Symposium on Reproductive Physiology of Fish. Olhão, Portugal (May 2014); 27th Conference of European Comparative Endocrinologists. Rennes, France (August 2014); V BCN-Chromatin and Epigenetics Meeting. Barcelona (March 2015); 7th International Symposium on Vertebrate Sex Determination. Hawaii (April 2015); Epiconcept Workshop. Dubrovnik (April 2015); Avenços en Recerca en Aqüicultura. Barcelona (June 2015); International Symposium on Genetics in Aquaculture. Santiago de Compostela (June 2015); IV International Symposium on Genomics in Aquaculture. Athens, Greece (April 2016).

During this thesis, I have stayed for a short period of time in foreign research institutions: 1) One week at the Department of Biochemistry, Genetics and Immunology, University of Vigo, under the supervision of Dr. Paloma Morán; 2) Three months at the Center for Molecular Medicine of the Austrian Academy of Sciences, Vienna, Austria, under the supervision of Dr. Christoph Bock; 3)

Three months at the Laboratory of Cell and Molecular Biology, University of Andrés Bello, at Viña del Mar, Chile, under the supervision of Dr. Marco Álvarez Santana.

The research carried out in this thesis was supported by the following research projects awarded to Dr. Francesc Piferrer:

EPIGEN-AQUA (AGL2010-15939): Efectos de factores externos sobre la reproducción y la proporción de sexos en los peces. Estudio de los mecanismos epigenéticos implicados en la respuesta al ambiente. From the Ministerio de Ciencia e Innovación.

Sea Bass Sex (Epi)genetics: Genetic and environmental variation in the epigenetic control of sex determination in sea bass. From AquaExcel (FP7 UE).

EPIFARM (AGL2013-41047-R): Regulación epigenética de la expresión génica en peces cultivados: Efecto de las condiciones de cultivo sobre el metiloma y transcriptoma de los peces. From Ministerio de Ciencia e Innovación.

I was supported by a PhD scolarship linked to the EPIGEN-AQUA project (BES–2011–044860) from the Spanish Government (Oct. 2011-Sep. 2015), by the INEM (Oct. 2015-Mar. 2016) and by a contract linked to the EPIFARM project (Apr. 2016-Sep.2016).

Table of contents

General introduction

1. Epigenetics

1.1. Brief historical perspective and definition

The first scientific evolutionary theory originated from the work of Jean-Baptiste Lamarck in 1809 and postulated the idea of inheritance of acquired traits. This inheritance was long considered false and opposing to the evolutionary theory of Charles Darwin, especially due to the genetic determinism dominating after the Modern Synthesis or Neo-Darwinism. Later in the '50s, Conrad Waddington, demonstrated the inheritance of acquired traits as response to environmental conditions and described the famous developmental epigenetic landscape (Fig. 1; Waddington, 1956), defining epigenetics for the first time as "above genetics". Only in recent years, research on epigenetic regulation and phenomena has expanded dramatically from model species to include ecological and/or evolutionary epigenetics, while an incorporation of the importance of epigenetic mechanisms is demanded into the evolutionary theory. Due to historical reasons, such as the connection of epigenetics with Lamarck's theory or political implications, there has been a controversy for the definition of epigenetics.

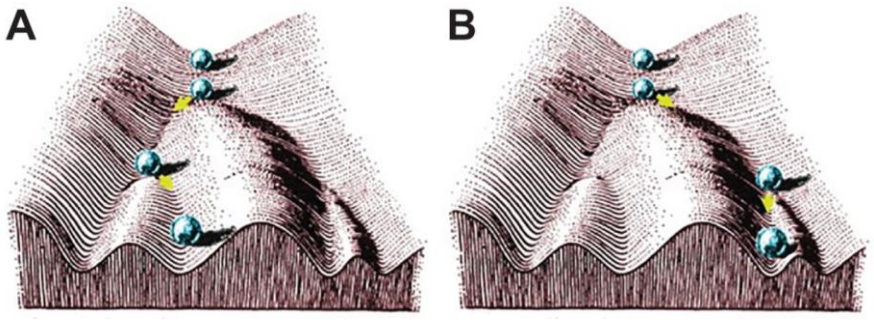


Figure 1. Waddington's developmental epigenetic landscape. This landscape represents the developmental decisions of an organism as "forks" and "valleys" resulting in one phenotypic outcome (A) or another (B). From Noble (2015) originally modified by K. Mitchell.

The definition of Russo et al. (1996) encompasses the most important properties of epigenetic mechanisms, therefore epigenetics is defined here as "the study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence". Thus, epigenetic mechanisms are being studied to answer two types of questions that cannot be explained by changes in the DNA sequence: 1) How two or more cell types can be produced by cells with identical genetic information? More specifically, the cells of an organism may have distinct phenotypic properties, while containing identical DNA sequence information, like for example, a muscle and an epithelial cell, or the same individual as a newborn and at an old age. 2) How the environment, being abiotic environmental factors or biotic, including the ascendants, facilitates the establishment of permanent, semi-permanent or reversible changes that condition the fitness of the phenotypic outcome? The two most stable epigenetic marks are methylation of the DNA and of the histones (Carlberg and Molnár, 2014a), while the epigenome is important for essential functions, such as gene regulation and development in eukaryotes (Lowdon et al., 2016). In addition, epigenomic evolution seems to be parallel to genomic evolution from comparative primate studies (Lowdon et al., 2016).

1.2. Epigenetic phenomena

The first discovered epigenetic phenomena included paramutation in plants, X-chromosome inactivation and genomic imprinting in mammals. In maize, a paramutational allele is able to alter the state of the heteromorph, a state inherited to the following generations (Brink, 1956), while a similar regulation of alleles was found mediated by small interfering RNAs in mice (Rassoulzadegan et al., 2006). The X-chromosome inactivation is required for dosage compensation of genetic material in female mammals which contain two X chromosomes compared to the one of males. The paternal X is inactivated by a complex network of epigenetic mechanisms working in a coordinated manner and is maintained stably inactive throughout life (Kiefer, 2007). This is an extreme case of a whole chromosome being parentally imprinted and inactivated. There are, nevertheless, cases of genomic imprinting of gene clusters that are

silenced in the offspring depending on the parent-of-origin (Feinberg et al., 2002). The genomic imprinting is regulated by epigenetic mechanisms tied to the X-inactivation mechanisms (Kiefer, 2007).

1.3. Epigenetic mechanisms

There are three major epigenetic molecular mechanisms widely accepted as such which are the methylation of cytosines in the DNA, the modifications of the histones and the regulatory non-coding RNA (Fig. 2). Here, we briefly present these epigenetic mechanisms, although in the rest of this thesis, we will mainly focus on DNA methylation.

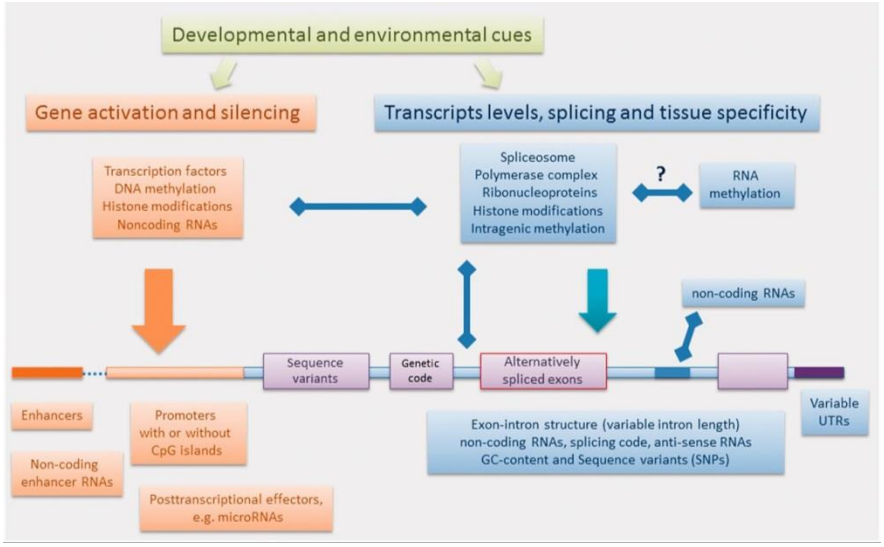


Figure 2. Overview of gene regulation in eukaryotes as response to developmental and environmental cues. Epigenetic modifications, such as DNA methylation, histone modifications and non-coding RNAs, are involved in gene activation and silencing but may also affect the gene expression by modifying the transcript levels and splicing. The central boxes schematically represent a gene and its main regulatory elements. Two-way arrows show two-way communications, feed-back and feed-forward loops. From Maleszka et al. (2014).

1.3.1. DNA methylation

One of the best studied epigenetic mechanisms is DNA methylation. Methylation can occur in two of the four nucleotides of DNA, cytosine and adenine. It is the process by which a methyl-group (CH_3) is transferred from a methyl donor, *S*-adenosyl-L-methionine (SAM), to the fifth position of a cytosine converting it to 5-methylcytosine (^{5m}C) or to the sixth position of an adenine converting it to N6-methyladenine (Grosjean, 2013; Pfeifer, 2016; Ratel et al., 2006). ^{5m}Cs are the most studied modifications and present in most species. The transfer of CH_3 is catalyzed by a group of enzymes called DNA (cytosine-5) methyltransferases (DNMTs) that in vertebrates belong to one of the two families: *dnmt1* or *dnmt3* (Goll and Bestor, 2005). *Dnmt1* is responsible for maintaining the patterns of DNA methylation across cell divisions by adding a methylgroup in the hemi-methylated CpGs (Goll and Bestor, 2005). *Dnmt3s* are responsible for *de novo* methylation, implicated in the process of maintenance of DNA methylation as well (Chen et al., 2003). There are several isoforms of *dnmt3a* and *dnmt3b* identified in vertebrates including fish (Campos et al., 2012; Chen et al., 2003).

1.3.2. Histone modifications

Chromatin modifications constitute a second major class of epigenetic mechanisms. The smallest unit of chromatin is the nucleosome, which consists of an histone octamer, including two of each histone types: H2A, H2B, H3 and H4, while a 147 bp DNA sequence surrounds the histone complex (Luger et al., 1997). The histones may contain post-translational modifications, the most common ones being acetylation, methylation, phosphorylation and ubiquitylation (Kouzarides, 2007; Turner, 2005). These modifications are able to alter gene expression, either by changing the chromatin structure or by recruiting enzymes with specific binding domains. Histone modifications are added by “epigenetic writers”, such as histone acetyltransferases, are deciphered by “epigenetic readers”, e.g., proteins with chromodomains, and are wiped out by “epigenetic erasers”, like histone deacetylases. In addition to regulating the transcription, chromatin modifications

participate in basic DNA functions, such as replication, repair and recombination (Bannister and Kouzarides, 2011). Importantly, histone modifications and DNA methylation patterns act in interplay to establish gene expression and this relationship is bidirectional (Cedar and Bergman, 2009). The enzymes that modify chromatin are sensitive to changes in environmental and metabolic cues, therefore act as sensors through which environmental conditions may influence gene expression (Turner, 2011).

1.3.3. Non-coding RNAs

Non-coding RNAs (ncRNAs) gain increasing significance in the regulation of gene expression, although less studied than DNA methylation and histone modifications. ncRNAs are transcribed from genomic sequences and do not encode for a protein. They generally have been assigned to two categories based on an arbitrary length threshold which are 1) the long ncRNAs and 2) the small ncRNAs, that include micoRNAs (miRNAs) and small interfering RNAs (Keller and Bühler, 2013; Mattick and Makunin, 2006). ncRNAs have been directly or indirectly implicated in epigenetic phenomena such as transcriptional and transposon silencing, parental imprinting, formation of heterochromatin and regulation of sex chromosomes for dosage compensation (Costa, 2008; Mattick, 2007). In a wide range of species, ncRNAs have been attributed important functional roles and suggested to form part of a complex network for fine developmental regulation (Mattick, 2007). Broadly, the ways ncRNAs act are by preventing DNA-binding factors to bind, as scaffolds for proteins to interact and as recruiters of chromatin modifying enzymes to the DNA (Carlberg and Molnár, 2014a). Even though ncRNAs constitute an independent mechanism of regulation of gene expression, they act synergistically with other epigenetic mechanisms as sensors of the environmental signals and sequence guiders of stable chromatin modifications. They also act as downstream regulators of gene expression being regulated themselves by DNA methylation (Yan, 2014).

1.4. DNA methylation and gene expression

In vertebrates, ^{5m}Cs are commonly found in a nucleotide context where the cytosine is followed by a guanine, that is a CpG context. Typically, vertebrate genomes are depleted in CpG dinucleotides due to the high deamination rates of methylcytosine being converted to thymine (Colot and Rossignol, 1999; Coulondre et al., 1978). This CpG depletion is evident in mammalian genomes, but less striking in fish genomes (Schübeler, 2015). However, there are genomic regions of variable length presenting high CpG density and low DNA methylation that are called CpG islands (Fig. 3a; Bird, 1980; Colot & Rossignol, 1999; Deaton & Bird, 2011). In a distance of 0-2 kb adjacent to the CpG islands are located the CpG shores (Irizarry et al., 2009) and in a distance of 2-4 kb adjacent to shores are located the CpG shelves, while the rest of the genome is called open sea in terms of CpG sites (Irizarry et al., 2009; Rechache et al., 2012; Sandoval et al., 2011). While the overall levels of genome-wide methylation are estimated to be 70-80% in mammals, the CpG islands are generally unmethylated and overlap with promoters in the 60-70% of genes of the human genome (Illingworth and Bird, 2009; Larsen et al., 1992; Law and Jacobsen, 2010; Saxonov et al., 2006; Weber et al., 2007). In addition to the promoter regions the DNA methylation in enhancer genomic regions, where the transcription factors bind and interact with the promoter, receive increasing importance (Bogdanović et al., 2016; Ong and Corces, 2011; Stone et al., 2015).

DNA methylation of cytosines is generally thought to regulate gene expression by two mechanisms: 1) by preventing the binding of transcription factors that activate gene expression, and 2) by recruiting methyl-binding proteins, which inhibit transcription through the stabilization of the nucleosomes (Gilbert and Epel, 2008; Moore et al., 2013). CpG islands overlap with the promoters of genes that are constitutively expressed and are linked to transcriptionally active genes (Fig. 3b), while they also overlap with 40% of genes expressed in a tissue-specific manner (Guenther et al., 2007; Illingworth and Bird, 2009; Larsen et al., 1992; Tazi and Bird, 1990; Weber et al., 2007; Zhu et al., 2008). Although this is the general model for CpG-rich genomic regions, evidence from whole-genome single-nucleotide recent studies have revealed more

complex relationships between DNA methylation of CpG sites and transcription factor binding in CpG-poor regulatory elements; relationships that still needs to be elucidated (Bock, 2012; Jones, 2012; Schübeler, 2015).

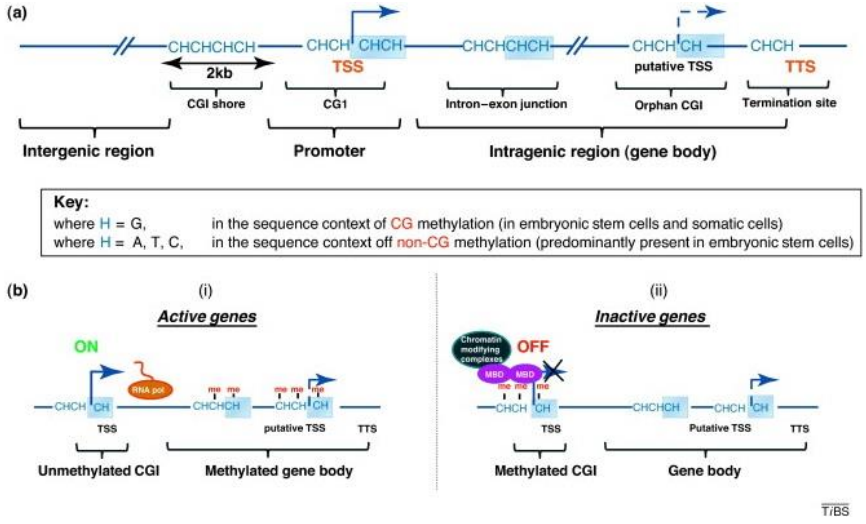


Figure 3. Overview of the regulation of gene expression by DNA methylation. Important gene and genomic features associated with this regulation are shown (a) as well as the relationship between DNA methylation in genomic features and the expression state of the gene (b). CGI,CpG islands; TSS, Transcription Start Sites; MBD, Methyl-CpG-Binding Domain. From Ndlovu et al. (2011).

1.5. Methods to study DNA methylation

The DNA methylation levels are assessed by technologies which can be categorized at three levels.

First level: the ^{5m}Cs have to be identified and distinguished from the Cs. There are three general approaches used for this identification and distinction that can be used isolated or in combination:

1) The enzymatic approach. Some restriction enzymes have differential sensitivity to ^{5m}Cs and can therefore be used in order to distinguish ^{5m}Cs from Cs. Typical examples are *Msp*I

and *Hpa*II, which are isoschizomers that both recognize the sequence 5'-CCGG-3', but *Msp*I only cuts when the internal C is methylated in both DNA strands and *Hpa*II only cuts when the external C is methylated in one of the two DNA strands.

2) The affinity approach. Specific antibodies against ^{5m}C or recombinant proteins that contain a methyl-CpG binding domain (MBD) can be used in order to enrich for the methylated fraction of chromatin.
3) The bisulfite treatment. Bisulfite treatment involves a chemical reaction that converts unmethylated Cs into uracils while ^{5m}Cs remain intact, allowing to record the DNA methylation status and, during the downstream steps of the process, to recall it.

Second level: there are variable degrees on the resolution of the subsequent methylation profiling, which can be globally grouped into:

1) Low resolution. Information on global ^{5m}Cs content can be obtained, however, the exact genomic locations of the ^{5m}Cs remain unknown.
2) Medium resolution. The ^{5m}Cs can be localized in an approximate context.
3) Single nucleotide resolution. The exact genomic location of both ^{5m}Cs and Cs is obtained.

Third level: the methylation profiling approaches may vary in the extent of the genome that can be interrogated which can roughly be:

1) Locus-specific. ^{5m}Cs are measured only in target regions of interest.
2) Genome-wide. ^{5m}Cs are profiled across a representative portion of the genome.
3) Whole-genome. All ^{5m}Cs of the entire genome are identified.

A summary of the most commonly used methods to measure DNA methylation is presented in Table 1 together with the assignment of each one to corresponding level of categorization. Even though the gold standard of next generation sequencing (NGS) methods remains the Whole Genome Bisulfite Sequencing (WGBS), its cost is still usually high for use in non-model vertebrates with a large genome, while the balance between genomic coverage and number

of samples to query often leans towards the second. Reduced Representation Bisulfite Sequencing (RRBS; Gu *et al.* 2011), recently modified for use without a reference genome (Klughammer et al., 2015), and Bisulfite RAD-seq (Trucchi et al., 2016) are both suitable methods for querying the DNA methylation status of cytosines at the genome-wide level in non-model species.

Table 1. Overview of commonly used techniques to measure DNA methylation

^{5m}Cs distinction		Resolution: Low	Resolution: Medium	Resolution: High	Genome coverage
	Restriction enzymes			*Hpa*II-PCR	**LS**
		MSAP, RLGS	CHARM	RRBS, MRE-seq	**GW**
					WG
	Affinity enrichment		MeDIP/MBD-PCR		**LS**
			MeDIP/MBD-chip		**GW**
		HPLC	MeDIP/MBD-seq		**WG**
	Bisulfite treatment		Infinium BeadChip	MSP, Bis-PCR, Sanger BS	**LS**
				RRBS, BisRAD-seq	**GW**
				WGBS	**WG**

Methods are divided based on the approach used to distinguish ^{5m}Cs from Cs (left), the level of resolution (right) and the genomic coverage (top) they provide. Abbreviations: MSAP, Methylation Sensitive Amplified Polymorphism; RLGS, Restriction Landmark Genomic Scanning; HPLC, High Performance Liquid Chromatography; CHARM, Comprehensive High-throughput Arrays for Relative Methylation; MeDIP, Methylated DNA ImmunoPrecipitation; MBD, Methyl-CpG-Binding Domain; RRBS, Reduced Representation Bisulfite Sequencing; MRE, Methyl-sensitive Restriction Enzyme; MSP, Methylation Specific PCR; BS, Bisulfite Sequencing; BisRAD, Bisulfite Restriction site Associated DNA; WGBS, Whole Genome Bisulfite Sequencing; LS, Locus Specific; GW, Genome-Wide; WG, Whole Genome.

The NGS approaches, such as WGBS or RRBS, produce a large amount of data that have to be treated bioinformatically in order to extract biologically meaningful conclusions. The NGS data processing starts with quality controls of the sequencing and filtering of the low quality sequences. Unique adapters have been added during library preparation in order to be able to distinguish samples, so at this bioinformatic step they have to be removed from the data because their presence can influence the downstream steps of the procedure. Following the adapter trimming, the sequences are aligned against the *in silico* bisulfite converted genome. Since bisulfite treatment converts the Cs into Ts after PCR amplification, many genomic sites stop matching with the reference genome. In addition, after PCR amplification, the complementary strand contains As instead of Gs in positions where Cs have been converted in the other strand. Therefore, alignments have to be cautiously performed in order to take these considerations into account against *in silico* bisulfite converted genomes. The last minimum step is the extraction of methylation information, or methylation calling, for each C position. In most cases, the data come from heterogeneous cell populations, therefore the methylation status of a C is calculated according to the proportion of ^{5m}Cs to the total number of times this genomic position is present, or the sum of ^{5m}Cs and Cs, and is expressed as a percentage.

1.6. Main intrinsic influences on the epigenome

1.6.1. Genetic variation (SNPs) and epigenetic variation

Epigenetic variation can be influenced by genetic variation. This effect is exerted by two general ways, one direct and one indirect. The single nucleotide polymorphisms (SNP) are naturally common across the genomes and the most abundant case of polymorphism is the C-to-T. These SNPs can directly affect DNA methylation patterns of the genome, since the CpG sequence context where methylation commonly occurs and has a regulatory role in animal genomes, is directly lost because the C is absent (Shoemaker et al., 2010). This nucleotide transition may also affect the binding motif of a transcription factor, therefore, indirectly alter gene expression

and DNA methylation in a complex- and context- dependent manner (Gutierrez-Arcelus et al., 2013). SNPs are also able to act in *cis* with CpG sites and influence their methylation status. These CpG sites have been named methylation quantitative trait loci (meQTLs) and are associated with variation of other epigenetic mechanisms as well as gene expression (Banovich et al., 2014; Hannon et al., 2016; Teh et al., 2014). The exact relationships between SNPs and methylation of CpGs are far from being well understood, but these studies provide evidence of complex interconnections between genetic and epigenetic variation.

1.6.2. Epigenetic changes across tissues and with age

During development, the cells acquire specific gene expression patterns that are maintained throughout cell divisions and define the cell identity. This flexibility of cells with the same genome to express and maintain various cellular phenotypes is an ideal ground for research on epigenetics. The best known mechanisms involved in the processes of cell differentiation and maintenance are the Polycomb group and the Trithorax group protein, first studied in *Drosophila* (Brock and Fisher, 2005). DNA methylation, however, has also recognized roles in cellular differentiation and subsequently tissue-specific phenotypes (Khavari et al., 2010). In addition, inter-individual profiles of DNA methylation are more similar than inter-tissue profiles in the same individual, suggesting tissue-specific signatures of DNA methylation (Choudhuri et al., 2010; De Bustos et al., 2009). A milestone was achieved recently by the NIH Roadmap Epigenomics Consortium, which described the tissue-specific signatures of 111 human epigenomes now available to serve as reference for studies in other organisms (Kundaje et al., 2015).

Another intrinsic factor during lifetime connected with epigenetic changes is the normal senescence. Despite the perception of epigenetic modifications as stable, it is now well established that epigenetic modifications are susceptible to age-related changes (Aguilera et al., 2010; Calvanese et al., 2009). Ageing of an organism has been linked to a drift of methylation patterns, comprised of whole-genome hypomethylation and site-specific

hypermethylation (Ghosh and Zhou, 2015; Issa, 2014; Jung and Pfeifer, 2015; Muñoz-Najar and Sedivy, 2011). These changes result from the coordinated action of factors both intrinsic, i.e., genetic, and extrinsic, i.e., environmental and developmental conditions, while they can be specific of the cell type, tissue and genomic context (Feil and Fraga, 2012; Huidobro et al., 2013). Although ageing is thus a complex process, Horvath (2013) has developed a tissue-independent predictor of biological age, which may differ from the chronological age, for humans, which in addition has been proved to function also in wild whale species (Polanowski et al., 2014).

1.7. Epigenetics during periconception and epigenetic inheritance

There are two types of epigenetic inheritance. One form occurs during cell divisions and is therefore a mitotic inheritance of gene expression regulation patterns. The most obvious example is related to cell identity, meaning that undifferentiated cells acquire a lineage-specific epigenetic pattern that is maintained throughout cell divisions, e.g., a muscle cell over cell divisions produces muscle and not liver cells (Probst et al., 2009). Epigenetic mitotic inheritance serves as a cellular memory of cell identity.

The second form of epigenetic inheritance is meiotic inheritance, meaning that an epigenetic state may be inherited to the next generation and after, a form of inheritance with paramount ecological and evolutionary implications (Day and Bonduriansky, 2011; Jablonka, 2005). Indeed, inheritance of epigenetic variation across generations has been demonstrated in plants (Bender and Fink, 1995; Johannes et al., 2009; Reinders et al., 2009; Verhoeven et al., 2010), prokaryotes (Lim and van Oudenaarden, 2007; Maamar et al., 2007), viruses (Stumpf et al., 2002) and eukaryotes (Acar et al., 2005; Levy et al., 2012), including vertebrates (Daxinger and Whitelaw, 2012; Greer et al., 2011).

In many eukaryotes, the parent-to-offspring transition involves at least one reprogramming event so that the zygote cells regain

totipotency (Fig. 4A). However, there are specific loci that are able to escape the reprogramming event and pass to the offspring. The exact events have been studied only in a few species, e.g., mouse (Messerschmidt et al., 2014), zebrafish (Jiang et al., 2013; Potok et al., 2013) and *Arabidopsis* (Kawashima and Berger, 2014; Yang et al., 2011), but are thought to be indicative of the general pattern in eukaryotes (Mills et al., 2015). In zebrafish, the paternal methylome is inherited by the sperm while the maternal methylome undergoes both methylation and demethylation in order to resemble the sperm pattern during early development (Ci and Liu, 2015; Jiang et al., 2013; Potok et al., 2013).

Although the prevalence and endurance of transgenerational epigenetic inheritance is still in debate, in bacteria, protists, fungi, plants and animals, environmentally-induced phenotypes inherited throughout multiple generations by means of DNA methylation, chromatin modifications and other epigenetic inheritance systems have been reported (Gilbert and Epel, 2008; Jablonka and Raz, 2009). These phenotypes are manifested even in the absence of the factor that initiated them in the first place.

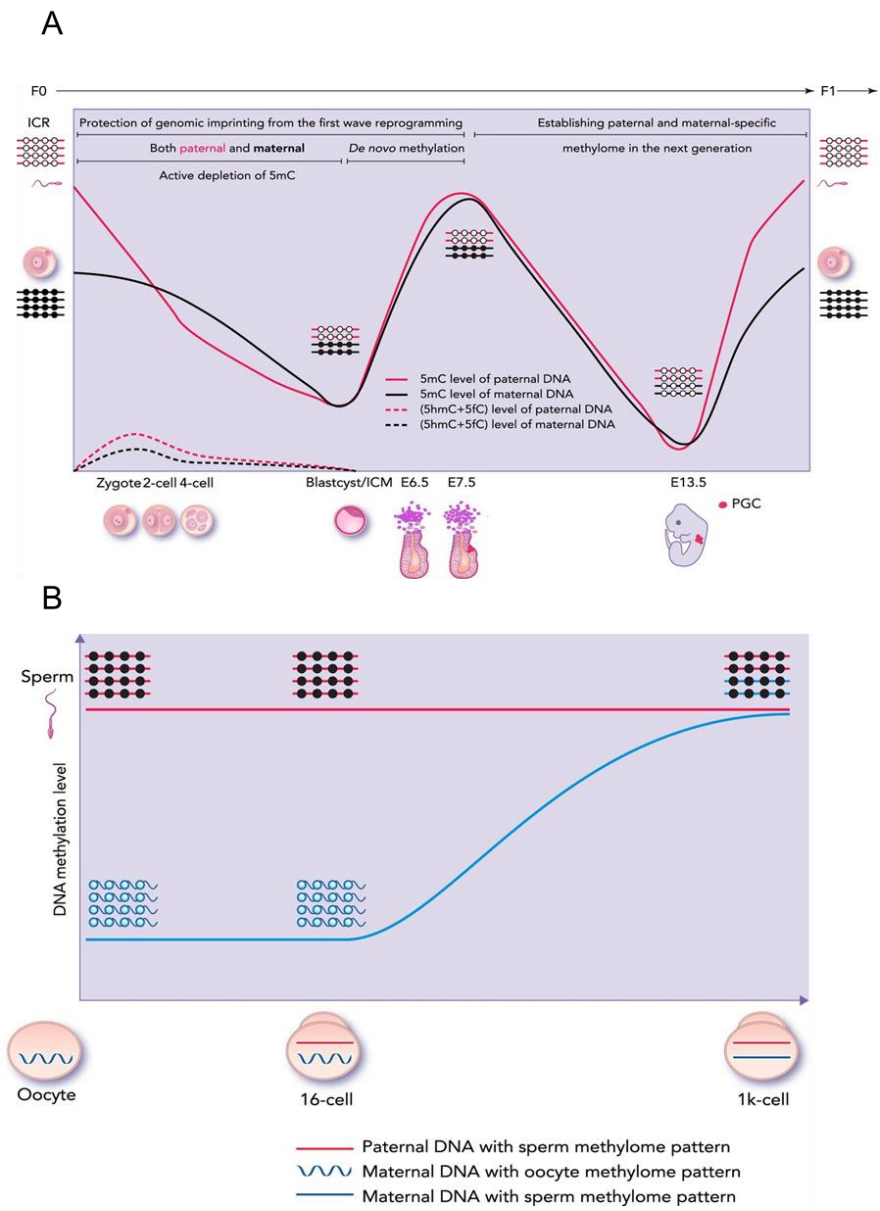


Figure 4. Patterns of reprogramming and inheritance of DNA methylation patterns in mouse (A) and in zebrafish (B). From Ci and Liu (2015).

2. Genetic and epigenetic variation and the rise of the phenotype

2.1. Phenotypic plasticity

Phenotypic plasticity, broadly defined, "is any change in an organism's characteristics in response to an environmental signal" as suggested by Schlicthing and Smith (2002). This response to environmental stimuli includes changes in biochemical, physiological, behavioural, morphological or life-history traits (Schlichting and Smith, 2002; Wright and Turko, 2016). The phenotype is therefore the result of the interaction between the genotype and the environment (GxE). However, the GxE interactions are not static but rather depend on the developmental history of an organism (Brandon and Nijhout, 2006; Gilbert and Epel, 2008). A plastic response occurs when the genotype is permissive to develop a range of phenotypes dependent on instructive environmental inputs. These plastic responses may differ in magnitude and in rank order and can be:

1) Continuous. The genotype defines a range of continuous trait properties called reaction norms, i.e., growth.

2) Discontinuous. Potential phenotypic outcomes are mutually exclusive, like in the case of sex determination, when an organism will develop either as male or as female.

Importantly, plasticity allows organisms to face a variety of environments while maintaining their performance and fitness (Beaman et al., 2016). Under the ecological and evolutionary context, i.e., in the long term, phenotypic plasticity is considered an adaptive response to environmental changes (Miner et al., 2005; Schlichting and Wund, 2014; Sultan, 1995).

2.2. Genetic, epigenetic and phenotypic variation

Phenotypic plasticity is rooted in genetic variation. There is, however, a part of phenotypic variation that cannot be explained by the genetic variation alone. Epigenetic variation, defined as variation in the epigenetic state or profiles at the interplay of GxE interactions has been suggested as candidate to explain part of the phenotypic variation (Fig. 5; Massicotte et al., 2011). Epigenetic variation is at the interplay of GxE interactions because it can be altered by the environment (Teh et al., 2014) and at the same time is regulated by the genetic variation and vice versa (Day and Bonduriansky, 2011; Furrow and Feldman, 2014). In addition, it affects the gene expression variation and thus ultimately the phenotypic variation. Due to its effects on phenotypic variation, epigenetic variation may directly be selected by natural selection (Herman et al., 2014), a case that has been shown in bacteria and plants (Adam et al., 2008; Balaban et al., 2004; Gilbert and Epel, 2008; Paun et al., 2010). In this context, the environment is also itself a source of variation that increases the fitness of an organism (Gilbert and Epel, 2008). Therefore, epigenetic mechanisms form part, together with the genome, of the molecular mechanisms that explain phenotypic plasticity. Epigenetic variation may be adaptive depending on the rate of environmental changes, the predictability of these changes and the costs associated with this variation (Furrow and Feldman, 2014; Herman et al., 2014; O'Dea et al., 2016). Epigenetically-mediated phenotypic plasticity is suggested to be also maladaptive in some cases, termed epigenetic trap, or to serve as a buffer in rapidly changing environments such as predicted by the anthropogenic climate change scenarios (Consuegra and Rodríguez López, 2016; O'Dea et al., 2016; Piferrer, 2016).

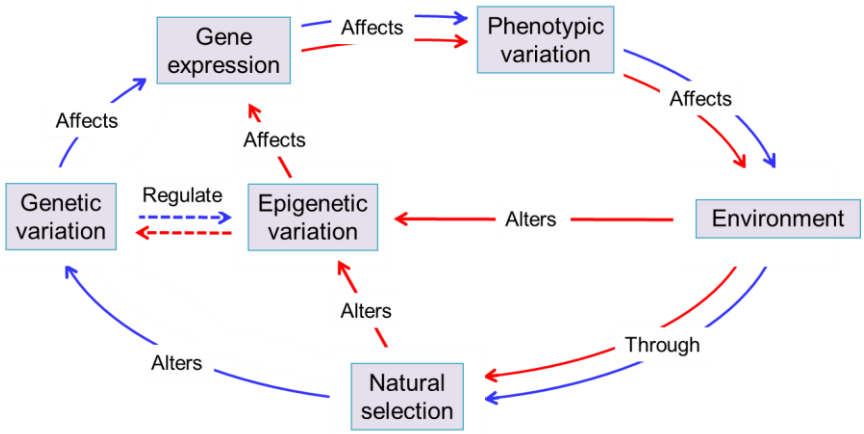


Figure 5. Integration of epigenetic variation in an evolutionary ecological context. Epigenetic variation at the interplay of GxE interactions may be regulated by genetic variation and vice versa, while it can be altered by the environment and affects the gene expression. Epigenetic variation may also be target of the natural selection. Modified from Bossdorf et al. (2008).

2.3. Environmental influences during early development – Developmental plasticity

Phenotypic plasticity observed during the early developmental stages of an organism is called developmental plasticity. The prezygotic and early postzygotic developmental environments are able to induce plastic responses that permanently alter the phenotype throughout adulthood (Beaman et al., 2016). Therefore, the developmental stages are crucial for the determination of the phenotypic outcome. The epigenetic marks are generally reversible and can be altered during life depending either on genetic or environmental factors. However, persistent epigenetic marks are established during the sensitive developmental stages, because the rates of cell division are high and these marks are transmitted to all daughter cells (Toraño et al., 2016). These persistent epigenetic marks stably define the phenotype along life (Faulk and Dolinoy, 2011). Indeed, epigenetics may explain at the molecular level the predictive adaptive responses, but in the case of phenotype-environment mismatch it may confer fitness disadvantage to the conditioned organism; a case increasingly studied to explain human

diseases under the Developmental Origins of Health and Disease framework (Hanson and Skinner, 2016).

2.4. Reversible plasticity—Acclimation

Phenotypic plasticity may also occur during juvenile or adult life stages and is called acclimation or reversible plasticity. Acclimation allows to maintain stable physiological rates in fluctuating environments over periods of time, compensating for the potential negative effects of these fluctuations (Beaman et al., 2016; Gabriel et al., 2005; Kelly et al., 2011). Nevertheless, developmental and reversible plasticity are not two clearly separated mechanisms of plasticity, but are rather interconnected. Actually, the capacity of acclimation may be conditioned during development, resulting in individuals being more or less capable to buffer the effects of changing environmental conditions later during life depending on the cues received early. For example, this plastic acclimation capacity was shown in mosquitofish that presented contrasting degrees of thermal acclimation when they were born in cold and warm temperatures (Seebacher et al., 2014). Seasonal variations of epigenetic marks exist in natural populations (Simonet et al., 2013) and epigenetic mechanisms have been suggested to mediate, at least in part, the acclimation procedures (Beaman et al., 2016; Horowitz, 2014).

2.5. Integration of environmental information

Berger et al. (2009) suggested an operational model on how the environmental information may get integrated into the genome and maintained by means of epigenetic mechanisms (Fig. 6). Their model implicates three subsequent levels of action:

1) An environmental signal that enters the nucleus, called epigenator, and is able to provoke a subsequent cascade of events. Epigenators can be environmental abiotic factors, such as temperature, or biotic signals, like differentiation signals.

2) A molecular agent, like a DNA-binding protein or ncRNAs, that stimulated by the epigenator, establishes a chromatin context at a precise genomic location.
3) Epigenetic maintenance, a process by which the signal gets permanently integrated into the genome and is maintained during subsequent cell divisions. These involve the typical epigenetic mechanisms, like DNA methylation, histone modifications or histone variants.

This process of integration may have temporal consequences on the phenotype, influencing the transcription status or the DNA repair mechanisms, or long-term consequences on the phenotype, establishing a permanent status of chromatin conformation. (Berger et al., 2009; Piferrer, 2013; Turner, 2007).

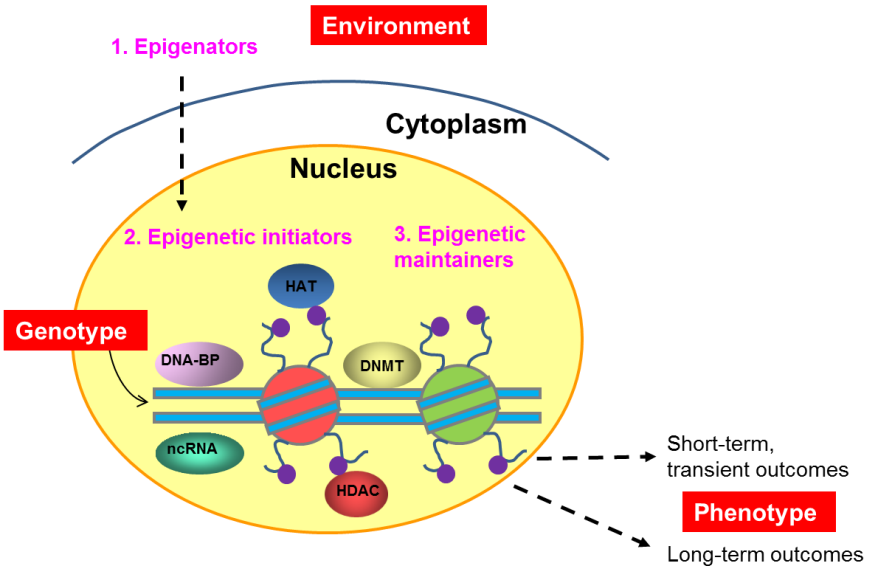


Figure 6. Schematic overview of the integration of the environmental information into the genome. Adapted from Piferrer (2013), in turn based on Berger et al. (2009) and Turner (2007).

2.6. Environmental influences on the phenotype through epigenetic mechanisms

Here, we provide some examples of environmental influences on the phenotype, focusing on nutrition, stress and temperature and only in relation to DNA methylation which is the focus of this thesis. Nutrition and stress include classic examples of epigenetic regulation while temperature is also of interest in this thesis. However, the list of environmental factors influencing the phenotype through epigenetics is not exhaustive, since other chemical and physical environmental agents such as pollutants or xenobiotic compounds also alter the epigenetic status, while the responses may implicate histone modifications or ncRNAs as well (Feil and Fraga, 2012).

Nutrients, especially the ones available during early development, have a decisive role on phenotype from insects to mammals. In honeybees, queens show a completely distinct phenotype than sterile workers, although their genetic information may be identical. These phenotypes are linked to lower levels of methylation in the brain of queens compared to workers (Lyko et al., 2010) and differential stage-dependent methylation between the castes' larvae (Shi et al., 2013). In mice, a dominant allele of the *Agouti* gene produces a shiny brown phenotype when the mother's nutrition includes methyl donors during pregnancy, while an obese yellow phenotype is produced if the mother's diet is not enriched with methyl donors (Waterland and Jirtle, 2003). Between these cases, there are differences in the methylation of *Agouti*, which can be inherited transgenerationally (Morgan et al., 1999). Other studies have shown that the diet of a mother is able to persistently alter the methylation status of genes related to metabolism and stress in rats (Burdge et al., 2007; Lillycrop et al., 2005, 2008). In humans, food restriction in women during pregnancy can lead to hypomethylation events related to obesity and diabetes, lasting up to her grandchildren, as it was shown in the famous study of the Dutch Hunger Winter of 1944-1945 (Heijmans et al., 2008).

Stress has also persistent effects on phenotype through epigenetic mechanisms. In the holly (*Ilex aquifolium*), browsing accessibility differentiates the prickly leaf type, which is highly methylated when

compared to the nonprickly leaf type (Herrera and Bazaga, 2013). In rats, grooming the pups by their mother during the first week of life is decisive for the establishment of the methylation status of the glucocorticoid receptor (GR), which influences the number of GRs in the adult brain and ultimately the capacity to deal with stressful situations (Weaver et al., 2004, 2007).

Temperature, long known to be crucial for the determination of diverse phenotypic outcomes, can exert its influences via epigenetic mechanisms across species. Changes in DNA methylation patterns, as sensors of temperature, participate, together with alterations of histone modifications, in the regulation of flowering in plants that require a period of cold in order to complete this process (Khan et al., 2013). In some animals, including species of turtles, reptiles and fish, a key discontinuous phenotype determined by temperature is sex. This is a process, called environmental sex determination, by which temperature during specific embryonic or larval stages functions as a threshold for gonadal differentiation into testis or ovary. Temperature-dependent sex determination is mediated by differential DNA methylation of the aromatase promoter, the enzyme that converts androgens into estrogens, between males and females as it was shown first in the European sea bass (Navarro-Martín et al., 2011), and also in alligators (Parrott et al., 2014) and turtles (Matsumoto et al., 2013). Between males and females, differential methylation was also shown in other fish species, like the Japanese flounder (Wen et al., 2014), the eel (Zhang et al., 2013) and the Nile tilapia (Chen et al., 2016; Sun et al., 2016). Importantly, temperature during early development of the half-smooth tongue sole (*Cynoglossus semilaevis*) is able to sex-reverse females through changes in the DNA methylation of the sex chromosomes and, further, these changes can be passed down to the next generation even in the absence of elevated temperature, with potential population consequences due to altered sex ratios (Shao et al., 2014).

3. Epigenetics in animals and the European sea bass model

3.1. Epigenetics in domesticated animals and prediction of phenotype

In domesticated animals, complex traits can be recorded across generations. The domesticated animals are therefore good models to study the epigenetic inheritance mechanisms and the translation of the influences of environmental and intrinsic factors into complex phenotypic traits. Genomic imprinting has been the most studied mechanism in domesticated animals, since paternally imprinted genomic loci could explain a significant part of the phenotypic variance observed (O'Doherty et al., 2015; Triantaphyllopoulos et al., 2016). The DNA methylomes of livestock are only recently starting to being assessed. There are now available the methylome of sheep (Cao et al., 2015; Couldrey et al., 2014), pig (Choi et al., 2015), chicken (Nätt et al., 2012), the half-smooth tongue sole (Shao et al., 2014), rainbow trout (Baerwald et al., 2016)(Baerwald et al., 2016) and the Pacific oyster (Gavery and Roberts, 2010, 2013, 2014).

In vertebrates, biomarkers have been used as indicators of environmental pollution (Monserrat et al., 2007) or animal health, including endocrine, immune, nutritional and metabolic processes (Warne et al., 2015). Epigenetic marks, and specifically DNA methylation marks, have been used across species, as diagnostic biomarkers of biological age (Benayoun et al., 2015). In domesticated animals, epigenetic biomarkers have only been suggested recently as candidates with extreme potential to predict the phenotypic outcome (Ibeagha-Awemu and Zhao, 2015; Moghadam et al., 2015; Piferrer, 2013).

3.2. Fish as models for epigenetic studies

Fish represent good animal models for the purposes of studying the biological and environmental influences on epigenetic mechanisms and the possible phenotypic responses during early development

and later in life (Labbé et al., 2016; Li and Leatherland, 2013). There are almost 34000 recognized fish species (Eschmeyer et al., 2016), while they are highly diverse in terms of physiology, behavior and life history, as well as are the habitats they occupy. Since they are ectothermic animals in contact with the water through their skin, gills and gastrointestinal tract, they are susceptible to direct environmental influences. Extremely susceptible are the fish eggs and larvae during all developmental stages, since they are completely exposed directly to the ocean environment, therefore being ideal models for developmental epigenetic studies. Well studied environmental factors influencing the fish phenotype include abiotic factors such as temperature, salinity, diet, pollutants, as well as biotic factors such as predator stress, parental care and population density (Jonsson and Jonsson, 2014; Pittman et al., 2013). In addition, fish exhibit phenotypic plasticity in a variety of easily measurable important functional traits, such as reproduction, metabolism, muscle growth and skeletal development. For example, temperature induces phenotypic plasticity of muscle growth in Senegalese sole (*Solea senegalensis*). The underlying mechanism is the hypo-methylation of the promoter of myogenin (*myog*) when larvae are raised at increasing temperatures, together with higher *myog* expression, muscle growth and lower expression of the DNA (cytosine-5-)-methyltransferase 1 (*dnmt1*; Campos et al., 2013a). Regarding global DNA methylation, the genome of fish is twice more methylated than the genome of birds, reptiles and mammals (Jabbari et al., 1997). An interesting feature highlighted by a latitudinal study of fish is that there is an inverse relationship between DNA methylation and body temperature, with decreasing levels of methylation from polar to tropical fish (Varriale, 2014; Varriale and Bernardi, 2006).

3.3. The European sea bass model

3.3.1. Systematics and general biology

The European sea bass (*Dicentrarchus labrax*) is a marine coastal fish which belongs to the class of Actinopterygii, order of Perciformes, suborder of Percoidei and family of Moronidae. The distribution of sea bass expands from the Northeastern Atlantic, including the range from Senegal to Norway, to the Mediterranean and the Black Sea (Kottelat and Freyhof, 2007). The sea bass lives in a broad range of environments, is euryhaline and eurythermic, tolerating a range of temperatures from 2ºC to 32ºC (Barnabé, 1980). It can be found at 100 m of depth but usually it frequents coastal inshore waters, estuaries, brackish water and freshwater to a depth of 10 m (Jensen et al., 1998). Sea bass has been reported to live up to 15–25 years (Freyhof and Kottelat, 2008).

3.3.2. Life cycle

An overview of the life cycle of sea bass is shown in Figure 7. Spawning happens annually and is pelagic in the open sea (Freyhof and Kottelat, 2008). It occurs at temperatures above 9ºC and therefore, in the Mediterranean Sea, it happens in winter and early development takes place in a temperature range between 13ºC and 17ºC. Sea bass eggs are pelagic and larvae planktonic, like in the majority of marine teleosts (Finn and Kristoffersen, 2007), and they float in the upper layers of the water column being exposed to higher temperature variation. The juveniles move towards brackish water and stay inshore until the second summer of their life is reached (Freyhof and Kottelat, 2008). Depending on the geographic region, the first sexual maturity is reached at the age of 2 to 4 years in the Mediterranean Sea (Bauchot, 1987) or at 4 to 7 years for females and 5 to 8 years for males in the Atlantic Ocean (Bauchot, 1987). In the wild and depending on water temperature, larvae hatch after 4 to 9 days after fertilization (Pickett and Pawson, 1994).

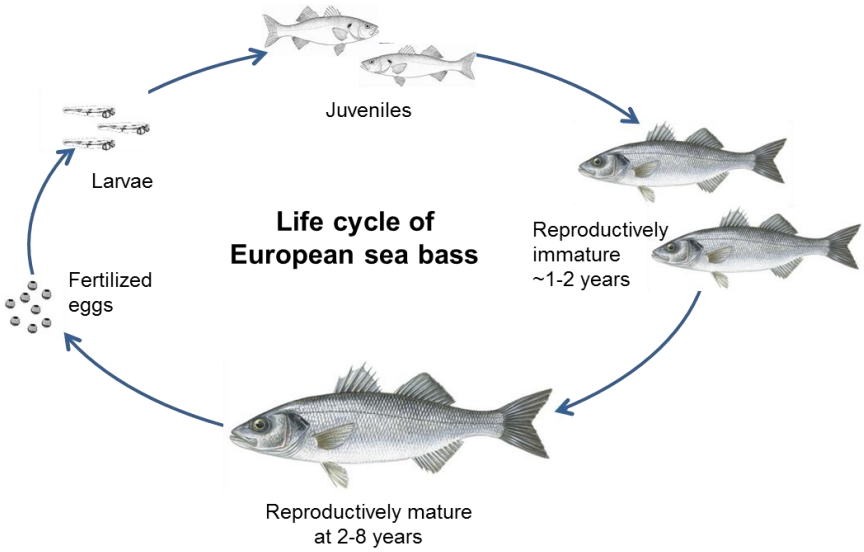


Figure 7. Schematic overview of the stages of life cycle of European sea bass.

The European sea bass is a gonochorostic species and has a polygenic sex determination system (Vandeputte et al., 2007) with temperature influences (Piferrer et al., 2005). The first 60 days of life constitute a critical phase, called the thermosensitive period (TSP), where the role of temperature is crucial for the determination of the phenotypic sex of the fish (Navarro-Martín *et al.* 2009). The sex differentiation occurs later during life, between 128 and 250 dpf (Saillant et al., 2003). During the TSP, higher temperature is linked to increased methylation of the promoter of *cyp19a1a*, the enzyme which converts androgens into estrogens, in the gonads. This is associated with less expression of this gene, which leads fish that would develop as females to differentiate as males (Navarro-Martín et al., 2011a).

3.3.3. Effects of environmental factors, particularly temperature, during early development

The sea bass has a highly plastic response to environmental temperature. Hence, despite being an eurythermic species, temperature can have various stage- and dose-dependent effects on its biology in addition to be a key sex determining factor. Thus,

temperature experienced early during life affects meristic and morphometric characteristics (Georgakopoulou et al., 2007a, 2007b; Koumoundouros et al., 2001), swimming capacity (Claireaux, 2006; Koumoundouros et al., 2009), growth rate (Hidalgo et al., 1987; Person-Le Ruyet et al., 2004) and metabolism (Claireaux and Lagardère, 1999; Moreira et al., 2008; Person-Le Ruyet et al., 2004). In addition, temperature is able to induce changes in muscle growth and cellularity (Ayala et al., 2000; Koumoundouros et al., 2009; López-Albors et al., 2003; Wilkes et al., 2001). The effects of temperature are mediated, at least in the case of sex differentiation, through long-lasting changes in gene expression (Díaz and Piferrer, 2015).

3.3.4. Sea bass aquaculture

Sea bass is a major aquaculture species with more than 150.000 tonnes produced in 2014 (http://www.fao.org/). The European sea bass was the first marine non-salmonid species to be commercially cultured in Europe and is now the most important cultured fish in the Mediterranean (http://www.fao.org/). Between wild and farmed sea bass there are differences in behaviour (Benhaïm et al., 2012, 2013), acute stress tolerance (Millot et al., 2011), morphology (Arechavala-Lopez et al., 2012, 2013), organoleptic characteristics (Arechavala-Lopez et al., 2013), as well as in muscle cellularity and other muscle properties (Periago et al., 2005). In aquaculture conditions, several reproduction-related traits are also altered. Sea bass mature earlier than their wild counterparts and induction of spawning can be achieved by artificial means, such as hormonal treatments or photoperiod manipulations (Bagni, 2005). In addition, under aquaculture conditions, ~75-95% of fish are males (Navarro-Martín et al., 2009) which grow slower than females (Díaz et al., 2013; Saillant et al., 2001).

3.3.5. Domestication programs and genetic markers

The European sea bass is one among the 30 species that are considered to have reached the higher level of domestication according to the criteria of Teletchea and Fontaine (2014). This

implies that the full life cycle of sea bass may take place under captivity conditions. Moreover, there are three selective breeding programs in place for improving traits related to growth, disease resistance, product quality and maturity (Migaud et al., 2013) and sea bass have been selected for up to 8 generations (Janssen et al., 2015). Despite these, usually in production farms the fish forming the broodstock are wild-caught sea bass or are part of the offspring of wild fish (F1), with some exceptional cases where the F2 or the F3 are used as broodstock (Chatain and Chavanne, 2009; Hillen et al., 2014; Novel et al., 2010; Teletchea, 2015; Vandeputte et al., 2012). The heritability of traits interesting for production has been estimated for growth, sex ratio, resistance to diseases, deformities and stress resistance (Hillen et al., 2014). In addition, Quantitative Trait Loci (QTL) have been identified for length, body weight, morphometric characteristics and stress response that could be used in selective breeding programs (Hillen et al., 2014). However, GxE interactions should be taken into account especially since epigenetics may explain part of the phenotypic variance which fails to be explained by the genetic variance.

3.3.6. Genomic tools

As a result of the commercial interest for sea bass, several genetic and genomics tools have been developed for this species (Louro et al., 2014). These range from transcriptomic resources, such oligo cDNA microarrays (Ferraresso et al., 2010; Geay et al., 2011) and *de novo* RNA-seq assemblies (Louro et al., 2010; Magnanou et al., 2014; Sarropoulou et al., 2012) to genetic scanning for QTLs (Chatziplis et al., 2007; Massault et al., 2010) and SNV (Kuhl et al., 2011a). In 2014, a high-quality assembled genome of a meiogynogenetic male European sea bass was published (http://seabass.mpipz.de/; dicLab v1.0c; Tine *et al.*, 2014) divided in 24 linkage groups according to the recognized karyotype of sea bass (Arefev, 1989). The haploid sea bass genome has a size of 676 Mb. In addition to the genome assembly, 26.719 genes were annotated and 234.148 SNPs were identified together with repetitive DNA sequences (Tine et al., 2014). Recently, a high-density linkage map derived from 6706 Single Nucleotide Polymorphisms (SNP) was published (Palaiokostas et al., 2015).

Aims

The **overall aim** of this thesis was to contribute to the undestanding of the phenotypic regulation by DNA methylation delineated by intrinsic and environmental factors. To that end, a non-model vertebrate species, the European sea bass (*Dicentrarchus labrax*), is used due to the availability of genomic resources and to the direct environmental influences an ectothermic animal experiences.

To achieve the overall objective, the research is organized around five topics: effects of ageing on candidate genes (*Chapter 1*); small temperature increases during early development on global DNA methylation (*Chapter 2*); effects of genetic background and early developmental temperature increases on key genes for sexual development (*Chapter 3*); genome-wide natural patterns of regulation of gene expression by DNA methylation in wild animals (*Chapter 4*); and effects of farming in early domesticates (*Chapter 5*).

The **specific objectives** were:

1. To identify age-associated DNA methylation tendencies in genomic loci of interest in somatic and gonadal tissues. This is achieved by comparing the DNA methylation status of muscle, ovaries and testis of fish from 3 ages including senescent fish using a targeted sequencing approach developed for this study.

2. To determine whether small temperature increases during different periods of early development affect global DNA methylation and expression of ecologically important genes. To achieve this, we exposed fish to moderate temperature changes (2-4ºC) during two sub-periods of larval development, early and late, and measured global DNA methylation and gene expression.

3. To study the contribution of paternally transmitted genetic background, temperature effects and their interactions on the epigenetic component of the sex determination system and ultimately sex ratios. To that end, progenies from sires known to produce low or high proportions of females were subjected to low or high temperatures during early development and the

DNA methylation status of genes related to sexual development were interrogated.

4. To evaluate the relationship of DNA methylation of specific gene features with gene expression patterns in tissues with high and low cellular heterogeneity. For this, integrative analysis of genome-wide DNA methylation and gene expression was performed in testis and muscle of wild sea bass.

5. To assess the epigenetic contribution to the first steps of domestication linked to Darwin's domestication syndrome. To reach this objective, comparisons of genetically similar wild fish and fish reared in farming conditions were performed at the genome-wide DNA methylation and gene expression level.

Results

Chapter 1

Age-related DNA methylation changes in somatic and gonadal tissues in the European sea bass –A piscine epigenetic clock?

Dafni Anastasiadi and Francesc Piferrer

Institute of Marine Sciences (ICM-CSIC), Passeig Marítim de la Barceloneta 37–49, 08003 Barcelona, Spain.

Abstract

In animals, age-related changes in DNA methylation occur. In general, there is global genomic hypomethylation accompanied by CpG-specific hypermethylation as a result of intrinsic and extrinsic influences. Mammalian epigenetic clocks have been suggested and used that estimate the biological age independently of the tissue tested. In fish, ageing studies based on epigenetic changes are scarce, while in fisheries and fish population studies there is a need for accurate estimation of age for which epigenetic biomarkers can be candidates. Here, we suggest CpG loci exhibiting decreasing or increasing methylation with age as potential piscine biomarkers of age. We used individuals from three different age classes comprising the reproductively immature and the senescent phase of a marine teleost fish, the European sea bass, to inquire the methylation status of the regulatory regions of important genes. For this inquiry, we successfully developed a low-cost high-resolution PCR-based NGS protocol applied in 22 genes and 36 samples. We found gene-specific ranges of DNA methylation that were affected by tissue identity and age. Furthermore, we identified individual CpG in the regulatory regions of four genes, two *sox19a*, *amh* exhibiting increasing and two *sox9*, *vasa*, exhibiting decreasing methylation with age across somatic and gonadal tissues (mean R^2 ~0.94). Importantly, CpGs with clear increasing or decreasing tendencies in DNA methylation with age were detected and the youngest fish mean methylation matched with ~1 year old fish from an independent experiment. The sum of the suggested CpGs can serve as a guide to focus the efforts for the development of an accurate intertissue piscine epigenetic clock.

1. Introduction

In fish, delayed senescence and extreme longevity have evolved (Reznick et al., 2002). This has been linked with indeterminate growth and age-increased fecundity. In contrast, in birds and mammals, senescence is suggested to have evolved as a by-product of determinate growth (Reznick et al., 2002). The first link between aging and DNA methylation was found in humpback salmon (*Oncorhynchus gorbuscha*), where a decrease of methylated cytosines with age was observed (Berdyshev et al., 1967). Since then, a decrease in DNA methylation has been shown in the rat, mouse and cow (Romanov and Vanyushin, 1981), and many recent studies have investigated the role of DNA methylation in aging in humans due to its importance for biomedical research. However, since across the tree of life the patterns of ageing are extensively diverse (Jones et al., 2013) and there are contrasting patterns of senescence and growth, it is of great interest to uncover the relationships of DNA methylation with ageing in other vertebrates as well.

In general, in animals there is an age-dependent drift of DNA methylation that may be summarized in global genomic hypomethylation (Heyn et al., 2012; Mugatroyd et al., 2010) accompanied by hypermethylation of specific CpG sites that are not associated with changes in gene expression (Jung and Pfeifer, 2015). However, age-related DNA methylation changes are influenced by intrinsic factors, such as genetic background, extrinsic factors, such as exposure to specific environmental agents, as well as by stochastic events (Jung and Pfeifer, 2015). Due to the complexity of interactions affecting DNA methylation with ageing, there are also tissue-specific responses with age and inter-individual differences (Issa, 2003; Jung and Pfeifer, 2015), in addition to gender-specific differences in this process (Marttila et al., 2015). The sex hormones have been suggested as candidates to test for the evolutionary theories of senescence because they are related to reproductive success early and decreased survival late in life (Gavrilov and Gavrilova, 2002). Thus, gonads represent good candidate tissues for comparing the effects of ageing on DNA methylation across the tree of life. On the other hand, muscle degenerates with age mostly due to a decrease in myofibers.

Recently, a genome-wide hypermethylation with age was shown in normal human ageing with the differences in methylation in skeletal muscle concentrated in the intragenic regions at the downstream part of genes (Zykovich et al., 2014).

Despite the tissue-specific variation in DNA methylation with age, recently, a tissue-independent epigenetic clock, extremely accurate, was discovered in humans, enabling the calculation of biological age which may differ from the chronological age due to both intrinsic and extrinsic factors (Horvath, 2013). This epigenetic clock included a total of 353 CpG sites that were sufficient to predict the age of an individual (Horvath, 2013). Apart from human tissues and cell lines, the epigenetic clock has been confirmed in chimpanzees (Horvath, 2013) and, more recently, in a wild mammalian non-model species, the humpback whale, *Megaptera novaeangliae* (Polanowski et al., 2014). In another study, only 3 CpG sites were identified as sufficient to predict the age of blood cells (Weidner et al., 2014). The epigenetic clock is extremely accurate. Thus, while the correlation between age and telomere length is less than 0.5 the epigenetic clock reaches 0.96 (Gibbs, 2014). However, epigenetic biomarkers of ageing have never been tested in non-mammalian vertebrates.

In fish, apart from the first study in humpback salmon, more recently a genome-wide study using zebrafish as a model showed an age-dependent hypomethylation of CpG sites identified by methylation-sensitive enzyme digestion (Shimoda et al., 2014). In addition, a very recent study using Chinook salmon (*Oncorhynchus tshawytscha*) and a candidate gene approach showed that there were changes to both directions of DNA methylation that were gene-, tissue- and age-dependent (Venney et al., 2016).

In parallel, while the epigenetic mechanisms start receiving greater importance in ecological studies, the techniques to accurately measure DNA methylation hold limitations. Most of the high-resolution methods available have been developed for use in biomedical research in model organisms with sequenced genome and genomic tools available. On the contrary, techniques available for use in non-model organisms without a reference genome provide low resolution data. These techniques include the High Performance

Liquid Chromatography (HPLC) or the Methylation-Sensitive Amplified Polymorphism (MSAP). Both are extremely useful for obtaining a global overview of DNA methylation, although they do not provide single nucleotide information.

Recently, two next generation sequencing (NGS) protocols have been adapted for non-model organisms, which can be applied even to organisms without a sequenced genome. The first one is a reference-free Reduced Representation Bisulfite Sequencing method, where after a typical RRBS experimental protocol, a bioinformatic pipeline which constructs *de novo* alignments without an available genome as scaffold is applied (Klughammer et al., 2015). The second is the bisulfite Restriction site Associated DNA sequencing (BisRAD-seq), which is a modification of the RAD-seq protocol involving bisulfite conversion to simultaneously collect information about the methylation status of cytosines (Trucchi et al., 2016). Still, the reference-free RRBS remains expensive for a large number of samples, as is often required in ecological studies. On the other hand, the bisulfite RAD-seq is the best alternative for exploring genome-wide alterations without *a priori* knowledge of target regions in a large number of samples. However, in many cases, *a priori* candidate regions of changes or genes of interest for the model system of study may be known, while a large sample size may be required to answer the scientific question.

In this study, the objective was to interrogate the DNA methylation status of cytosines in the regulatory regions of selected genes of known functional importance. Therefore, we first developed a protocol for constructing targeted bisulfite NGS libraries, named hereafter Multiplex Bisulfite Sequencing (MBS). MBS was designed to analyse the DNA methylation status of individual cytosines in 22 genomic regions associated with the promoter and/or the first exon in 36 samples, customizable according to specific needs. We applied MBS in the ovaries, testes and muscle from fish of three distinct ages and reproductive-statuses: one-year-old reproductively immature fish, three-year-old fish at the early reproductively mature phase, and twelve-year-old senescent fish. The genomic targets included genes related to important gonadal and muscle functions, as well as to growth and DNA methylation.

2. Materials and Methods

2.1. Animals and rearing conditions

Fish from 3 age groups were used in this experiment. All fish were reared at the aquarium facilities of the Institute of Marine Sciences following the standard rearing procedures (Díaz et al., 2013; Morretti, 1999). The first age group originated from eggs fertilized after a natural spawning of our broodstock maintained in the facilities in April 2013 and sacrificed in July 2014, at 468 days post fertilization (dpf; ~1 year old; referred thereafter as 1Y). These fish came from the same batch used in Chapter 2 as juvenile fish reared at high temperature between 20 and 60 dpf. The second age group originated from fertilized eggs transported to our aquarium facilities from the Hatchery Base Viva de Sant Pere Pescador at 5 dpf (June 2010). These fish were sacrificed at 1122 dpf (June 2013; ~3 years old; referred thereafter as 3Y) and are the same as the fish used in Chapter 5 as the farmed natural temperature group. For the sampling of both age groups, an overdose of 2-phenoxyethanol (2PE) was used in order to sacrifice the fish. The third age group was maintained in our experimental facilities for 12 years. Fish from the third group died from natural causes between March and July of 2014 (~12 years old; referred thereafter as 12Y). Tissue samples were dissected from moribund fish or within 24 hours after death.

For each age group, testis and muscle samples were dissected from 4 males and ovarian tissue from 4 females, so testis and muscle samples came from the same fish; a total of 36 samples were used in this experiment. In all cases dissection of tissue samples was followed by immediate immersion into liquid nitrogen. Individual measurements of body length and weight were obtained (Table 1).

Table 1. Weight and length data from the three age groups

Age	n	Weight (g)	Standard length (cm)
1 year	8	89.3 ± 10.7	16.98 ± 0.62
3 years	8	555.3 ± 59.4	32.63 ± 1.13
12 years	8	5133.3 ± 392.7	63.73 ± 1.56

Data as mean ± SEM.

The aquarium facilities of the Institute of Marine Sciences (Spanish National Research Council; CSIC, Barcelona, Spain) are authorized for experiments with animals by the Ministry of Agriculture and Fisheries certificate number 08039–46–A) according to Spanish legislation (Real Decreto 223 of March 1988). The treatment of animals was in accordance to the European Convention for the Protection of Vertebrate Animals used for Experimental and Other Scientific Purposes (ETS Nu 123, 01/01/91) and the protocol was licensed by the Ethics Committee of the CSIC under the project AGL2013–41047–R.

2.2. DNA extractions and bisulfite conversion

An overview of the multiplex bisulfite sequencing design is shown in Figure 1. DNA extraction was performed by the standard phenol-chloroform-isoamyl alcohol (PCI; 25:24:1) protocol. One microgram of proteinase K (Sigma-Aldrich) and 0.5 μg of ribonuclease A (PureLink RNase A; Life Technologies) were added to the digestion buffer to ensure the absence of proteins and RNA respectively. Quantity and purity of DNA were measured by the ND-100 spectrophotometer (NanoDrop Technologies). Two micrograms of DNA for each sample were bisulfite converted using the EZ DNA Methylation-Direct™ Kit (Zymo Research; D5020) according to the instructions of the manufacturer, except from the desulphonation time which was prolonged to 30 min. Bisulfite converted DNA was eluted from the columns using the same volume of 40 μl of Milli-Q autoclaved H_2O twice. Then, 120 μl of Milli-Q autoclaved H_2O were added to each sample to reach a concentration of 12.5 ng/μl of bisulfite converted DNA.

2.3. Genes selection and primers design

Previous knowledge about the gene functions related to muscle, testis, ovary or gonads was used for selecting the target genes. Genes were ranked according to the following criteria unordered: 1) number of CpGs per 100 bp aiming for more than 5 CpG/100 bp in the region around the transcription start site (TSS) and 2) literature

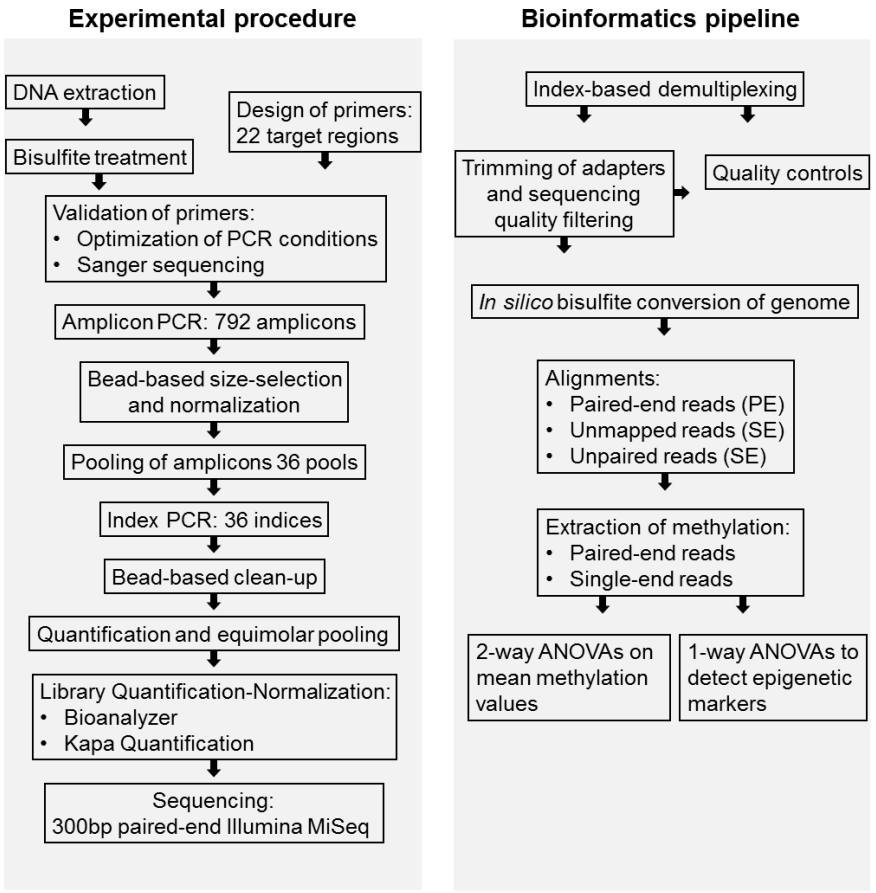


Figure 1. Overview of the multiplex bisulfite sequencing (MBS) workflow comprised of the experimental procedure and the bioinformatics pipeline.

survey of relative functional importance. We, therefore, selected 22 target genes of interest (Table 2, p. 56).

Primers were designed for bisulfite converted DNA using MethPrimer (Table S1; Li and Dahiya, 2002). Primers were further validated using Primer3Plus (Untergasser et al., 2012) after *in silico* bisulfite conversion of the target sequence using Bisulfite Primer Seeker (Zymo Research). Amplicons were designed so that they never exceeded 550 bp in length in order to ascertain the acquisition of overlapping paired-end reads using the 300 bp Illumina

sequencing protocol. An ideal amplicon's range was considered between 450 and 500 bp encompassing as many CpGs as possible. The target regions included as much as possible from the first exon and the promoter, in this order of priority due to the importance of the first exon in the regulation of gene expression (Brenet et al., 2011), of each target gene. The following adapters were added to the 5' ends of the region-specific primers as in Illumina's protocol for 16S metagenomic library preparation:

Forward: 5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3'

Reverse: 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3'.

2.4. Amplicons PCR

Amplifications of targeted regions were performed in a total volume of 25 µl containing: 25 ng of DNA (2 µl), 4 mM $MgCl_2$, 0.8 mM dNTPs , primers at 0.8 µM (Life Technologies), 2.5 U of GoTaq G2 Hot Start polymerase (Promega) and its corresponding 5X Green GoTaq Flexi Buffer (Promega). PCR conditions were as follows: 7 min at 95ºC, followed by 40 cycles of 95ºC for 1 min, annealing temperature dependent on each primer pair (Table 2, p. 56) for 2 min and 65ºC for 2 min, with a final step at 65ºC for 10 min. The presence and size of bands were confirmed by agarose gel electrophoresis. The presence of amplicons for the 22 targeted regions was visually confirmed by agarose gel electrophoresis. Sequences identities were confirmed by Sanger sequencing. Amplicon PCRs were performed in 96-well plates, each plate containing 2 target regions (2 primer pairs). In total, we obtained 792 PCR products divided in 96-well plates.

2.5. Size-selection and normalization

After PCR amplification of the target regions, we performed size-selection and normalization of DNA quantities across PCR products following a customized version of the bead-based normalization of Hosomichi et al. (2014). The working solution of serapure magnetic beads was prepared by washing 2 ml of Sera-mag SpeedBeads (Fisher 09981123) with Tris-EDTA (TE; 10 mM Tris; 1 mM

EDTA) and adding the beads in a total volume of 50 ml containing 20% PEG-8000, a concentration of 2.5 M NaCl, 500 mM of Tris-Hcl, 1 mM EDTA and 0.00055% Tween 20 (Adapted from*:* Rohland and Reich, 2012). In brief, 8 µl of PCR product and 42 µl of Milli-Q autoclaved H_2O were incubated for 5 min at room temperature with 20 µl of beads. Following 2 min incubation on the magnetic stand (3D-printed 96-well magnetic rack designed by http://www.thingiverse.com/acadey/ and realized by MAKE Creative Spaces), supernatants were transferred to new wells and incubated with 0.8x of magnetic beads for 5 min. After discarding the supernatant, a single wash with 70% freshly prepared ethanol was performed and DNA was eluted in 20 µl Milli-Q autoclaved H_2O. Size-selected PCR products were incubated in equal volumes with 20-fold diluted magnetic beads (PEG 20% and 2.5 M NaCl) and isopropanol. After incubation for 5 min at room temperature and washing with 70% freshly prepared ethanol, PCR products were eluted in 10 µl Milli-Q autoclaved H_2O. Since each amplicon contained theoretically equal DNA amounts, identical volumes of each amplicon were pooled for each biological sample, resulting in 36 tubes.

2.6. Index PCR and size-selection

Sample-specific indices were incorporated into the amplicons following a dual-index strategy with i7 indices from Nextera XT index Kit SetA and i5 indices from Nextera XT index Kit SetD (Illumina; FC-131–2001 and FC-131–2004). PCR reactions were performed using the 2x KAPA HiFi HotStart ReadyMix, 5 µl of each primer and 5 µl of template pooled amplicons DNA in a total volume of 50 µl with the following conditions: 95ºC for 3 min, 8 cycles of 95ºC for 30 s, 55ºC for 30s and 72ºC for 30s and a final step at 72ºC for 5 min, according to Illumina's protocol for 16S metagenomic library preparation. Bead-based size-selection was carried out using 0.6x of magnetic beads and eluting in 15 µl Milli-Q autoclaved H_2O. DNA quantities were double measured by the Qubit dsDNA HS Assay Kit (ThermoFisher Scientific) and the Agilent DNA 1000 chip (Agilent) by which the distribution of fragments was also visualized. The samples were pooled in equal quantities. The final library was sequenced two independent times,

with the amplicons originating from the first and the second MBS library of Chapter 3. The multiplexed single final library was quantified both times by real-time qPCR using the Kapa system and sequenced on two separate MiSeq (Illumina) instruments using always the paired-end 300 bp protocol.

2.7. Multiplex bisulfite sequencing library bioinformatics

Samples were demultiplexed based on the dual-indices by the instrument's software. Adapters and linker sequences were trimmed for paired-end reads by Trim Galore! (Babraham Bioinformatics), while filtering for low quality bases (Phred score < 20). Quality controls of the data were carried out before and after trimming using FastQC (Babraham Bioinformatics). Trimmed reads were aligned against the *in silico* bisulfite converted sea bass genome (Tine et al., 2014) using Bismark (Krueger and Andrews, 2011) for both *in silico* bisulfite conversion and alignments. Mappings were done in three steps in a non-directional way: 1) paired reads were aligned, 2) unmapped reads from the first step were aligned as single reads, and 3) unpaired reads were aligned like the unmapped reads. An alignment was considered valid if the score attributed was above f(x) = 0 + -0.6 * read length. The alignments of unmapped (2) and unpaired (3) reads were merged using samtools (Li et al., 2009) and treated as single-end reads. Alignments were visually inspected using the Integrated Genome Browser (Nicol et al., 2009) and the genomic boundaries and amplicon sizes were confirmed. Methylation calling was performed by the bismark_methylation_extractor of Bismark separately for paired-end and for single-end reads. Paired and single reads were merged for each sample in a single file and the rest of the analysis were carried out using R (R Core Team, 2015; RStudio Team, 2015) and Bioconductor (Huber et al., 2015), unless stated otherwise. The genomic coordinates of CpGs were obtained at the whole-genome level using a custom-made BSgenome package (Pagès, 2016) of the sea bass genome and intersected with the amplicons' target region using bedtools (Quinlan and Hall, 2010). For each CpG per sample, counted cytosines and thymines were summed up from paired-end and single-end reads of both sequencing runs and percent of methylation was calculated as 100*(Cs/Cs+Ts). Mean methylation

per gene was calculated as mean of methylation of individual CpGs for each sample. These values were used to calculate the mean of the four biological replicates per tissue and per age.

2.8. Statistical analysis

The mean methylation was calculated per sample and per gene. Binomial generalized linear models with the link logit function were used to calculate the deviance explained by each of the 3 factors considered: gene, tissue and age, using the summary of methylated cytosines and the summary of unmethylated cytosines as dependent variables. For each gene, the effects of tissue and age and their interactions were tested using 2-way ANOVAs on the arscine transformed mean methylation values per sample. The Tukey Honest Significance Differences test was applied following the 2-way ANOVA. Differences in methylation of individual CpGs associated with age were tested using testis and muscle of the male individuals to detect intertissue changes by one-way ANOVAs followed by Tukey Honest Significance Differences tests. Linear regression models were fitted to log-transformed mean methylation of individual CpGs per age group using testis and muscle of male fish as replicates. Males and females reared at low temperature during early development from Chapter 3 for which data were available were used to calculate the mean of individual CpGs of *amhr2*, *fhsr*, *er-b2* and *nr3c1* in an independent set of 0.88 year old sea bass.

3. Results

The selected genes for the MBS had a CpG density from 0.6 to 7.6 CpGs in 100 bp and amplicons' target sizes from 160 to 548 (Table 2). Seven candidate genes contained annotated CpG islands in their target genomic regions and one, *myod*, contained binding sites for the zinc finger protein 263 transcription factor (ZNF263) in its target region (Table 2).

Table 2. Target genes of the MBS approach

Gene symbol	Gene name	No. of CpGs	No. of CpG/100 bp	Presence of CpG island	Presence of TFBS	Amplicon size (bp)
3bhsd	3 beta-hydroxysteroid dehydrogenase delta 5	4	0.8			515
amh	anti-Mullerian hormone	13	2.5			515
amhr2	anti-Mullerian hormone receptor, type II	4	1.3			300
cyp11a	cytochrome p450 family 11 subfamily A	13	2.5			524
cyp17a1	cytochrome p450 family 17 polypeptide 1	15	2.8			531
cyp19a1a	cytochrome p450 aromatase	7	1.3			522
cyp26a1	cytochrome p450 26a1-like	26	4.7	yes		548
dnmt3a	dna (cytosine-5)-methyltransferase 3a	27	5.0	yes		539
erb1	estrogen receptor beta	20	3.8	yes		528
er-b2	estrogen receptor beta 2	5	2.1			240
fshr	follicle stimulating hormone receptor	4	1.8			219
igf1	insulin-like growth factor i	14	2.6			546
lhr	luteinizing hormone receptor	14	5.4	yes		257
myf6	myogenic regulatory factor 4	21	5.7	yes		367
myhm86-1	myosin heavy chain	8	2.6			308
mylz2	myosin light chain 2	12	2.5			488
myod	myogenic factor 1	40	7.6			526

nr3c1	glucocorticoid receptor	21	5.5	yes	ZNF263	384
sox19a	transcription factor sox-19a-like	21	5.8	yes		364
sox9	transcription factor sox9	7	1.9			362
vasa	vasa	3	0.6			533
wnt1	protein wnt-1-like	6	3.8			160

Abbreviation: TFBS, transcription factor-binding site.

The MBS for 22 target genes resulted in 757 valid amplicons according to our filtering criteria out of 792 expected. The average coverage per gene and per sample was 4592.3±874.9. The coverage was sufficient but variable according to the gene (Fig. S1) and uniform according to sample (Fig. S2).

We, then, used the valid amplicons to calculate the mean methylation per gene and per sample. Considered together, the chosen genes had a high dynamic range in methylation values, essentially from 0 to 100% (Fig. 2A). Consequently, the deviance of the methylation levels was mostly explained by the gene (Fig. 2A), followed by the tissue (1.28%; Fig. 2B) and the age (0.76%; Fig. 2C).

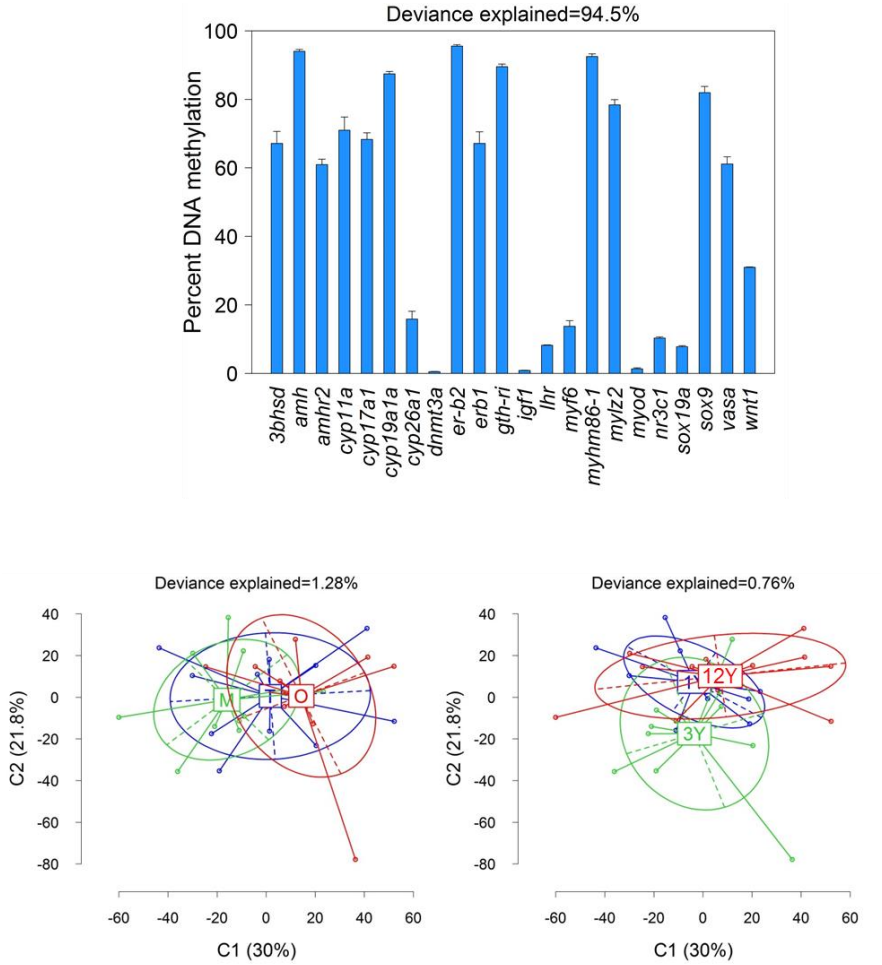


Figure 2. DNA methylation levels per gene and their similarities between tissues and ages. A) Mean methylation per gene and the deviance of methylation explained by gene as factor. Principal coordinates analyses based on dissimilarity euclidean matrices by tissue (M=muscle, O=ovary, T=testis) in B and age (1Y=1 year, 3Y=3 years, 12Y= 12 years) in C, and the deviance of methylation explained by each factor alone. The first two components of the PCoA are represented and the percentage of variance explained by each one is shown inside parenthesis.

Since the methylation levels were mostly dependent on the gene, we grouped the genes into three arbitrary categories of low, intermediate and high methylation and explored the effects of tissue and age on each gene. In total, age had an effect on DNA methylation levels in 4 of our target genes, tissue had an effect on 2 of the genes and in one there was an effect of tissue and of the interactions between age and tissue (genes with shaded panels in Figs. 3–5; Table 3).

In 3 out of the 10 lowly methylated genes, there was a significant effect of age ($p<0.05$; Fig. 3, Table 3, see next pages). More specifically, in *amhr2*, there was significantly lower methylation in 1Y and 3Y vs 12Y fish ($p=0.005$ and $p=0.007$ respectively; Table S2). In *igf1*, there was significantly higher methylation in 1Y than in 3Y fish ($p=0.034$; Table S2). In *wnt1*, there was significantly higher methylation in 1Y compared to 12Y fish ($p=0.014$; Table S2). In one lowly methylated gene, *myf6*, there was a significant effect of tissue ($p=0.005$; Fig. 3, Table 3, see next pages) with higher methylation in muscle than in gonads (muscle vs ovary, $p=0.006$; muscle vs testis, $p=0.021$; Table S2).

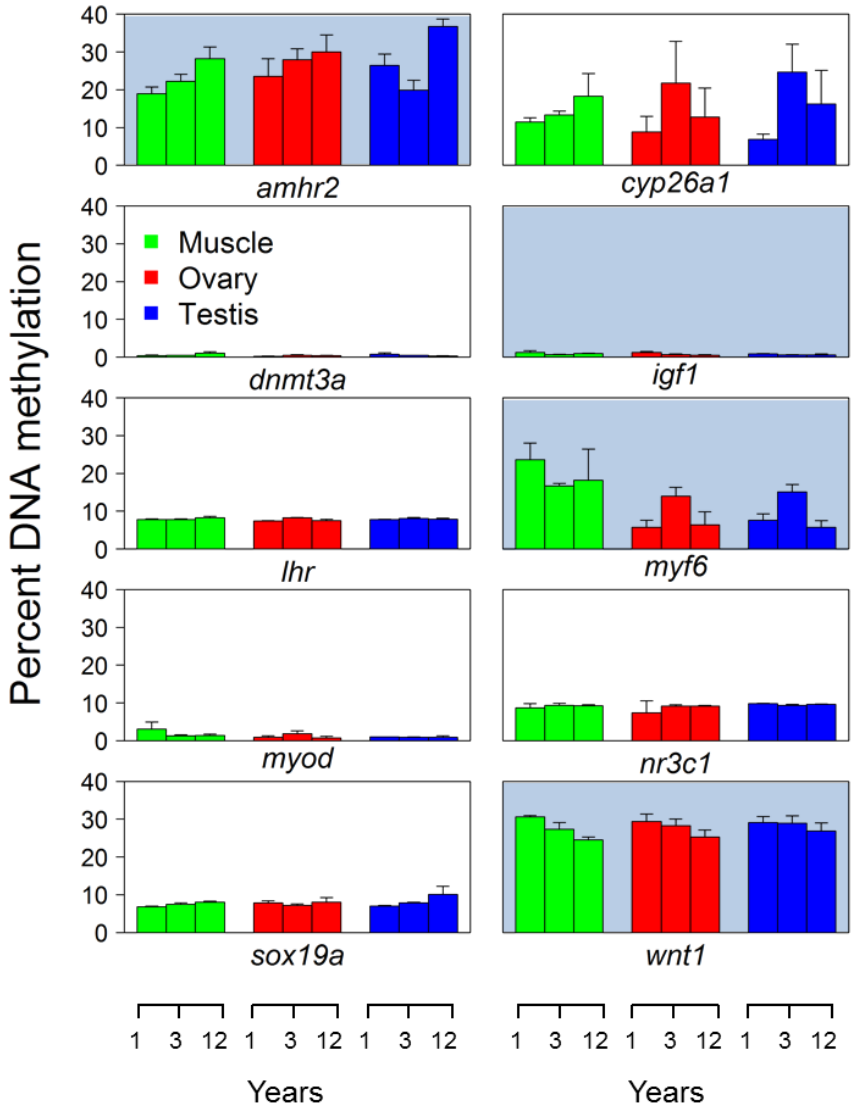


Figure 3. DNA methylation per tissue and per age for 10 genes with low methylation levels. The blue background indicates significant differences based on 2-way ANOVAs with tissue, age and their interactions as factors. Significant effects of age on *amhr2* (p=0.002), *igf1* (p=0.03), *wnt1* (p=0.017) and tissue on *myf6* (p=0.005).

Table 3. Effects of age, tissue and their interactions on DNA methylation of each gene tested by 2-way ANOVA

Gene		d.f.	Sums of Squares	Means of Squares	F value	*p*-value	Significance
3bshd	Age	2	0.055	0.028	0.548	0.586	
	Tissue	2	0.088	0.044	0.872	0.432	
	Interaction of Age-Tissue	4	0.380	0.095	1.891	0.148	
	Residuals	22	1.106	0.050			
amh	Age	2	0.026	0.013	2.346	0.120	
	Tissue	2	0.009	0.005	0.848	0.443	
	Interaction of Age-Tissue	4	0.003	0.001	0.148	0.962	
	Residuals	21	0.115	0.005			
amhr2	Age	2	0.074	0.037	7.659	0.002	**
	Tissue	2	0.019	0.009	1.948	0.162	
	Interaction of Age-Tissue	4	0.035	0.009	1.831	0.152	
	Residuals	27	0.131	0.005			
cyp11a	Age	2	0.839	0.420	8.055	0.002	**
	Tissue	2	0.013	0.007	0.126	0.882	
	Interaction of Age-Tissue	4	0.220	0.055	1.054	0.402	
	Residuals	23	1.199	0.052			
cyp17a1	Age	2	0.043	0.022	1.410	0.262	
	Tissue	2	0.095	0.047	3.088	0.063	.

	Interaction of Age-Tissue	4	0.040	0.010	0.651	0.631	
	Residuals	26	0.399	0.015			
cyp19a1a	Age	2	0.008	0.004	0.716	0.498	
	Tissue	2	0.001	0.000	0.059	0.943	
	Interaction of Age-Tissue	4	0.040	0.010	1.816	0.155	
	Residuals	27	0.150	0.006			
cyp26a1	Age	2	0.121	0.061	1.513	0.239	
	Tissue	2	0.016	0.008	0.205	0.816	
	Interaction of Age-Tissue	4	0.069	0.017	0.428	0.787	
	Residuals	26	1.041	0.040			
dnmt3a	Age	2	0.001	0.001	0.881	0.427	
	Tissue	2	0.003	0.001	2.122	0.140	
	Interaction of Age-Tissue	4	0.006	0.001	2.267	0.089	.
	Residuals	26	0.017	0.001			
er-b2	Age	2	0.001	0.000	0.130	0.878	
	Tissue	2	0.006	0.003	1.059	0.361	
	Interaction of Age-Tissue	4	0.006	0.002	0.540	0.708	
	Residuals	27	0.078	0.003			
erb1	Age	2	0.100	0.050	1.429	0.259	
	Tissue	2	0.479	0.239	6.862	0.004	**

	Interaction of Age and Tissue	4	0.274	0.068	1.964	0.131	
	Residuals	25	0.872	0.035			
fshr	Age	2	0.070	0.035	1.172	0.325	
	Tissue	2	0.001	0.000	0.015	0.985	
	Interaction of Age-Tissue	4	0.107	0.027	0.897	0.479	
	Residuals	27	0.806	0.030			
igf1	Age	2	0.004	0.002	4.025	0.030	*
	Tissue	2	0.001	0.001	1.101	0.348	
	Interaction of Age-Tissue	4	0.001	0.000	0.692	0.604	
	Residuals	26	0.014	0.001			
lhr	Age	2	0.000	0.000	1.616	0.217	
	Tissue	2	0.000	0.000	0.690	0.510	
	Interaction of Age-Tissue	4	0.000	0.000	1.538	0.219	
	Residuals	27	0.002	0.000			
myf6	Age	2	0.088	0.044	3.027	0.065	.
	Tissue	2	0.191	0.095	6.594	0.005	**
	Interaction of Age-Tissue	4	0.066	0.016	1.135	0.361	
	Residuals	27	0.391	0.014			
myhm86-1	Age	2	0.012	0.006	0.649	0.532	
	Tissue	2	0.006	0.003	0.334	0.720	

	Interaction of Age-Tissue	4	0.028	0.007	0.778	0.550	
	Residuals	24	0.213	0.009			
mylz2	Age	2	0.004	0.002	0.156	0.857	
	Tissue	2	0.023	0.012	0.945	0.402	
	Interaction of Age-Tissue	4	0.023	0.006	0.466	0.760	
	Residuals	25	0.306	0.012			
myod	Age	2	0.004	0.002	0.810	0.458	
	Tissue	2	0.005	0.003	0.977	0.392	
	Interaction of Age-Tissue	4	0.005	0.001	0.456	0.767	
	Residuals	22	0.060	0.003			
nr3c1	Age	2	0.003	0.001	0.781	0.470	
	Tissue	2	0.002	0.001	0.745	0.486	
	Interaction of Age-Tissue	4	0.004	0.001	0.604	0.664	
	Residuals	23	0.038	0.002			
sox19a	Age	2	0.004	0.002	2.125	0.139	
	Tissue	2	0.001	0.001	0.676	0.517	
	Interaction of Age-Tissue	4	0.002	0.001	0.659	0.626	
	Residuals	27	0.025	0.001			
sox9	Age	2	0.049	0.024	2.992	0.067	.
	Tissue	2	0.272	0.136	16.617	0.000	***

	Interaction of Age-Tissue	4	0.100	0.025	3.071	0.033	*
	Residuals	27	0.221	0.008			
vasa	Age	2	0.098	0.049	2.747	0.082	.
	Tissue	2	0.005	0.002	0.129	0.880	
	Interaction of Age-Tissue	4	0.058	0.014	0.807	0.531	
	Residuals	27	0.483	0.018			
wnt1	Age	2	0.014	0.007	4.720	0.017	*
	Tissue	2	0.001	0.000	0.190	0.828	
	Interaction of Age-Tissue	4	0.002	0.001	0.359	0.835	
	Residuals	27	0.039	0.001			

Among the genes of intermediate methylation, there were 2 out of 6 where a significant effect of the factors tested was observed. In *cyp11a*, there was an age effect (p=0.002; Fig. 4, Table 3) with lower methylation levels in 3Y fish vs 1Y (p=0.008) and 12Y (p=0.005; Table S2). In *erb1*, there was a tissue effect (p=0.004; Fig. 4, Table 3) with ovaries having significantly more methylation than testis (p=0.041) and muscle (p=0.004; Table S2).

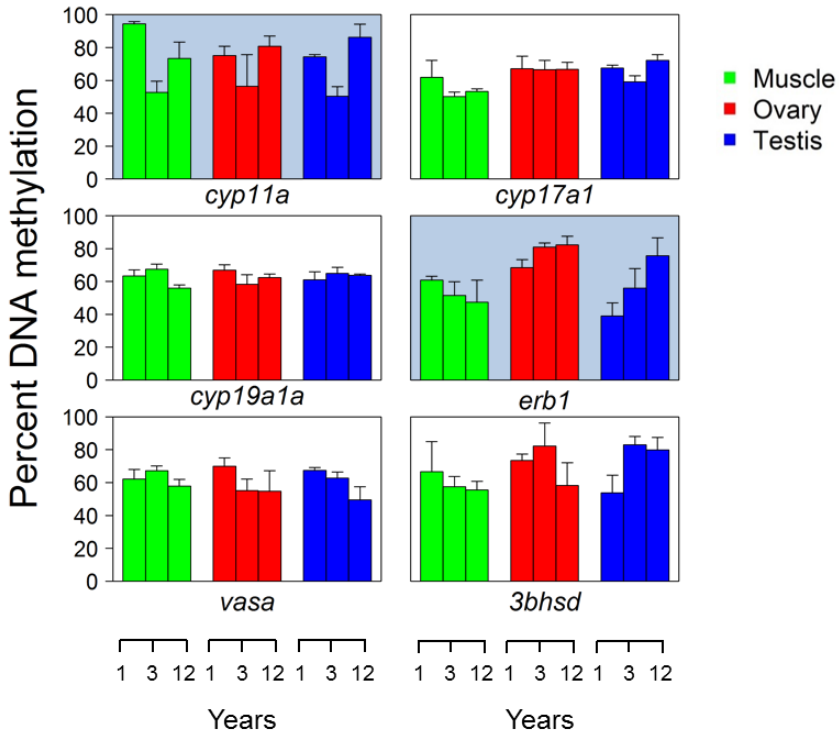


Figure 4. DNA methylation per tissue and per age for 6 genes with intermediate methylation levels. The blue background indicates significant differences based on 2-way ANOVAs with tissue, age and their interactions as factors. Significant effects of age on *cyp11a* (p=0.002) and tissue on *erb1* (p=0.005).

In one gene that belonged to the high methylation group, *sox9*, there was an effect of tissue (p=0) and effects of tissue and age interactions (p=0.033; Fig. 5, Table 3). There was significantly higher methylation in the gonads than in the muscle (ovary vs muscle, p=0; testis vs muscle, p=0; Table S2). The effects of the interactions between tissue and age were for their majority observed between the muscle of 12Y fish and the gonads of fish from all ages (Table S2).

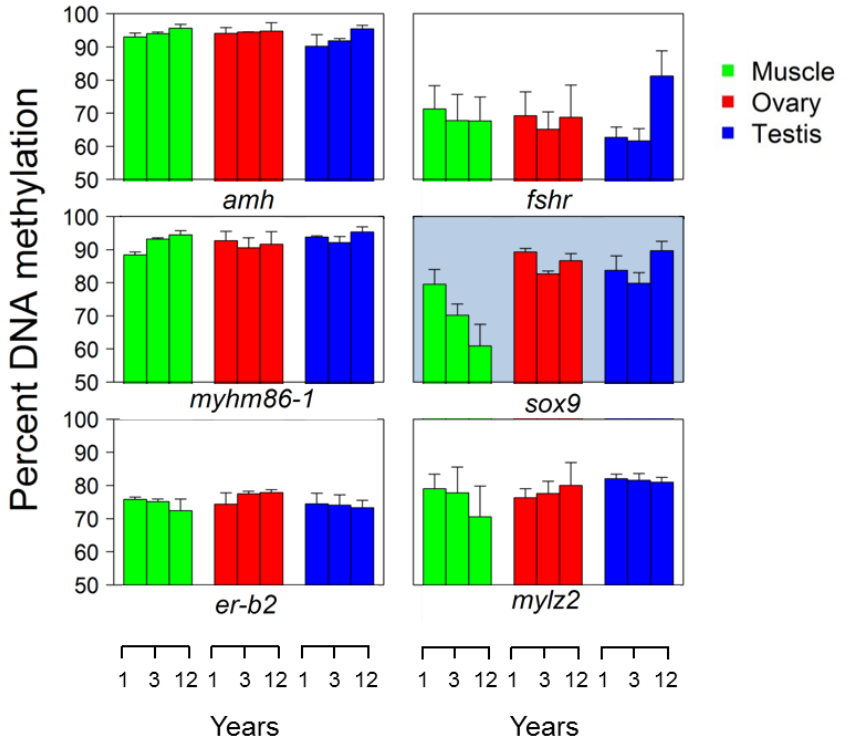


Figure 5. DNA methylation per tissue and per age for 6 genes with high methylation levels. The blue background indicates significant differences based on 2-way ANOVAs with tissue, age and their interactions as factors. Significant effects on *sox9* of tissue (p=1.97e-05) and age-tissue interaction (p=0.03).

Next, we focused on individual CpGs, using only muscle and testis since they originated from the same male fish to investigate the effect of age on potential inter-tissue changes in methylation. We detected two CpG that had constantly increasing methylation with age in *sox19a* (Fig. 6, see next page) and one of them (CpG2; LG3:78596–78597) showed significant increase between 1Y and 12Y (p=0.008) and 3Y vs 12Y (p=0.04). In *vasa*, there was decreasing methylation with age and one of the two CpGs detected (CpG1; LG24:13525513–13525514; Fig. 6, see next page showed significant decrease in 1Y vs 12Y (p=0.01) and in 3Y vs 12Y (p=0.028). CpGs with constantly increasing methylation were also detected in *amh* and with decreasing methylation with age also in *sox9* (Fig. 6, see next page). Other CpGs showed constant levels of methylation with age, i.e., *lhr* and *myhm86-1* (Fig. 6, see next page).

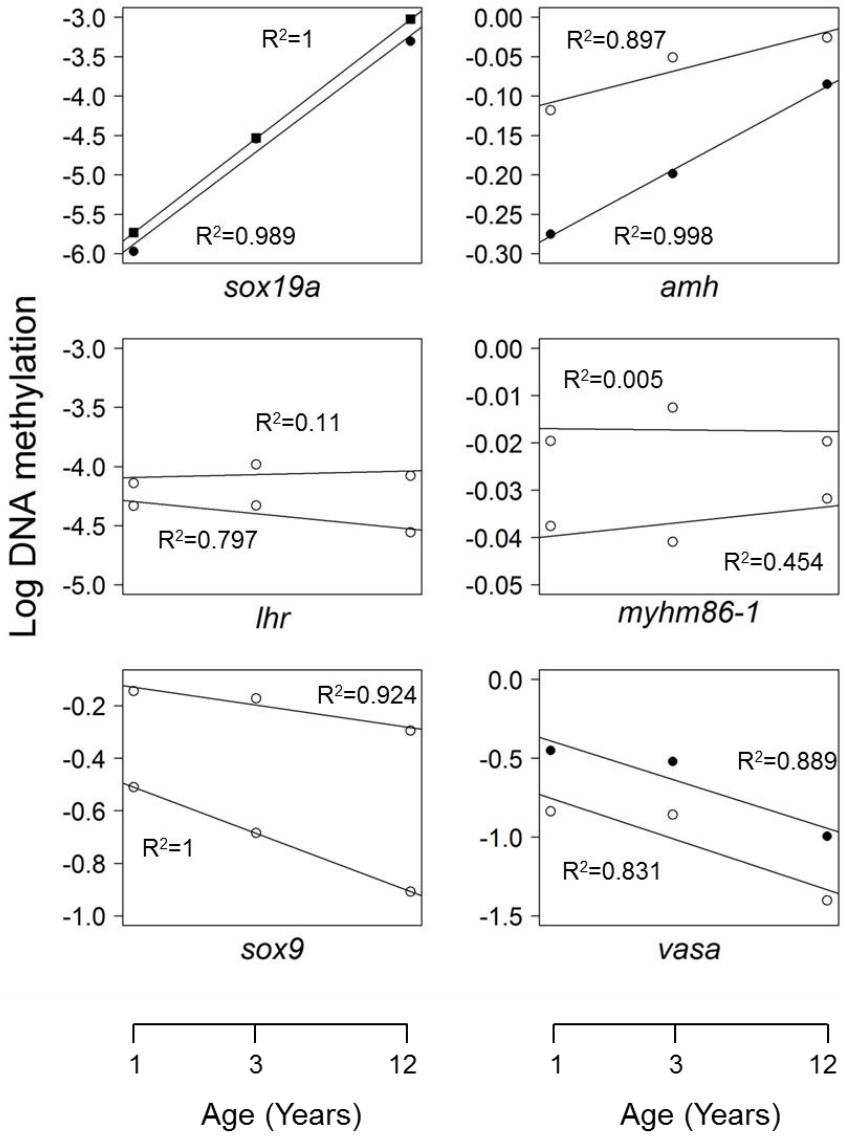


Figure 6. Tendencies of methylation levels of indicative individual CpGs with age. DNA methylation values are shown as log-transformed 0–1 methylation and the x-axis is shown at the log scale. Regression lines and R^2 values calculated for each of the 2 CpGs per gene presented are obtained from fitting linear regression models on mean methylation per CpG per age based on muscle and testis of male individuals (n=8). Significant differences between methylation detected by ANOVA are indicated with closed circles for 1Y vs 12Y fish and closed squares for 1Y vs 12Y and 3Y vs 12Y.

Four of the genes interrogated in this study—by the way, none of them among the best ones in terms of age-related changes as just seen above—coincided with the target genes of Chapter 3, the methylation status of which was assessed in 99 samples of 0.88 years old sea bass. Selected CpGs showed increasing methylation with age in the target region of *amhr2* and *fshr*, while selected CpGs of *er-b2* and *nr3c1* showed decreasing methylation with age (Fig. 7). The mean methylation values of 0.88 years old sea bass samples from Chapter 3 was similar to the methylation values of 1Y sea bass (Fig. 7).

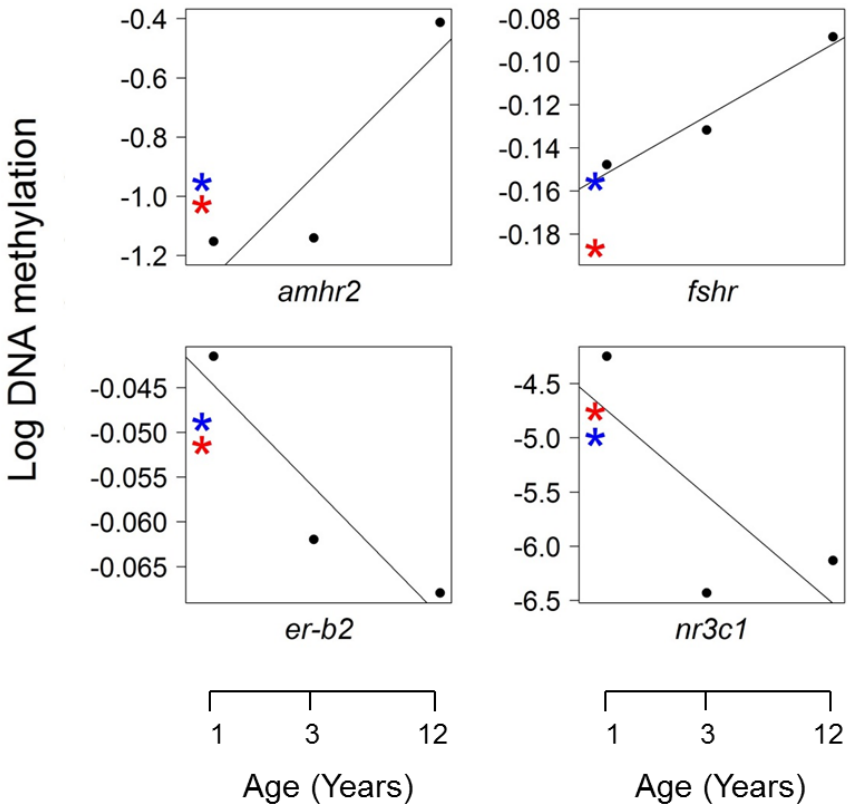


Figure 7. Methylation changes in individual CpGs of four genes (*amrh2, fshr, er-b2* and *nr3c1*). The red asterisk indicates the corresponding values of the same genes in females and the blue one in males of fish tested in Chapter 3 with unrelated 0.88-year-old sea bass. Log-transformed 0–1 methylation values are shown in the y-axis and the x-axis is represented in the log scale. Regression lines were fitted on mean methylation values per CpG per age using muscle and testis from male fish (n=8).

4. Discussion

In this study, we successfully developed and applied a low-cost version of targeted multiplex bisulfite sequencing to simultaneously interrogate the methylation status of several genes in a large sample size at single nucleotide resolution. We showed that there are clear age-related changes in the DNA methylation of known regulatory regions of specific genes. Importantly, we identified individual CpGs that showed intertissue decreasing or increasing changes in their DNA methylation status along the normal ageing process.

MBS was developed for the specific needs of this study; nevertheless it is a customizable protocol for any species and any genomic region of interest. For example, in Chapter 3, we have applied the MBS protocol for measuring methylation of one gene in 200 samples and 6 genes in 99 samples. Few studies have recently been published using similar technical approaches to measure DNA methylation (Bernstein et al., 2015; Korbie et al., 2015; Masser et al., 2013; Venney et al., 2016). In 3 out of the 4 published protocols, the number of samples and amplicons tested was lower than in our study (Table S3; Bernstein et al., 2015; Korbie et al., 2015; Masser et al., 2013). In one case (Masser et al., 2013), the NGS library construction is transposome-mediated and in another (Korbie et al., 2015) exonuclease treatment is used to clean-up the PCR products. Bernstein et al. (2015) used proprietary kits to perform the PCRs, clean-up the products and quantify the amplicons in order to perform poolings, while Venney et al. (2016) used proprietary magnetic beads; in both cases the cost of library construction was increased. In addition, Venney et al. (2016) sequenced on an Ion Personal Genome Machine which is less widely available than the Illumina MiSeq and they performed the pooling steps before cleaning-up the PCR products which is highly likely to lead to unbalanced PCR product concentrations and ultimately to highly variable numbers of reads.

Here, we have considerably reduced the cost of size-selections and clean-ups, while taking into account the normalization steps. We combined: 1) the size-selection and normalization protocol of Hosomichi et al. (BeNUS; 2014) which allows for simultaneous normalization of PCR quantities across wells and plates using only

magnetic beads with a 2) home-made version of Serapure magnetic beads together with 3D-printed magnetic stands. Both steps lead to a significant reduction of library construction costs, making this protocol more accessible to more research groups that aim to analyze a considerable amount of samples. Furthermore, we avoided the use of commercial kits at all steps of the protocol, except when the use of kits compensated the amount of lab work needed, like in the case of bisulfite conversion. Further cost reduction could be achieved by designing custom index sequences compatible with the Illumina MiSeq instrument.

The coverage obtained per gene was sufficient but variable; a situation that can be expected in any PCR-based approach since the primers may show variable efficiencies which also depends on the target sequence properties. Our custom version of the BeNUS protocol, although optimized and robust, may have also affected the normalization of genes, since there are several steps performed on PCR plates using multichannel pipettes, something susceptible to technical errors. In our case, the reduction of library construction cost compensated for the variable coverage since the target amplicons were sufficiently covered.

Another important consideration of the approach is the amplicon size. Amplicons were designed with sizes less than 550 bp and the maximum capacities of the MiSeq instrument, which are 300 bp of length, were used to ensure sufficient overlapping of the paired reads. However, in practice, there were reads shorter than 300 bp due to commonly observed decreased quality towards the end of the reads. Shorter reads were then aligned as single reads. Taking this consideration into account, one may want to design amplicons of smaller size to ensure reads mating and reduce the alignments steps.

In this study, we identified genes that showed significant differences in their methylation levels with age in a marine teleost fish. We compared sea bass at three different ages covering the whole lifespan and reproductive statuses. Reports of sea bass ages in the wild vary between 15 to 25 years maximum age (Kottelat and Freyhof, 2007). Even though the 12 Y fish used in this study were younger in absolute year number, they were senescent because reproduction had stopped and they died from natural causes.

Our approach allowed us to interrogate a very small and specific part of the genome. Nevertheless, taking into account that in elderly human, only 20% of the differentially methylated regions overlap with promoters and exons (Heyn et al., 2012) and that changes in DNA methylation with age are enriched in other parts of the genome, such as enhancers and intragenic regions (Weidner and Wagner, 2014), it is noticeable that we identified an age effect in our target genes.

We detected 4 genes (*igf1*, *wnt1*, *amhr2* and *cyp11a*) which exhibited changes in the DNA methylation levels of their known regulatory elements, namely promoter and first exon. Importantly, an age-related hypomethylation was found in *igf1*, which is in accordance with the age-related hypomethylated DMRs found in the promoters and exons of members of the IGF signaling pathway in humans (Heyn et al., 2012). The *wnt1*, also found hypomethylated in senescent compared to the immature fish, is part of the Wnt signaling pathway, which has a protective role of the cell in situation of oxidative stress; a cellular condition that has been linked to ageing in the oxidative stress theory of aging and empirically supported by observations in some model organisms *C. elegans*, *Drosophila* or mice (Maiese et al., 2008; Muller et al., 2007). The other two genes affected by age were *amhr2* and *cyp11a*, both involved in the adult reproductive cycle (Piferrer and Guiguen, 2008) suggestive of a link between their regulation and the age-dependent reproductive status of the fish.

The tissue type affected the DNA methylation of 3 genes, *myf6, erb1* and *sox9*. The target regions of *myf6* and *erb1* contained a CpG island, a case found only in ~30% of our target regions. *Myf6* is involved in muscle differentiation, while *erb1* and *sox9* are involved in sexual development and the reproductive cycle (Blázquez et al., 2008; Viñas and Piferrer, 2008). The differential methylation of these genes depending on the tissue may have regulatory consequences at the gene expression level. Even though the patterns observed oppose the standard model of gene expression regulation by DNA methylation, e.g., higher methylation of *erb1* in ovaries or higher methylation of *myf6* in the muscle, recent studies have found that there are tissue-specific differentially methylated regions that

may positively or negatively regulate gene expression (Wan et al., 2015).

Globally we observed that the main effects on DNA methylation were associated with the gene, followed by the tissue and the age. Thus, genes exhibited a range of possible DNA methylation levels that did not necessarily overlap. After this major gene effect then we had the effect of age and tissue. Age affected the DNA methylation to both directions, hypo- and hyper-methylation, and even if the differences were not significant, there were genes where a trend was evident. In a recent study using fish and a candidate gene approach, DNA methylation changes were observed with age to both directions as well and each gene exhibited a limited spectrum of DNA methylation levels (Venney et al., 2016). In the study of Venney et al. (2016), DNA methylation was compared between larvae and juveniles of Chinook salmon. Although these results cannot be directly compared with our study comparing three life stages expanding through the whole lifespan and including senescent fish, they further support our findings regarding gene-dependent effects of tissue and age.

Importantly, we identified CpGs that were increasingly or decreasingly methylated with age in early reproductively mature and senescent fish in the regulatory region of *sox19a, amh*, *sox9* and *vasa*. These CpGs were detected across tissues of male individuals, suggesting a tissue-independent direction of change. Taking into account the tissue-independent epigenetic clock predicting with extreme accuracy the age in mammals (Horvath, 2013; Polanowski et al., 2014), as well as the blood cell epigenetic clock comprising only 3 CpGs (Weidner et al., 2014), a potential similar clock based on these detected CpGs may be suggested to identify the age of fish. In fact, age prediction based on four CpGs of four genes that were not among the best in terms of age-dependent methylation changes approximately predicted age in unrelated 0.88-years-old sea bass from a different experiment. In particular, the CpG in position LG1B: 11628807 of *fshr* and the CpG in position LG2: 10393440 in *nr3c1* predicted the age of the European sea bass males and females, respectively, with an accuracy of a few months. These findings are remarkable and hold much promise especially if one takes into account that the clock was not calibrated and that muscle

was one of the tissues with the poorest prediction accuracy in human probably due to high stem cells content (Horvath, 2013). Correct age estimates are crucial for population studies and fisheries science. Currently fish age is estimated using mainly scales and otoliths (Campana, 2001). The identified CpGs need to be further confirmed in other tissues, intermediate ages and other species, in order to obtain an accurate intertissue piscine epigenetic clock. However, these first results could serve as a guide for focusing the efforts.

In conclusion, since there is increasing interest in ecological and evolutionary studies for incorporating the epigenetic level of information but methodological tools are limited for non-model species, the low cost version of targeted multiplex bisulfite sequencing developed in this study could be useful in other studies with wild or non-model species. The piscine epigenetic clock based on the suggested CpGs with increasing or decreasing methylation identified here remains to be further investigated, but if adequately adjusted it holds great potential for use in fish population studies and fisheries science to accurately estimate fish age.

Acknowledgments

We thank Elvira Martínez for assistance with fish husbandry, Marta Lombó and Margarita Metallinou for suggestions regarding the PCRs, Sílvia Joly for sample processing and Federico Quattrocchi for help with the statistical models. DA was supported by a Ph.D. scholarship from the Spanish Government (BES–2011–044860). Research supported by Spanish Ministry of Economy and Competitiveness “Epifarm” (ref. AGL2013–41047–R) to FP.

Supplementary Figures

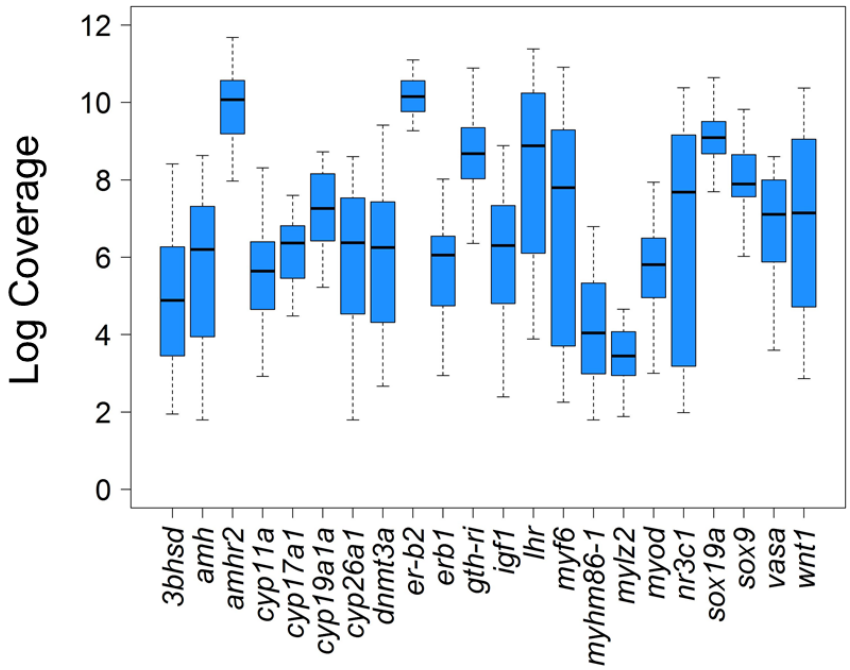


Figure S1. Coverage distribution along the target genes. Shown are $\log_2$-transformed mean values of number of cytosines covered per each gene. Boxes represent the distribution of values between the lower and upper quartiles, the upper whisker = min(max(x), Q3 + 1.5 * IQR), the lower whisker = max(min(x), Q1 – 1.5 * IQR), where IQR= third quartile (Q3) – first quartile (Q1) and the median is indicated by a black horizontal line in the box.

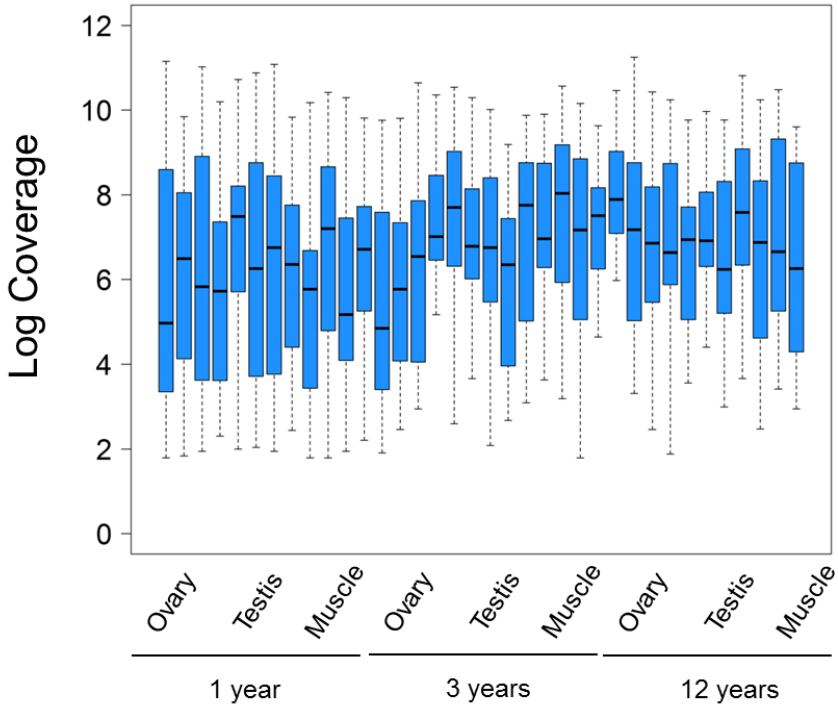


Figure S2. Coverage distribution along the samples. Shown are Log_2-transformed mean values of number of cytosines covered per each type of sample. Boxes represent the distribution of values between the lower and upper quartiles, the upper whisker = min(max(x), Q3 + 1.5 * IQR), the lower whisker = max(min(x), Q1 – 1.5 * IQR), where IQR= third quartile (Q3) – first quartile (Q1) and the median is indicated by a black horizontal line in the box.

Supplementary Tables

Supplementary Table 1. Primers used in MBS

Gene symbol	Forward (5' to 3')	Reverse (5' to 3')
3bhsd	GATTTGTTTTGTTTTTATATTAAGAGAGAA	TTTAAATCTCAATAAAACTCCCCTA
amh	TTTGATTTATTTAAAATAAGTGATTGTG	AATTTATAAATTCGCCGTCGTACCC
amhr-2	GTGGGAAATTTTTTTTATATTTTTAGGA	ACAACGACCTAAACCCTTTACTACA
cyp11a	ATTAGTTTTTTTGTTTTGTATTAGA	ATAAATTATCAATCACAAATATTTTTC
cyp17a1	TTAGTTTTTGTTGAATTAGATTTTTTT	CGCTATAATCTCCAAACGCGATATC
cyp19a1a	TTAGTTTTTCGTTGTTTGTTTTT	ACCTACAAAAATCATTACCCGTTCA
cyp26a1	GGAGGGAATTATTATTTTTTT	TCCCACTAATATCAAACATCAAACA
dnmt3a	GTTTGTTTGGGGTTTGTTTTTATAG	AAACACAAAAACTAATAACTCTTTC
erb1	GTGAATTTGATATTAAAGGGAAAAA	ACAACATACAACAATCAAAAAATAAC
er-b2	ATTATATTTTTATTTTGGTATTTTTTAGTT	ACCGACATTAAAAATTCCAACTTCCT
fshr	AATATAGAGGGAAATAATAGTGAGAGAGTG	AACAAAAACTCAAAATTCGTTTAACCAAAC
igf1	TTTGAATATGTGTTTAAAATTTTTAATGAA	ATACAAAACAATTACTCACTAACTAAATAA
lhr	AGTAAAAGAGATTTTAAAGAAGGGTTAGAG	TCTTTTATCTCTATAATCCATCATCAACGC
myf6	TTTTATTGTTAAAATAGAGAGAGGG	CTTAAACCTCCTCCTCTCCCTAAAC

myhm86-1	TATTGTGTATAAATTTTAAATAATAATTGA	TAACCTCCTCCGTATCCAAAACTTT
mylz2	AAGTTAATTGATTTATTTGAAGATTGTA	AATATCAAAAATTTCAAAACTCCTCC
myod	TTTTTTTTGTAAAAATTTGTGT	CTCCACTTTAAACAACCTCTAATTC
nr3c1	TTATTGTAGGGATTGGAGGATTAAA	ACCGCTAACTATCGATCCAATAACA
sox19a	GTATGTGCGTTTGGAGGTGATTTTA	CCTAAAACTTCCACCTTAACCTTAC
sox9	TTTATATATATTTTTATATAATAATGATA	AATTATTAACCAACAAATAAACCTCCTCCA
vasa	TTTTTGTTTTTTTTAGGGTTATTTA	TTACCCAATTATTATAATTACGCATAACCT
wnt1	AATTTATGATATAAAAAGTTAAAGAATTGT	AACGACCCCTATACCTAAAAACGAA
Adapters	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG

Supplementary Table 2. Tukey Honest Significant Differences

Gene	Factor	Difference	Lower	Upper	*p*-adjusted
amhr2	1Y-12Y	-0.099	-0.169	-0.028	**0.005**
	3Y-12Y	-0.094	-0.164	-0.023	**0.007**
	3Y-1Y	0.005	-0.066	0.075	0.983
cyp11a	1Y-12Y	-0.005	-0.262	0.252	0.999
	3Y-12Y	-0.337	-0.575	-0.098	**0.005**
	3Y-1Y	-0.332	-0.584	-0.080	**0.008**
erb1	O-M	0.278	0.083	0.472	**0.004**
	T-M	0.077	-0.122	0.275	0.607
	T-O	-0.201	-0.395	-0.007	**0.041**
igf1	1Y-12Y	0.021	-0.003	0.045	0.100
	3Y-12Y	-0.004	-0.028	0.019	0.890
	3Y-1Y	-0.025	-0.048	-0.002	**0.034**
myf6	O-M	-0.165	-0.287	-0.044	**0.006**
	T-M	-0.141	-0.262	-0.019	**0.021**
	T-O	0.025	-0.097	0.147	0.869
sox9	O-M	0.193	0.101	0.284	**0.000**
	T-M	0.175	0.083	0.266	**0.000**
	T-O	-0.018	-0.109	0.074	0.881
	1Y:M-12Y:M	0.210	-0.006	0.425	0.061
	3Y:M-12Y:M	0.095	-0.120	0.311	0.849
	12Y:O-12Y:M	0.301	0.086	0.517	**0.002**
	1Y:O-12Y:M	0.339	0.124	0.555	**0.000**
	3Y:O-12Y:M	0.242	0.027	0.457	**0.019**
	12Y:T-12Y:M	0.353	0.138	0.569	**0.000**
	1Y:T-12Y:M	0.267	0.052	0.482	**0.007**
	3Y:T-12Y:M	0.209	-0.006	0.424	0.061
	3Y:M-1Y:M	-0.114	-0.329	0.101	0.691
	12Y:O-1Y:M	0.092	-0.123	0.307	0.873
	1Y:O-1Y:M	0.130	-0.085	0.345	0.538
	3Y:O-1Y:M	0.032	-0.183	0.248	1.000
	12Y:T-1Y:M	0.144	-0.071	0.359	0.406
	1Y:T-1Y:M	0.057	-0.158	0.273	0.991

	3Y:T-1Y:M	0.000	-0.216	0.215	1.000
	12Y:O-3Y:M	0.206	-0.009	0.421	0.068
	1Y:O-3Y:M	0.244	0.029	0.459	0.018
	3Y:O-3Y:M	0.146	-0.069	0.362	0.382
	12Y:T-3Y:M	0.258	0.043	0.473	**0.010**
	1Y:T-3Y:M	0.172	-0.044	0.387	0.201
	3Y:T-3Y:M	0.114	-0.101	0.329	0.694
	1Y:O-12Y:O	0.038	-0.177	0.253	1.000
	3Y:O-12Y:O	-0.060	-0.275	0.156	0.989
	12Y:T-12Y:O	0.052	-0.163	0.267	0.996
	1Y:T-12Y:O	-0.035	-0.250	0.181	1.000
	3Y:T-12Y:O	-0.092	-0.307	0.123	0.871
	3Y:O-1Y:O	-0.097	-0.313	0.118	0.835
	12Y:T-1Y:O	0.014	-0.201	0.229	1.000
	1Y:T-1Y:O	-0.072	-0.288	0.143	0.964
	3Y:T-1Y:O	-0.130	-0.345	0.085	0.535
	12Y:T-3Y:O	0.111	-0.104	0.327	0.717
	1Y:T-3Y:O	0.025	-0.190	0.240	1.000
	3Y:T-3Y:O	-0.033	-0.248	0.183	1.000
	1Y:T-12Y:T	-0.086	-0.302	0.129	0.906
	3Y:T-12Y:T	-0.144	-0.359	0.071	0.403
	3Y:T-1Y:T	-0.058	-0.273	0.158	0.991
wnt1	1Y-12Y	0.047	0.009	0.085	**0.014**
	3Y-12Y	0.030	-0.009	0.068	0.156
	3Y-1Y	-0.018	-0.056	0.021	0.504

Abbreviations: Y, year; O, ovary; T, testis; M, muscle.

Supplementary Table 3. Published protocols based on the same principle as MBS

	Number of samples	Number of regions	Total number of PCRs
Masser et al., 2013	14	2	34
Korbie et al., 2015	13	8	13
Bernstein et al., 2015	10	5	60
Venney et al., 2016	160	22	3680
This study	36	22	792

Bernstein, D.L., Kameswaran, V., Le Lay, J.E., Sheaffer, K.L., and Kaestner, K.H. (2015). The BisPCR2 method for targeted bisulfite sequencing. Epigenetics Chromatin *8*, 27.

Korbie, D., Lin, E., Wall, D., Nair, S.S., Stirzaker, C., Clark, S.J., and Trau, M. (2015). Multiplex bisulfite PCR resequencing of clinical FFPE DNA. Clin. Epigenetics *7*, 28.

Masser, D.R., Berg, A.S., and Freeman, W.M. (2013). Focused, high accuracy 5-methylcytosine quantitation with base resolution by benchtop next-generation sequencing. Epigenetics Chromatin *6*, 33.

Venney, C.J., Johansson, M.L., and Heath, D.D. (2016). Inbreeding effects on gene-specific DNA methylation among tissues of Chinook salmon. Mol. Ecol. n/a-n/a.

Chapter 2

Anastasiadi D, Díaz N, Piferrer F. Small ocean temperature increases elicit stage-dependent changes in DNA methylation and gene expression in a fish, the European sea bass. Sci Rep. 2017 Dec 29;7(1):12401. DOI: 10.1038/s41598-017-10861-6

Chapter 3

The genetic make-up and temperature together affect the DNA methylation of gonadal aromatase and other genes important for European sea bass sexual development

Dafni Anastasiadi[1], Marc Vandeputte[2,3], Núria Sánchez-Baizán[1], Béatrice Chatain[2] and Francesc Piferrer[1,4]

1Institute of Marine Sciences (ICM-CSIC), Passeig Marítim de la Barceloneta 37–49, 08003 Barcelona, Spain.

[2]Ifremer, Chemin de Maguelone, F-34250 Palavas-les-Flots, France.

[3]INRA, UMR1313 Génétique animale et Biologie intégrative, F-78350 Jouy-en-Josas, France

Abstract

Sex is a phenotypically plastic trait, and in many piscine and reptilian species can be influenced by temperature. The epigenetic link of sex determination to temperature involves the inhibition of gonadal aromatase (*cyp19a1a*) expression, the enzyme converting androgens into estrogens, by higher DNA methylation levels of the gene promoter. In systems of mixed genetic and environmental sex determination, however, the genetic influences on epigenetic variation should also be taken into account. The combined effects of genetic background and temperature on the regulation of *cyp19a1a* and other genes related to gonad development by DNA methylation remain elusive. Here, using the European sea bass, a fish whose sex ratio is determined by both genetic and environmental factors, we show for the first time that the genetic component of sex determination influences the epigenetically mediated response to temperature. These influences are evident in genes related to gonad formation (*fshr* and *nr3c1*), ovarian differentiation (*cyp19a1a* and *foxl2*) and male differentiation (*amh*-*r2*, *er*-*β2* and *dmrt1*). In addition, DNA methylation levels are upregulated by temperature in genes related to female development and downregulated in genes related to male development. Indication of father-to-daughter

epigenetic inheritance is also present. The expression of *cyp19a1a* decreases in the presence of higher methylation levels, but, importantly, only below a sex-specific methylation threshold. In addition, a sex-specific direction of association of DNA methylation with gene expression is evident, suggestive of sex-specific genetic influences on DNA methylation. Thus, plastic responses to temperature may be genetically influenced and epigenetically inherited through the paternal methylome, contributing to sex in the European sea bass and probably also in other temperature-sensitive species.

1. Introduction

The sex ratio is an important trait which impacts the population fitness since it is related to population growth and viability (Lynch and O'Hely; Ospina-Álvarez and Piferrer, 2008). Sex ratio is the combined result of sex determination and sex differentiation (Penman and Piferrer, 2008; Piferrer and Guiguen, 2008; Piferrer et al., 2012). In many species of reptiles and fish, sex is a phenotypically discontinuous trait which is plastic and can be seen as an adaptive response to thermal fluctuation (Consuegra and Rodríguez López, 2016).

Sex determination mechanisms are divided into genotypic sex determination (GSD) and environmental sex determination (ESD), or a combination of both systems (Bull, 1983; Devlin and Nagahama, 2002; Penman and Piferrer, 2008; Valenzuela, 2008; Valenzuela and Lance, 2004). GSD mechanisms vary from monofactorial systems with sex chromosomes to multifactorial and polyfactorial systems where many loci contribute with minor additive effect to the phenotypic outcome (Penman and Piferrer, 2008; Vandeputte et al., 2007). In some fish species, master sex determining genes have been described: *sdY*, *amhr2*, *amhy*, *dmy*, *sox3* and *gsdf* (Heule et al., 2014; Kikuchi and Hamaguchi, 2013; Takehana et al., 2014), like in mammals and birds. In other fish species such as the medaka (*Oryzias minutillus*) or Lake Malawi cichlids of the genus *Metriaclima*, the sex determining loci comprise genomic regions often located in different chromosomes (i.e., Nagai et al., 2008; Ser et al., 2010).

In ESD systems, environmental conditions during a sensitive developmental period are able to induce plastic responses and determine the long-term phenotype (Devlin and Nagahama, 2002). The most common environmental factor associated with sex determination is temperature, which exerts its effects on various fish species. In contrast to reptiles, which exhibit three sex ratio patterns in response to elevated temperature (Valenzuela and Lance, 2004), in fish higher temperature is linked to male-biased populations (Ospina-Álvarez and Piferrer, 2008). The underlying molecular mechanism linking temperature to sex ratios in both fish and reptiles has been the subject of much debate (Lance, 2009). Furthermore, the sex determining systems may act in combination when a GSD system is influenced by external factors, like temperature. In these systems, even temperature response itself may be subject to genetic variation control. This is the case of Nile tilapia (*Oreochromis niloticus*), where quantitative trait loci (QTL) have been identified between families that respond differently to temperature (Lühmann et al., 2012).

The epigenetically mediated sensitivity to temperature has been shown to be mediated by differential methylation of the aromatase promoter first in fish (Navarro-Martín et al., 2011), and then in turtles (Matsumoto et al., 2013) and alligators (Parrott et al., 2014) and. The aromatase is encoded by the cytochrome P450 aromatase (*cyp19a1a*) gene in the gonads and aromatizes the A ring of androgens converting them to estrogens (Guiguen et al., 2010). This balance of androgens and estrogens is crucial for the differentiation of gonads into testes or ovaries (Guiguen et al., 2010). The aromatase promoter contains regulatory genomic elements controlling the transcription of the gene (Piferrer and Blázquez, 2005) and it also contains CpG sites that can be methylated. In male-producing temperatures, the methylation of CpG sites in the promoter increases, which prevents the binding of the transcription factor forkhead box L2 (*foxl2*) leading to a decrease of aromatase expression, in turn connected to male sex differentiation (Navarro-Martín et al., 2011).

DNA methylation also mediates the effects of temperature at the whole-genome level in the case of environmental sex reversal. In the half-smooth tongue sole (*Cynoglossus semilaevis*) and in Nile tilapia (*Oreochromis niloticus*), both fish with sex chromosomes

and sex reversal potential, temperature alters the genome-wide DNA methylation patterns (Shao et al., 2014; Sun et al., 2016). These alterations are evident in several genes of the sex determination and sex differentiation pathway, e.g., *dmrt1*, *amrh2* and *foxl2*. Importantly, in the half-smooth tongue sole the male patterns of the Z chromosome are evident in the offspring of sex reversed fish, even in the absence of male-inducing temperature, suggesting an epigenetic inheritance mechanism explanatory of these parental effects (Shao et al., 2014).

The European sea bass (*Dicentrarchus labrax*) is a gonochoristic fish and has a polygenic sex determination system (>3 QTLs; Vandeputte et al., 2007) with temperature influences (Piferrer et al., 2005). There are between families variations in the offspring sex ratios (Saillant et al., 2002; Vandeputte et al., 2007) and identified possible QTLs associated with the sex determining loci (Palaiokostas et al., 2015). In parallel, the period during egg to juvenile development (0-60 days-post-fertilization; dpf) is thermosensitive, with temperatures above 17ºC masculinizing fish that otherwise would develop as females (hereafter called genotypic females for convenience; Navarro-Martín et al., 2009). This masculinization occurs via differential methylation of the CpGs of the aromatase promoter and fish that have experienced high temperature during development carry over the epigenetic memory of temperature in the gonads throughout adulthood (Navarro-Martín et al., 2011). Under the polygenic hypothesis for sex determination, sex is determined by an underlying sex tendency influenced by the genotype and the environment (Vandeputte et al., 2007, 2012). However, the relative contributions of each, their potential interactions although suggested (Saillant et al., 2002) and the potential multigenerational inheritance of sex determining epigenetic patterns remain unknown.

Here, the objective was to study whether the epigenetically-mediated effect of temperature on sex determination of sea bass is linked to the parentally inherited genetic component, as well as the genotype with temperature interactions and the potential non-genetic inheritance of the DNA methylation patterns. We, therefore, obtained progenies from four sires known to produce low or high proportions of females and we subjected them to masculinizing temperatures. We collected biometric data, sex ratio data and

measured the expression and the methylation levels of *cyp19a1a*, as well as the methylation levels of 6 more genes related to sexual development in females and males raised at either low or high temperature. A better understanding of these mechanisms is important to disentangle the effects of genotype and early rearing conditions in sea bass on the sex determination system and ultimately to the sex ratios of the populations.

2. Materials and Methods

2.1. Fish and general rearing conditions

The experiment started by crossing four wild West Mediterranean males with a known tendency to produce, from a previous crossing, low (2 males: sires a and b) or high (2 males: sires c and d) proportions of females in the offspring with two randomly chosen females, thus producing 8 families. The offspring of each male had been recorded in two independent test experiments, where the male a had given 20.5% and 0% females in the offspring, male b 8.80% and 7.50% females, male c 43.2% and 25% females and male d 58.6% and 25.9% females. After the crossing, the females used to obtain the eggs were sacrificed to collect gonadal tissue in liquid nitrogen. Cryopreserved sperm was also kept from the four males for further evaluation of the DNA methylation in the parent fish. After fertilization, the batches corresponding to the four males were incubated separately at ~14.55°C, then equalized at 48 hours post fertilization and mixed in two groups, a male-prone group (sires a and b) and a female-prone group (sires c and d; Fig. S1). A second crossing of eleven females and twenty male albino sea bass was done the same day of the experimental crossing.

2.2. Temperature treatments

The hatched larvae from each group (the male- and female-prone groups) were split in two groups at the age of 13 days-post-fertilization (dpf), and one group was reared at 16.5°C (low temperature, LT) and the other group at 20°C (high temperature, HT) to induce environmentally-mediated masculinization. Thus,

four combinations were available: male-prone (sires a and b) at LT, male-prone at HT, female-prone (sires c and d) at LT and female-prone at HT. Each combination was replicated in two experimental tanks, so that 8 tanks were used during this period (Fig. S1). In each tank, an equal number of albino fish was included, in order to be able to ascertain that the expected distorted sex ratios in high temperature were indeed due to the temperature treatments. Temperature treatments lasted until 65 dpf, the end of the thermosensitive period (TSP), when temperatures were raised to 21°C to allow sufficient growth of the LT groups in order to facilitate sexing of the fish at the end of the experiment. From day 135 onwards, temperature followed natural fluctuations until sampling.

As soon as the fish reached an average of ~1.5 mg (85 dpf for HT and 105 dpf for LT), meristic measurements were taken and fish were individually tagged with nano-tags (Nonatec) to follow their growth rate, which is highly linked to sex determination in the juvenile sea bass (Blázquez et al., 1999). At 115 dpf for HT and 135 dpf for LT, meristic measurements were taken a second time. The fish were individually measured for body weight and length at 219 dpf a third time and a fourth one at sampling.

2.3. Sampling

At 323 dpf, fish were sedated in their rearing tank by adding a mild dose of anesthetic, captured and then euthanized with excess anesthesia. A total of 790 experimental and 868 control albino fish were measured for body weight and length and were sexed by visual inspection. Gonad samples for further molecular biology experiments were preserved in liquid nitrogen for 200 fish in total: 10 females and 10 males from each one of the four LT tanks (80 fish) and 10 females and 20 males from each one of the four HT tanks (120 fish) to account for masculinized females in the HT groups. Additionally, three samples from each sex and each tank (in total 49 samples) were collected for histological confirmation of sex. For a summary of the experimental design see Fig. S1.

2.4. Genotyping

The 790 experimental fish were genotyped for 12 microsatellites markers by Labogena-DNA (Jouy-en-Josas, France). Seven hundred sixty eight of those (97.2%) gave adequate markers' amplification, and 764 (96.7%) were traced back to a single parent pair, using VITASSIGN (Vandeputte et al., 2006) with 1 mismatch tolerated. Almost all fish were assigned to be derived from only one female, so the remaining fish from the second female (only 7 in total) were not included in the analyses, which then comprised 757 fish with known pedigree. The number of fish available for analysis of sex ratios from the LT groups were: 42 from sire a, 147 from sire b, 100 from sire c, and 87 from sire d. The number of fish available for analysis from the HT groups were: 59 from sire a, 134 from sire b, 77 from sire c, and 112 from sire d.

2.5. Quantitative real-time PCR (qRT-PCR)

RNA was extracted from the 200 fish gonads using the TRIzol® Reagent (ThermoFisher Scientific) according to manufacturer's instructions after homogenization of the tissues using a pistil immersed in TRIzol® solution. RNA was quantified using a ND-100 spectrophotometer (NanoDrop Technologies). Five hundred nanograms of RNA were reverse transcribed into cDNA using the SuperScript III Reverse Transcriptase (ThermoFisher Scientific) and 100 µM of random hexamers (ThermoFisher Scientific). Primers for cytochrome p450 aromatase (*cyp19a1a*) were previously validated in sea bass (Díaz and Piferrer, 2015; Navarro-Martín et al., 2011), as well as primers for the two reference genes used, the elongation factor-1 (*ef-1a*) and the 40S ribosomal protein S30 (Mitter et al., 2009). All primers targeted regions between two exons to avoid amplification of possible traces of genomic DNA. Primers efficiency was estimated using serial dilutions (1, 1:5, 1:10, 1:50, 1:100, 1:500) of a pool of 1 µl from each sample (200 µl in total) as $E=10^{(-1/slope)}$, with slope derived from the log-linear regression of the calibration curve. qRT-PCR reactions were carried out in triplicate including negative controls without cDNA in a total volume of 10 µl using the EvaGreen dye (Biotium) under the following conditions: 50ºC for 2 min, 95ºC for 10 min and 40 cycles of 95ºC for 15 s and 60ºC for 1 min. The specificity of the

amplification was evaluated using melting curve with the following conditions: 95ºC for 15 s, 60ºC for 15 s and 95ºC for 15 s. qRT-PCR reactions were performed on an ABI 7900HT machine (Applied Biosystems).

2.6. Quantitative real-time PCR (qRT-PCR) data analysis

Cq values were exported from a multiple plate analysis which included five 384-well plates using the RQ Manager 1.2.1 (Applied Biosystems). The mean Cq values and standard deviations (<0.3) were calculated for each technical triplicate. Relative expression was calculated using the $2^{-\Delta\Delta Cq}$ method (Livak and Schmittgen, 2001). The geometric mean of the two reference genes was subtracted from the *cyp19a1a* Cq values to obtain the normalized dCq for statistical analysis.

2.7. Multiplex bisulfite sequencing (MBS) library preparation and bioinformatics

Two separate MBS libraries were constructed: the first library (MBS1) included *cyp19a1a* in 200 samples and the second (MBS2) included *amh-r2*, *dmrt1*, *er-β2*, *foxl2*, *fsh-r* and *nr3c1* in a representative subset of 99 samples. DNA was extracted by the standard phenol-chloroform-isoamyl alcohol (PCI; 25:24:1) protocol. Treatment with 1 μg of proteinase K (Sigma-Aldrich) and 0.5 μg of ribonuclease A (PureLink RNase A; Life Technologies) were used to eliminate the presence of proteins and RNA respectively. For cryopreserved sperm samples, two PBS washings, followed by a 1:5 dilution in PBS preceded the incubation with proteinase K.

Five hundred nanograms of DNA per sample were bisulfite converted using the EZ DNA Methylation-Direct™ Kit (Zymo Research; D5023) in two batches of 96-well plates, following the manufacturer's instructions with extended desulphonation time to 30 min. Elution of bisulfite converted DNA was performed with 20 μl of Milli-Q autoclaved H_2O passing the same volume twice through the column by centrifugation.

The target region of *cyp19a1a* included the 7 CpGs studied previously in sea bass (Navarro-Martín et al., 2011) at positions -431, -56, -49, -33, -13, 9 and 60 relative to the transcription start site (TSS), encompassing parts of the promoter, 5' UTR and first exon. The primers for *cyp19a1a*, *er-β2*, *nr3c1*, *fshr* and *amh-r2* are described in Chapter 1. The same principles of design were applied for primers targeting the regulatory regions of *dmrt1* and *foxl2* (Table S1). Amplification of the target regions was performed as described in Chapter 1, section 2.4. Annealing temperature was set to 55ºC for *dmrt1* and *foxl2*. The presence and size of the bands were confirmed by agarose gel electrophoresis in a subset of samples. For MBS1, size-selection of PCR was performed by the 2-step size-selection procedure described in Chapter 1, using Serapure magnetic beads. Size-selected PCR products were eluted in 20 µl Milli-Q autoclaved H_2O. For MBS2, 0.6x of magnetic beads were used for the second step of size-selection, normalization followed as described in Chapter 1, elution in 20 µl Milli-Q autoclaved H_2O and pooling of equal volumes per biological sample, resulting in 99 tubes each containing 6 amplicons.

The incorporation of sample-specific indices followed the same strategy as described in Chapter 1. For MBS1, the Nextera XT index Kit SetA for 94 samples and the Nextera XT index Kit SetD (Illumina; FC-131-2001 and FC-131-2004 accordingly) for 95 samples were used. For MBS2, a combination of indices from the same Nextera XT index kits was used. Size-selection and normalization of DNA quantities across samples was carried out after the index PCR according to the customized version of the bead-based normalization of Hosomichi et al (2014) described in Chapter 1. PCR products were eluted in 15 µl Milli-Q autoclaved H_2O and 2 µl of each sample were pooled together. Therefore, we obtained a single multiplexed library with *cyp19a1a* for 200 samples and another single multiplexed library with 6 genes (*amh-r2*, *dmrt1*, *er-β2*, *foxl2*, *fshr* and *nr3c1*) for 99 samples. After pooling, extra clean-up steps were performed using 0.5x magnetic beads in order to ensure the absence of primers.

DNA quantity of final libraries was triple measured by the Qubit dsDNA HS Assay Kit (ThermoFisher Scientific) and the Agilent DNA 1000 chip and DNA High Sensitivity (Agilent) by which the size of bands were also visualized. The multiplexed final libraries

were additionally quantified by real-time qPCR using the Kapa system prior to sequencing on a MiSeq (Illumina) using the paired-end 300 bp protocol.

Bioinformatic analysis were performed as described in Chapter 1 with the following modifications. For *cyp19a1a*, 4 bp from read 1 and 50 bp from read 2 were trimmed from the first step of trimming based on quality controls. Filtering for bases with less than 20 read coverage was performed during the methylation calling procedure. The genomic coordinates of CpGs were intersected with the target region of *cyp19a1a* ± 30 bp.

2.8. General statistical analysis

All statistical analysis were performed using R (R Core Team, 2015; RStudio Team, 2015). A binomial logistic regression with the logit function was used to test for the effects of sire type (female- or male-prone) and temperature on resulting sex ratios. Departures from Fisherian sex ratios and effects of temperature on sex ratios were assessed by Pearson's χ^2 test with Yates' continuity correction. The effects of sex and father were evaluated on body weight, fork length and condition factor (K), the latter defined as $K = 100*(W/L^{3.02})$, where W = weight in g and L = length in cm, of the offspring by two-factor ANOVA after ensuring for homogeneity of variances by Levene's test and normality of the residuals by the Shapiro-Wilk normality test using log-transformed values for body weight and fork length and sin-transformed values for condition factor (K).

2.9. Statistical analysis of DNA methylation data

Association between DNA methylation levels of the CpGs was estimated using Pearson's product-moment correlation coefficients. Mean methylation levels were calculated by averaging the methylation percentages of each CpG per gene per sample and subsequently averaging the overall methylation per grouping factor depending on the comparison. Differences on methylation levels were tested by the Wilcoxon signed rank test applying a continuity

correction (test statistic = W) when needed using arcsine transformed values. For *cyp19a1a* specifically, mean methylation levels were calculated using only the 5 central CpGs (-56, -49, -33, -13 and +9) per sample because the two extreme CpGs (-431 and +60) turned out to be always 100% methylated. A multivariate ANOVA based on euclidean distance matrices with 9999 permutations was used to assess the effect of sex, temperature and sire on DNA methylation levels of the 5 CpGs of *cyp19a1a,* after ensuring for multivariate homogeneity of groups dispersions using the R package vegan (Oksanen et al., 2016).

2.10. Statistical analysis of DNA methylation and gene expression data for cyp19a1a

A multifactorial ANOVA was used to assess the effects of sex, male parent, temperature and methylation levels, as well as their possible interactions, on the expression of *cyp19a1a* using log-transformed $2^{\Delta Cq}$ values. The Shapiro-Wilk normality test was used to confirm normality of the residuals' distribution. The methylation levels in this ANOVA model were integrated as categorical variables, being hypomethylation the first 33.3% (between 5.49% and 68.87% methylation) of the total distribution of values, intermediate methylation the values between 33.3% and 66.6% (between 68.87% and 78.74% methylation) of the total distribution and hypermethylation values above the 66.6% (between 78.74% and 97.88% methylation) of the total distribution. Differences in *cypa19a1a* expression were tested using Wilcoxon signed rank test applying a continuity correction when needed after log-transformation of the $2^{\Delta Cq}$ values to account for normality of the distribution checked by the Shapiro-Wilk normality test. Associations between DNA methylation levels and *cypa19a1a* expression were estimated using Spearman's rank correlation coefficients.

3. Results

3.1. Sex ratios

The progeny of wild males previously recorded to give more female offspring showed indeed a higher percent of females than the progeny of males known to give less female offspring ($p<0.0001$). Sires from a to d gave increasing percentages of female offspring: 11.9% of the offspring of sire a were females, 25.2% of sire b, 35% of sire c and 48.3% of sire d (Fig. 1). Progeny sex ratios in species with PSD are usually skewed. The sex ratios of the progeny of sires a, b and c significantly departed from the Fisherian sex ratio (Table S2). In addition, sex ratios were independent of the replicate tank in which fish were raised since similar sex ratios were observed in replicate tanks (Fig. S2). From the other hand, high temperature, in the progeny of the four sires, masculinized a subset of the genotypic females (Fig. 1) and this masculinization was parallel to that of the albino control fish mixed in each tank (Fig. S3). However, the percentages of fish that would have developed as females at low temperare that were masculinized by elevated temperature were dependent on the father. Thus, the percent of masculinization of these fish was 72% in the offspring of sire a, 50% in the offspring of sire b, 22% in the offspring of sire c and 45% in the offspring of sire d (Fig. 1). The effect of temperature on sex ratio was significant in the progeny of sires b and d, but not in the progeny of sires a and c (Table S2).

The effect of the male parent was evident also on body weight (Fig. S4A), fork length (Fig. S4B), sex dimorphic growth (Fig. S4C) and K (condition factor; Fig. S4A). Sex affected as well body weight, fork length and condition factor (Fig. S4; for the effects of sex and sire see also Table S3).

3.2. DNA methylation levels of cyp19a1a

Of the seven CpGs measured in total, DNA methylation was close to 100% in CpG at positions -431 (Fig. 2A and Fig. S5) and +60 (Fig. 2B and Fig. S5) relative to the TSS in all fish measured. For this reason and to avoid unnecessary noise, we decided to exclude

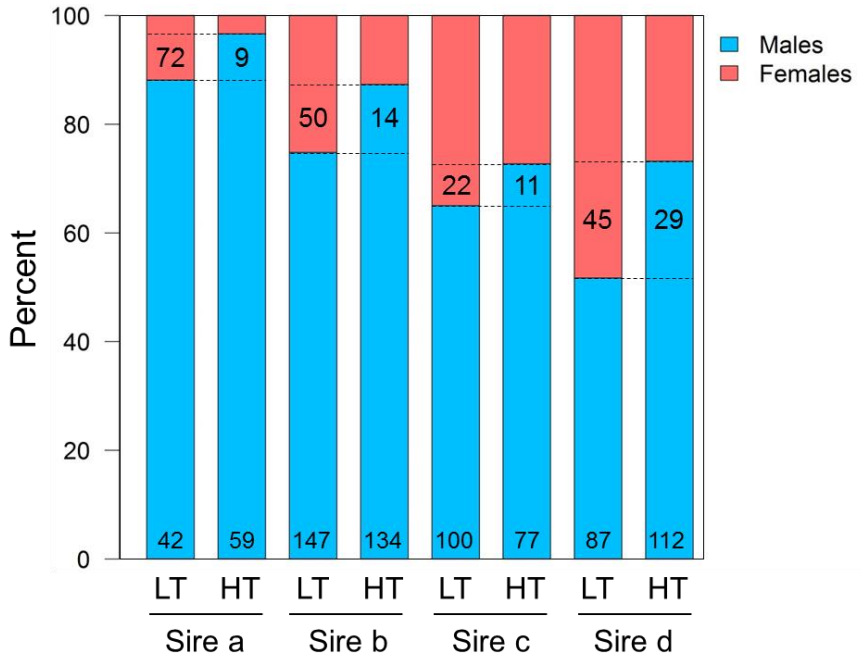


Figure 1. Paternal effects on European sea bass sex ratio and effects of temperature. Percent of female (red) and male (blue) offspring of each sire (a, b, c and d) raised at low (LT) or high (HT) temperature. Absolute numbers of fish are shown in the bottom of each bar. The dotted lines point out the cutoff between the LT and HT female and male percentages. The numbers between the dotted lines in the LT bars indicate the percent of fish that would have developed as females in LT and that are sex-reversed in HT, while the numbers in the HT bars indicate the percentage of presumed neomales among total males.

these two CpGs from further analysis. Among the remaining five CpGs, CpG at position 9 showed a strong positive correlation with the rest of the CpGs at positions -56, -49, -33 and -13 (ρ=-0.77, p<2.2e-16***; Fig. 2C and Fig. S5). Even stronger was the positive correlation of the CpGs at positions -56, -49, -33 and -13 (an example of the four possible combinations is shown in Fig. 2D and all of them in Fig. S5).

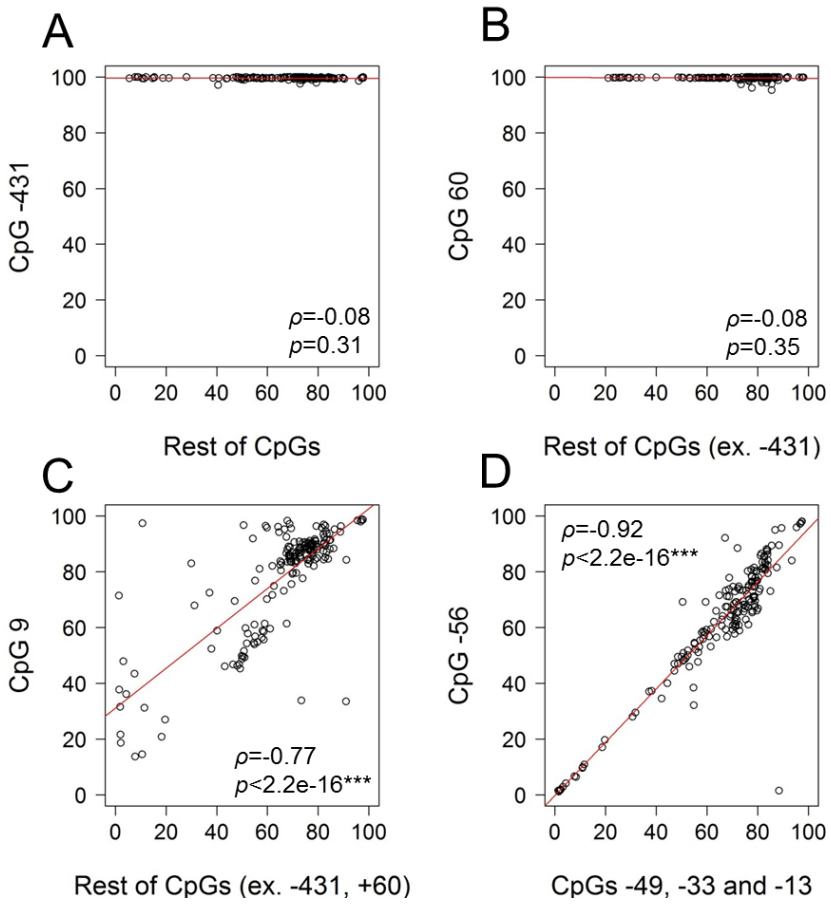


Figure 2. Correlations of DNA methylation among the different CpGs of the *cyp19a1a* promoter. Correlations of DNA methylation are shown for the CpG at position -431 vs. the rest CpGs (A), the CpG at position 60 vs. the rest CpGs excluding the position -431 (B), the CpG at position 9 vs. the rest CpGs excluding positions -431 and 60, and the CpG at position -56 vs the CpGs at positions -49, -33 and -13. Pearson's product-moment correlation coefficients (ρ) and *p*-values of correlation significance are shown.

The mean methylation of the five central CpGs was globally higher in males than in females (W=2984, p=0.04; Fig. 3A) and in the progeny of sire b (W=396, p=0.011; Fig. 3B). In addition, fish raised at high temperature had higher mean methylation levels than fish raised at low temperature (W=4298.5, p=0.02557; Fig. 3C), while the effect of temperature on mean methylation levels was also sire-specific, with a significant increase in the progeny of sire b reared at high temperature (W=927.5, p=0.0003; Fig. 3D).

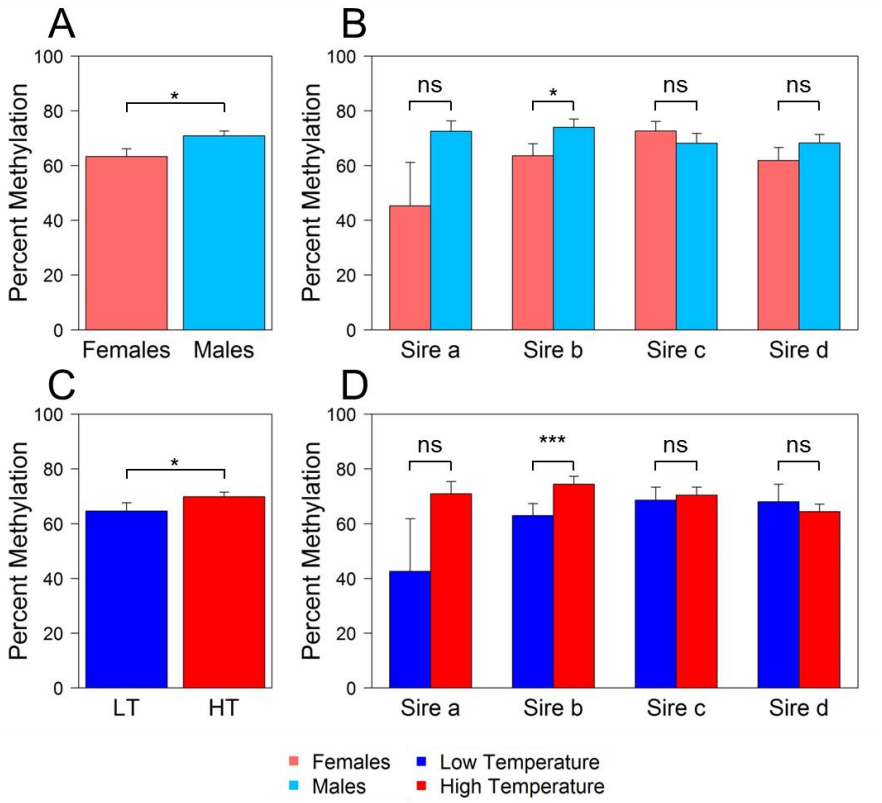


Figure 3. DNA methylation levels of the five central CpGs of the *cyp19a1a* promoter. A) Overall DNA methylation levels per sex. B) DNA methylation levels per sire and per sex. C) Overall DNA methylation levels per temperature during the thermosensitive period. D) DNA methylation levels per sire and per temperature. Mean DNA methylation values measured ± S.E.M. are shown. Data as mean ± SEM. Asterisks indicate significant differences: * = p<0.05; *** = p <0.001.

The factor most significantly affecting the DNA methylation levels of individual CpGs was sex (F=5.391, *p*-value=0.017), followed by temperature (F=2.876, *p*-value=0.077; Table 1). The mean methylation levels of the five central CpGs in the sires' sperm samples were generally high, roughly around 80%, except in sire in which it was around 30% (Fig. 4A). Interestingly, the female offspring of sire a had similarly low methylation levels in fish raised at low temperature (Fig. 4B). However, in the female progeny of sire a that was raised at high temperature, the methylation levels were high (Fig. 4B). In the male offspring of sire a low methylation levels were not observed neither at low nor at high temperature (Fig. 4B).

Table 1. Effects of intrinsic (sex, male parent) and environmental (temperature) factors on DNA methylation of individual CpGs of the European sea bass aromatase promoter as evaluated using multivariate ANOVA using euclidean distance matrices with 9999 permutations

	df	SS	MS	F model	R^2	*p*
Sex	1	11897	11896.6	5.391	0.029	**0.017**
Temperature	1	6346	6346	2.876	0.016	0.077
Male parent	3	8809	2936.4	1.331	0.022	0.249
Interaction Sex:Temperature	1	538	538.1	0.244	0.001	0.722
Interaction Sex:Male parent	3	10888	3629.3	1.645	0.027	0.167
Interaction Temperature:Male parent	3	12289	4096.5	1.856	0.030	0.128
Interaction Sex:Temperature: Male parent	3	5435	1811.7	0.821	0.013	0.496
Residuals	160	353059	2206.6	0.863		
Total	175	409261	1			

Abbreviations: d.f., degrees of freedom; SS, Sums of Squares; MS, Means of Squares; *p*, *p*-value.

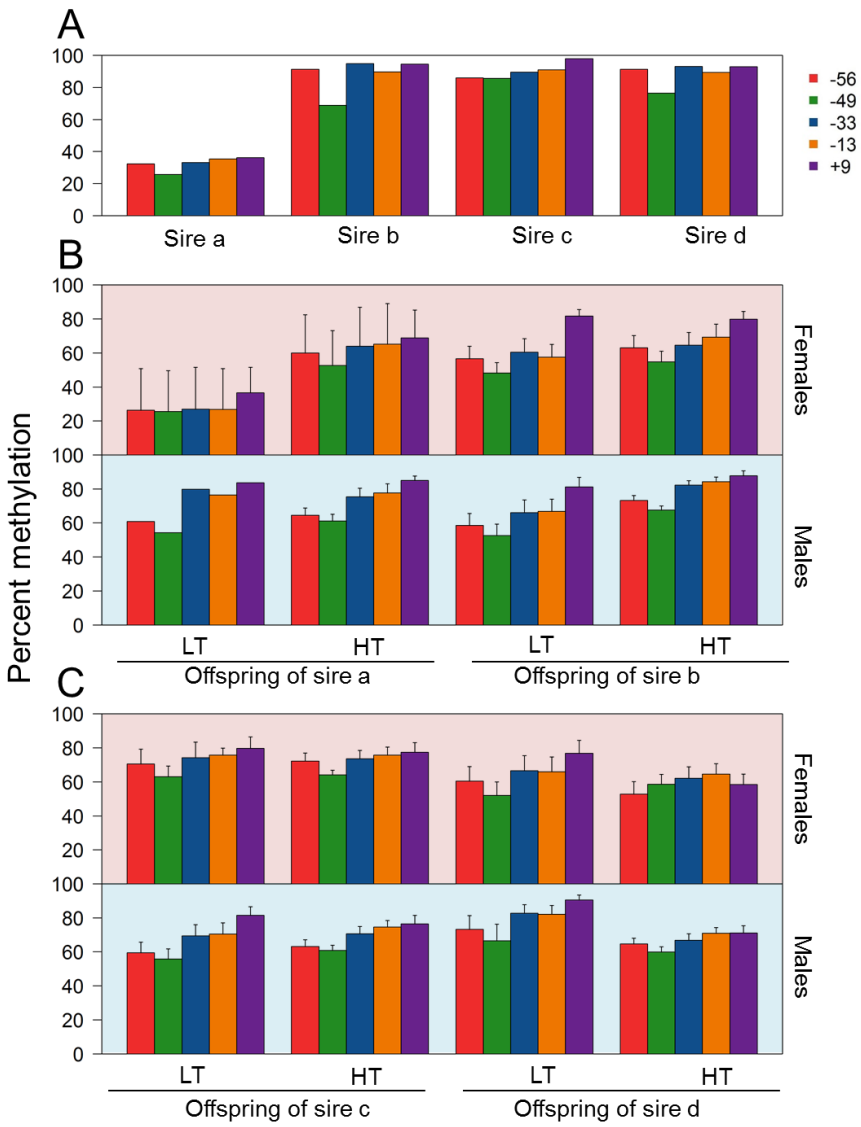


Figure 4. Mean DNA methylation levels of CpGs -56 to 9 of the *cyp19a1a* promoter. DNA methylation levels are shown in each of the four sires (A). Mean methylation ± S.E.M. in female (pink background) and male (lightblue background) offspring of sires a and b (B) and female (pink background) and male (lightblue background) offspring of sires c and d (C) raised at low (LT) or high (HT) temperature.

3.3. DNA methylation of other genes involved in sexual development

The overall DNA methylation profile of the six genes related to sexual development was gene-specific showing a wide range of methylation profiles from 0.3 to 100%. Across the set of genes associated to male development, in *amh-r2* (Fig. 5A) and *er-β2* (Fig. 5C) no significant differences between sexes were observed. In contrast, *dmrt1* had lower levels of DNA methylation in males than in females (p <1.8e-10***; Fig. 5B). A clear decrease of methylation levels in fish reared at high temperature was observed in all genes related to male development, with lower methylation percentages in fish reared at high temperature (*amh-r2*, p < 2.72e-4***; *er-β2*, p < 2.85e-5***; and p < 0.018*, *dmrt1*; Fig. 5D, E, F).

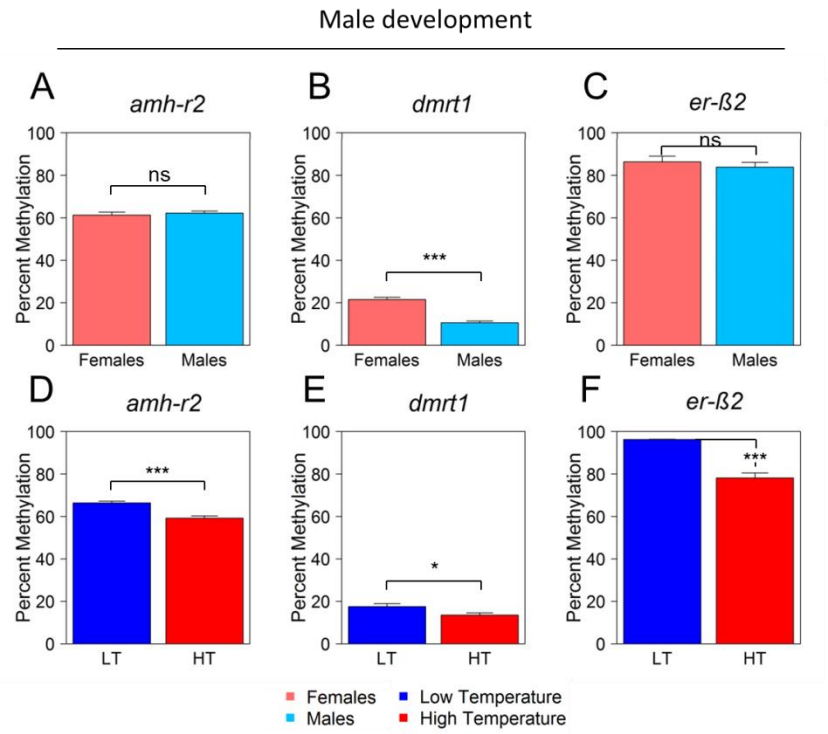


Figure 5. DNA methylation levels of genes related to male development. Overall DNA methylation levels per sex in A) *amh-r2*, B) *dmrt1* and C) *er-β2*. Overall DNA methylation per temperature during the thermosensitive period in D) *amh-r2*, E) *dmrt1* and F) *er-β2*. Data as mean ± SEM. Asterisks indicate significant differences: * = p<0.05; *** = p<0.001; ns=not significant.

Regarding the genes not associated with the development of a particular sex, *fsh-r* DNA methylation was lower in fish reared at high temperature (p < 3.4e-4***; Fig.6D), while no significant differences were found by sex nor temperature in *nr3c1* (Fig. 6B; E). Nevertheless, in *nr3c1*, there were higher levels of methylation in females reared at high temperature when compared to females reared at low temperature (p<0.02). In *foxl2*, a gene related to ovarian development, mean methylation was higher in males than in females (p < 0.034; Fig. 6C) and in fish reared at high temperature (p < 0.037; Fig. 6F). A summary of direction of DNA methylation changes in all genes involved in sexual development according to sex and temperature is shown in Table 2.

Table 2. Summary of genetic and environmental effects on the DNA methylation in the promoter region of candidate genes associated with European sea bass sexual development

Gene	Sex-associated	Females vs. males	High temp. vs. low temp.
cyp19a1a	Female development		
foxl2	Female development		
dmrt1	Male development		
amh-r2	Male development	ns	
er-β2	Male development	ns	
fsh-r	Both sexes	ns	
nr3c1	Both sexes	ns	

DNA methylation relative values color codes: Blue = lower; Red = higher.

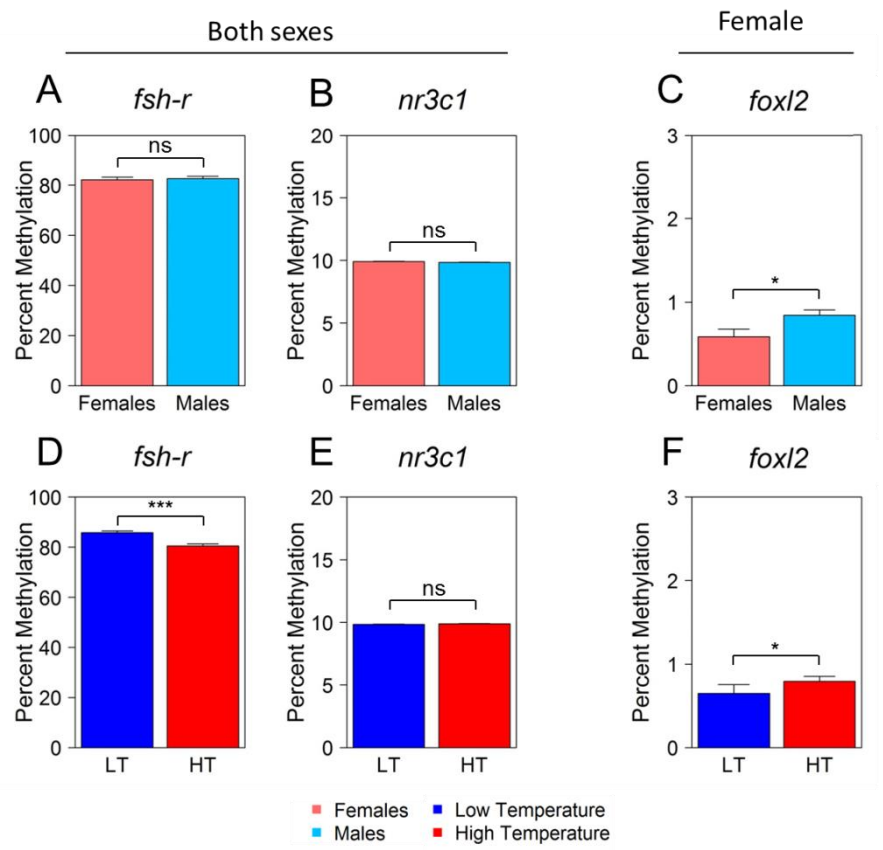


Figure 6. DNA methylation levels of genes related to development of both sexes and female development. Overall DNA methylation levels per sex in A) *fshr*, B) *nr3c1* and C) *foxl2*. Overall DNA methylation levels per temperature during the thermosensitive period in D) *fhsr*, E) *nr3c1* and F) *foxl2*. Data as mean ± SEM. Asterisks indicate differences: *** = $p<0.001$; * = $p<0.05$; ns=not significant.

The effects of temperature on DNA methylation of the progeny were sire-dependent, with the effect of temperature evident in the offspring of sires c and d in *amhr-2*, *er-β2*, *fshr* and only of sire d in *folx2* and *nr3c1* (Table 3).

Table 3. The effects of sex, temperature by sire on mean DNA methylation tested by Wilcoxon signed rank test

Factor	Gene	Sire	*p*-value	Significance level
Sex	*amh-r2*	a	0.785	ns
		b	0.988	ns
		c	0.837	ns
		d	0.295	ns
	dmrt1	a	0.079	ns
		b	2.60E-06	***
		c	0.01	**
		d	0.002	**
	er-β2	a	0.412	ns
		b	0.211	ns
		c	0.507	ns
		d	0.731	ns
	foxl2	a	0.066	ns
		b	0.61	ns
		c	0.436	ns
		d	0.701	ns
	fshr	a	1	ns
		b	0.479	ns
		c	0.423	ns
		d	0.945	ns
	nr3c1	a	1	ns
		b	0.168	ns
		c	0.477	ns
		d	0.731	ns
Temperature	*amh-r2*	a	0.432	ns
		b	0.116	ns
		c	0.004	**
		d	0.008	**
	dmrt1	a	0.767	ns
		b	0.22	ns
		c	0.759	ns
		d	0.366	ns
	er-β2	a	0.362	ns
		b	0.962	ns
		c	1.10E-06	***

	d	0.001	**
foxl2	a	0.271	ns
	b	0.61	ns
	c	0.09	ns
	d	0.027	*
fshr	a	0.953	ns
	b	0.582	ns
	c	4.50E-06	***
	d	0.001	**
nr3c1	a	0.361	ns
	b	0.386	ns
	c	0.926	ns
	d	0.001	**

3.4. DNA methylation and gene expression relationships in cyp19a1a

Sex was also indeed the factor affecting the most the expression of *cyp19a1a* (F=40.603, p <0.001), followed by the male parent (F=2.79, p =0.043; Table 4 and Fig. 7A and 7B and for the distribution of expression values see Fig. S6), while there was also a significant interaction between sex and male parent (F=3.014, p=0.032). The effect of temperature on *cyp19a1a* expression was not significant (F=0.005, p=0.946; Table 4), although in fish raised at low temperature there was a slightly lower expression (Fig. 7C and 7D).

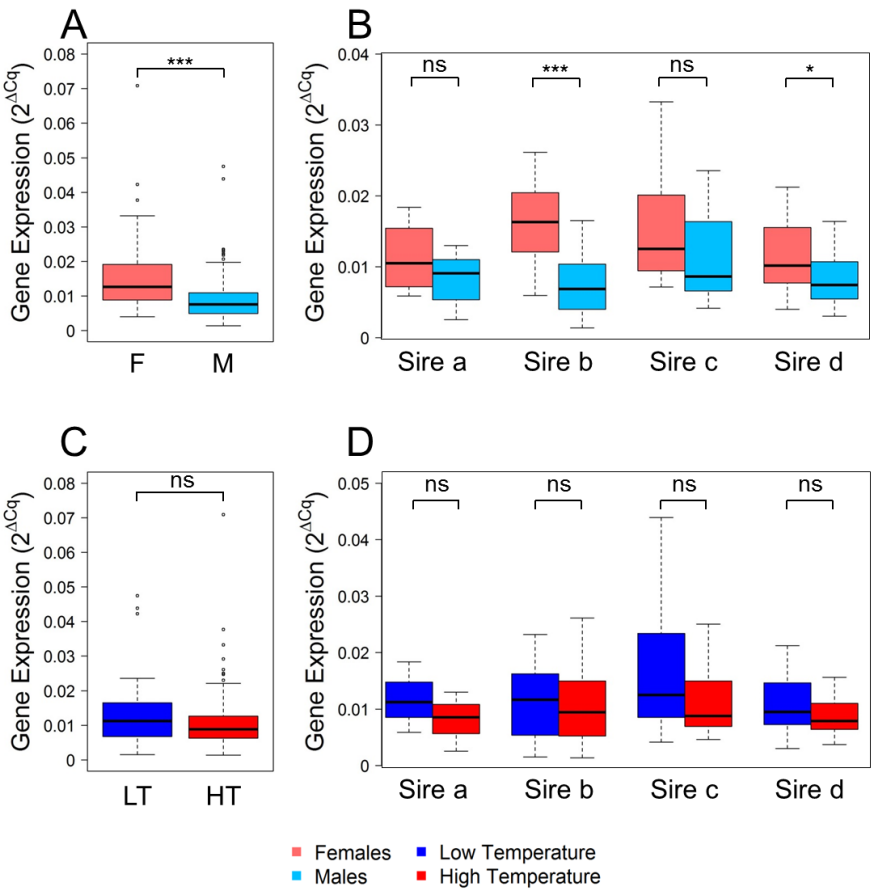


Figure 7. Expression of *cyp19a1a*. Expression is shown in females (F) and males (M; A), in female and male offspring of each sire (B), in fish raised at low (LT) and high (HT) temperature (C) and in the offspring of each sire reared at low and high temperature (D). Expression is displayed as $2^{\Delta Cq}$ values by boxplots where the boxes include the values distributed between the lower and upper quartiles, the upper whisker = min(max(x), Q3 + 1.5 * IQR), the lower whisker = max(min(x), Q1 – 1.5 * IQR), where IQR= third quartile (Q3) – first quartile (Q1), the black horizontal line in the box indicates the median and the points outside the boxes represent values higher than the upper whisker. Asterisks indicate significant differences: *** = $p<0.001$; *=$p<0.05$.

Table 4. Contribution of intrinsic (sex, male parent, promoter methylation) and environmental (temperature) factors to the expression of *cyp19a1a* as assessed by multifactorial ANOVA

		d.f.	SS	MS	F-value	*p*-value	Significance level
Factors	Sex	1	11.41	11.409	40.603	2.97E-09	***
	Sire	3	2.35	0.784	2.79	0.0432	*
	Methylation level	2	0.78	0.392	1.396	0.2513	
	Temperature	1	0	0.001	0.005	0.9457	
Interactions	Sex:Sire	3	2.54	0.847	3.014	0.0324	*
	Sex:Methylation level	2	7.16	3.579	12.737	8.88E-06	***
	Sire:Methylation level	6	1.5	0.25	0.889	0.5048	
	Sex:Temperature	1	0.84	0.836	2.975	0.0869	.
	Sire:Temperature	3	1.23	0.409	1.457	0.2294	
	Methylation level:Temperature	2	0.16	0.081	0.289	0.7493	
	Sex:Sire:Methylation level	5	1.51	0.301	1.072	0.3787	
	Sex:Sire:Temperature	3	0.73	0.243	0.865	0.4611	
	Sex:Methylation level:Temperature	2	0.29	0.143	0.508	0.6027	
	Sire:Methylation level:Temperature	4	0.63	0.156	0.556	0.6947	
	Sex:Sire:Methylation level:Temperature	4	1.47	0.368	1.308	0.2704	
	Residuals	130	36.53	0.281			

Abbreviations: d.f., degrees of freedom; SS, Sums of Squares; MS, Means of Squares

In addition, there was a significant effect of DNA methylation level on *cypa19a1a* expression depending on the sex (F=12.737; $p<0.001$). More specifically, when the promoter of *cypa19a1a* was hypermethylated in both sexes, there was no difference in neither methylation nor expression of *cypa19a1a* between females and males (W=274, *p*-value=0.05767 and $t=1.3026$, $p=0.1982$ respectively; Fig. 8). However, when there was intermediate methylation or hypomethylation there was significantly more methylation in males than females (W=198, $p=0.00059$ and W=188, $p=0.00048$ respectively), as well as significantly less expression in males than in females ($t=3.1649$, $p=0.002496$ and $t=6.0019$, $p=3.191e-07$ respectively). Furthermore, the differences of *cyp19a1a* between females and males were higher when there was hypomethylation than intermediate methylation.

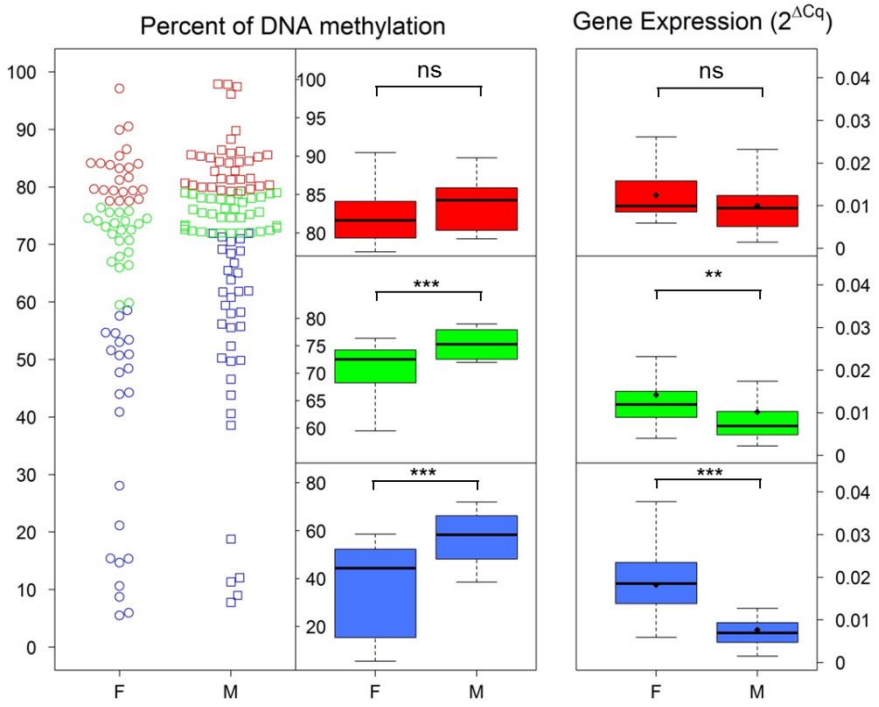


Figure 8. Expression of *cyp19a1a* in females and males with low, intermediate or high DNA methylation levels. In the left part, the distribution of DNA methylation values is shown by individual points in females (F) and males (M), where in blue (low) are the points below the 33.3% of the total distribution, in green (intermediate) are the points between the 33.3% and the 66.6% of the total distribution and in red (high) are the points above the 66.6% of the total distribution. The central boxplots represent low (blue), intermediate (green) and high (red) DNA methylation levels in females and males. The right part displays the distribution of *cyp19a1a* expression depending on the level of DNA methylation in females and males. Data as mean ± SEM. Asterisks represent the level of significance of Student's *t*-test between females and males: **=p<0.01 and ***=p<0.001. Notice the increase in inverse relationship between DNA methylation and gene expression.

Overall, there was a negative correlation of *cypa19a1a* expression with the mean DNA methylation (ρ=-0.14, p =0.03672), as well as with the DNA methylation levels of the CpGs at positions -49 (ρ=-0.14, p=0.03266), -33 (ρ=-0.13, p=0.04445) and -13 (ρ=-0.17, p=0.01303; Fig. 9). However, separating males from females, negative correlations of *cyp19a1a* expression were observed in females (ρ=-0.31, p=0.00448) and positive correlations were observed in males (ρ=0.23, p=0.009907; Fig. 10). The correlation coefficients varied depending on the male parent and the temperature.

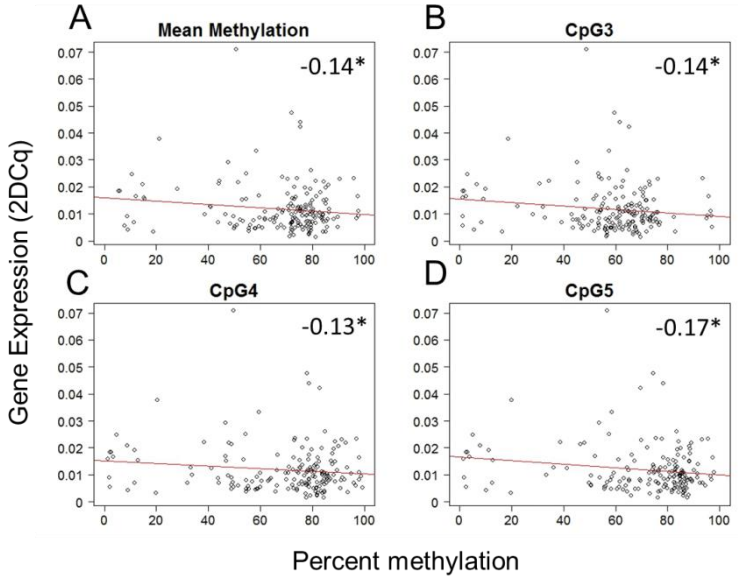


Figure 9. Relationship between *cyp19a1a* DNA methylation and gene expression. Correlation of *cypa19a1a* expression with the mean methylation levels (A), the CpG at position -49 (B), the CpG at position -33 (C) and the CpG at position -13 (D). Associations between *cypa19a1a* expression and DNA methylation are shown with Spearman's rank correlation coefficient (ρ) with the level of significance: *=p<0.05.

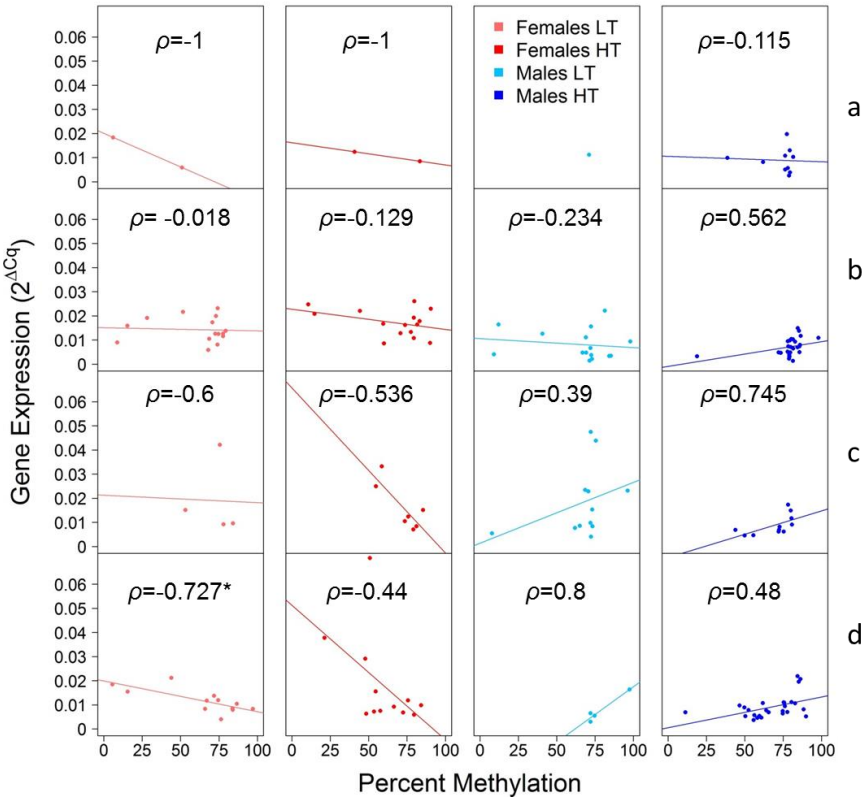


Figure 10. Correlation of expression of *cyp19a1a* and mean DNA methylation of the five central CpGs of its promoter. Correlations are shown per sex (A), per temperature (B) and per father (C). Expression is shown as 2^{DCq} values. Associations between *cypa19a1a* expression and DNA methylation are shown with Spearman's rank correlation coefficient (ρ) with the level of significance: *=p<0.05.

4. Discussion

In this study, we show for the first time that there is a genetic and environmental component, the former being at least sire-dependent, to the DNA methylation of the promoter of *cyp19a1a* and other key genes involved in the sexual development of the European sea bass. The higher levels of methylation in the *cyp19a1a* promoter in males and in fish subjected to high temperature during early development are dependent on the male parent. We demonstrate that these methylation levels negatively regulate *cyp19a1a* expression in females but correlate positively with *cyp19a1a* expression in males. In addition, six more genes of the sexual development pathways show gene-specific hyper- or hypo-methylation depending on sex and temperature with sire-specific contribution.

The sex ratios were male-biased in the offspring of sires previously recorded to produce male-prone progeny, according to expectations. In addition, we estimated that 22–72% (average ~50%) of fish that would develop as females when subjected to high temperature during the TSP differentiated as males in all offspring, as previously reported under the same temperature regime (Navarro-Martín et al., 2009). However, the masculinizing effect of temperature was significant in the progeny of two sires, b and d, confirming the family-specific response to temperature shown previously (Saillant et al., 2002). Indeed, in Nile tilapia, family-specific QTLs have been identified that associate with the temperature-dependent sex (Lühmann et al., 2012). In this study, we used the offspring of one female. This allowed us to attribute any parental effect to the male parent, but from the other hand prevents the evaluation of potential dam-specific effects on the sex ratios. Thus, the male parent and the temperature contributed to the sex ratios of the progeny.

Temperature influences on sex determination are mediated by an epigenetic mechanism involving DNA methylation changes (Navarro-Martín et al., 2011). However, sex determination in a polygenic system, even if influenced by temperature, has genetic components. Here, to disentangle the genetic from the environmental effects on sex ratios and DNA methylation, we used 4 sires and 1 dam, produced full-sib families and subjected them to low or high temperature during the TSP. Seven hundred ninety fish

from the offspring were sexed and genotyped and 200 fish were used for evaluating the DNA methylation of 7 genes related to sex determination and differentiation, among them *cyp19a1a* for which gene expression was also measured. Equal numbers of albino fish included in the experimental tanks ascertained for temperature-mediated sex ratio distortions and double the amount of males was assessed for each HT group taking into account the fish that would potentially differentiate as females at LT, but developed as males due to temperature during the TSP. Therefore, this is the largest, most complete and carefully designed study to date focusing on the genetic and environmental components of the epigenetic contribution to sex ratios.

Higher DNA methylation levels in the promoter of *cyp19a1a* were found in males, as well as in fish subjected to high temperature during the TSP, as previously described in sea bass (Navarro-Martín et al., 2011). However, compared to the previous study we obtained globally higher levels of DNA methylation. Since these DNA methylation levels could be partitioned according to the male parent, it became evident that they were sire-dependent. Compared to the study of Navarro-Martín et al. (2011), similar global and sex-specific DNA methylation levels, as well as temperature-responses, were apparent in the progeny of sire a. In sire b, males and fish subjected to high temperature had higher methylation in *cyp19a1a* promoter, but globally higher methylation percentages than in the study of Navarro-Martín et al. (2011). However, the offspring of the other two sires did not show differential methylation according to sex and/or temperature, suggestive of the importance of the genetic component for DNA methylation, as well as of the temperature-responsiveness. Indeed, genetic variation is known to influence epigenetic variation by the presence of single nucleotide polymorphisms (SNP) which may influence the methylation of CpG sites by acting in *cis* (Banovich et al., 2014; Bell et al., 2011; Hannon et al., 2016; Teh et al., 2014). In the promoter of *cyp19a1a*, three alleles have been identified (Galay-Burgos et al., 2006) which may be associated with the methylation status of CpG sites in *cis*.

DNA methylation alterations in regulatory gene regions have generally functional consequences through gene expression changes and ultimately to the phenotypic outcome. Higher methylation of the *cyp19a1a* promoter in males and in females subjected to high

temperature during early development has been associated with lower *cyp19a1a* expression (Navarro-Martín et al., 2011), regardless of whether there is a causal relationship. Here, there was lower *cyp19a1a* expression in males as expected and the magnitude of differences in expression between sexes was sire-specific. However, although a tendency for lower *cyp19a1a* expression levels was evident in fish reared at high temperature during the TSP, there was no direct significant effect of temperature. Indeed, a more complex regulation of *cyp19a1a* expression is apparent from this study. The model of lower *cyp19a1a* expression and higher DNA methylation of the promoter in males than in females holds only when the overall methylation percentages are low or intermediate, but not when there is sex-independent hypermethylation, indicative of a methylation threshold below which *cyp19a1a* expression is inversely related to methylation. Since DNA methylation is itself influenced by temperature in a sire-specific manner, these results imply an indirect effect of temperature on *cyp19a1a* expression through DNA methylation.

Interestingly, there is negative correlation of DNA methylation with gene expression in females, but in males the correlation is positive. Several recent studies have revealed that a positive correlation of DNA methylation with gene expression is present in the genome (Bell et al., 2011; van Eijk et al., 2012; Gibbs et al., 2010). Importantly, there are *cis* acting genetic loci associated with DNA methylation of CpGs, especially outside CpG islands, as well as with the expression of nearby genes in which positive correlation of DNA methylation and expression is evident (Gibbs et al., 2010). The positive correlation observed in males could be, therefore, resulting from the genetic component of the sex determination system. This would be the case of male-specific genetic variants in genomic proximity to *cyp19a1a* that influence the methylation and the expression of the gene.

The DNA methylation of other genes related to sex determination and differentiation showed gene-specific levels affected also by the male parent and the temperature that were sex-dependent. One of the most important transcription factors regulating the expression of *cyp19a1a* is Foxl2 (Pannetier et al., 2006; Wang et al., 2007) which presented similar patterns of DNA methylation in its gene promoter. The expression of *foxl2* is known to be correlated with the

expression of *cyp19a1a*, is a marker of ovarian differentiation (Wang et al., 2007) and downregulated by high temperature (Piferrer and Guiguen, 2008; Poonlaphdecha et al., 2013; Yamaguchi et al., 2007). According to its characteristics and despite the overall lower levels of DNA methylation compared to *cyp19a1a*, *foxl2* had higher methylation in males and in fish subjected to high temperature during the TSP, indicative of similar epigenetic regulation of both genes studied here related to female development.

The methylation levels of three genes associated with male development: *dmrt1* (Deloffre et al., 2009), *er-β2* (Blázquez et al., 2008) and *amh-r2* (Rocha et al., 2016) were measured. *Dmrt1* is a key gene involved in testis differentiation in vertebrates and its expression shows a contrasting pattern to *cyp19a1a* according to sex (Deloffre et al., 2009; Piferrer and Guiguen, 2008). As expected by the model of negative regulation of gene expression by DNA methylation of the promoter and according to what has been found in the half-smooth tongue sole (Shao et al., 2014) and in the Japanese flounder (*Paralichtus olivaceus*; Wen et al., 2014), DNA methylation levels of *dmrt1* promoter were higher in females than in males independently of the male parent. The sires exhibited male-type methylation levels and in the female offspring, there was sire-specific increase of DNA methylation percentages. In the half-smooth tongue sole, *dmrt1* methylation is maintained low in testis and gains methylation in ovaries, which is likely to be the case in sea bass as well (Shao et al., 2014). Furthermore, *er-β2* has higher expression levels in testis and has been suggested as important for male development (Blázquez et al., 2008), but did not show differential patterns of DNA methylation by sex. However, DNA methylation was affected by temperature with sire-specific lower methylation at high temperature. In the half-smooth tongue sole, *amh-r2* is less methylated and more expressed in the testis (Shao et al., 2014). Nevertheless, in this study *amh-r2* followed the same pattern as *er-β2*, exhibiting no sex-specific differences, but decreased methylation in fish subjected to high temperature, which was independent of the male parent. The higher methylation of *amh-r2* and *er-β2* in fish reared at low temperature could be associated to the inhibition of their expression at the specific developmental stage of this study. Conversely, the lower

methylation of these genes in fish reared at high temperature is suggestive of an epigenetic activation inducing testis development.

Fshr and *nr3c1* are genes related to gonadal development of both sexes (Rocha et al., 2007, 2009). *Fshr* is known to exhibit decreased expression in fish subjected to high temperature (Díaz and Piferrer, 2015; Yamaguchi et al., 2007), in accordance to the lower levels of methylation we detected in fish reared at high temperature. The temperature response of *fshr* seems to follow the DNA methylation patterns of the other genes related to male development. On the contrary, the expression of *nr3c1* increases with high temperature treatment during early development (Díaz and Piferrer, 2015; Fernandino et al., 2012), but there was no significant effect detected in this study associated to temperature, while *nr3c1* seems to follow the DNA methylation patterns of the other genes associated with female development. Other epigenetic mechanisms could be involved in the regulation of *nr3c1* expression resulting in long-lasting higher expression levels in fish reared at high temperature during early development.

Importantly, we measured the DNA methylation levels in the *cyp19a1a* promoter of each sires separately. In the sire that followed the model of Navarro-Martín et al. (2011), sire a, there was low DNA methylation and similar levels were observed in the female offspring reared at low temperature, but higher levels were present in female offspring as well as in males independently of temperature. These results suggest an epigenetic inheritance of DNA methylation with potential of conversion by temperature and possibly other factors related to male development.

Epigenetically mediated responses of phenotypically plastic traits, such as sex, can be adaptive or maladaptive depending on the speed of environmental change. It has been suggested that these responses may have implications on population sex ratio in species responding to climate change, as well as for farmed species (Consuegra and Rodríguez-López, 2016; Piferrer, 2016). In sea bass aquaculture, male-biased stocks are often still present despite of thermal protocols applied to control sex. This could be due to an epigenetic maladaptive response or epigenetic trap leading to male-biased offspring of male-biased stocks. Thus, in a sort of "epigenetic programming" applied to fish farming (Moghadam et al., 2015), the

identification of broodstock fish with a particular methylation profile holds promise because these animals would pass to their offspring specific DNA methylation marks. These epigenetically inherited DNA methylation profiles would provide offspring with desired features. For example, by conferring them resistance to the masculinizing effect of elevated temperature

In conclusion, the genetic component of the sex determination system controls the epigenetic link to temperature influences. The inhibitory role of temperature on *cyp19a1a* expression through methylation increases is not linear, but rather affected by the genetic component and other factors. Other genes related to male and female sex differentiation also respond to high temperature during early development in a parent-specific manner. In addition, there seems to be indication of father-to-daughter inheritance of epigenetic marks, although this aspect needs confirmation. Taking these into account, a complex epigenetic layer contributing to sex determination and differentiation is revealed, adding to the better understanding of the shaping of population sex ratios. The sea bass is a vertebrate where a major plastic phenotypic trait, sex, is under the control of genetic and environmental influences with approximate strength each. This study clearly illustrates how the epigenome is crucial for the integration of genomic and environmental information. It also shows variation in this epigenetic component and calls for further studies to gain a better picture of the interplay between these different regulatory components that bring, as Waddington said, “the phenotype into being”.

Acknowledgments

We thank Alain Vergnet, François Allal, Marie-Odile Vidal and François Ruelle for raising the fish and help with the samplings; Sílvia Joly for performing the RNA extractions and qRT-PCR; Dr. Catherine Labbé and Alexandra Depince for help with the sperm DNA extraction protocol. DA was supported by a Ph.D. scholarship from the Spanish Government (BES–2011–044860). Research was supported by the Spanish Ministry of Economy and Competitiveness "Epifarm" ref. AGL2013–41047–R to FP and by AquaExcel (FP7 UE) "Sea Bass Sex (Epi)genetics" ref 0102/06/07/20 to MV and FP.

Supplementary Figures

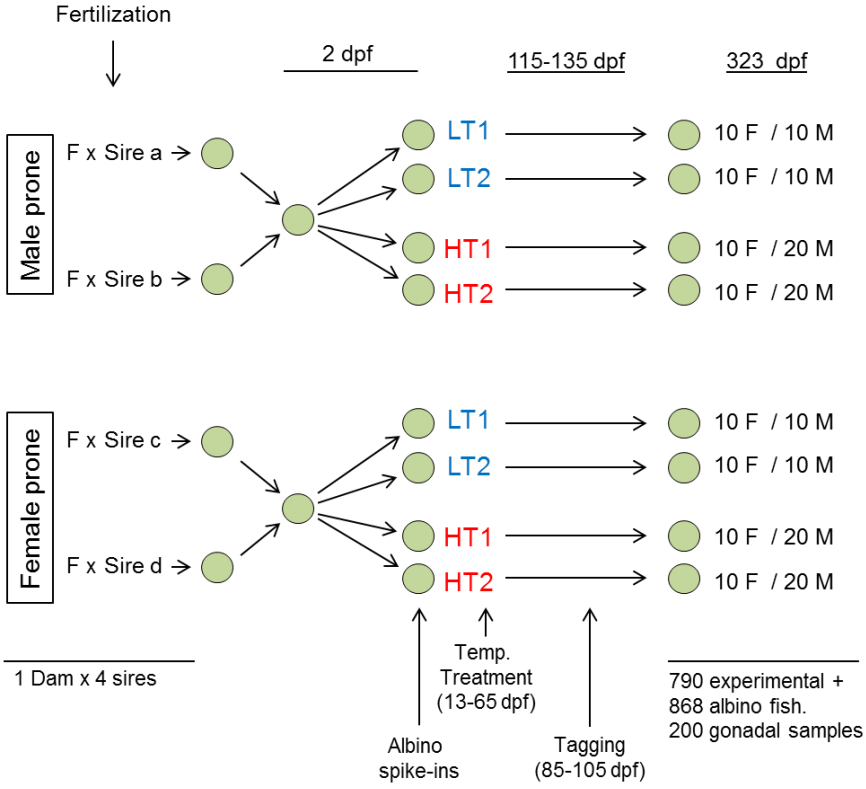


Figure S1. Experimental set-up of crossings and temperature treatments. Eggs were obtained from crossing males known from previous experiments to produce offspring with lower (Male prone, sires a and b) or higher (Female prone, sires c and d) percentage of females. Two days post fertilization (dpf) eggs were mixed according to the male prone and female prone groups. On day 13, larvae from the two groups were divided into four with half of the fish being raised at low temperature (LT; 16.5ºC) and the other half at high temperature (HT; 21ºC) until 65 dpf, the end of the thermosensitive period. Fish were sampled at one year of age (323 dpf) and samples for molecular biology experiments were taken from 10 females and 10 males for the LT groups and 10 females and 20 males from the HT groups.

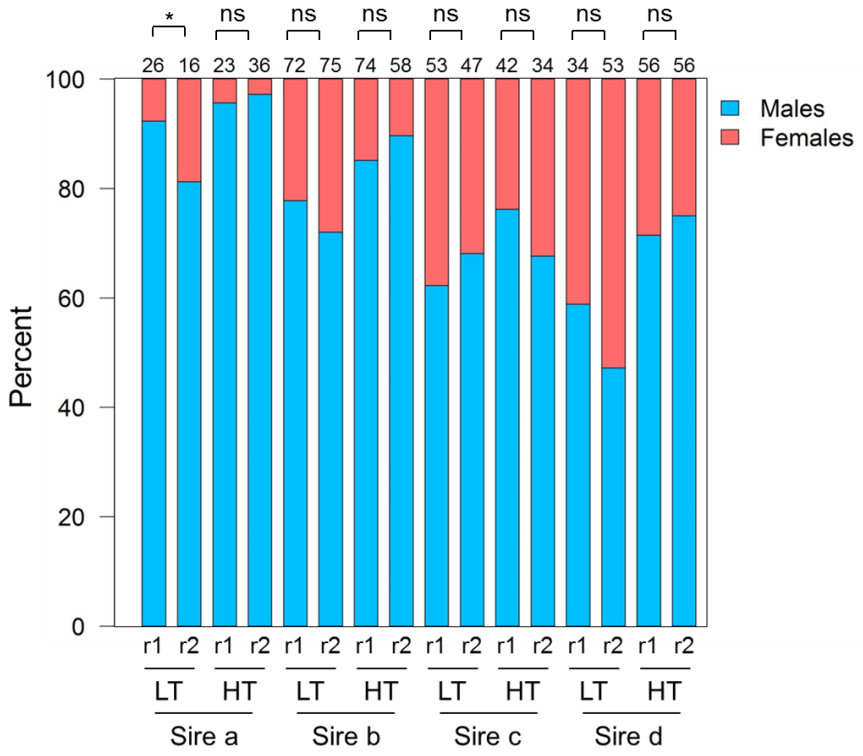


Figure S2. Percent of males and females per experimental tank and group. Two replicate tanks (r1 and r2) were used for low (LT) and high (HT) temperature per each sire. Total number of fish per group are indicated above the bars.

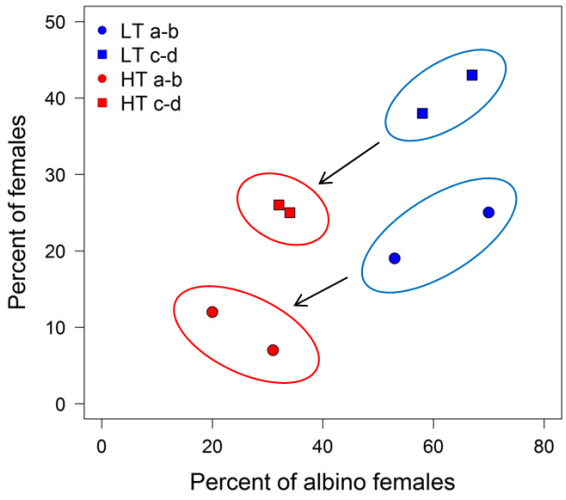


Figure S3. Sex ratio of experimental and albino fish. Points indicate the percent of females of experimental and albino fish in the same tank and circles surround the two replicate tanks for each experimental group: offspring of sires a and b (circles) at low (blue) and high (red) temperature and offspring of sires c and (squares) at low (blue) and high (temperature).

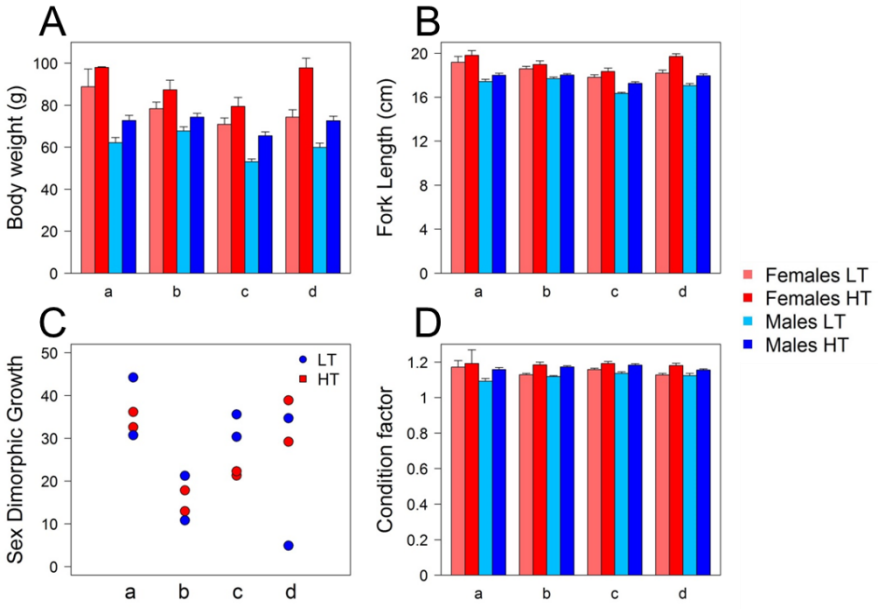


Figure S4. Weight, length, sexual growth dimorphism and condition factor (K). The mean body weight (g; A), fork length (cm; B) and condition factor (K; D) ± S.E.M. are shown per sire (a, b, c and d) for females (red) and males (blue) at sampling. Sexual growth dimorphism (SGD) is shown per sire (a, b, c and d) for low (black circles; LT) and high (white squares; HT) temperature.

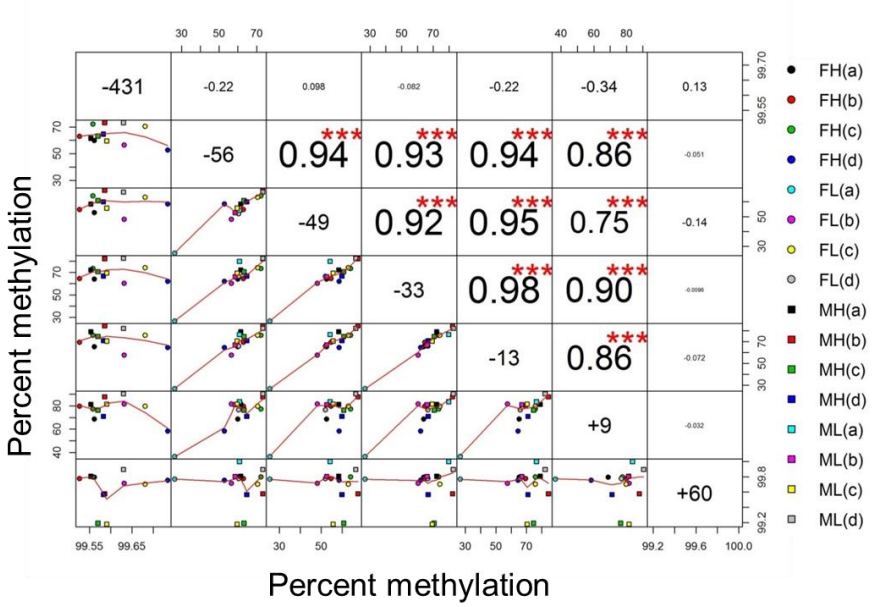


Figure S5. Pairwise correlations of methylation levels of the 7 CpGs of the *cyp19a1a* promoter and first exon. In the diagonal the position of the CpG relative to the transcription start site (TSS) is shown. The lower part of the figure displays scatterplots of the mean methylation of the vertical CpG (x-axis) and the horizontal CpG (y-axis) per temperature, sire and sex (females indicated by circles and males by squares). Loess smoothers are shown in red in each scatterplot. Pearson's correlation coefficients are displayed in the upper part of the diagonal with the significance level of the correlation denoted (***=p<0.001).

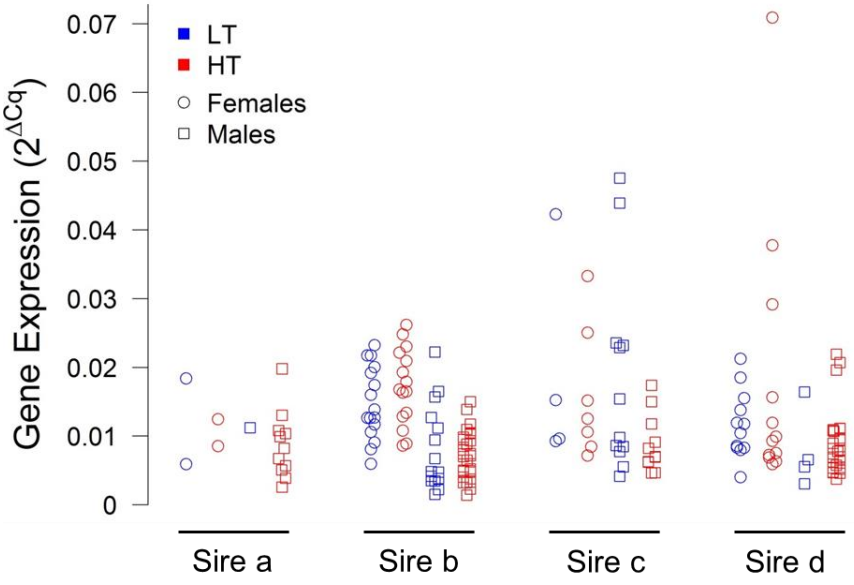


Figure S6. Distribution of *cyp19a1a* expression. Four groups of dotplots are shown for each sire (a, b, c and d) indicating females (circles) and males (squares) reared at low (LT) or high (HT) temperature. Expression is shown as $2^{\Delta Cq}$ values.

Supplementary Tables

Supplementary Table 1. Primers for *dmrt1* and *foxl2* used in MBS2 and regions of target

	dmrt1	*foxl2*
Gene name	doublesex and mab-3 related transcription factor 1	forkhead box l2
Forward (5' to 3')	TGAGAGTGGGTGTATGTTATTGTTT	TAGTTTGTGAGGATATGTTTGAGAAG
Reverse (5' to 3')	ACTAACAATCCCTCCAATTACAAAA	TTCCCAATAAAAACAATACATCATC
Position	LG20: 18505581-18506025	LG13: 7386529-7386830

Supplementary Table 2. Pearson's χ^2 test with Yates' continuity correction for sex ratios

	Sire	χ^2	*p*-value
Departure from Fisherian sex ratio	a	12.53300	0.00040
	b	18.27300	0.00002
	c	4.01020	0.04522
	d	0.00575	0.93960
Effect of temperature	a	1.59570	0.20650
	b	6.25570	0.01238
	c	0.87030	0.35090
	d	8.88480	0.00288

Supplementary Table 3. ANOVAs for body weight, length and condition factor

Variable	Factor	d.f.	Sums of Squares	Means of Squares	F-value	*P*-value	Significance level
Body weight	Sex	1	4.530	4.530	67.510	9.05E-16	***
	Sire	3	2.940	0.979	14.590	3.01E-09	***
	Interaction Sex:Sire	3	0.230	0.076	1.139	3.32E-01	
	Residuals	758	50.870	0.067			
Fork length	Sex	1	0.455	0.455	79.700	2.00E-16	***
	Sire	3	0.380	0.127	22.210	8.64E-14	***
	Interaction Sex:Sire	3	0.015	0.005	0.887	4.47E-01	
	Residuals	755	4.311	0.006			
Condition factor	Sex	1	0.003	0.003	4.142	4.22E-02	*
	Sire	3	0.012	0.004	5.685	7.49E-04	***
	Interaction Sex:Sire	3	0.002	0.001	0.972	4.05E-01	
	Residuals	754	0.537	0.001			

Chapter 4

Methylation of the first intron shows genome-wide negative correlation with gene expression in tissues of different transcriptomic complexity

Dafni Anastasiadi[1], Anna Esteve-Codina[2], and Francesc Piferrer[1,3]

[1]Institute of Marine Sciences (ICM-CSIC), Passeig Marítim de la Barceloneta 37-49, 08003 Barcelona, Spain.

[2]Centre Nacional d'Anàlisi Genòmica (CNAG-CRG), Centre de Regulació Genómica, Parc Científic de Barcelona, Barcelona 08028, Spain.

Abstract

In the standard model of gene expression regulation, DNA methylation of the promoter region and also the first exon are generally associated with transcriptional silencing. Nevertheless, studies concerning specific genes in different cell types, including cancer cell lines, fetal and adult tissues, CD4+ and multiple myeloma cell lines, have shown that DNA methylation of the first intron can have both positive and negative correlation with gene expression. However, the relationship between DNA methylation of the first intron and gene expression at the genome-wide level has not been explored. Here, using tissues with different cellular heterogeneity from a fish, we show, for the first time in any organism, a clear genome-wide, quasi-linear, inverse relationship between DNA methylation of the first intron and gene expression. Remarkably, we find more tissue-specific differentially methylated regions (tDMRs) located in the first intron than in the promoter or the first exon. Part of these tDMRs show positive correlation with gene expression, indicative of tissue-specific mechanisms of gene expression regulation. Furthermore, DNA methylation in the first intron, together with the promoter and the first exon, show gene class-dependent patterns, with housekeeping genes exhibiting constantly low DNA methylation. In contrast, genes at the extremes of the expression distribution of each tissue show variation in DNA methylation dynamics at their regulatory gene features. Our

findings challenge the preeminence of the promoter region and highlight the importance of the first intron in the regulation of gene expression by DNA methylation. The link between first intron tDMRs and gene expression regulation in normal and diseased tissues should be further studied.

1. Introduction

DNA methylation is one of the main epigenetic mechanisms for the regulation of gene expression in eukaryotes (Carlberg and Molnár, 2014b; Lowdon et al., 2016). The standard model of gene expression regulation by DNA methylation in vertebrates is based on mammalian genomes, despite some recent comparative studies. Under this model, the methylation of cytosines in the promoter regions of genes represses transcription directly, by blocking the access of transcription factors (TFs), or indirectly, by recruiting other repressive proteins with methyl-binding domains (Gilbert and Epel, 2008; Moore et al., 2013). As a general rule, the CpG dinucleotides in the mammalian genome are underrepresented and methylated; this is linked to overall CpG depletion in the genome that is indicative of methylation extent in an organism (Illingworth and Bird, 2009; Straussman et al., 2009). However there are regions rich in CpGs that typically span 200-1000bp, called CpG islands (CGI), which remain unmethylated, typically overlap with gene promoters and are associated to gene transcription (Illingworth and Bird, 2009; Straussman et al., 2009).

Recent technological advances have resulted in the widespread use of next generation sequencing (NGS) methods for the identification of the DNA methylation status of individual cytosines at the genome-wide or whole-genome level (Bock et al., 2010). This methodological progress has clarified some aspects of the DNA methylome's properties and functions, while it has opened a variety of previously unthought questions, generating the so-named "top of the iceberg" picture of what was previously known (Ndlovu et al., 2011).

The genomic regions relative to CGIs are called shores (0-2 kb distance), shelves (2-4 kb) and open sea (Edgar et al., 2014),

according to the order of increasing distance. The CpG shores show more dynamic methylation than the CGIs (Irizarry et al., 2009; Park et al., 2011), while there are cases of contrasting methylation patterns between the CGI and the neighboring shore, where a positive correlation between methylation and gene expression may be present (Edgar et al., 2014). Regions of low CpG density and low methylation distal to promoters, called low methylated regions (LMRs), are also linked to high gene expression (Stadler et al., 2011). Furthermore, non-methylated islands (NMI) have been divided in classes, expanding the standard definition of regulatory CGI (Long et al., 2013). One class of NMIs is located in a large distance from the promoter, contains low CpG density and exhibits tissue-specific differences, while another class of NMIs covers the whole gene and is linked to developmental TFs (Long et al., 2013). In addition to promoters, methylation of the genomic regions that bind the TFs and interact with the promoter for the activation of the gene, that is the enhancers, often located inside gene bodies, receives increasing attention as a regulatory mechanism. There is widespread hypomethylation of enhancers during development (Bogdanović et al., 2016), while in cancer cells enhancers also exhibit dynamic changes in methylation (Blattler et al., 2014; Tomazou et al., 2015). These studies suggest that the regulation of gene expression by DNA methylation is not restricted to the unmethylated CGIs of the promoters, but also to distal or intragenic regulatory elements with different degrees of CpG density.

Moreover, a positive correlation has been demonstrated between gene body methylation and gene expression (Ball et al., 2009). DNA methylation has been suggested to play a role in alternative transcript splicing and promoter use, since abrupt transitions of methylation levels have been observed in exon-intron junctions, as well as in transcription termination sites (Laurent et al., 2010; Lister et al., 2009). In mammalian cells, a clear distinction was shown between the methylation levels of the first exon and the rest of exons (Brenet et al., 2011), indicating discrete patterns of methylation in gene features. In the same study, the gene expression levels were more inversely correlated with the methylation of the first exon than with that of the upstream promoter (Brenet et al., 2011). These data are indicative of dual roles of DNA methylation, both inhibitory and activating, depending on the genomic region, as well as of discrete regulatory roles of each gene feature.

DNA methylation has, together with other epigenetic mechanisms, a key role in cell differentiation, so the tissue-specific signatures of DNA methylation and the mode of tissue-specific gene expression control have received much attention. Only on the basis of genomic sequences, there are identified characteristics of tissue-specific (TS) and housekeeping (HK) genes (Schug et al., 2005), with TS genes being more compact, smaller in length, with less number of exons and having less conserved promoter regions (Zhu et al., 2008). Gene promoters with intermediate CpG density are linked to methylation specific to somatic cells (Ndlovu et al., 2011; Weber et al., 2007), while CpG density is more linked to expression breadth between tissues than to expression levels (Park et al., 2011). Another distinctive characteristic is the number of TF-binding sites being higher in genes with high expression breadth (Hurst et al., 2014). Between tissues, there are differentially methylated regions (DMRs) inside the gene body and in the regions surrounding it, both upstream the TSS and downstream, which are specific for each tissue (Wan et al., 2015). These tissue-specific DMRs (tDMRs) show either positive or negative correlation with gene expression, contain binding sites for different TFs and are identified in genes implicated in different processes (Wan et al., 2015). tDMRs overlap with regions of variable CpG density and their hypomethylation is suggested to be related to tissue-specific functions (Lokk et al., 2014). These studies imply that tissue-specific differential methylation is associated with the tissue-specific transcriptome and cell phenotype in mouse and humans. However, open questions remain regarding the generality of the phenomenon, as well as the regulatory roles of gene body tDMRs and TF-binding sites.

The above mentioned specific characteristics of genome-wide patterns of gene expression regulation by DNA methylation are deduced mainly from biomedical research, using mostly cell lines and human tissues or model mammals, such as mice and rats, and in some cases zebrafish. Comparative epigenomic studies have shown that epigenetic divergence follows the genetic phylogenetic patterns across species (Hernando-Herraez et al., 2015; Zhong, 2016), but have not still widely focused in the conservation of genomic features of gene expression regulation by DNA methylation. More specifically for non-model vertebrates, research has been carried out at the genome-wide or whole-genome level in sheep (Cao et al., 2015; Couldrey et al., 2014), pig (Choi et al., 2015), the great tit

(Derks et al., 2016), primates (Hernando-Herraez et al., 2013, 2015), chicken (Nätt et al., 2012), baboons (Lea et al., 2016), dogs and wolves (Janowitz Koch et al., 2016). However, in most of these studies the objective was to correlate DNA methylation patterns with a specific phenotypic trait rather than to explore and compare the patterns of gene regulation. In fish, whole methylomes of the half-smooth tongue sole (*Cynoglossus semilaevis*), one pufferfish (*Tetraodon nigrovirids*), rainbow trout (*Oncorhynchus mykiss*) and zebrafish (*Danio rerio*) are available (Baerwald et al., 2016; Jiang et al., 2013; Potok et al., 2013; Shao et al., 2014; Zemach et al., 2010). In these studies, some features of gene regulation by DNA methylation in fish are elucidated; nevertheless, a thorough comparison with the latest advancements in the mammalian literature is missing.

In this study, we used two tissues with high and low cellular heterogeneity, testis and muscle, to evaluate the association of gene features with patterns of gene expression. We explored the natural patterns of DNA methylation in genomic features, such as repetitive elements, promoters and gene bodies and also the transcriptomic properties of two tissues with different transcriptomic complexity. We, then, performed integrative analysis of DNA methylation and gene expression data in order to investigate their relationships with important gene features, such as the first exon/intron, and also to elucidate tissue-specific patterns.

2. Materials and Methods

2.1 Animals and rearing conditions

Wild European sea bass (*Dicentrarchus labrax*) adults with body weight of 1000 ± 109.5 g (mean ± SEM), standard length = 39.3 ± 1.4 cm and gonadosomatic index = 0.076 ± 0.009, the latter calculated as in Navarro-Martín et al. (2009), were captured by speargun at the Montgrí, Medes Islands and Baix Ter Natural Reserve (NE Spain) during the non-reproductive season (June 2013). Tissues were dissected immediately upon capture and stored in RNAlater® (ThermoFisher Scientific).

2.2. DNA isolation

Genomic DNA was extracted from 3 muscle and 3 testis fragments by phenol/chloroform/isoamyl alcohol (PCI). In brief, testis and muscle samples were dried out from RNAlater® and immersed into digestion buffer (100 mM NaCl, 10 mM Tris-HCl, 25 mM EDTA pH 8, 0.5% SDS), proteins were digested by 1 µg of proteinase K (Sigma-Aldrich) and RNA by 0.5 µg of ribonuclease A (PureLink RNase A; Life Technologies). DNA was precipitated by 95% ethanol, eluted in Milli-Q® water (Merck, Millipore) and cleaned-up with 2x AMPure XP beads (Beckman Coulter) to ensure purity. DNA was quantified three times by independent means, being by ND-spectrophotometer (NanoDrop Technologies) or Qubit™ fluorometric quantitation (ThermoFisher Scientific), each time followed by dilutions with nuclease-free water in order to normalize DNA quantities across samples.

2.3. RRBS libraries preparation

RRBS libraries were prepared as in Klughammer et al. (2015). One hundred nanograms of genomic DNA were digested by 20 units of *Msp*I (NEB) overnight at 37ºC. Five units of Klenow Fragment (3'→5'exo-; NEB) and dNTP mix (final concentration: 300 µM dATP, 30 µM dCTP and 30 µM dGTP) were added to the reaction. End fill-in was performed for 20 min at 30ºC, A-tailing for 20 min at 37ºC and inactivation of the enzyme for 20 min at 75ºC. Ligation of Illumina TruSeq Adapters v2 was performed by Quick Ligase (NEB) for 20 min at 25ºC, followed by heat inactivation of the enzyme for 10 min at 65ºC. Libraries were size-selected by 0.75x 1:5 diluted AMPure XP beads, quantified by qPCR, pooled based on qPCR values and cleaned-up with 2.5x 1:5 diluted AMPure XP beads. Samples were subjected to bisulfite conversion using the EZ DNA Methylation-Direct kit (Zymo Research) with 0.9x CT Conversion Reagent, 20 cycles of 95ºC for 1 min and 60ºC for 10 min and desulphonation time extended to 30 min. Libraries were enriched by the PfuTurbo Cx HotStart Polymerase (Agilent Technologies) with the following cycling parameters: 95ºC for 2 min, followed by the optimal number of cycles of 95ºC for 30 s, 65ºC for 30 s and 72ºC for 45 s, and a final step at 72ºC for 7 min.

The optimal number of cycles for the enrichment PCR was calculated based on qPCR values. A final cleanup step was performed by 1x AMPure XP beads. The quantity of the libraries was measured by Qubit High Sensitivity assays (ThermoFisher Scientific) and the quality was evaluated by Experion DNA 1k assays (BioRad). Sequencing of RRBS libraries was performed on an Illumina HiSeq 2000 platform in 50 bp single-end mode.

2.4. RNA isolation

Total mRNA was isolated from testis and muscle of 5 fish, including the same 3 fish used to prepare the RRBS libraries from a fragment contiguous to the one used for DNA extraction. Tissues were removed from RNAlater®, immersed into TRIzol® Reagent (ThermoFisher Scientific) and homogenized by the Polytron PT 1200 CL (Kinematica AG). RNA extraction was performed according to the manufacturer's instructions. RNA was quantified by the Qubit® RNA BR Assay Kit (ThermoFisher Scientific) and RNA quality was evaluated by the Agilent RNA 6000 Nano Kit (Agilent). Samples with RNA Integrity Number (RIN)>8 were used for library construction.

2.5. RNA-seq

The libraries were prepared using the mRNA-Seq sample preparation kit (Illumina Inc., Cat. # RS-122-2001x2) according to the manufacturer's protocol. Briefly, 0.5 μg of total RNA were used for poly-A based mRNA enrichment selection using oligo-dT magnetic beads followed by fragmentation by divalent cations at elevated temperature resulting into fragments of 80-250 nt, with the major peak at 130 nt. First strand cDNA synthesis by random hexamers and reverse transcriptase was followed by the second strand cDNA synthesis. Double stranded cDNA was end-repaired, 3´adenylated and the 3´-"T" nucleotide at the Illumina adapter was used for the indexed adapters ligation. The ligation product was amplified using 15 PCR cycles. Each library was sequenced using the TruSeq SBS Kit v3-HS, in 76 bp paired-end mode on an Illumina HiSeq2000 instrument following the manufacturer's

protocol. Images from the instrument were processed using the manufacturer's software to generate FASTQ sequence files.

2.6. DNA methylation analysis

RRBS reads were quality trimmed by Trim Galore!(Babraham Bioinformatics) based on a Phred quality threshold of 20 and minimum length after trimming 16 bp and adaptors sequences were removed from the reads. Trimmed reads were aligned to the reference genome of sea bass (Kuhl et al., 2011b; Tine et al., 2014) using BSMAP (Xi and Li, 2009) and a minimum coverage of 5 reads. Methylation calling was performed by Bis-SNP (Liu et al., 2012). All subsequent bioinformatics analysis were performed using R and Rstudio (R Core Team, 2015; RStudio Team, 2015), unless stated otherwise, and Bioconductor packages (Gentleman et al., 2004). The package *methylKit* (Akalin et al., 2012) was used for DNA methylation analysis. Called bases with less than 10 reads or more than the 99.9^{th} percentile of coverage distribution were filtered out. Coverage values were normalized as in default and bases were united in order to retain the ones that were covered in all samples. The sea bass genome is one of the best *in silico* annotated fish genomes (Tine et al., 2014). Repetitive elements were detected based on the RepeatMasker track of the sea bass genome. Annotations of gene features were based on the COMBINED ANNOTATION track of the seabass genome. Promoters were defined as 1000 bp upstream the *in silico* annotated Transcription Start Sites (TSSes) from the COMBINED ANNOTATION track. Differentially methylated cytosines (DMC) were defined as CpGs with more than 15% methylation differences and q-value < 0.01 after applying logistic regression using the SLIM method for p-value adjustment. Differentially methylated regions (DMR) were identified using the weighted optimization algorithm for empirically based DMRs using the package *edmr* (Li et al., 2013) with default parameters, except for DMC differences cutoffs which were set to 15% and DMR differences cutoffs set to 10%. A BSgenome package (Pagès, 2016) was created for use when required using the full genome and masks from the UCSC server (dicLab v1.0c, Jul. 2012). Genomic overlaps of features were identified using the GenomicRanges package (Lawrence et al., 2013).

2.7. RNA-seq analysis

RNA-seq reads were aligned with the GEMtools RNAseq pipeline v1.7 (http://gemtools.github.io/), which is based on the GEM mapper (Marco-Sola et al., 2012). The pipeline aligns the reads in a sample in three phases, mapping against the reference genome (dicLab v1.0c, Jul. 2012), against a reference transcriptome (COMBINED ANNOTATION track) and against a *de novo* transcriptome, generated from the input data to detect new junction sites. After mapping, all alignments were filtered to increase the number of uniquely mapped reads. The filtering criteria included a minimum intron length of 20 bp, a maximum exon overlap of 5 bp and a filter step against a reference annotation checking for consistent pairs and junctions where both sites align to the same annotated gene. The same pipeline was used to quantify gene expression. Subsequently, the TMM method (Robinson and Oshlack, 2010) was used for gene expression normalization and the EdgeR robust method (Zhou et al., 2014) was used for differential expression analysis. Genes with p-adjusted < 0.05 were considered significant.

2.8. Combined analysis of DNA methylation and gene expression

The vioplot package was used for visualizing methylation data by expression decile (Adler, 2005). For positive and negative correlations of methylation differences and gene expression, first we identified the DMRs located in genomic regions encompassing the whole gene bodies and 4 kb both upstream and downstream. Then, DMRs overlapping with promoters, first exons or first introns were identified. Only genes with $\log_2$ FC>|1.5| and FDR<0.05 were considered. Transcription-factor binding sites were identified using the fimo tool (Grant et al., 2011) of the MEME suite (Bailey et al., 2009) against the JASPAR CORE 2016 database (Mathelier et al., 2013). Individual genes were visualized using the Integrated Genome Browser (Nicol et al., 2009). GO-terms enrichment was performed using the GO.db (Carlson) and topGO (Alexa and Rahnenfuhrer, 2016) R packages and significantly enriched GO-terms were summarized by REVIGO (Supek et al., 2011).

Housekeeping genes were defined based on common tissue housekeeping genes of human in the PaGenBase (Pan et al., 2013). Tissue-specific genes were considered as the genes that were expressed in only one of the two tissues, regardless of the actual expression level.

2.9. Statistical analysis of the data

Statistical analyses of the data were performed by R and Rstudio (R Core Team, 2015; RStudio Team, 2015). Correlations between DNA methylation data were measured using Pearson's product-moment correlation coefficient. In all other cases Spearman's rank correlation coefficient was used because the relationship of DNA methylation with gene expression data is not necessarily expected to be linear. Homogeneity of variances was checked by Levene's test. In case of homogeneous variances, the Student's *t*-test was applied. Otherwise, the Kruskal-Wallis rank sum test was used.

3. Results

3.1. DNA methylation

Pairwise comparisons of DNA methylation values for testis and muscle showed good correlation between biological replicates within each tissue although higher for testis than for muscle (Pearson's correlation scores: testis $\geq$ 0.97; muscle $\geq$ 0.78). Clustering based on Pearson's correlation distances clearly separated sample based on tissue of origin (Fig. S1).

Overall, DNA methylation levels were similar between the two tissues, being either fully methylated or unmethylated and showing a strong positive correlation (ρ=0.94***; white to blue scale in Fig. 1A). However, among the 500 top-differentially methylated CpG (DMC) sites, more were fully methylated in testis and unmethylated

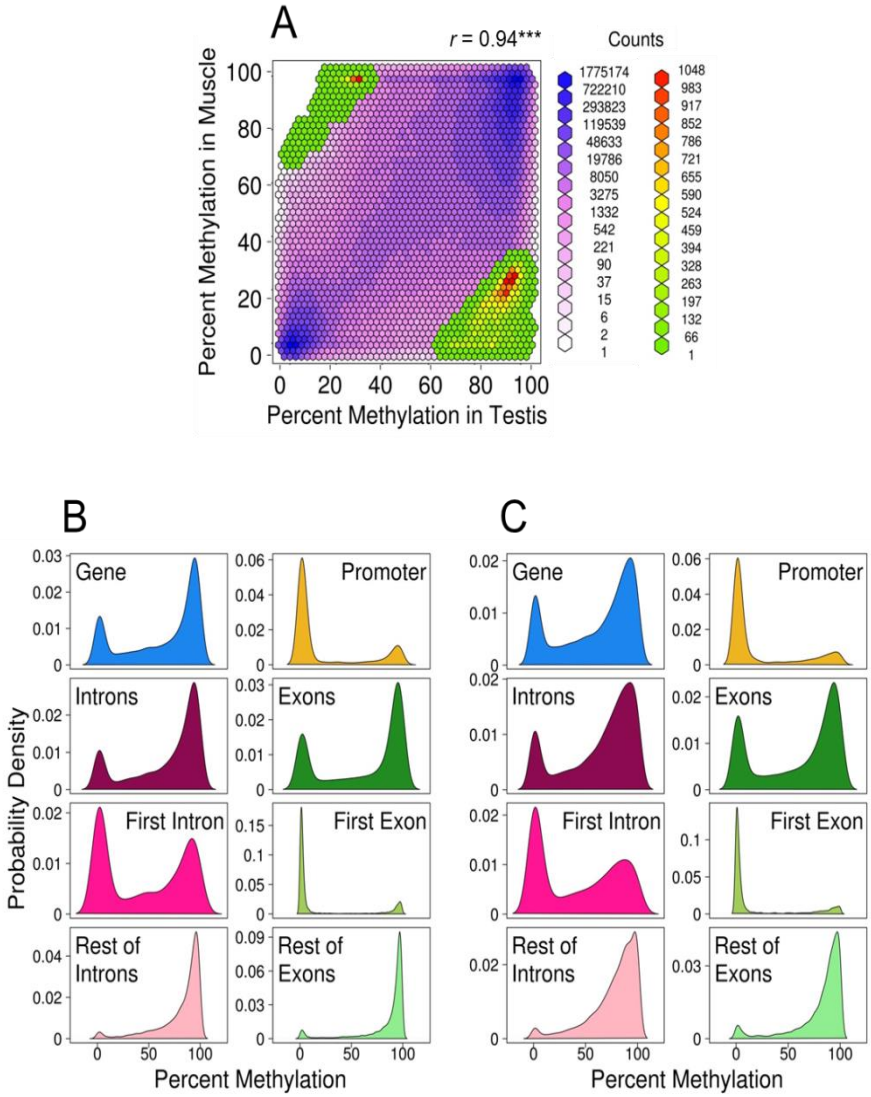


Figure 1. Overview of genome-wide DNA methylation in the European sea bass. A) . Scatterplots of DNA methylation levels in testis and in muscle divided in 40 bins containing from 1 (white) to 1775174 (blue) data points (Pearson's correlation r = 0.80; n=278368) and overlayed scatterplots of top 500 differentially CpGs divided in 40 bins containing from 1 (green) to 1048 (red) data points. B-C)). Kernel density plots for DNA methylation distribution in gene features in testis (B) and in muscle (C in genes, promoters (-1000 bp), all introns and all exons. Separation of exons in first exon and rest of exons and of introns in first intron and rest of introns.

in muscle, while there were no CpG sites with >90% methylation in muscle and <10% methylation in the testis (green to red scale in Fig. 1A). There were also 33 genes that contained one or more of the top 500 differentially methylated cytosines (DMCs) in the promoter, first exon or first intron, most of which were related to regulation of macrophage activation or zygotic specification of dorsal/ventral axis.

Being aware of the global genome-wide picture of DNA methylation, we then focused on specific genomic features. In whole genes, DNA methylation patterns followed a binomial distribution, with high (>80%) or low (~10%) levels of DNA methylation in the majority of CpG sites (Fig. 1B-C, see next page). Separating the whole gene in specific gene features exposed distinct patterns. A similar binomial DNA methylation pattern was observed in introns and exons. However, in promoters, most of CpG sites were unmethylated. In addition, by partitioning data from exons into first exon and the rest of exons, a contrasting pattern was revealed, with the majority of unmethylated cytosines restricted to the first exon and methylated cytosines almost exclusively localized in the rest of exons. Likewise, partitioning the introns showed a vast majority of highly methylated cytosines in all except the first intron. In the first intron the distribution was still binomial but skewed towards the unmethylated sites and smoother than in the first exon (Fig. 1B-C, see next page). Thus, regardless of tissue and cellular heterogeneity, the majority of CpG sites are unmethylated in the promoters and first exon and, to a lesser degree, also in the first intron.

DNA methylation levels in repetitive elements, independently of the class, were above 76% for three quarters of the CpGs and above 89% in half of the CpG sites in both tissues. In addition, they showed strong correlation (Pearson's correlation, ρ=0.80) between the two tissues (Fig. 2A, see next page). In specific repetitive elements classes, DNA methylation was usually higher in testis than in muscle and above 81% in most classes, except in low complexity repeats, simple repeats, snRNA and tRNA, where methylation levels were below 64%. Thus, essentially there were no repetitive elements with DNA methylation between 60 and 80% or lower than 50% (Fig. 2B, see next page).

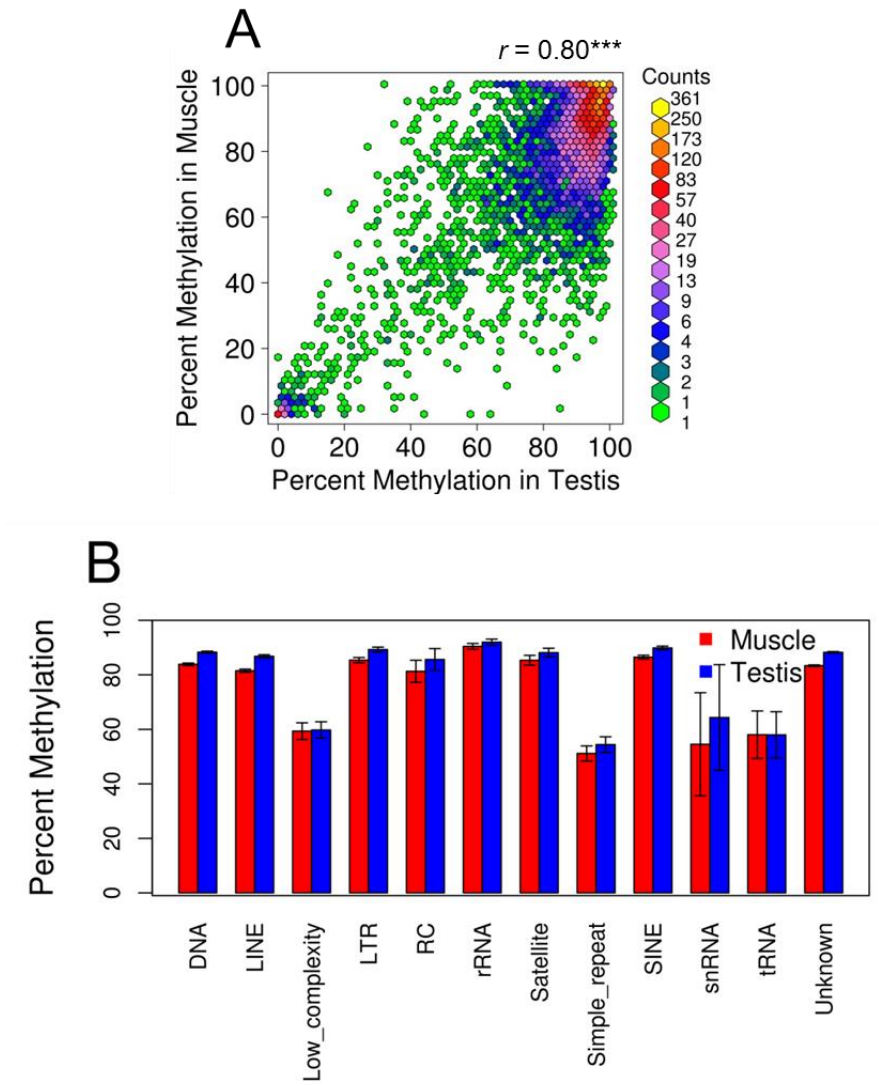


Figure 2. DNA methylation levels in genomic repetitive elements in the European sea bass. A) Scatterplots of DNA methylation in muscle vs. testis divided in 40 bins containing from 1 (green) to 361 (yellow) data points (Pearson's correlation r = 0.80; n=8156). B) Barplots of mean DNA methylation levels in each type of repetitive elements class for muscle (red) and testis (blue). Standard error of the mean is shown for each bar.

The comparison between the two tissues revealed 20298 (7.29% of total CpGs with methylation reads) hypermethylated sites in the testis and 7762 (2.79% of total CpGs with methylation reads) in the muscle. Thus, there were ~2.6 times more hypermethylated CpGs in the testis than in the muscle (Fig. S2).

3.2. Gene expression

Regarding RNA-seq data, Principal Component Analysis showed that the two first components explained > 97% of the variance, with good separation based on tissue, although muscle samples had more dispersion that testis samples (Fig. S3). Analysis of gene expression data revealed that gene expression levels were more dispersed in muscle (Levene's test; $p < 0.001$) but had higher expression mean in testis (Kruskal-Wallis test by ranks; $p < 0.001$; Fig. 3A, see next page). There were 9449 genes with higher expression in the testis and 6220 with higher expression in the muscle (FDR<0.05). Furthermore, most of the genes commonly expressed in both tissues had higher expression levels in the testis than in the muscle (Fig. 3B, see next page). In muscle, among the total number of transcripts detected a small number of genes accounted for 50% of the transcripts, i.e., four transcripts of myosin heavy chain (*myhm86-1*), muscle-type creatine kinase (*ckma*), myosin light chain 2 (*mylz2*), fast white muscle troponin t embryonic isoform, myosin-binding protein fast-type-like (*mybpc2*) and tropomyosin α-4 chain isoform 1 (*tpm1*). On the contrary, in testis among the total number of transcripts detected there was a higher number of genes with lower expression levels. Thus, ~20 and ~2000 genes accounted for half of the number of transcripts detected in the muscle and testis, respectively (Fig. 3C, see next page). However, when considering only tissue-specific genes, we found that the testis-specific genes had lower mean expression (Kruskal-Wallis test by ranks; $p < 0.001$) than the muscle-specific genes, in contrast to the relationships of all transcripts detected (Fig. 3D, see next page).

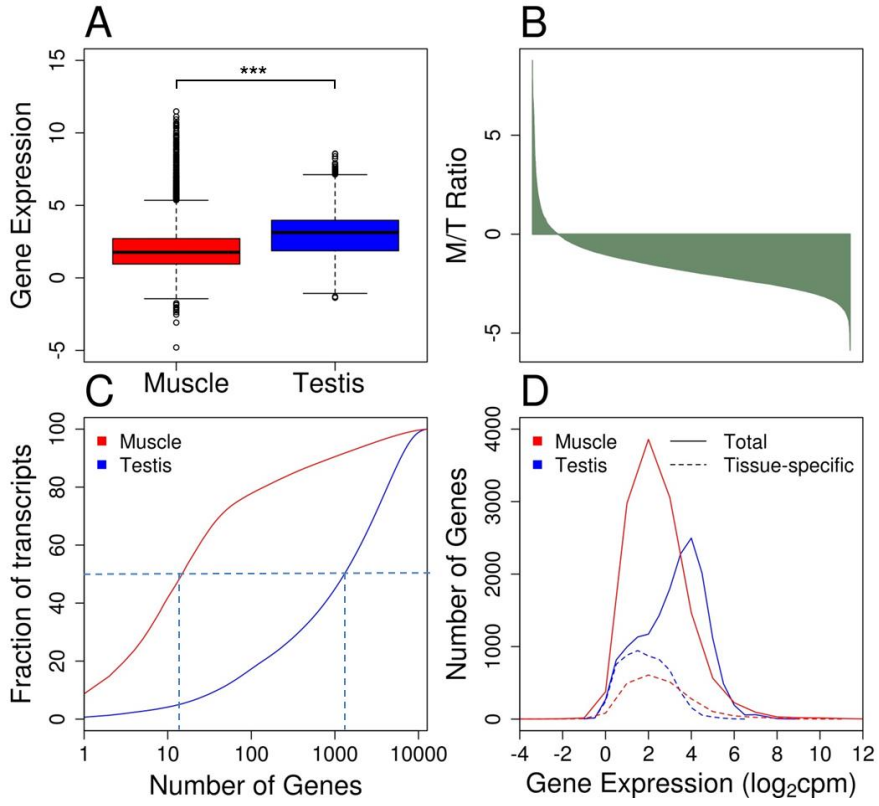


Figure 3. Distribution of gene expression. A) Boxplots of gene expression measured as log_2-transformed copy million number (cpm) values in muscle (red) and in testis (blue). Kruskal-Wallis rank sum test indicates significant differences between the two tissues (Kruskal-Wallis χ^2=3491, ***=p<0.001). The boxes show the interquartile distribution (IQR) of values, the upper whisker equals to min(max(x), Q3 + 1.5 * IQR) and the lower whisker equals to max(min(x), Q1 – 1.5 * IQR), the black lines indicate the median and the black dots expanding outside the boxes are considered outliers. B) Histogram of muscle (M)-to-testis (T) ratio of log_2-transformed copy million number (cpm) values in decreasing order. C) Cumulative distribution of transcripts contribution to total transcript number in muscle (red) and testis (blue). The dashed line indicates 50% contribution to transcription and the associated number of genes. D) Distribution of total transcripts detected (continuous lines) and tissue-specific genes (dotted lines) in muscle (red) and in testis (blue).

3.3. Combined DNA methylation and gene expression analysis

In order to relate the gene expression levels with the DNA methylation levels of specific gene features, we divided the gene expression levels in deciles based on the increasing distribution of log_2-transformed cpm values. In muscle, DNA methylation was low regardless of gene expression in promoter and first exon (Fig. 4, see next page). By contrast, in the first intron DNA methylation levels decreased with increasing expression levels (ρ=-0.144; p-value < 0.001). In the rest of exons and introns, DNA methylation levels were high independently of gene expression.

In testis, in the promoter and first exon, median DNA methylation levels were low in all expression deciles (Fig. 4). However, in the genes belonging to the first and second expression deciles, there were variable levels of DNA methylation, in contrast to the rest of expression deciles and to the situation in muscle. DNA methylation levels of the first intron were clearly inversely correlated (ρ=-0.24; p-value < 0.001) to gene expression, similarly to what was observed in muscle. In the rest of exons and in the rest of introns, median DNA methylation levels were high regardless of gene expression levels. Thus, expression levels clearly inversely correlated with DNA methylation levels across the two tissues only in the first intron.

Next, we searched for functional relationships of the genes that contained DNA methylation inside their first intron. For this, we performed GO-term enrichment analysis for the genes that belonged to the first expression decile and their DNA methylation levels were either below the first quartile or above the median of the total distribution of the decile, as well as for the genes that belonged to the tenth expression decile and their DNA methylation levels were either above the third quartile or below the median of the total distribution of the decile. There were common GO-terms enriched between the two tissues for genes of the first expression decile that had low methylation and also for genes of the tenth expression decile that had low methylation (Table 1).

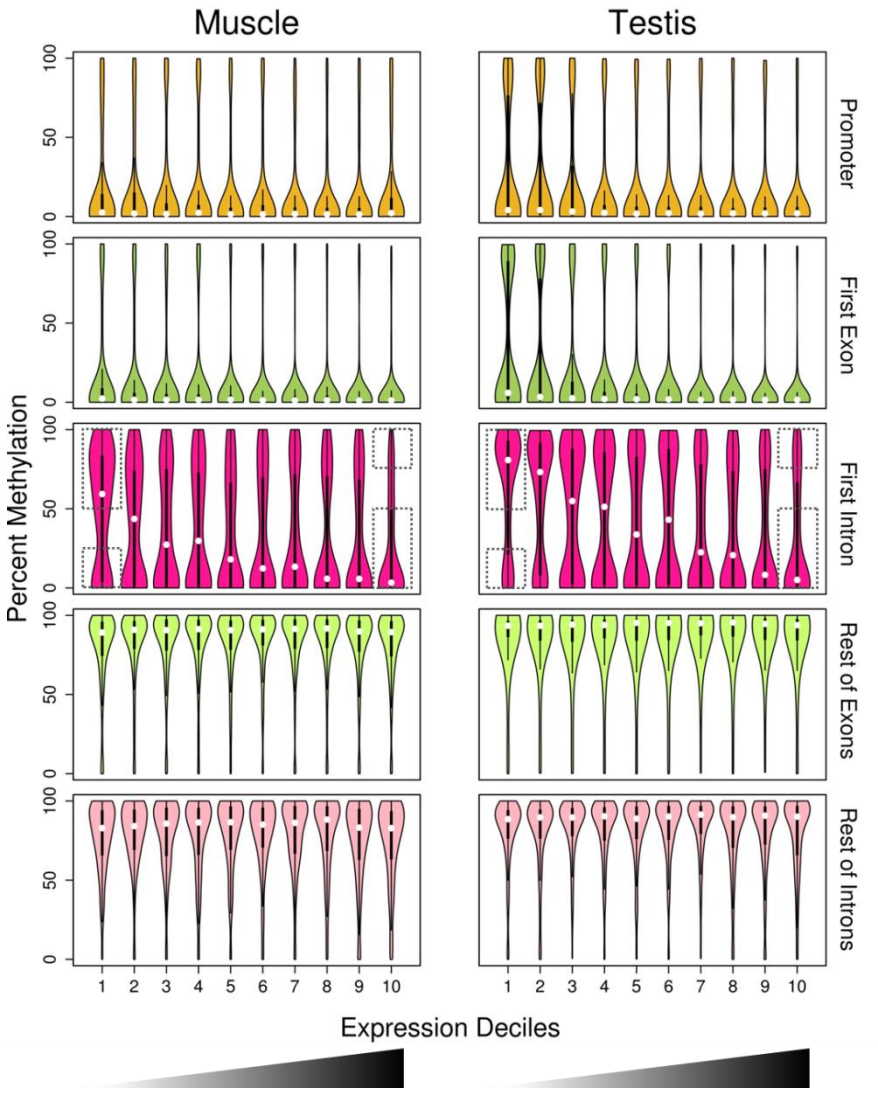


Figure 4. Distribution of DNA methylation in gene features by expression deciles in muscle and in testis. Violin plots of DNA methylation in promoter, first exon, first intron, rest of exons and rest of introns divided in deciles based on increasing ranking of gene expression measured as $\log_2$-transformed copy million number (cpm) values. Box plots with rotated kernel density plots at both sides indicate the interquartile range and white central dots the median of the distribution. Dotted boxes surrounding the first and tenth expression decile in the first intron indicate the groups of genes used for GO-term enrichment.

Table 1. Common enriched GO-terms between muscle and testis. GO-terms are associated with genes that belong to the first or the tenth expression decile and contain methylation in their first intron

Decile	GO-term
First	positive regulation of receptor biosynthesis
	receptor metabolism
	medium-chain fatty acid transport
	L-serine biosynthesis
	ectodermal placode development
Tenth	mRNA metabolism
	RNA localization
	cellular metabolism
	macromolecular complex subunit organization
	organic substance metabolism
	primary metabolism
	metabolism
	biosynthesis
	macromolecule metabolism

There are distinct patterns in general between the tissue-specific genes and the non-tissue-specific genes that were obvious from the gene expression data. Therefore, we targeted tissue-specific genes within each expression decile in order to explore their behavior under this context. For this, we filtered the genes of the expression deciles with DNA methylation in the promoter, first exon or first intron and we observed that the majority (>80%) of testis-specific genes were lowly expressed, i.e. belonged to the first four expression deciles (Fig. 5, see next page). In contrast, the muscle-specific genes showed a linear relationship, i.e., a homogeneous distribution of gene numbers across expression deciles.

Since we observed these distinct patterns in tissue-specific genes, we extended the analysis by comparing the profiles of both gene expression and DNA methylation in promoter, first exon and first intron of (1) housekeeping genes and (2) genes that were common in both tissues but at the limits of gene expression values. Housekeeping genes had lower methylation in the promoter, first exon and first intron in testis than in muscle, while the expression

tendencies of the two tissues resembled the profiles of overall gene expression (Fig. 6A, see next page). The top 10% expressed genes in muscle had low DNA methylation and similar expression levels in both tissues, although with a higher dispersion in testis (Fig. 6B). On the contrary, the top 10% expressed genes in testis had low DNA methylation in both tissues, but their expression level was lower in muscle than in testis (Fig. 6C). The least 10% genes expressed in muscle, showed more variation in their DNA methylation levels with the median DNA methylation remaining low in both tissues, while their expression in testis was higher than in muscle (Fig. 6D). However, the least 10% expressed genes in testis showed even higher variation, not only in testis but also in muscle, while the median DNA methylation levels were lower in muscle than in testis and the expression values more dispersed (Fig. 6E). Together, these data show clear differences between high vs. low expressed genes within and between tissues and DNA methylation ranges depending on the gene classes considered.

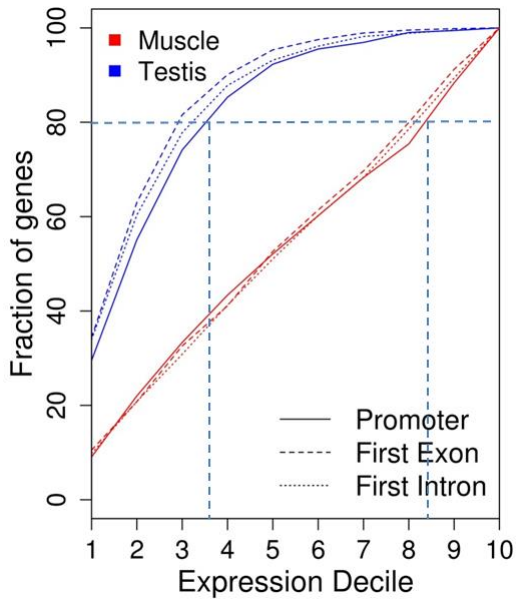
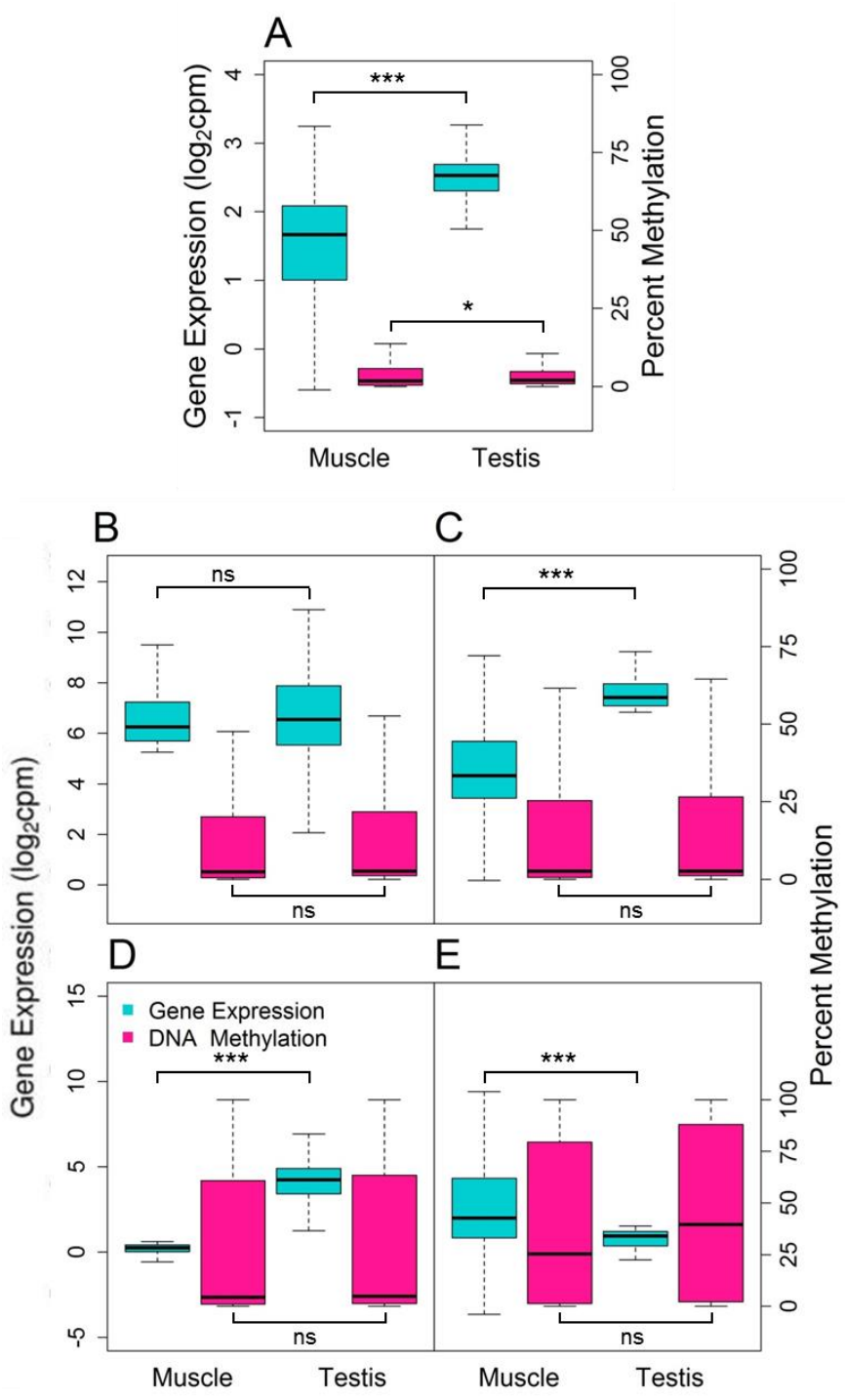


Figure 5. Cumulative distribution of tissue-specific genes in muscle (red) and in testis (blue) with DNA methylation in the promoter, first exon and first intron based on expression decile.


A
Gene Expression (log2cpm)
Percent Methylation
***
*
Muscle
Testis
B
C
D
E
ns
***
Gene Expression
DNA Methylation
Muscle
Testis

3.4. Tissue-specific differentially methylated regions (tDMRs)

Next, we focused on tDMRs between testis and muscle and explored the relationships between differential DNA methylation and differences in gene expression. Both directions of correlation were evident between DNA methylation and gene expression. There was strong negative correlation (ρ=-0.723) for genes that contained hyper-methylated, in either testis or muscle, tDMRs inside their gene body or 4 kb upstream of the TSS or downstream of the 3' UTR and were less expressed in both tissues (squares in Fig. 7). Nonetheless, there was also weaker positive correlation (ρ=0.235) for genes that were hyper-methylated and more expressed in both tissues (circles in Fig. 7). The same two categories of genes showing either negative or positive correlation between DNA methylation and gene expression were obvious in genes that contained tDMRs only in the promoter, first exon or first intron (Fig. 7).

Figure 6 (previous page). DNA methylation and gene expression per tissue for gene categorical fractions. A) Housekeeping genes in muscle and in testis. B-C) Common genes in both tissues that represent the top 10% expressed genes in muscle (B) and the top 10% expressed in testis (C) ranked in increasing order based on $\log_2$-transformed copy million number (cpm) values. D-E). Common genes in both tissues that represent the lowest 10% expressed genes in muscle (D) and the lowest 10% expressed in testis (E) ranked in increasing order based on $\log_2$-transformed copy million number (cpm) values. Gene expression values are shown in blue (left y-axis) and DNA methylation values referring only to promoter, first exon and first intron are shown as percentage in pink (right y-axis). Boxes indicate the interquartile distribution (IQR) of values, the upper whisker is the min(max(x), Q3 + 1.5 * IQR) and the lower whisker is the max(min(x), Q1 – 1.5 * IQR), the median is showed by black central lines. Asterisks indicate significance level of Kruskal-Walls rank sum tests with the following equivalence: *=p<0.05 and ***=p<0.001.

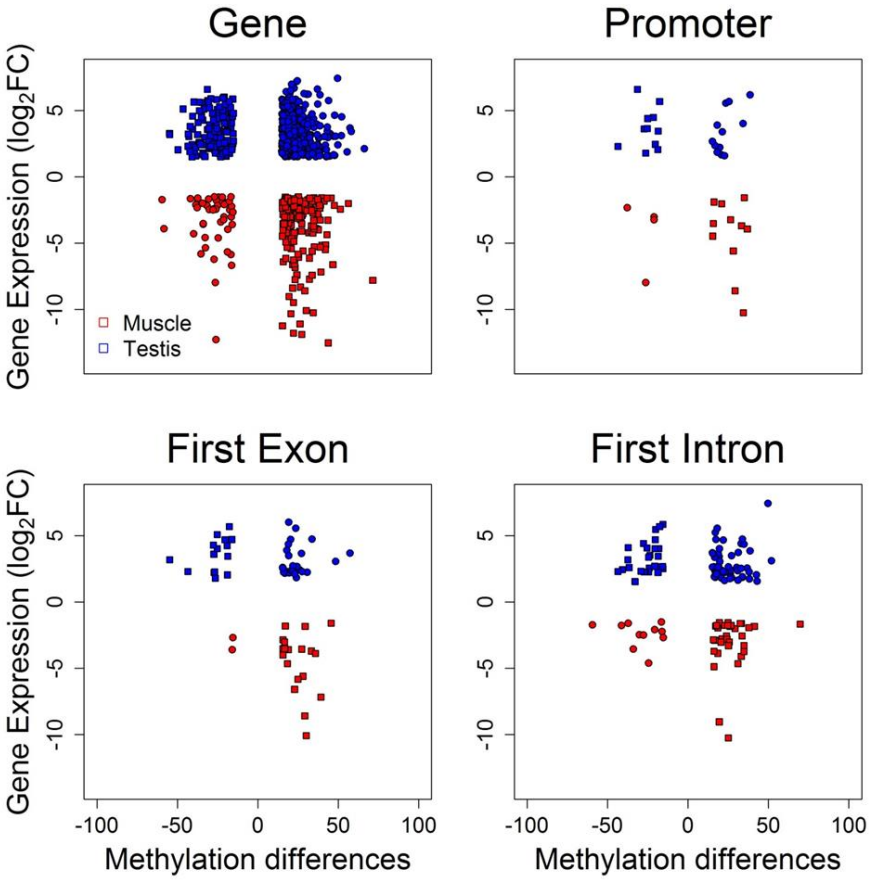


Figure 7. Differentially expressed genes between tissues with differentially methylated regions (tDMR). tDMRs overlap with the gene body and/or ±4 kb (Gene), the promoter, the first exon or the first intron. Positive (circles) and negative (boxes) correlation is shown for muscle (red) and testis (blue). Hyper-methylated tDMRs and up-regulated DEGs in testis (blue squares), hyper-methylated tDMRs and down--regulated DEGs in testis (blue circles), hyper-methylated tDMRs and up-regulated DEGs in muscle (red circles) and hyper-methylated tDMRs and down-regulated DEGs in muscle (red squares).

Next, we scanned the tDMRs present inside genes or 4 kb upstream or downstream for TF binding sites in an attempt to identify features that characterize the type of correlation. There were three motifs of TF-binding that were common between tDMRs associated with genes showing either positive or negative correlation. However, there were also two correlation-specific motifs in each case, with

the negative correlation-specific motifs encompassing a central CpG, absent from the positive-correlation specific ones (Fig. 8). The negative-correlation specific motifs were targets for SP4 and KLF14, both members of the three-zinc finger Krüppel-related factors family, while the positive-correlation specific motifs were targets for RREB1 (factors with multiple dispersed zinc fingers) and ESR1 (steroid hormone receptor –NR3– family).

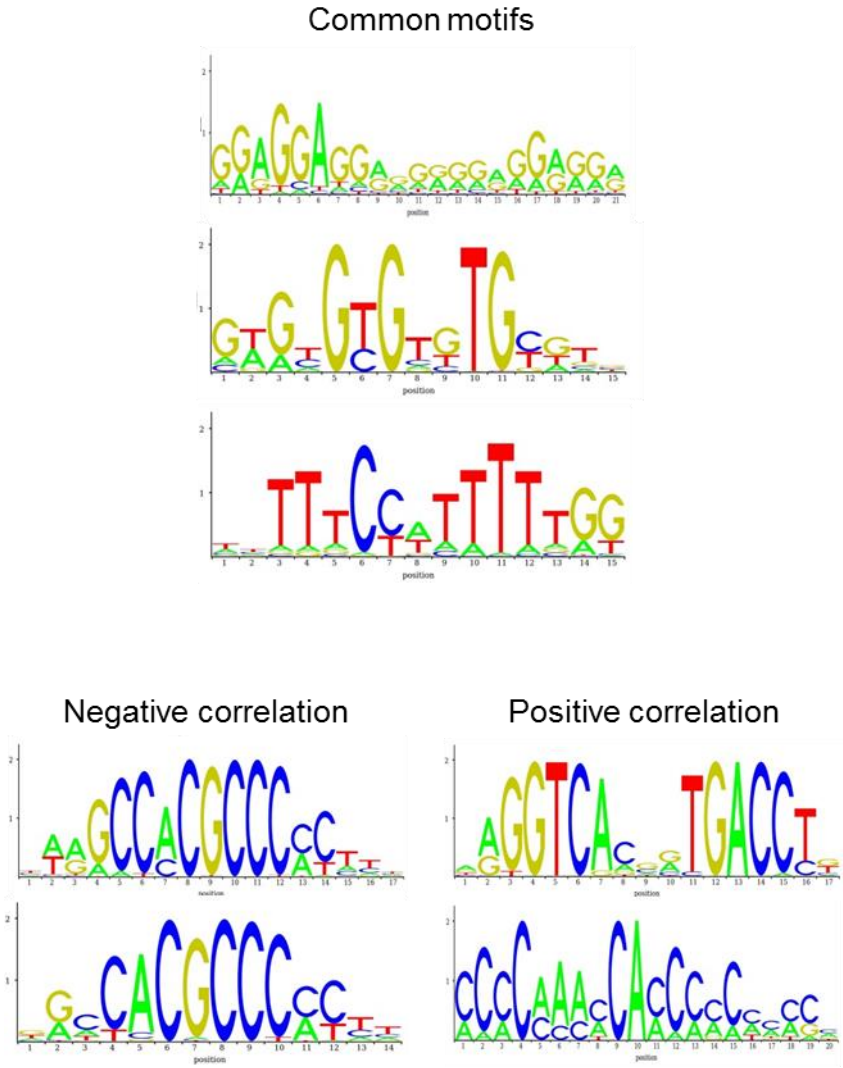


Figure 8. Transcription-factor binding motifs present in the tDMRs that are common between positive and negative correlation of DNA methylation with gene expression or that are correlation-specific (negative or positive).

Examples of CGIs, DMCs and DMRs distributed in genes known to be important for the function of each tissue were found in our dataset. Estrogen receptor β2, essential receptor for several reproductive functions and present in European sea bass testis during early stages of spermatogenesis (Viñas and Piferrer, 2008), was significantly more expressed in testis than in muscle and showed lower DNA methylation in the first intron of the testis, although no DMCs or DMRs were detected. In contrast, myogenic factor 5 (*myf5*), a key gene for muscle differentiation and myogenesis, was significantly more expressed in muscle and contained DMCs and DMRs in the first and last exon (Fig. 9).

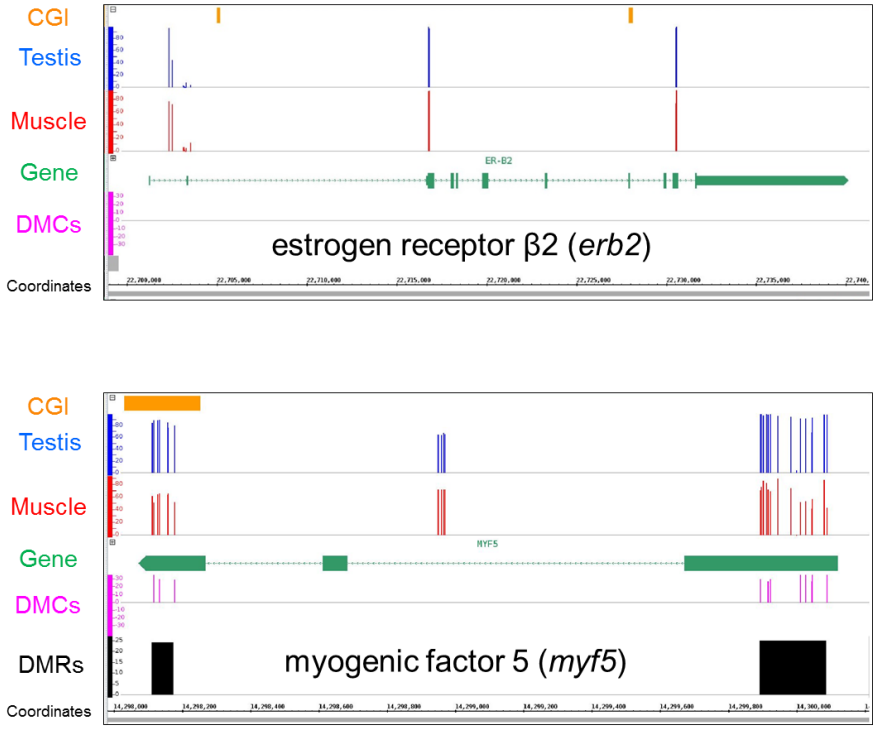


Figure 9. Schematic representation of estrogen receptor B2 and myogenic factor 5 with known important function in testis and muscle, respectively. Genomic coordinates (grey), exons (thick green boxes), introns (green lines with arrows), CpG islands (orange), DNA methylation in testis (blue) and muscle (red), differentially methylated cytosines (DMCs; pink) and differentially methylated regions (DMRs; black) are shown.

4. Discussion

This is the first study on patterns of gene expression regulation by DNA methylation in tissues with different levels of cellular heterogeneity carried out in a non-mammalian vertebrate, a modern teleost fish. We illustrate the general distribution patterns of DNA methylation in genomic features and the transcriptomic complexity of testis and muscle. We also provide evidence of the importance of the first exon in complex tissues while indicating, for the first time in any organism, a key role for the first intron in the standard inverse relationship of DNA methylation and gene expression. We investigate the properties of housekeeping and tissue-specific genes, while we identify tDMRs in different genomic features.

4.1. Overall DNA methylation patterns in specific genomic features

In vertebrate genomes, the majority of cytosines in the CpG context are methylated except in the CGIs. Likewise, in the sea bass genome, the same pattern was found for both target tissues with highly similar profiles between them. Focusing only on genes, as defined by the sum of promoters ~1000 bp upstream the TSS and gene bodies, encompassing 5'-UTRs, exons, introns and 3'-UTRs, the global pattern is similar. However, when we partition the whole gene region per genomic feature unique patterns are revealed: the majority of unmethylated cytosines are concentrated only in promoters, first exons and first introns, while the rest of exons and introns are methylated. This partition suggests a special role for the first intron overlooked before. Still, there are more methylated cytosines in the first intron in comparison with the promoter and first exon, suggestive of a less clear role of this methylation.

The repetitive genomic elements form part of the constitutive heterochromatin and are heavily methylated in vertebrate genomes, possibly to ensure chromosomal stability and genome integrity (Carlberg and Molnár, 2014b; Donnelly et al., 1999). As expected, in both tissues the repetitive elements were methylated for the majority and independently of the class at levels similar to the human (Su et al., 2012) and pufferfish genomes (Zemach et al.,

2010). However, the constraints of RRBS enriching only for a representative part of the genome, heterochromatin generally escaping sequencing and a lack of deep studies of sea bass repetitive elements, permits us only to perceive these results as a glimpse of the situation.

4.2. Association of intragenic methylation with gene expression levels

The recent explorations of the emerging complex relationships of DNA methylation and gene expression have revealed a key role for the first exon in transcriptional silencing (Brenet et al., 2011). Brenet et al. (2011) showed in mammalian cell lines that the methylation of the region surrounding the TSS, encompassing the promoter (~1000 bp upstream) and the first exon, is linked to gene expression, while in the rest of exonic and intronic regions the methylation was dissociated from the gene expression levels. In addition, the first exon showed even more pronounced negative correlation between DNA methylation and gene expression than the promoter region itself.

In two complex tissues, consisting of different proportions of cell types, we present the same pattern previously observed in vertebrate cell lines. In the sea bass genome, the CpGs remain unmethylated for their majority in the promoter regulatory regions, but also importantly in the first exon. There was a transitory region of DNA methylation between the first exon and the rest of gene body which is the first intron. The median methylation of the first intron showed the most clear inverse relationship with gene expression among all gene features in both tissues, although methylation depending on the gene could vary from 0% to 100% independently of the expression level.

The methylation of the first intron has been shown by functional studies to have both positive and negative correlation with gene expression in specific genes in cancer cell lines, fetal and adult tissues (Unoki and Nakamura, 2003), CD4+ cell lines isolated from mice (Hashimoto et al., 2013), in multiple myeloma cell lines (Hayami et al., 2003) and blood sample isolated from children (Kim

et al., 2012). Furthermore, the first introns have been suggested to contain distinct properties from the rest of introns and to be linked to transcriptional regulation (Li et al., 2012; Majewski and Ott, 2002). However, to the best of our knowledge, an association at the genome-wide scale with methylation data has not been demonstrated to date. The function of this association could be partially explained by the presence of intronic enhancers interacting with the promoters of their corresponding genes. In fact, silencing of intragenic enhancers is considered to play a role even more significant than promoter methylation in the silencing of their target genes (Blattler et al., 2014).

Indeed, the classic promoter methylation-gene expression model seems to hold only in extreme cases and specific genes, while a "triple-inverse" model was suggested, where the methylation of the promoter and the gene body exerts separate influences on gene expression; in other words, low expression may be indicated either by high promoter methylation or by high gene body methylation (Lou et al., 2014). In our case, a general pattern would be a negative association of the methylation of the first intron with gene expression in the majority of genes but with subclasses of genes escaping this pattern and exhibiting positive patterns of correlations. DNA methylation in positive correlation with gene expression has been suggested to appear either as cause or consequence of transcription (Rountree and Selker, 1997).

4.3. Tissue-specific characteristics of DNA methylation and gene expression

Muscle and testis have obvious distinct functional roles, since muscle is a specialized tissue in contraction, while in testis a variety of functions take place, mainly steroidogenesis and spermatogenesis. In addition, the cell type composition of each one is different, with few cell types performing a specific function in muscle and many cell types in testis (Vickaryous and Hall, 2006). At the same time, a transcriptome is characterized by the gene expression level, i.e., how many transcripts of a gene are present, and the gene expression breadth, i.e., in how many cell types a transcript is present (Park et al., 2011). The overall higher detected

transcipt levels found in the testis as compared with the muscle, even for the genes that are common between the two tissues, is in accordance with what has been shown before in human (GTEx Consortium, 2015; Melé et al., 2015; Ramsköld et al., 2009). Fewer genes are estimated to contribute relatively to the total number of transcripts detected, although the exact number is unknown with our RNA-seq data, in the highly specialized muscle tissue in sea bass and in the Indonesian coelacanth (*Latimeria menadoensis*; Pallavicini et al., 2013). However, the tissue-specific genes showed lower expression in the testis, probably because these are expressed in few cells among the heterogeneous cell population of this tissue.

Regarding the DNA methylation per gene feature and gene expression levels, despite global similarities, there was higher variation in DNA methylation of the promoter and the first exon of the lower expressed genes in testis. This could be due to the cellular heterogeneity of testis and linked to changes connected with the formation of gametes that require chromatin rearrangements (Pallavicini et al., 2013; Tanaka and Baba, 2005).

The intrinsic differences between HK and other genes deduced by sequence data are also reflected on the DNA methylation-gene expression relationship. HK genes followed the model of low methylation in the promoter, first exon and first intron, which occurred as well in the set of the highest expressed genes in both tissues. However, the lowest expressed genes, even though not tissue-specific, had different levels of methylation, indicating a more complex relationship of expression inhibition by DNA methylation of the regulatory features. These results suggest a permissive state of gene expression linked with low methylation, but not a linear inhibitory link with high methylation.

tDMRs located in the whole gene showed both positive and negative correlation with gene expression as in human tissues (Wan et al., 2015). Here, in addition to confirm this in a phylogenetically distant species, we partition the genomic localization of tDMRs in three important gene features: promoter, first exon and first intron. tDMRs are distributed across all these gene features and exhibit both directions of correlation, but no enrichment of correlation type depending on the location. This is in accordance with the latest findings in human tissues (Wan et al., 2015) and in contrast with the

standard model of gene regulation by DNA methylation. However, in the first intron there are more tDMRs, in agreement to our finding of the importance of the first intron in the regulation of gene expression by DNA methylation.

The different sets of TF binding sites in tDMRs with positive and negative correlation suggest that different processes are regulated in different manners. The importance of TF binding to methylated cytosines has been highlighted as a widespread phenomenon to be further investigated (Hu et al., 2013; Wan et al., 2015). Here, the presence of a central CpG in both unique TF-binding sites identified in tDMRs with negative correlation contradicts the findings in humans, where the 78% of tDMRs with positive correlation contained one CpG (Wan et al., 2015). The TF-binding motifs in tDMRs with positive correlation are targets of two transcription factors, both related to signal transduction pathways, suggesting that processes related to a cellular response to an environmental stimulus may be characterized by a positive correlation of DNA methylation and gene expression in a tissue-specific manner.

In conclusion, there are similarities in the overall patterns of DNA methylation and gene expression between mammalian and fish tissues. However, there are properties such as a genome-wide negative correlation of the DNA methylation of the first intron in gene expression regulation that we have demonstrated in this study that are probably present in other vertebrate genomes but overlooked until now. In addition, further studies are needed to elucidate the roles and properties of positive and negative correlation of the tDMRs with gene expression and the specific processes involved in a tissue-dependent way. The results presented here should stimulate research along these directions.

Acknowledgments

We would like to thank Oscar Sagué and Álex Lorente for assistance with the capture of the wild fish; Dr. Noelia Díaz and Sílvia Joly for help with the samplings; Dr. Christoph Bock, Dr. Matthias Farlik, Paul Datlinger, Johanna Klughammer and Dr. Angelo Nuzzo for help with the RRBS libraries preparation and

basic bioinformatics; and Dr. Esteban Ballestar for helpful comments. DA was supported by a Ph.D. scholarship from the Spanish Government (BES–2011–044860). AEC is funded by the RED-BIO project of the Spanish National Bioinformatics Institute (INB) under grant number PT13/0001/0044. The INB is funded by the Spanish National Health Institute Carlos III (ISCIII) and the Spanish Ministry of Economy and Competitiveness (MINECO). Research supported by the Spanish Ministry of Economy and Competitiveness "Epifarm" (ref. AGL2013–41047–R) to FP.

Supplementary Figures

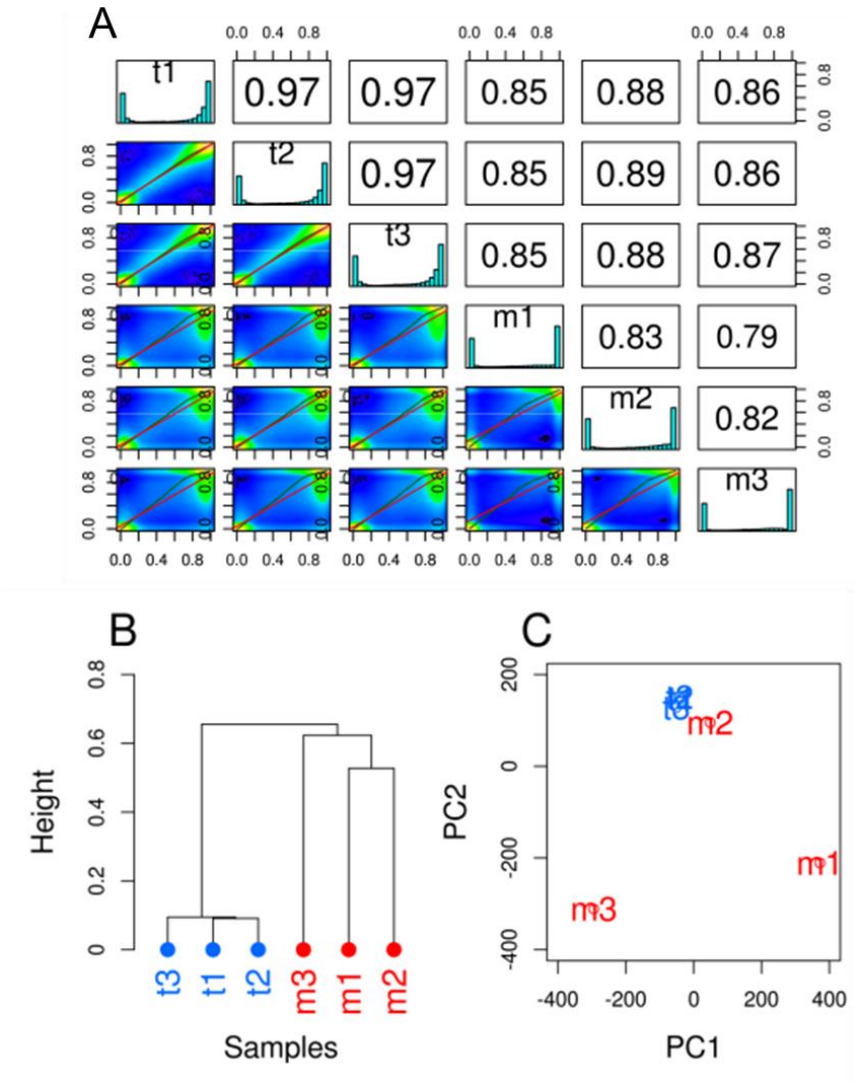


Figure S1. Similarity between RRBS samples. A) Pairwise comparisons of DNA methylation values for testis and muscle. The distribution of DNA methylation values is indicated by scatterplots of percent methylation values for each pair and diagonal histograms per sample. In addition, Pearson's correlation scores are given for each pairwise comparison. B) Clustering of samples based on Pearson's correlation distances using the Ward's method. C) Principal Component Analysis (PCA) of DNA methylation data.

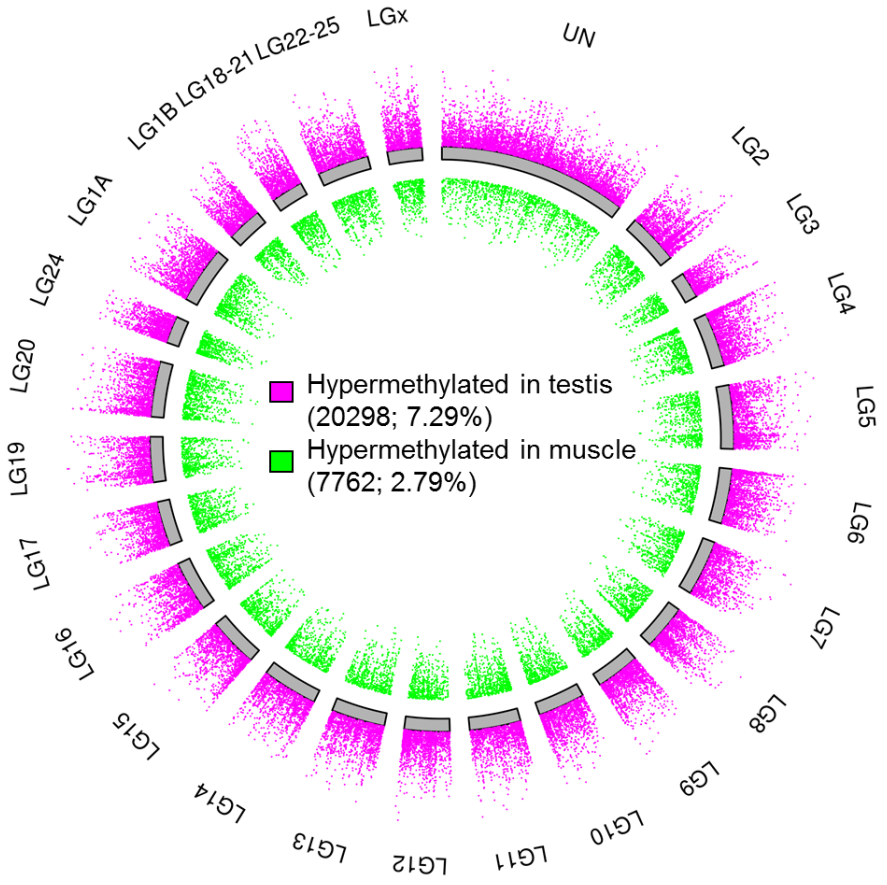


Figure S2. Differentially methylated CpGs between testis and muscle across the European sea bass genome. Distribution of hypermethylated sites in testis (magenta; outer circle) and in muscle (green; inner circle) per chromosome.

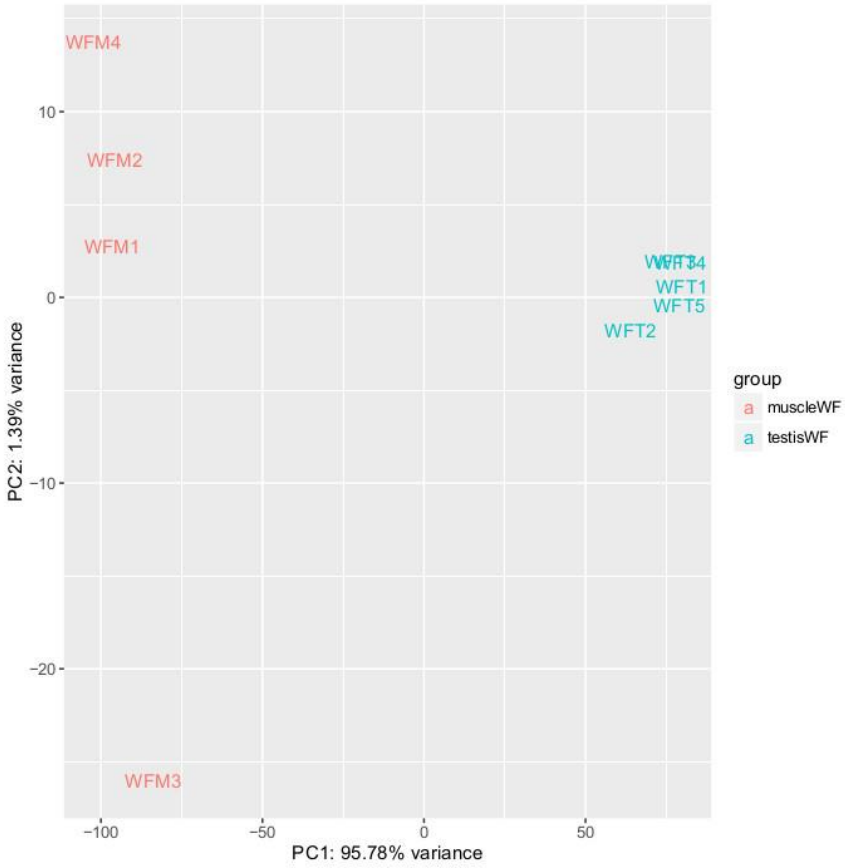


Figure S3. Similarity between RNA-seq samples as shown by Principal Component Analysis (PCA).

Chapter 5

The first steps to domestication in fish support the neural crest cell deficit hypothesis to explain Darwin's domestication syndrome

Dafni Anastasiadi and Francesc Piferrer

Institute of Marine Sciences (ICM-CSIC), Passeig Marítim de la Barceloneta 37–49, 08003 Barcelona, Spain.

Abstract

Domesticated animals present certain common altered phenotypic characteristics that constitute Darwin's Domestication Syndrome (DS). Despite much research, the underlying molecular mechanism responsible for the DS is still not clear. Recently, it was hypothesized that the DS results from a disrupted neural crest cell developmental program, the neural crest cell (NCC) deficit hypothesis, but clear support for it is hampered by difficulties of disentangling pure domestication effects from pre-existing genetic differences between farmed and wild counterparts in mammalian and avian models. Epigenetic modifications may play an important role during these first steps due to their responsiveness to environmental conditions that if they occur during early development can set lifelong phenotypic consequences. Here, we used a fish as a model to study the potential implications of DNA methylation alterations in the first steps to domestication. We compared the methylome and transcriptome of a somatic and a gonadal tissue between wild fish and genetically similar fish reared in a farming environment. We found genome-wide differences in DNA methylation associated with the farming environment in functional gene elements. Remarkably, the enrichment of biological processes linked with DNA methylation differences revealed that the most affected were related to the development of the nervous system, including NCC migration, differentiation and development. These DNA methylation marks were detected in somatic and gonadal tissues considered as derivatives of the neuroectodermal lineage. In addition, some DNA methylation differences were related to gene expression alterations in the adult organism. These findings show that the first steps to domestication in the absence of

yet genetic changes include dynamic alterations in DNA methylation of specific loci in genes related with the NCC. Thus, our results constitute, to the best of our knowledge, the first empirical demonstrations incorporating epigenetic mechanisms in support of the NCC deficit hypothesis to explain Darwin's DS.

1. Introduction

Domestication is an ancient process by which an animal species is ultimately "bred in captivity and thereby modified from its wild ancestors in ways making it more useful to humans who control its reproduction and its food supply" (Diamond, (2002). Domestication involves a series of changes in morphological, physiological, behavioral and reproductive traits (Hemmer, 1990); in mammals these changes result in the "Domestication Syndrome" (DS), a phenomenon highlighted by Darwin himself, including depigmentation, smaller teeth, ears and brain, more frequent reproductive cycles, earlier sexual maturity and tameness; some of these traits being shared also by birds and fish (Darwin, 1868; Gross, 1998; Wilkins et al., 2014).

Most livestock animals have been domesticated for hundreds of years (Diamond, 2002). Artificial selection through selective breeding has led to the establishment of a genetic basis on domesticated traits and domesticated animals have been models for the study of selection processes (Larson and Fuller, 2014). Most livestock animals share common phenotypic characteristics, such as the DS mentioned above. In some cases, phenotypic traits have been identified to be the result of a major gene affected by domestication, as for example the coat color in chicken or milk production in cow (Wright, 2015). In other cases, there is a small effect of multiple genetic loci potentially acting pleiotropically, identified as Quantitative Trait Loci (QTL), like in meat quality in cow and disease resistance in pigs (Wright, 2015). However, no single gene has yet been identified that can explain the phenotype of the DS as a whole (Wilkins et al., 2014). In old domesticates, in addition, it is difficult to recognize the wild ancestor and to identify gene flow events between wild and domestic animals (Larson and Fuller, 2014).

The farming or domestication environment is necessarily very different from the natural environment. The environment is able to influence the phenotypic variation by means of epigenetic mechanisms (Jablonka and Raz, 2009). In Mongolian sheep, for example, body size variation has been linked to differences in DNA methylation (Cao et al., 2015). In addition, DNA methylation patterns reflect the phylogenetic relationships of animals, a relationship shown in sheep (Cao et al., 2015), humans and the great apes (Hernando-Herraez et al., 2013). Thus, in domesticated livestock animals, disentangling genetic vs. epigenetic contributions on the phenotype is complicated by the unknown early domestication environment and the non-reproducibility of historic events. In this regard, the initial events in the DS have been suggested to be epigenetic changes (Wilkins et al., 2014).

In marine species, however, the situation is quite different. The first domesticated fish were the common carp (*Cyprinus carpio*), which was domesticated by the Romans ~2000 years ago (Balon, 2004), the Nile tilapia (*Oreochromis nilocitus*) in Egypt ~1500 years ago (Teletchea and Fontaine, 2014), and the goldfish (*Carassius aulatus*) in China ~1000 years ago (Balon, 2004). Nevertheless, only in the 20th century, marine species domestication has expanded dramatically. Teletchea and Fontaine (2014) suggested a classification system comprising five levels of domestication after capture fisheries, which is level 0. These include 1) first efforts to acclimatize fish to culture, 2) partly closed life cycle in captivity, 3) life cycle closed in captivity but wild inputs are required and 4) life cycle closed without need of wild inputs. At the 5th level of domestication, the entire life cycle can be reached in captivity independently of wild stocks, and in addition selective breeding programs for specific traits may already be in place (Bilio, 2007; Teletchea and Fontaine, 2014). However, only 30 fish species are considered to have reached the 5th level of domestication, and in these species, in practice production is not necessarily completely independent of wild stocks because the broodstock is often renewed with wild fish (Teletchea and Fontaine, 2014). In addition, it has been recently shown in fish that only one generation of domestication is able to alter the gene expression in a heritable manner (Christie et al., 2016). Thus, fish considered to be in the 3rd to 5th levels of domestication and raised in an aquaculture environment distinct from their native place provide excellent

models to study the effects of farming in the absence of true domestication on the phenotype and epigenetic patterns in nearly absence of genetic differentiation, as long as wild inputs are involved and selective breeding programs are not.

The farming environment as a whole comprises a set of indistinguishable biotic and abiotic factors on which to attribute the possible effects on phenotypic or epigenetic variation. One key single abiotic factor in the aquatic environment is temperature, known to influence fish from the molecular to the phenotypic level of organization (e.g., Díaz and Piferrer, 2015; Feidantsis et al., 2009; Kyprianou et al., 2010; Macqueen et al., 2008; Veilleux et al., 2015). The effects of temperature are established early during life and have lost-lasting consequences in various important functional processes, such as reproduction, metabolism, muscle growth and skeletal development (Jonsson and Jonsson, 2014; Pittman et al., 2013).

Temperature during sensitive developmental windows has a known sex determining role in species with environmental sex determination (Devlin and Nagahama, 2002) and is able to define the sexual fate of undifferentiated gonads towards the male or the female pathway. These effects are permanent and detectable in adult gonads as gene expression patterns differ between fish reared at natural and at high temperature in the pejerrey (Fernandino et al., 2012), the European sea bass (Blázquez et al., 2011; Díaz and Piferrer, 2015) and the half-smooth tongue sole (Shao et al., 2014). Persistent temperature effects in gonads at the molecular level are mediated, at least in part, by epigenetic mechanisms such as DNA methylation changes evident later in life as observed both at the gene (Navarro-Martín et al., 2011) or at the genome-wide level (Shao et al., 2014).

Plastic responses to early developmental temperature have been observed in the myogenic phenotype of several fish species (Johnston et al., 2011). In Atlantic salmon, for example, early temperature set alterations of the muscle fiber number, diameter and size distribution detected later in adult life (Macqueen et al., 2008). Persistent effects on muscle growth have also been linked with long-lasting gene expression alterations in the muscle of fish that experience high temperature during a sensitive developmental

period; patterns are altered in gilthead sea bream (Serrana et al., 2012) or pacu fish (Gutierrez de Paula et al., 2014). In the case of plastic response of the muscle, DNA methylation changes have also been suggested to be part of the underlying mechanism of maintenance since in the Senegalese sole, there is hypo-methylation of the promoter of myogenin (*myog*) when larvae are raised at increasing temperatures, together with higher *myog* expression, muscle growth and lower expression of the DNA (cytosine-5-)-methyltransferase 1 (Campos et al., 2013a). Furthermore, non-coding RNAs have been found permanently altered in the muscle of Atlantic cod (Bizuayehu et al., 2015) and Senegalese sole (Campos et al., 2014).

The European sea bass (*Dicentrarchus labrax*) is a good teleost model to study the effects farming in absence of true domestication and early temperature on epigenetic mechanisms. The sea bass is a marine coastal fish distributed in the Atlantic ocean from Senegal to Norway, the Mediterranean and the Black Sea (Kottelat and Freyhof, 2007). Sea bass tolerates diverse environments with varying degrees of salinity and temperature. Population genetic studies have identified two major populations : the Atlantic and the Mediterranean contacting each other in the Almeria-Oran front (Naciri et al., 1999; Quéré et al., 2012; Souche et al., 2015; Tine et al., 2014). Further subpopulations exist within the Mediterranean Sea, structuring the Mediterranean population into Western and Eastern Mediterranean (Bahri-Sfar et al., 2000; Castilho and Ciftci, 2005; Quéré et al., 2012). In the Atlantic populations no clear divergence has been found (Coscia and Mariani, 2011; Coscia et al., 2012; Quéré et al., 2010).

The sea bass is one of the species that has reached the fifth level of domestication (Teletchea and Fontaine, 2014), while there are three selective breeding programs in place (Migaud et al., 2013). However, even in this case, usually the broodstock of production farms originates from wild fish or the F1 generation of wild fish and in very scarce cases the F2 or the F3 (Chatain and Chavanne, 2009; Hillen et al., 2014; Novel et al., 2010; Teletchea, 2015; Vandeputte et al., 2012). It has been noted in sea bass that only one generation is sufficient to observe changes in a given trait (Vandeputte et al., 2014). Between wild and farmed sea bass there are differences in behaviour (Benhaïm et al., 2012, 2013), acute stress tolerance

(Millot et al., 2011), morphology (Arechavala-Lopez et al., 2012, 2013), organoleptic characteristics (Arechavala-Lopez et al., 2013), as well as in muscle cellularity and other muscle properties (Periago et al., 2005). Temperature during a sensitive developmental period lasting from 0 to 60 days post fertilization, called the thermosensitive period, induces plastic responses in the gonads and defines the sexual fate of sea bass, mediated at least in part by permanent changes in gonadal gene expression (Díaz and Piferrer, 2015; Navarro-Martín et al., 2009). Temperature during larval development is also able to affect muscle growth and body length in the sea bass (Ayala et al., 2000; Koumoundouros et al., 2009). Epigenetic mechanisms are good candidates to explain rapid changes in phenotypic traits in the case of farmed sea bass in the absence of domestication or domesticated sea bass, as well as permanent changes in gene expression and phenotypic traits induced by early life temperature.

In the present study, the first goal was to evaluate the effects of farming on the genome-wide DNA methylation patterns and gene expression in a recently domesticated species which may explain the rapid morphological and phenotypic changes observed between wild and farmed fish. The second aim was to survey the long-lasting effects of a single key environmental factor, temperature, experienced during early development on the methylome and the transcriptome of adult fish. To achieve these goals, we used testis and muscle obtained from wild fish, farmed fish raised at "natural" or low temperature and farmed fish raised at high temperature during development. We applied next generation sequencing techniques, e.g., Reduced Representation Bisulfite Sequencing (RRBS) to measure DNA methylation levels and RNA sequencing (RNA-seq) to measure gene expression, in order to gain single nucleotide information at the genome-wide level.

2. Materials and Methods

2.1. Animals and rearing conditions

The wild fish (WF) used to study the domestication effects are the same as in Chapter 4. The wild fish were caught at the Montgrí, Medes Islands and Baix Ter Natural Reserve (NE Spain), in order to have fish grown under natural environmental conditions. The Natural Reserve is located a few kilometers away from the commercial hatchery that provided us the farmed fish, which in turn derived from wild specimens caught in the same area. In this way we ensured that the fish used in this study came from the same area and thus from the same sea bass populations of Western Mediterranean. Futher crossings were not performed, since the aim was to study the effects of farming environment during the first steps to domestication.

Farmed European sea bass were obtained from the Hatchery Base Viva de Sant Pere Pescador. Broodstock fish typically were >3-year-old adults. In practice, no more than one generation of tank-raised fish were used as broodstock, and the older fish in each tank were occasionally renewed with younger wild sea bass caught by local fishermen of the area and acclimated at least for one year at the hatchery facilities. Farmed fish were thus at the early stages of domestication. The European sea bass naturally spawns in captivity. Fertilized float in the surface, are collected, incubated and hatch three days later. Five days-post-fertilization (dpf) larvae were transported to our the aquarium facilities of the Institute of Marine Sciences. The fish were reared under standard raising conditions for this species as described in Morretti (1999) and Díaz et al. (2013) and were divided in two groups subjected to distinct thermal treatments: farmed fish maintained at "natural" low temperature (16.5–17°C; FLT group) or farmed fish subjected to increasing (0.5°C/day) temperature from day 7 and maintained at high temperature (21°C; FHT group) until day 68 dpf, when temperature in both cases started to follow natural fluctuations. After that, fish were maintained under natural conditions of photoperiod and temperature and fed pelleted food of the appropriate size (EFICO, BioMar). The aquarium facilities are authorized for experimentations with animals by the Ministry of Agriculture and

Fisheries (certificate number 08039–46–A) according to the Spanish law (Real Decreto 223 of March 1988). Fish treatments were in accordance to the European Convention for the Protection of Vertebrate Animals used for Experimental and Other Scientific Purposes (ETS Nu 123, 126, 01/01/91). The experimental protocol was also authorized by the Spanish National Research Council (CSIC) Ethics Committee within the project AGL2013–41047–R.

2.2. Biometry

The biometric data of the fish used in this study is summarized in Table 1. The condition factor (*k*) was calculated as $100*(W/L^{3.02})$, where W= body weight in g and L = length in cm, and the gonadosomatic index (GSI) was calculated as $100*(W_g/W)$, where W_g = gonadal weight and W = body weight in g. Although we did our best to size-match the fish, the fish groups used in this study ranked in size in this order: WF>FHT>FLT. Based on the length/weight and age relationship in sea bass (Pickett and Pawson, 1994), we estimate that WF fish were older than FLT or FHT fish.

Table 1. Summary of the biometric data of fish used in this study

	WF	FLT	FHT
Sample size (n)	12	13	13
n males/ n females	8/4	9/4	11/1
GSI (mean ± SD)	0.08 ± 0.01	0.01 ± 0.03	0.07 ± 0.01
Body weight (g)			
Range	800–1400	326–624	540–664
Mean ±SD	1000 ± 110	441 ± 42	604 ± 22
Standard length (cm)			
Range	38.0–41.0	27.2–34.6	31.8–36.0
Mean ± SD	39.3 ± 0.6	30.44 ± 1.2	33.7 ± 0.7
Condition factor			
Range	1.36–1.89	1.31–1.58	2.32–1.57
Mean ± SD	1.51 ± 0.10	1.43 ± 0.03	1.48 ± 0.05

GSD: gonadosomatic index.

2.3. Sampling

Wild and farmed fish were sampled in June 2012 and June 2013, respectively, as described in Chapter 4. At sampling farmed fish were more than three years old (1122 dpf). We sampled all groups during the non-reproductive season in order to eliminate any possible effect in the gonadal tissue of physiological changes due to the annual reproductive cycle and of gonadal recrudescence itself. Fish were sacrificed using an overdose of 2-phenoxyethanol (2PE) and tissues were immediately frozen in liquid nitrogen following dissection.

DNA isolation, RRBS libraries preparation, RNA isolation, RNA-seq were carried out as described in Chapter 4, sections 2.2–2.5. The lists of genes were compared using the transcript ID by Venn diagrams drawn by Venny (Oliveros, 2007).

2.4. Genetic differentiation

Single nucleotide polymorphism (SNP) data were extracted simultaneously with DNA methylation data by BisSNP. The package VCFtools (Danecek et al., 2011) was used to prepare the files and calculate the Weir and Cockerham (Weir and Cockerham, 1984) F_{st} population estimates using only diploid sites.

2.5. Data analysis

DNA methylation analysis and RNA-seq analysis were carried out as described in Chapter 4, sections 2.6–2.8.

3. Results

Overall, pairwise correlations of DNA methylation data between technical replicates were good in testis and in muscle, although lower in the latter (Pearson's correlation scores: testis ≥ 0.97; muscle ≥ 0.78; Fig. S1). There was a clearer clustering of samples from WF than samples of FLT or FHT (Fig. S2). RNA-seq data of testis and muscle indicated the same pattern of clear separation of samples from WF from samples of FLT and FHT which were both more dispersed (Fig. S3). In all cases, the Weir and Cockerham mean F_{st} estimates were low: 0.013 between WF and FLT fish and 0.008 between FLT and FHT fish.

3.1. Overview of differentially methylated sites

Differentially methylated CpGs (DMCs) were detected across the whole genome in the four comparisons we considered: 1) in muscle of FLT vs. WF fish (36756; 21.41% of DMCs over all covered bases), 2) in testis of FLT vs WF fish (15894; 3.5%), 3) in muscle of FHT vs. FLT fish (33306; 17.63%) and 4) in testis of FHT vs. FLT fish (12806; 2.73; Fig. S4).

We, then, focused on differentially methylated regions (DMRs) since these are thought to be crucial for the regulation of gene transcription (Bock, 2012). In the muscle of FLT fish, 1011 DMRs were identified, of which 885 (88%) were inside CpG islands (CGI) or CpG shores (CGS) and 126 (12%) were outside of either CGI or CGS. Eight hundred seventy (86%) DMRs were inside gene bodies and/or promoters, 141 (14%) were outside of gene bodies and/or promoters, and 247 overlapped with repetitive elements (Fig. 1, see next page). Among the DMRs inside gene bodies and/or promoters, there were 47 inside promoters, 78 inside first exons and 98 inside first introns (Table 2, see next page).

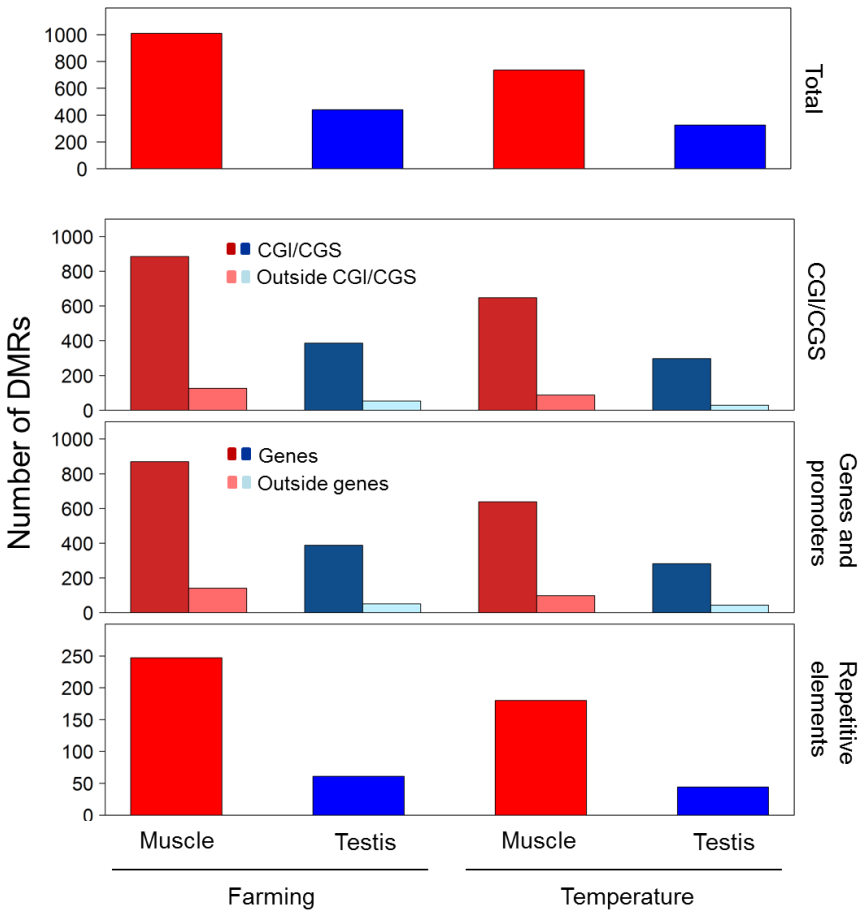


Figure 1. Number of differentially methylated regions (DMRs) in muscle and testis of farmed (FLT) vs wild fish (WF) and fish reared at high temperature during early development (FHT) vs fish reared at low temperature (FLT). Numbers of DMRs for each comparison are shown in total; inside CpG islands (CGI) and CpG shores (CGS) and outside CGI/CGS; inside gene bodies and/or promoters and outside gene bodies and/or promoters and inside repetitive elements. Numbers of DMRs in muscle are shown inside genomic features (dark red) and outside genomic features (light red). Numbers of DMRs in testis are shown inside genomic features (dark blue) and outside genomic features (light blue).

In the testis of FLT fish, we detected 439 DMRs, among which 386 (88%) overlapped with CGI/CGS and 53 (12%) were outside CGI/CGS, 388 (88%) overlapped with gene bodies and/or promoters, 51 (12%) were outside gene bodies and/or promoters, and 61 overlapped with repetitive elements (Fig. 1). The DMRs inside gene bodies and/or promoters were further partitioned in 22 inside promoters, 42 inside first exons and 44 inside first introns (Table 2).

In the muscle of FHT fish, 736 DMRs were identified. Six hundred forty eight (88%) DMRs overlapped with CGI/CGS, while 88 (12%) showed no overlap with CGI/CGS. Among the identified DMRs, 638 (87%) overlapped with gene bodies and/or promoters, 98 (13%) were located outside gene bodies and/or promoters, and 180 overlapped with repetitive elements (Fig. 1).Twenty six of the DMRs inside gene bodies and/or promoters overlapped with promoters, 50 with first exons and 65 with first introns (Table 2).

In the testis of FHT fish, 325 DMRs were detected, among which 297 (91%) overlapped with CGI/CGS and 28 (9%) were located outside CGI/CGS. In gene bodies and/or promoters 208 (87%) DMRs were assigned, while 43 (13%) DMRs showed no overlap with gene bodies and/or promoters. In the testis of FHT fish, 44 DMRs overlapped with repetitive elements (Fig. 1). Among the DMRs overlapping with gene regions, 20 overlapped with promoters, 36 with first exons and 27 with first introns (Table 2).

Table 2. Number of DMRs inside promoter, first exon or first intron in muscle and testis of farmed fish (FLT vs WF) and fish reared at high temperature (FHT vs FLT)

Comparison	Tissue	Promoter	First exon	First intron	Total
Farming	Muscle	47	78	98	192
	Testis	22	42	44	99
Temperature	Muscle	26	50	65	121
	Testis	20	36	27	74

Total refers to the total number of DMRs in any of these gene elements even if they overlap in some genes.

We, then, made an attempt to detect common DMRs between comparisons. This is limited by the definition of DMR being comparison-specific, leading to loose boundaries not necessarily coinciding at the base pair level. Despite this fact, we identified 24 robust DMRs defined by the exact same genomic coordinates in both muscle and testis of both comparisons (FLT vs. WF and FHT vs. FLT). Next, we used the identity of genomic features, and not the genomic coordinates, to uncover common signatures of differential methylation. With this approach, we detected 6 repetitive elements that contained DMRs across all comparisons (Table 3).

Table 3. Repetitive elements containing DMRs common in muscle and testis of farmed fish (FLT vs WF) and fish subjected to high temperature during early development (FHT vs FLT).

Name	Class	Family
TE-X-5_DR	Satellite	Satellite
DLA_family-5505	DNA	DNA
DLA_family-1373	LTR	Ngaro
DLA_family-1816	LINE	Rex-Babar
DLA_family-298	Satellite	Satellite
DLA_family-222	Unknown	Unknown

3.2. DMRs in gene bodies and promoters

Next, we focused on DMRs overlapping with gene bodies and/or promoters since these can be directly compared between tissues and environmental factors and may imply functional consequences. The signatures of farming were evident in 109 genes (53+17+24+15=109) that contained DMRs in both muscle and testis (Fig. 2). GO-enrichment of these 109 genes (Table S1) and visualization of the GO-terms by REViGO (Fig. 3A) revealed their association with biological processes. The most significantly enriched GO-terms were associated with nervous system development, including neural crest cell migration, followed by cell-type specific apoptotic process, multicellular organismal

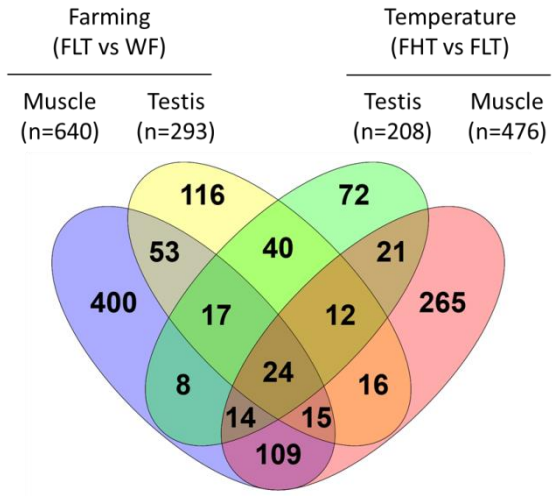


Figure 2. Number of overlapping differentially methylated regions (DMRs) in gene bodies and/or promoters. Numbers of DMRs are shown in muscle (blue) and testis (yellow) of farmed (FLT) vs. wild fish (WF), and in testis (green) and muscle (red) of fish subjected to high (FHT) vs. low temperature (FLT) during early development. The total number of DMRs is reported in each case inside parenthesis.

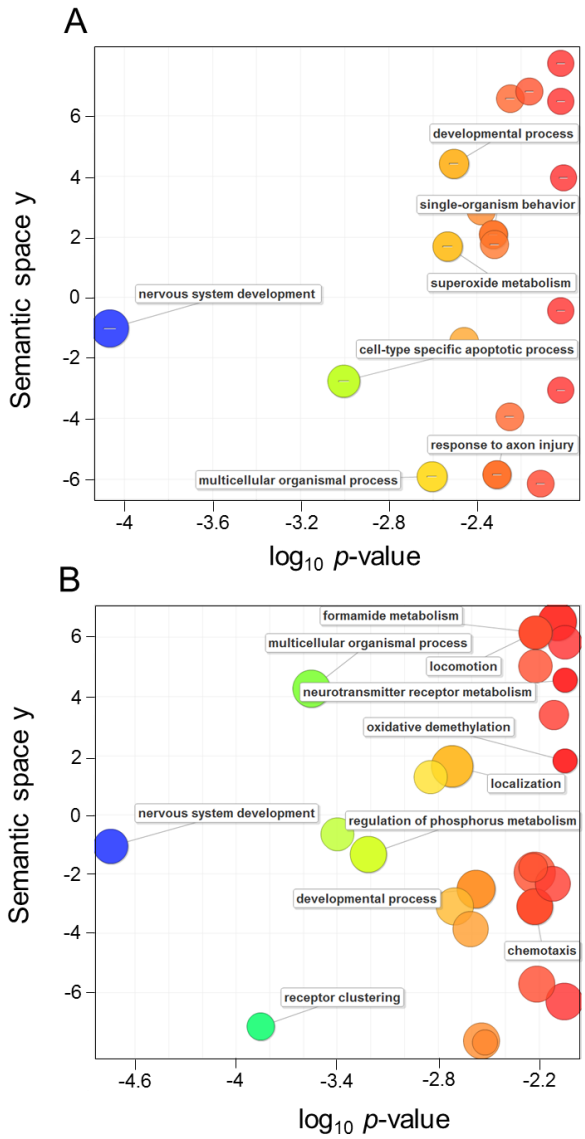


Figure 3. Reduction and visualization of enrichment of GO-terms associated with genes containing differentially methylated regions (DMRs) in their gene bodies and/or promoters. Visualization of GO-terms associated with common genes between muscle and testis in farmed (FLT vs WF) fish (A) and fish reared at high temperature during early development (FHT vs. FLT) (B). For each GO-term, the color saturation indicates the $\log_{10}$ Fisher's *p*-value of enrichment from smaller (blue) to higher (red) along the x-axis and the semantic space y after multidimensional scaling of GO-terms semantic similarities in the y-axis.

process, superoxide metabolism, developmental process, response to axon injury and single-organism behavior. On the other hand, the signatures of temperature were evident in 71 genes (21+12++24+14=71) that had common DMRs between testis and muscle (Fig. 2). GO-enrichment analysis (Table S2) and visualization of GO-terms (Fig. 3B) associated with these genes showed that the most significantly enriched were related also to nervous system development, followed by regulation of phosphorus metabolism, multicellular organismal process, developmental process, localization, formamide metabolism, locomotion, neurotransmitter metabolism, oxidative demethylation and chemotaxis.

In muscle, both farming and temperature during early development affected 162 genes (109+14+15+24=162) that were detected to contain DMRs (Fig. 2). The functional analysis of these genes (Table S3, Fig. 4A) indicated their involvement in biological processes, of which the most significantly enriched were associated with developmental process and also neural crest cell migration and differentiation (Table S3), multicellular and single-organism process, regulation of transcription from RNA polymerase III promoter, intraspecies interaction between organisms, habituation, signaling, biological regulation and cellular component organization of biogenesis. In testis, 93 genes (40+17+24+12=93) contained DMRs in their gene body and/or promoter and were affected by both farming and temperature during early development (Fig. 2). GO-term analysis of these genes (Table S4, Fig. 4B) revealed that the most significantly enriched were associated with neural crest cell migration, differentiation and development and also with cell-cell signaling, negative regulation of adenylate cyclase activity and cellular phosphate ion homeostasis.

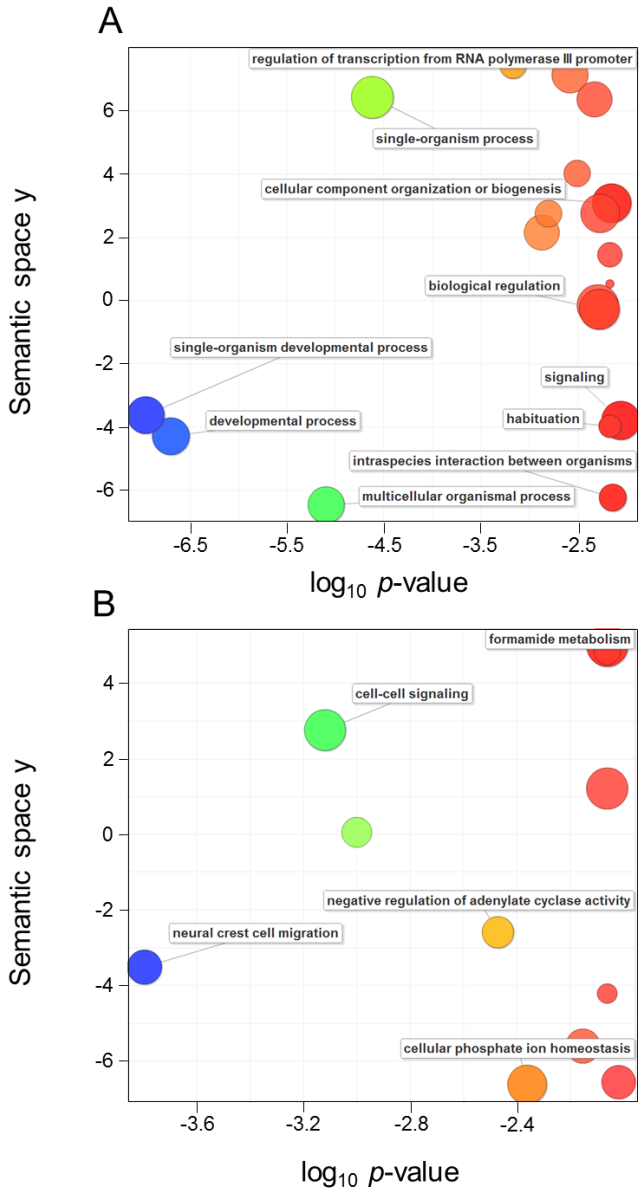


Figure 4. Reduction and visualization of GO-terms of genes that contain differentially methylated regions (DMR) in their gene body and/or promoter. GO-terms of common genes between farmed (FLT vs WF) fish and fish reared at high temperature (FHT vs FLT) are shown in muscle (A) and testis (B). The bubble color saturation indicates the $\log_{10}$ Fisher's *p*-value of GO-term enrichment from smaller (blue) to higher (red) along the x-axis and the semantic space y after multidimensional scaling of GO-terms semantic similarities in the y-axis.

Under all conditions, there were 24 common genes that contained DMRs in their gene body and/or promoter (Fig. 2, Table 4). Functional analysis based on GO-terms of these genes (Table 5, Fig. 5) showed that the most significantly enriched were involved in formate and formamide metabolism, followed by neurotransmitter receptor metabolism, positive regulation of neuromuscular junction development, mRNA cleavage, prenylation, cell-cell signaling and signal release.

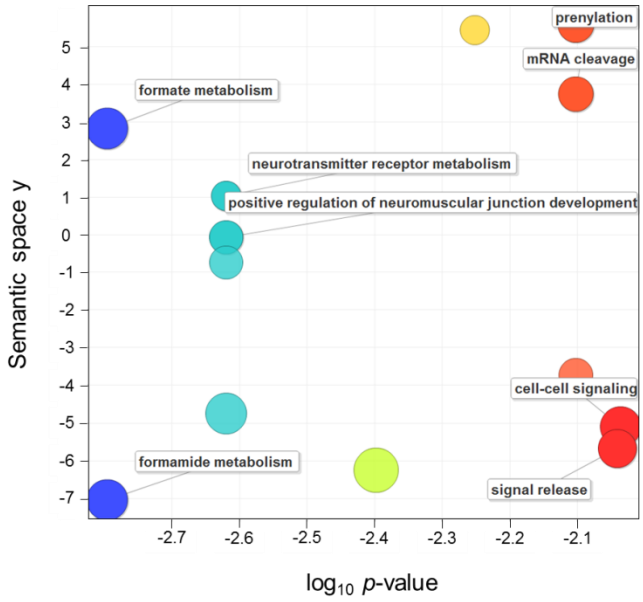


Figure 5. Enrichment of GO-terms associated with common genes overlapping with differentially methylated regions in their gene body and/or promoter. Reduction and visualization of GO-term enrichment for common genes across muscle and testis of farmed fish (FLT vs. WF) and fish subjected to high temperature (FHT vs. FLT) during early development. For each GO-term, the color saturation indicates the $\log_{10}$ Fisher's *p*-value of enrichment from smaller (blue) to higher (red) along the x-axis and the semantic space y after multidimensional scaling of GO-terms semantic similarities in the y-axis.

Table 4. Genes containing DMRs in their gene body and/or promoter common in muscle and testis of farmed fish (FLT vs WF) and fish subjected to high temperature during early development (FHT vs FLT)

Gene symbol	Gene name	Transcript ID
kcnb2	potassium voltage-gated channel subfamily b member 2	00000130
rims1	regulating synaptic membrane exocytosis protein partial	00010040
syne2	nesprin-2	00027200
bcor	bcl-6 corepressor	00053140
dlgap2	disks large-associated protein 2-like	00070410
spag16	sperm-associated antigen 16 protein	00079470
helq	helicase polq-like	00080520
rap2ip	run domain-containing protein 3a-like	00101750
ebf3	early b-cell factor 3	00113240
cit	citron rho-interacting kinase-like	00123920
p4htm	transmembrane prolyl 4-hydroxylase-like	00129910
piezo2	piezo-type mechanosensitive ion channel component 2-like	00133110
agrn	agrin precursor	00134210
mpp4	maguk p55 subfamily member 4	00136640
zpd	zona pellucida protein d	00164910
fam57b	protein fam57b-like	00194910
amdhd1	probable imidazolonepropionase-like	00209790
acvr2b	activin receptor iib	00226000
camta2	calmodulin-binding transcription activator 2 isoform 4	00232180
znf234	zinc finger protein 709-like	00232190
pop4	ribonuclease p protein subunit p29	00245100
pkhf1	pleckstrin homology domain-containing family f member 1	00245120
ankrd9	ankyrin repeat domain-containing protein 9-like	00250290
taf4	transcription initiation factor tfiid subunit 4	00263970

Table 5. Enrichment of GO-terms associated with genes that contain DMRs in muscle and testis of farmed (FLT vs WF) and high temperature treated fish (FHT vs FLT). Percentage of frequency of indispensable GO-term in human proteins of UniProt and the $\log_{10}$ of Fisher's *p*-value after enrichment are shown

GO description	Frequency	$\log_{10}$ *p*-value
formamide metabolic process	0.07%	-2.7959
neurotransmitter receptor metabolic process	0.00%	-2.6198
prenylation	0.01%	-2.1024
formate metabolic process	0.08%	-2.7959
lymphatic endothelial cell differentiation	0.00%	-2.2518
positive regulation of neuromuscular junction development	0.00%	-2.6198
protein geranylgeranylation	0.00%	-2.6198
mRNA cleavage	0.00%	-2.1024
G-protein coupled acetylcholine receptor signaling pathway	0.00%	-2.1024
imidazole-containing compound metabolic process	0.46%	-2.3979
imidazole-containing compound catabolic process	0.08%	-2.6198
cell-cell signaling	0.06%	-2.0362
signal release	0.02%	-2.041
blood vessel remodeling	0.00%	-2.0223
embryonic foregut morphogenesis	0.00%	-2.1024
positive regulation of activin receptor signaling pathway	0.00%	-2.0605

3.3. Differentially expressed genes

Differentially expressed genes (DEG) were found in the four comparisons performed, although none was common between all four (Fig. 6). There were differentially expressed genes (DEG) between the WF and the FLT fish in both testis (172; FDR<0.05) and muscle (1399; FDR<0.05). The DEGs in muscle were related to GO-terms most significantly enriched for aspartate family amino acid metabolism, somite specification, response to reactive oxygen species, B cell activation, regulation of cell growth, response to abiotic and to endogenous stimulus, one-carbon metabolism, dephosphorylation, entrainment of circadian clock, synaptic transmission, succinyl-CoA metabolism, developmental process and biological adhesion (Table S5). The DEG genes in testis were mostly related to immune system processes, among which antigen processing and presentation, regulation of leukocyte apoptoptic processes, positive regulation of interleukin-17 production, as well as to response to other organism and brown fat cell differentiation (Table S5).

DEGs were also identified between FLT and FHT (Fig. 6). In muscle, sixty two genes were differentially expressed and were related to glutamine biosynthesis, B-cell proliferation, response to stress, generation of precursor metabolites, cellular glucan metabolism, negative regulation of sequence-specific DNA binding transcription factor activity and cellular response to stimulus (Table S6). In testis, seven genes were differentially expressed (FDR<0.05) between FLT and FHT fish. GO-enrichment indicated that at least one gene of these was related to negative regulation of receptor activity (Table S6).

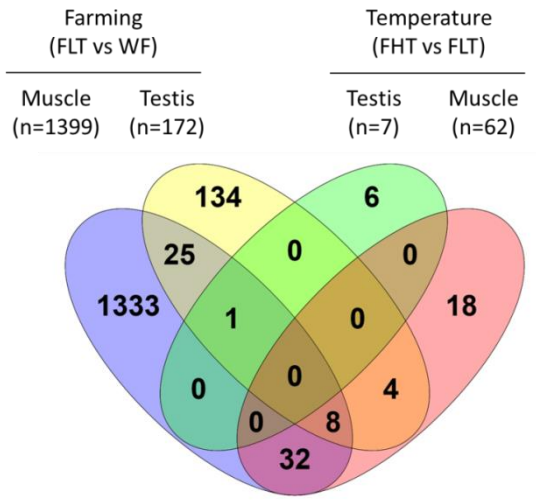


Figure 6. Overlapping differentially expressed genes (DEGs). Numbers of overlapping DEGs are indicated in muscle (blue) and testis (yellow) of farmed fish (FLT vs. WF) and in testis (green) and muscle (red) of fish that experienced high temperature during early development (FHT vs. FLT). The total number of DEGs is reported in each case inside parenthesis.

3.4. Integrative analysis of DMRs and DEGs

In order to explore the possible functional consequences of the presence of DMRs in gene bodies and/or promoters, we sought to identify DEGs that contained DMRs in each performed comparison. In the muscle of FLT fish, there were 45 DEG that contained DMRs (Table 6). GO-enrichment analysis of these genes (Table 7) followed by reduction and visualization by REViGO (Fig. 7) showed that there is enrichment in biological processes like endothelial cell proliferation, multicellular organism and developmental process, regulation of phosphorus and phosphate metabolism, succinate metabolism, galactolipid metabolism, gamma-aminobutyric acid catabolism and response to endogenous stimulus.

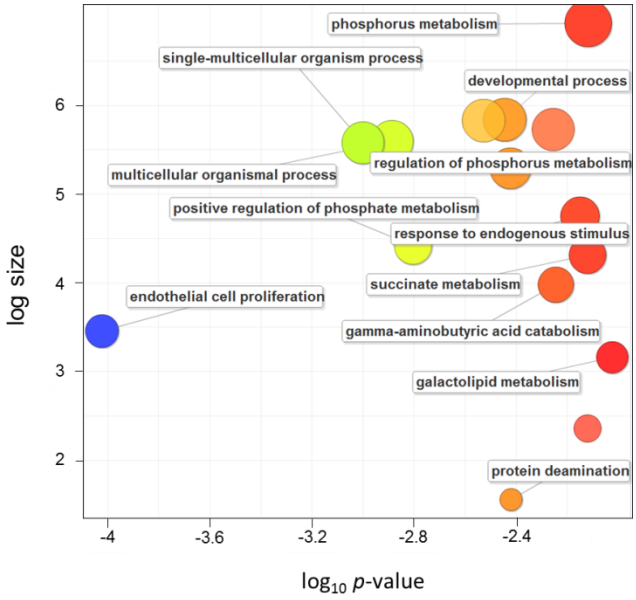


Figure 7. Enrichment of GO-terms associated with differentially expressed genes containing differentially methylated regions in their gene body and/or promoter in muscle of farmed fish. For each GO-term, the color saturation indicates the $\log_{10}$ Fisher's *p*-value of enrichment from smaller (blue) to higher (red) along the x-axis and the log size in the y-axis.

Among these genes there were 12 that they contained DMRs in their promoter, first exon and/or first intron. Transcription-factor binding sites of three transcription factors, *znf263*, *foxj3* and *daf-12* were identified inside the DMRs overlapping with promoters, first exons and/or first introns. In the testis of FLT, only one gene was differentially expressed and contained DMRs: the laminin subunit gamma-2 (*lamc2*; Fig. 8A). *Lamc2* contained a hypomethylated DMR with a total number of 10 CpGs, 7 among which were DMCs, in an internal exonic region (-21.01%; *q*-value=6.81e-35) and had lower expression in farmed fish ($\log_2$FC=-3.36; FDR=0.0006). In both tissues we could detect both positive and negative correlations between gene expression and DNA methylation differences (Fig. 9A-B).

Table 6. Genes differentially expressed containing DMRs in gene body and/or promoter in muscle of farmed fish (FLT vs WF). The presence of DMRs in the promoter (PR), first exon (FE) or first intron (FI) is indicated for each gene

Gene symbol	Gene name	PR/FE/FI	Transcript ID
mgea5	bifunctional protein ncoat isoform a		00013210
mmrn2	low quality protein: multimerin-2		00016970
mecom	mds1 and evi1 complex locus protein evi1-like		00030940
c1qtnf5	complement c1q tumor necrosis factor-related protein 5-like		00032180
stim1	stromal interaction molecule 1 isoform x1		00034820
ncor1	nuclear receptor corepressor 1-like		00039930
map3k13	mitogen-activated protein kinase kinase kinase 13-like		00049920
arxa	aristaless-related homeobox protein		00075480
tmem108	transmembrane protein 108	yes	00077810
uap1l1	udp-n-acetylhexosamine pyrophosphorylase-like protein 1		00079840
ckmt2	creatine mitochondrial 2		00081100
zswim6	zinc finger swim domain-containing protein 6	yes	00081450
slc41a1	solute carrier family 41 member 1		00088720
fgd5	and ph domain-containing protein 5-like		00090000
acap3	arf-gap with coiled- ank repeat and ph domain-containing protein 3-like		00091150
flnb	filamin-b isoform x3	yes	00097520
mapk8b	mitogen-activated protein kinase 8b-like		00098760

tnmd	tenomodulin		00107010
m1ip1	mid1-interacting protein 1		00107330
dlg3	disks large homolog 3-like		00107730
pip5k1b	phosphatidylinositol-4-phosphate 5-kinase type-1 beta-like		00116990
rtkn	rhotekin isoform x1		00117630
lipg	endothelial lipase-like		00120950
loxl2	lysyl oxidase homolog 2b-like		00124140
ctgf	connective tissue growth factor		00125930
nr4a1	nuclear receptor subfamily 4 group a member 1-like		00126930
bmpr2	bone morphogenetic protein receptor type-2-like	yes	00138170
midn	low quality protein: midnolin-like	yes	00151190
-	sits-binding		00165980
cbx8a	chromobox protein homolog 8-like		00189590
hoxb6a	homeobox protein hox-b6a-like	yes	00196380
hoxb4aa	homeobox protein hox-b4a-like		00196400
zhx2	zinc fingers and homeoboxes protein 2-like	yes	00197320
hoxa3ab	homeobox protein 3ab		00199680
hoxa2aa	homeobox protein 2aa		00199690
yrk	proto-oncogene tyrosine-protein kinase yrk-like		00200780
ppara2	peroxisome proliferator-activated receptor alpha		00207400

arhgap12	rho gtpase-activating protein 12 isoform 1	yes	00221650
acvr2b	activin receptor iib		00226000
aldh5a1	succinate-semialdehyde mitochondrial		00227150
-	---NA---	yes	00234250
olgc7	atrial natriuretic peptide receptor 1-like		00235880
-	nuclease harbi1-like	yes	00243940
ankrd9	ankyrin repeat domain-containing protein 9-like	yes	00250290
hspb11	heat shock protein beta-11	yes	00251030

Table 7. Enrichment of GO-terms associated with genes differentially expressed which contain DMRs in their gene body and/or promoter in muscle of farmed fish (FLT vs WF). Percentage of frequency of indispensable GO-term in human proteins of UniProt and the $\log_{10}$ of Fisher's *p*-value after enrichment are shown

GO description	Frequency	$\log_{10}$ *p*-value
endothelial cell proliferation	0.01%	-4.0223
multicellular organismal process	0.79%	-2.8861
developmental process	1.39%	-2.4449
positive regulation of phosphate metabolic process	0.05%	-2.8041
protein deamination	0.00%	-2.4202
single-multicellular organism process	0.76%	-3
gamma-aminobutyric acid catabolic process	0.02%	-2.2449
galactolipid metabolic process	0.00%	-2.0237
phosphorus metabolic process	16.89%	-2.118
single-organism developmental process	1.35%	-2.5272
negative regulation of DNA biosynthetic process	0.00%	-2.1203
succinate metabolic process	0.04%	-2.1203
regulation of phosphorus metabolic process	0.39%	-2.4225
signal transduction by protein phosphorylation	1.07%	-2.2549
response to endogenous stimulus	0.11%	-2.15
phosphorylation	6.30%	-2.7878

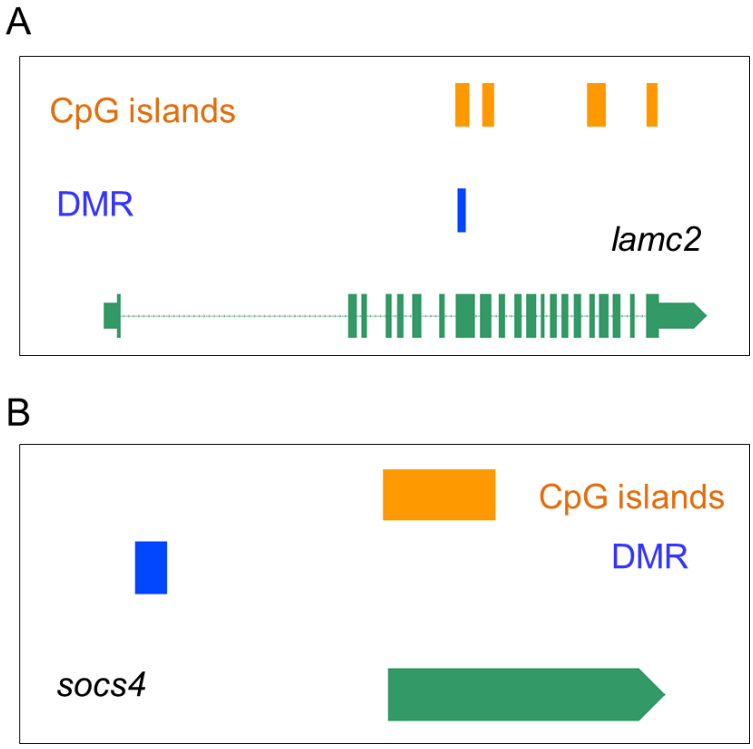


Figure 8. Schematic representation of the differentially expressed genes which contained differentially methylated regions (DMRs) in testis of farmed fish (A; FLT vs WF) and testis of fish reared at high temperature during early development (B; FHT vs FLT). Exons (thick green boxes), introns (green lines with arrows), CpG islands (orange) and differentially methylated regions (DMRs; blue) are shown. The direction of transcription is indicated by the arrow of the last exon.

In the muscle of FHT fish, no DEG with DMRs was identified. In the testis of these fish, however, one DEG with a DMR was detected: the suppressor of cytokine signaling 4 (*socs4*; Fig. 8B) which contained a hypomethylated DMR (-26.86%; *q*-value=1.35e-30), with a total number of 6 CpGs, 5 of which were DMCs, in the promoter and was more expressed (log_2FC=4.5; FDR=0.049) in high temperature fish. Positive and negative correlation of gene expression levels with DNA methylation differences was evident in both tissues (Fig. 10A-B).

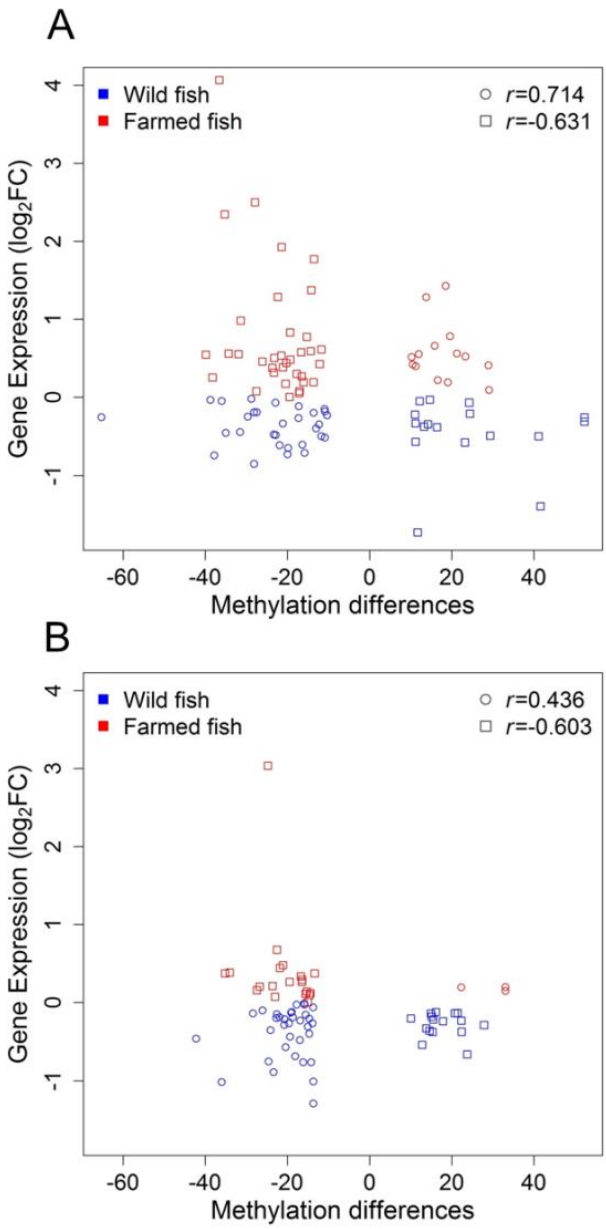


Figure 9. Expressed genes in wild and farmed fish with differentially methylated regions (DMR). DMRs overlap with the promoter, first exon or first intron. Positive (circles) and negative (boxes) correlation of DMR methylation differences and gene expression are shown for wild fish (blue) and farmed fish (red).

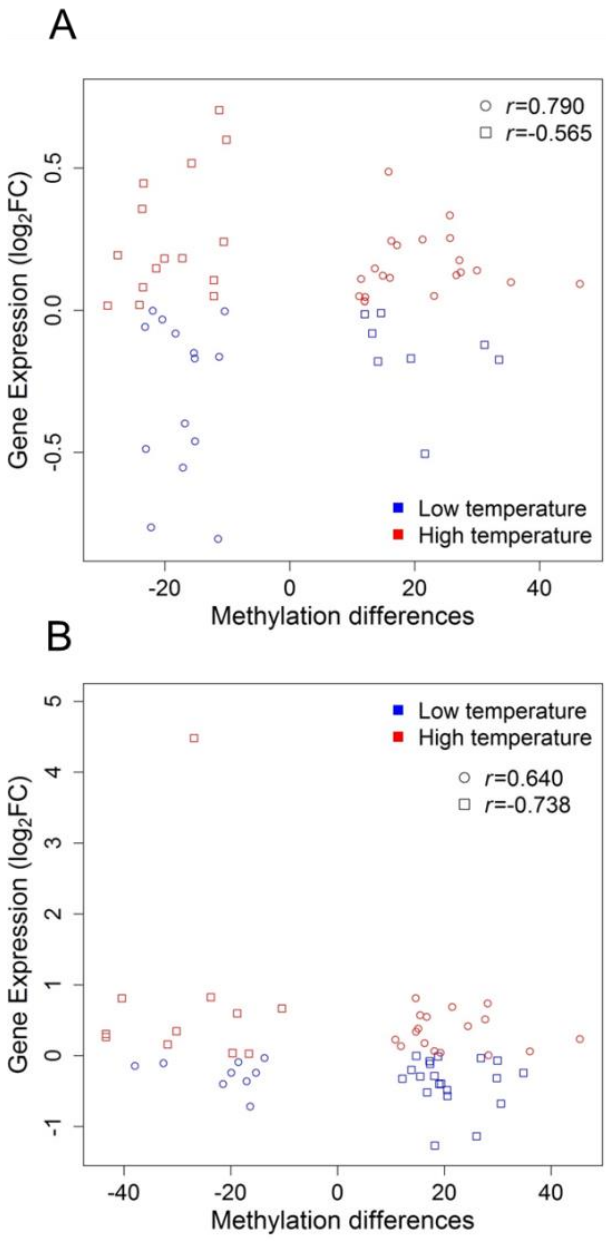


Figure 10. Expressed genes in fish raised at low and high temperature with differentially methylated regions (DMR). DMRs overlap with the promoter, first exon or first intron. Positive (circles) and negative (boxes) correlation of DMR methylation differences and gene expression are shown for "natural" or low (blue) and high (red) temperature.

4. Discussion

In the present study, we show for the first time that the farming environment in absence of true domestication affects the DNA methylation profiles at the whole genome level and the gene expression patterns of somatic and gonadal tissues. We also show that temperature increases of just 4°C experienced during sensitive developmental stages result in genome-wide changes in DNA methylation and gene expression patterns that are still evident in tissues of adult fish three years after the end of the exposure period.

Epigenetic modifications such as DNA methylation are, by some degree, defined by the genetic context of an organism (Banovich et al., 2014; Hannon et al., 2016; Teh et al., 2014) and are responsive to environmental cues (Berger et al., 2009; Feil and Fraga, 2012; Piferrer, 2013; Turner, 2009). Our experiment was designed in order to ensure that wild and farmed fish came from the same population. Moreover, we used the sequencing data to calculate the degree of genetic differentiation between the FLT and WF fish, with a Fst = 0.013 and thus at the lowest level of the range reported for sea bass natural populations (Fst = 0.011–0.035; Loukovitis et al., 2015), and similar to the value expected due to error (Waples, 1998). On the other hand, the FLT and FHT fish were randomly divided from the same batch of fertilized eggs and the only difference between them was the 4°C of temperature increase during early development (0–60 days). Accordingly, the degree of genetic differentiation was very low between the two groups of farmed fish (Fst = 0.008). Thus, although the exclusion of any genetic variant influencing the DNA methylation status of specific CpG loci is impossible, one may confidently consider that the DNA methylation and gene expression changes should be associated with the farming environment in the case of the FLT vs WF comparison, and with temperature in the case of the FHT vs FLT comparison.

The aquaculture farming environment is necessarily different from the natural environment in many abiotic (temperature, light intensity, substrate, etc.) and biotic factors (e.g., absence of predators, competence with conspecifics for food, etc.). Thus, the influences of the farming environment can be considered multifactorial (Gross, 1998; Lloret et al., 2013). Therefore, it is not

possible to attribute observed changes in DMRs and DEGs in the FLT vs. WT comparison to a single environmental factor, but, rather, they should be considered as the result of the combined action of interacting factors. On the contrary, the DMRs and DEGs observed in FHT vs FLT fish can confidently be attributed to temperature. The higher number of affected loci in the first comparison is thus likely to reflect the effect of many factors of the farming environment.

Although WF were older than FLT fish, the general pattern with ageing is genome-wide hypomethylation and hypermethylation of specific CpG loci in vertebrates (Heyn et al., 2012; Jung and Pfeifer, 2015; Romanov and Vanyushin, 1981), including fishes (Berdyshev et al., 1967; Shimoda et al., 2014). However, the DMRs detected in the FLT fish did not follow this pattern, but showed differences in both directions, i.e., hypo- and hyper-methylation. Thus, despite that an age effect on DNA methylation and gene expression cannot be completely excluded it is very unlikely to be the case. On the other hand, growth differences between the FHT and FLT fish is well documented in sea bass (Díaz et al., 2013; Person-Le Ruyet et al., 2004; Pope et al., 2014; Russell et al., 1996). Thus, since FHT were bigger than FLT fish, the observed DMRs and DEGs could be also due to higher growth rates related to temperature.

Fish muscle is a highly plastic tissue reflecting multiple reversible changes over the lifetime of a fish (Johnston et al., 2011). Thus, changes in swimming and diet composition, for example, are reflected in muscle alterations (Valente et al., 2013). Further, albeit reversible plasticity has been observed and the muscle is highly responsive to environmental changes, the conditions experienced during early life of an individual may have long-lasting consequences on the muscle phenotype (Johnston, 2006; Johnston et al., 2011). In gonads, early life experiences are also able to induce plastic responses including the irreversible differentiation as testis or ovaries (Piferrer et al., 2012; Pittman et al., 2013). In this study, we found more DMRs and DEGs in the muscle than in the testis in both comparisons. The plasticity of the muscle, particularly during early development may explain the higher number of affected loci in muscle than in the gonad, since once is differentiated, as a testis in our case, has to perform a crucial function related to the perpetuation of the species.

Differentially methylated loci were detected across the whole genome independently of the chromosome, suggesting that the early stages of domestication do not target specific genomic loci. Similar generalized effects associated with early temperature exposure have been demonstrated in the half-smooth tongue sole (Shao et al., 2014). In our study, the majority of DMRs detected were located inside functional elements in both tissues and comparisons. In the case of CGI/CGS, this may be due to the technique used, which enriches for CpG-rich regions and lacks information for the CpG-poor regions. The technique used cannot explain the localization of DMRs inside and outside of gene bodies and/or promoters since, first, there are both low and high CpG density promoters in vertebrate genomes (Elango and Yi, 2008). Second, DMRs were detected outside CpG-rich promoters, since we included exons and introns in our analysis that presented a comparable number of DMRs. Recently, it has been demonstrated that, in addition to CGI in promoters, as predicted by the standard regulation model, regions of low CpG density and low methylation outside promoter regions, enhancers located inside the gene bodies and other distal elements, are associated with the regulation of gene expression by DNA methylation (Blattler et al., 2014; Brenet et al., 2011; Edgar et al., 2014; Laurent et al., 2010; Long et al., 2013; Stadler et al., 2011). The DMRs detected, therefore, may well be part of the regulatory machinery and involved to gene expression differences. Although few differentially expressed genes contained DMRs, the methylation differences of the important regulatory elements, i.e., promoter, first exon and first intron, showed correlation with the gene expression levels. The correlations were positive or negative depending on the gene, as shown recently that may be the case of regulation of gene expression by DNA methylation in other species (Jin et al., 2014; Wan et al., 2015).

One of the major discrepancies between the natural ecosystems and the farming environment is food availability and composition affecting fish condition and resulting in different proximate composition and higher fat percentage in farmed fish (Lloret et al., 2013). The artificial (pellet) diet may explain the presence of differentially methylated and differentially expressed genes in the muscle of farmed fish related to metabolic process, like phosphorus, succinate and galactolipid metabolic processes. Phosphorus has very important physiological roles, is supplied by diet and is related

to growth (Nutrition National Research Council, 1984), while it has been associated with skeletal muscle exercise (McCully et al., 1988). Succinate metabolism is connected with the oxidative capacity of muscle fibers and muscle healthy functioning, as it is part of the citric acid cycle (Haller et al., 1991; Vladutiu and Heffner, 2000). Galactolipids, also provided by diet, are present in the myelin sheath of nerves (Fuller, 2004; Marcus et al., 2002) and as part of the extracellular matrix play an important role in the formation of muscle tissue, the innervation of which is dependent on them (Agrawal et al., 2009). Energy metabolism processes, such as citric acid cycle and oxidative phosphorylation have been found affected at the gene expression between wild and farmed Atlantic salmon (Bicskei et al., 2016; Roberge et al., 2005). In addition, the differential expression and methylation of genes related to developmental processes leads us to deem the developmental period as very important for the phenotypic outcome in a farming environment.

In the testes of farmed fish, one gene was differentially expressed and contained a DMR inside the gene body, the laminin subunit gamma-2 (*lamc2*). This gene followed the model of positive regulation of gene expression with methylation of the gene body. *Lamc2* belongs to the family of laminins which are extracellular matrix glycoproteins and is responsible for the organization of tissues, by mediating the cell migration and organization during development. In human, 13% of the testicular proteome is composed by extracellular proteins (Baert et al., 2015). The testicular extracellular matrix has an important structural role in spermatogenesis and facilitates the cellular communication (Baert et al., 2015). On the other hand, in testes of fish reared at high temperature during early development, the suppressor of cytokine signaling 4 (*socs4*) had an hypomethylated DMR in the promoter and was less expressed, following the standard model of regulation of gene expression by DNA methylation. This gene also contained an hypermethylated DMR inside the gene body in the testis of farmed fish, although it was not differentially expressed. *Socs4* is a member of the SOCS family of proteins which are implicated in cytokine signal transduction. In addition, *socs4* participates in protein ubiquitination and therefore in the protein modification pathway. In yellow catfish, members of the laminin and the SOCS family have been identified as involved in testis development and

spermatogenesis from gene expression data and as predicted targets of a microRNA related to testis development and maturation (Wu et al., 2015)

In domestication research, a major challenge is to understand the molecular mechanisms underlying the DS. A proposed explanation is the Neural Crest Cell (NCC) hypothesis (NCCH; Wilkins et al., 2014; Wright, 2015). NCCs emerge during early development, migrate to various locations and differentiate into multiple tissue cell types (Sánchez-Villagra et al., 2016). The NCCH is concerned with artificial selection for tameness and predicts that the other characteristics of the DS emerge as a by-product of this artificial selection (Sánchez-Villagra et al., 2016; Wilkins et al., 2014). Here it is important to mention that we found enriched GO-terms associated with NCC in our dataset. Remarkably, the genes with common DMRs between muscle and testis of farmed fish were significantly associated with the development of the nervous system, including NCC migration. NCC migration, differentiation and migration was also found commonly affected by the farming environment and the high developmental temperature in muscle and in testis. It is well established that the peripheral nervous system influences muscle during late development, but it has recently been shown, that also during early development, migrating NCC and the NCC lineage regulate the differentiation of skeletal muscle progenitor cells (Kalcheim, 2011; Van Ho et al., 2011). Furthermore, regarding the Leydig cells of the testis in addition to the neuroendocrine substances that they present, there is increasing evidence that they have neuroectodermal origin and derive from the neural crest (Davidoff et al., 2009; Ivell and Holstein, 2012), while more testicular cell types, i.e., Sertoli cells and intertubular cells, have been suggested to derive from NCC (Ivell and Holstein, 2012). Therefore, it seems plausible that an epigenetic memory established in the NCC during early development is carried in adult life after cell differentiation without necessarily direct translation to gene expression changes in muscle and in testis.

There is increasing evidence for supporting the NCCH to explain the domestication syndrome. It was recently shown that genes under selection in domesticates were primarily related to the nervous system in the bison (Gautier et al., 2016) and had important roles in NCC migration and other functions in cats (Montague et al., 2014).

In addition, in *Drosophila*, even in the absence of selection for tameness, it was shown that the genomic basis to domestication was enriched in neurogenetic genes further supporting the NCCH (Stanley and Kulathinal, 2016). Under this context, our findings imply that the first steps to domestication, in the absence of genetic signatures, pass through epigenetic differences of genes related to the neural crest.

Although the association of the genes highlighted in this study strongly suggests an association with the nervous system development, especially the neural crest, therefore a DNA methylation memory mitotically inherited without obligatory implications of gene expression differences, a part of gene expression differences could be attributed to other factors. Other epigenetic mechanisms regulating the gene expression (Cedar and Bergman, 2009; Turner, 2011; Yan, 2014), life stage and seasonal variations may explain a part of gene expression.

Persisting DNA methylation changes in adult animals reflect early environmental conditions. Thus, DMRs between farmed and wild fish can be good candidates as biomarkers. Epigenetic biomarkers based on the DMRs identified in this study could be extremely useful for a range of applications. First, for wild populations conservation purposes, to identify fish that have escaped from aquaculture facilities but that are not genetically different from their wild counterparts and thus cannot be distinguished by conventional genetic markers Second, the distinction between wild and farmed fish could also be extremely useful at the fish market for fish labeling in order to detect potential concealment of the fish origin. Finally, DMRs could be used as indicators of suboptimal environmental conditions such as the presence of pollution, as already have been suggested for example for methallothioneins' expression in fish as a biomarker for metal pollution (Monserrat et al., 2007).

In conclusion, we show that DNA methylation marks associated with the nervous development system are present in a farming environment and affect both somatic and gonadal tissues. These epigenetic marks are very likely to be established during early development and be mitotically inherited in the NCC lineage. In addition, there are functional consequences, as measured by gene

expression, associated with the DNA methylation changes evident in just a couple of generations of living in a farming environment and in the absence of major genetic differences. Thus, our results constitute, to the best of our knowledge, the first empirical demonstrations incorporating epigenetic mechanisms in support of the NCC deficit hypothesis to explain Darwin's DS.

Acknowledgments

We would like to thank Oscar Sagué and Álex Lorente for assistance with the capture of the wild fish; Dr. Antonio Mateos for providing the farmed fish; Dr. Noelia Díaz and Sílvia Joly for help with the samplings; Dr. Christoph Bock, Dr. Matthias Farlik, Paul Datlinger, Johanna Klughammer and Dr. Angelo Nuzzo for help with the RRBS libraries preparation and basic bioinformatics; Dr. Anna Esteve-Codina for help with differential gene expression. DA was supported by a Ph.D. scholarship from the Spanish Government (BES–2011–044860). Research supported by Spanish Ministry of Economy and Competitiveness "Epifarm" (ref. AGL2013–41047–R) to FP.

Supplementary Figures

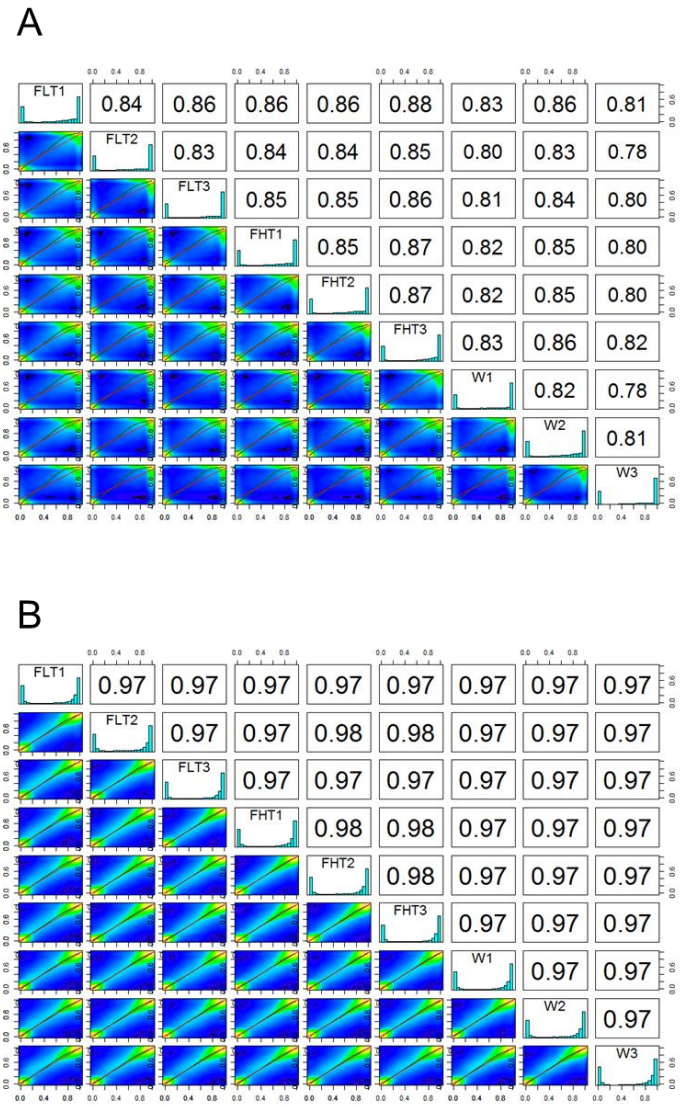


Figure S1. Pairwise comparisons of DNA methylation values between RRBS samples for muscle (A) and testis (B). The distribution of DNA methylation values is indicated by scatterplots of percent methylation values for each pair and diagonal histograms per sample. Pearson's correlation scores are given for each pairwise comparison. Replicates include wild fish (WF1-WF3), farmed fish raised at low temperature (FLT1-FLT3) and farmed fish raised at high temperature during early development (FHT1-FHT3).

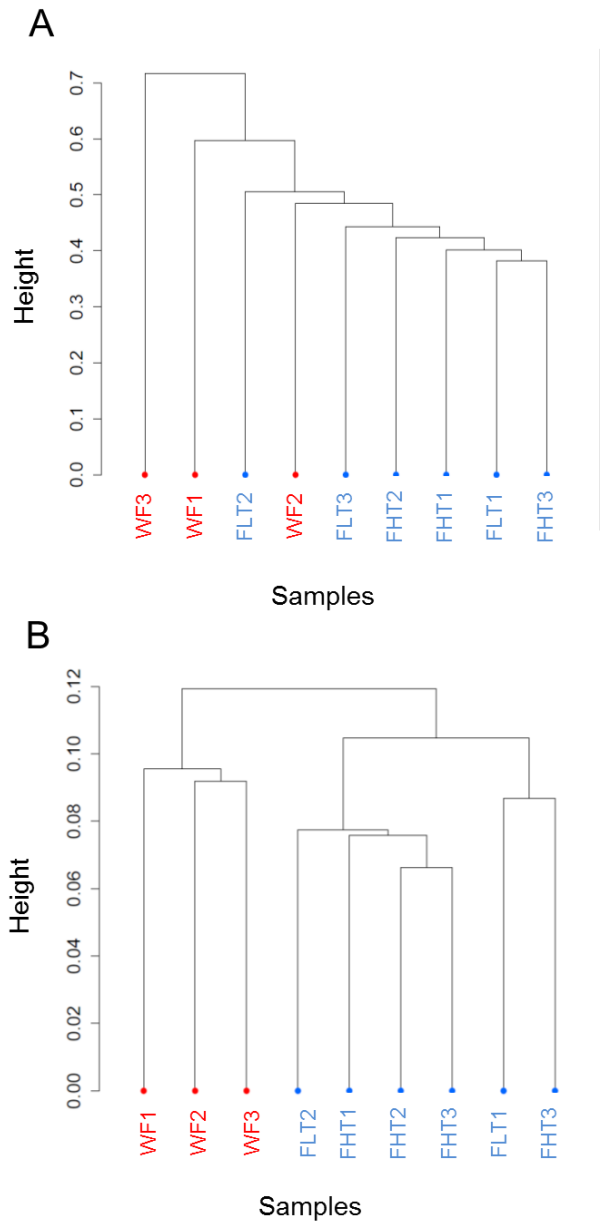


Figure S2. Similarity between RRBS samples visualized by clustering of samples based on Pearson's correlation distances using the Ward's method in muscle (A) and testis (B). Clustering is shown for biological replicates of wild fish (WF1-WF3), farmed fish reared at low temperature (FLT1-FLT3) and farmed fish reared at high temperature during early development (FHT1-FHT3).

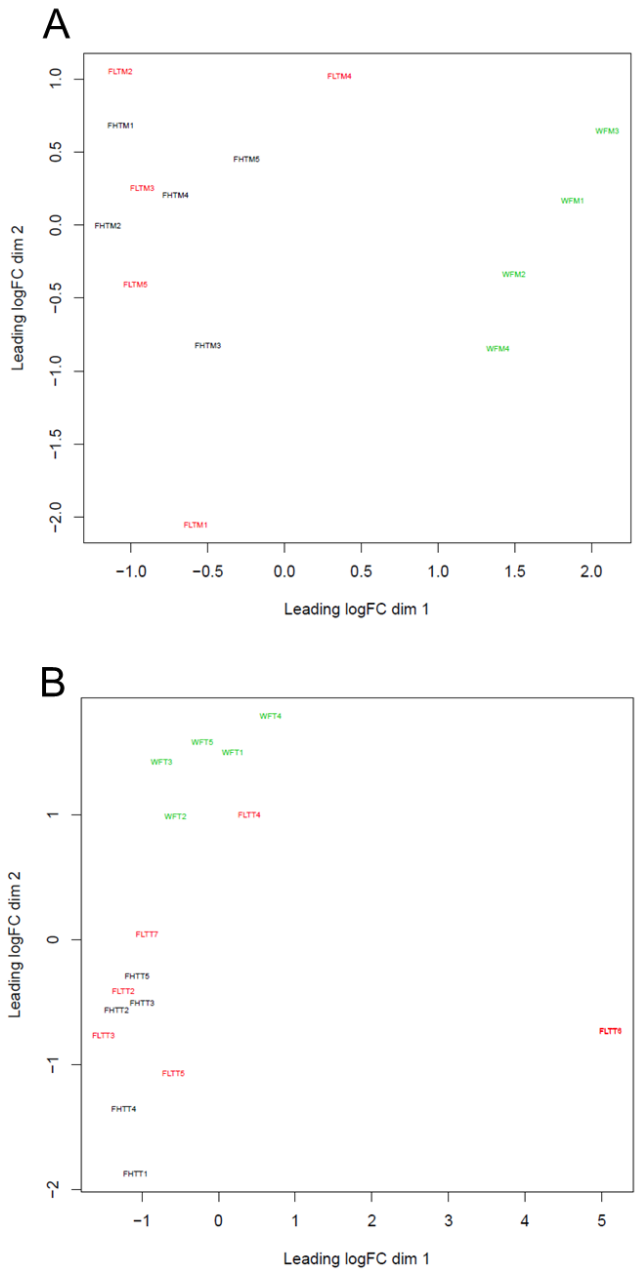


Figure S3. Similarity between RNA-seq samples. Multidimensional scaling plot of gene expression in muscle (A) and testis (B). Biological replicates are shown for wild fish (green), farmed fish reared at low temperature (red) and farmed fish reared at high temperature during early development (black). The "leading log FC dim n" represents the Fold Change of the n dimension.

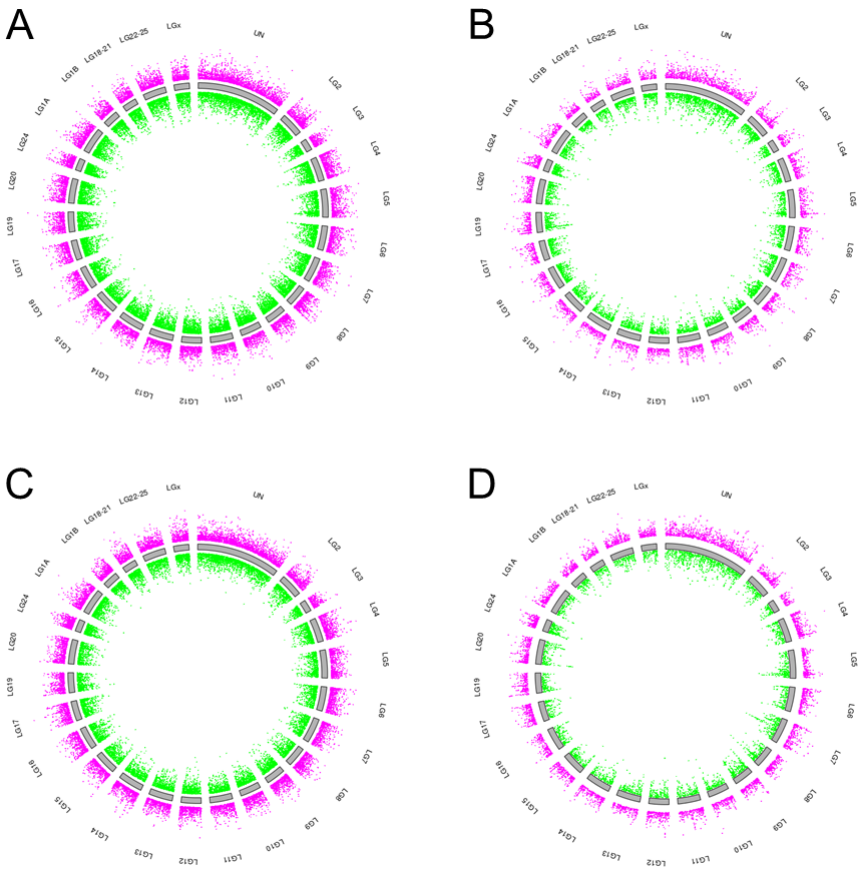


Figure S4. Differentially methylated CpGs along each chromosome of the sea bass genome. Hypermethylated (magenta) and hypomethylated (green) sites are shown for the muscle (A) and the testis (B) of FLT vs WT fish and the muscle (C) and the testis (D) of FHT vs FLT. Methylation differences ≥ 15% and with a *q*-value ≤ 0.05 were considered.

Supplementary Tables

Supplementary Table 1. Enrichment of GO-terms associated with genes that contain DMRs and that are common in testis and muscle of farmed fish (FLT vs WF)

GO ID	GO term	Annotated	Significant	Expected	Fisher's *p*-value
GO:0007399	nervous system development	1674	20	8.01	8.60E-05
GO:0031175	neuron projection development	615	10	2.94	0.00068
GO:0007275	multicellular organismal development	3957	32	18.93	0.00078
GO:0022008	neurogenesis	1145	14	5.48	0.00097
GO:0097285	cell-type specific apoptotic process	159	5	0.76	0.00099
GO:0061564	axon development	448	8	2.14	0.00133
GO:0043523	regulation of neuron apoptotic process	100	4	0.48	0.00134
GO:1901214	regulation of neuron death	104	4	0.5	0.00155
GO:0044707	single-multicellular organism process	4868	36	23.29	0.0017
GO:0051402	neuron apoptotic process	112	4	0.54	0.00204
GO:0030182	neuron differentiation	978	12	4.68	0.00225
GO:0044767	single-organism developmental process	4389	33	21	0.00232
GO:0070997	neuron death	117	4	0.56	0.00239
GO:0031290	retinal ganglion cell axon guidance	56	3	0.27	0.00243
GO:0032501	multicellular organismal process	4967	36	23.76	0.00249
GO:0006801	superoxide metabolic process	17	2	0.08	0.00293
GO:0043524	negative regulation of neuron apoptotic process	60	3	0.29	0.00296
GO:0032502	developmental process	4465	33	21.36	0.00313
GO:1901215	negative regulation of neuron death	62	3	0.3	0.00325
GO:0031103	axon regeneration	18	2	0.09	0.00329
GO:0048666	neuron development	761	10	3.64	0.00329

GO:0001755	neural crest cell migration	63	3	0.3	0.0034
GO:0007154	cell communication	4308	32	20.61	0.00348
GO:0031102	neuron projection regeneration	20	2	0.1	0.00406
GO:0034654	nucleobase-containing compound biosynthetic process	1339	14	6.41	0.00416
GO:0007409	axonogenesis	430	7	2.06	0.00446
GO:0044708	single-organism behavior	228	5	1.09	0.00472
GO:0007416	synapse assembly	71	3	0.34	0.00477
GO:0048699	generation of neurons	1073	12	5.13	0.00478
GO:0036301	macrophage colony-stimulating factor production	1	1	0	0.00478
GO:0042351	'de novo' GDP-L-fucose biosynthetic process	1	1	0	0.00478
GO:0042668	auditory receptor cell fate determination	1	1	0	0.00478
GO:0060913	cardiac cell fate determination	1	1	0	0.00478
GO:0071611	granulocyte colony-stimulating factor production	1	1	0	0.00478
GO:0071655	regulation of granulocyte colony-stimulating factor production	1	1	0	0.00478
GO:0071657	positive regulation of granulocyte colony-stimulating factor production	1	1	0	0.00478
GO:1901256	regulation of macrophage colony-stimulating factor production	1	1	0	0.00478
GO:1901258	positive regulation of macrophage colony-stimulating factor production	1	1	0	0.00478
GO:0001709	cell fate determination	22	2	0.11	0.0049
GO:0048678	response to axon injury	22	2	0.11	0.0049
GO:0007422	peripheral nervous system development	75	3	0.36	0.00555
GO:0030030	cell projection organization	821	10	3.93	0.00561
GO:0007267	cell-cell signaling	567	8	2.71	0.00564
GO:0048731	system development	3362	26	16.08	0.00592

GO:0019438	aromatic compound biosynthetic process	1405	14	6.72	0.00636
GO:0018130	heterocycle biosynthetic process	1415	14	6.77	0.00676
GO:0030073	insulin secretion	81	3	0.39	0.00688
GO:0051965	positive regulation of synapse assembly	27	2	0.13	0.00734
GO:0071396	cellular response to lipid	163	4	0.78	0.00774
GO:0048667	cell morphogenesis involved in neuron differentiation	484	7	2.32	0.00837
GO:0030154	cell differentiation	2253	19	10.78	0.00867
GO:0021536	diencephalon development	89	3	0.43	0.00891
GO:0014032	neural crest cell development	90	3	0.43	0.00918
GO:0048812	neuron projection morphogenesis	493	7	2.36	0.00922
GO:0045765	regulation of angiogenesis	91	3	0.44	0.00946
GO:0003139	secondary heart field specification	2	1	0.01	0.00955
GO:0015942	formate metabolic process	2	1	0.01	0.00955
GO:0019556	histidine catabolic process to glutamate and formamide	2	1	0.01	0.00955
GO:0019557	histidine catabolic process to glutamate and formate	2	1	0.01	0.00955
GO:0021524	visceral motor neuron differentiation	2	1	0.01	0.00955
GO:0021564	vagus nerve development	2	1	0.01	0.00955
GO:0021767	mammillary body development	2	1	0.01	0.00955
GO:0021855	hypothalamus cell migration	2	1	0.01	0.00955
GO:0032604	granulocyte macrophage colony-stimulating factor production	2	1	0.01	0.00955
GO:0032610	interleukin-1 alpha production	2	1	0.01	0.00955
GO:0032645	regulation of granulocyte macrophage colony-stimulating factor production	2	1	0.01	0.00955
GO:0032650	regulation of interleukin-1 alpha production	2	1	0.01	0.00955
GO:0032725	positive regulation of granulocyte macrophage colony-stimulating factor production	2	1	0.01	0.00955

GO:0032730	positive regulation of interleukin-1 alpha production	2	1	0.01	0.00955
GO:0043606	formamide metabolic process	2	1	0.01	0.00955
GO:0060689	cell differentiation involved in salivary gland development	2	1	0.01	0.00955
GO:0060911	cardiac cell fate commitment	2	1	0.01	0.00955
GO:0061373	mammillary axonal complex development	2	1	0.01	0.00955
GO:0061374	mammillothalamic axonal tract development	2	1	0.01	0.00955
GO:0061378	corpora quadrigemina development	2	1	0.01	0.00955
GO:0061379	inferior colliculus development	2	1	0.01	0.00955
GO:0061381	cell migration in diencephalon	2	1	0.01	0.00955
GO:0090074	negative regulation of protein homodimerization activity	2	1	0.01	0.00955
GO:2000539	regulation of protein geranylgeranylation	2	1	0.01	0.00955
GO:2000541	positive regulation of protein geranylgeranylation	2	1	0.01	0.00955
GO:0006366	transcription from RNA polymerase II promoter	892	10	4.27	0.00983

Shaded terms indicate relation to the nervous system.

Supplementary Table 2. Enrichment of GO-terms associated with genes that contain DMRs and that are common in testis and muscle of fish subjected to high temperature during early development (FHT vs. FLT)

GO ID	GO term	Annotated	Significant	Expected	Fisher's *p*-value
GO:0007399	nervous system development	1674	16	5.03	1.80E-05
GO:0008582	regulation of synaptic growth at neuromuscular junction	5	2	0.02	8.80E-05
GO:1904396	regulation of neuromuscular junction development	5	2	0.02	8.80E-05
GO:0043113	receptor clustering	34	3	0.1	0.00014
GO:0051124	synaptic growth at neuromuscular junction	8	2	0.02	0.00024
GO:0048699	generation of neurons	1073	11	3.22	0.00027
GO:0032501	multicellular organismal process	4967	27	14.93	0.00028
GO:0048731	system development	3362	21	10.1	0.00035
GO:0010562	positive regulation of phosphorus metabolic process	457	7	1.37	0.0004
GO:0045937	positive regulation of phosphate metabolic process	457	7	1.37	0.0004
GO:0007275	multicellular organismal development	3957	23	11.89	0.00045
GO:0022008	neurogenesis	1145	11	3.44	0.00048
GO:0030182	neuron differentiation	978	10	2.94	0.00056
GO:0044707	single-multicellular organism process	4868	26	14.63	0.00057
GO:0019220	regulation of phosphate metabolic process	810	9	2.43	0.0006
GO:0051174	regulation of phosphorus metabolic process	812	9	2.44	0.00061
GO:0035476	angioblast cell migration	15	2	0.05	0.00091
GO:0007267	cell-cell signaling	567	7	1.7	0.00143
GO:0007409	axonogenesis	430	6	1.29	0.00174
GO:0030154	cell differentiation	2253	15	6.77	0.00188
GO:0051179	localization	4082	22	12.27	0.00192
GO:0065009	regulation of molecular function	1156	10	3.47	0.00199
GO:0044767	single-organism developmental process	4389	23	13.19	0.00208
GO:0061564	axon development	448	6	1.35	0.00214

GO:0031175	neuron projection development	615	7	1.85	0.00227
GO:0031401	positive regulation of protein modification process	456	6	1.37	0.00234
GO:0044093	positive regulation of molecular function	624	7	1.88	0.00247
GO:0050954	sensory perception of mechanical stimulus	90	3	0.27	0.00249
GO:0032502	developmental process	4465	23	13.42	0.00266
GO:0030030	cell projection organization	821	8	2.47	0.00287
GO:0006335	DNA replication-dependent nucleosome assembly	1	1	0	0.00301
GO:0048667	cell morphogenesis involved in neuron differentiation	484	6	1.45	0.00314
GO:0043525	positive regulation of neuron apoptotic process	28	2	0.08	0.00318
GO:1901216	positive regulation of neuron death	29	2	0.09	0.00341
GO:0048812	neuron projection morphogenesis	493	6	1.48	0.00344
GO:0043085	positive regulation of catalytic activity	495	6	1.49	0.00351
GO:0048869	cellular developmental process	2401	15	7.22	0.00353
GO:0001944	vasculature development	501	6	1.51	0.00373
GO:0048856	anatomical structure development	4081	21	12.26	0.00474
GO:0051128	regulation of cellular component organization	892	8	2.68	0.00476
GO:0016477	cell migration	703	7	2.11	0.00477
GO:0032940	secretion by cell	372	5	1.12	0.00504
GO:0050790	regulation of catalytic activity	904	8	2.72	0.00515
GO:0006796	phosphate-containing compound metabolic process	2256	14	6.78	0.00533
GO:0051960	regulation of nervous system development	382	5	1.15	0.00563
GO:0048013	ephrin receptor signaling pathway	38	2	0.11	0.00581
GO:0051963	regulation of synapse assembly	38	2	0.11	0.00581
GO:0006793	phosphorus metabolic process	2283	14	6.86	0.00594
GO:0006935	chemotaxis	387	5	1.16	0.00594
GO:0032270	positive regulation of cellular protein metabolic process	552	6	1.66	0.00596
GO:0060429	epithelium development	927	8	2.79	0.00598
GO:0003149	membranous septum morphogenesis	2	1	0.01	0.006

GO:0015942	formate metabolic process	2	1	0.01	0.006
GO:0019556	histidine catabolic process to glutamate and formamide	2	1	0.01	0.006
GO:0019557	histidine catabolic process to glutamate and formate	2	1	0.01	0.006
GO:0034723	DNA replication-dependent nucleosome organization	2	1	0.01	0.006
GO:0035513	oxidative RNA demethylation	2	1	0.01	0.006
GO:0035553	oxidative single-stranded RNA demethylation	2	1	0.01	0.006
GO:0043606	formamide metabolic process	2	1	0.01	0.006
GO:0071340	skeletal muscle acetylcholine-gated channel clustering	2	1	0.01	0.006
GO:2000539	regulation of protein geranylgeranylation	2	1	0.01	0.006
GO:2000541	positive regulation of protein geranylgeranylation	2	1	0.01	0.006
GO:0001934	positive regulation of protein phosphorylation	388	5	1.17	0.00601
GO:0048870	cell motility	736	7	2.21	0.00611
GO:0051674	localization of cell	736	7	2.21	0.00611
GO:0042330	taxis	390	5	1.17	0.00614
GO:0007528	neuromuscular junction development	40	2	0.12	0.00642
GO:0048589	developmental growth	397	5	1.19	0.00661
GO:0007417	central nervous system development	748	7	2.25	0.00666
GO:0042327	positive regulation of phosphorylation	402	5	1.21	0.00696
GO:0033674	positive regulation of kinase activity	254	4	0.76	0.00704
GO:0007166	cell surface receptor signaling pathway	1162	9	3.49	0.00708
GO:0048666	neuron development	761	7	2.29	0.0073
GO:0007411	axon guidance	259	4	0.78	0.00753
GO:0097485	neuron projection guidance	259	4	0.78	0.00753
GO:0031399	regulation of protein modification process	767	7	2.3	0.00761
GO:0044087	regulation of cellular component biogenesis	260	4	0.78	0.00763
GO:0046068	cGMP metabolic process	44	2	0.13	0.00772
GO:0051247	positive regulation of protein metabolic process	584	6	1.76	0.00778
GO:0051347	positive regulation of transferase activity	262	4	0.79	0.00783

GO:0040011	locomotion	976	8	2.93	0.00809
GO:0060840	artery development	46	2	0.14	0.00842
GO:0008045	motor neuron axon guidance	47	2	0.14	0.00877
GO:0001932	regulation of protein phosphorylation	600	6	1.8	0.00883
GO:0000904	cell morphogenesis involved in differentiation	601	6	1.81	0.0089
GO:0043549	regulation of kinase activity	427	5	1.28	0.00891
GO:0046903	secretion	427	5	1.28	0.00891
GO:0001661	conditioned taste aversion	3	1	0.01	0.00899
GO:0001999	renal response to blood flow involved in circulatory renin-angiotensin regulation of systemic arterial blood pressure	3	1	0.01	0.00899
GO:0002001	renin secretion into blood stream	3	1	0.01	0.00899
GO:0003150	muscular septum morphogenesis	3	1	0.01	0.00899
GO:0006548	histidine catabolic process	3	1	0.01	0.00899
GO:0007614	short-term memory	3	1	0.01	0.00899
GO:0018344	protein geranylgeranylation	3	1	0.01	0.00899
GO:0045213	neurotransmitter receptor metabolic process	3	1	0.01	0.00899
GO:0045887	positive regulation of synaptic growth at neuromuscular junction	3	1	0.01	0.00899
GO:0052805	imidazole-containing compound catabolic process	3	1	0.01	0.00899
GO:0070989	oxidative demethylation	3	1	0.01	0.00899
GO:1904398	positive regulation of neuromuscular junction development	3	1	0.01	0.00899
GO:0048858	cell projection morphogenesis	603	6	1.81	0.00904
GO:0044089	positive regulation of cellular component biogenesis	145	3	0.44	0.00941

Shaded terms indicate relation to the nervous system.

Supplementary Table 3. Enrichment of GO-terms associated with genes that contain DMRs and are common in muscle of farmed fish (FLT vs WF) and fish reared at high temperature during early development (FHT vs WF)

GO ID	GO term	Annotated	Significant	Expected	Fisher's *p*-value
GO:0044767	single-organism developmental process	4389	56	29.88	1.10E-07
GO:0007275	multicellular organismal development	3957	52	26.94	1.70E-07
GO:0032502	developmental process	4465	56	30.4	2.00E-07
GO:0007399	nervous system development	1674	29	11.4	1.60E-06
GO:0032501	multicellular organismal process	4967	56	33.81	8.00E-06
GO:0048856	anatomical structure development	4081	49	27.78	8.40E-06
GO:0048731	system development	3362	43	22.89	8.90E-06
GO:0048869	cellular developmental process	2401	34	16.35	1.50E-05
GO:0044707	single-multicellular organism process	4868	54	33.14	2.30E-05
GO:0044699	single-organism process	11905	99	81.05	2.40E-05
GO:0030154	cell differentiation	2253	32	15.34	2.80E-05
GO:0030098	lymphocyte differentiation	96	5	0.65	0.0005
GO:0048468	cell development	1422	21	9.68	0.00055
GO:0006359	regulation of transcription from RNA polymerase III promoter	6	2	0.04	0.00068
GO:0009653	anatomical structure morphogenesis	2237	28	15.23	0.00084
GO:0030183	B cell differentiation	29	3	0.2	0.00099
GO:0030030	cell projection organization	821	14	5.59	0.00135
GO:0043113	receptor clustering	34	3	0.23	0.00158
GO:0050794	regulation of cellular process	6404	59	43.6	0.00204
GO:0032507	maintenance of protein location in cell	39	3	0.27	0.00235
GO:0048583	regulation of response to stimulus	1502	20	10.23	0.00264
GO:0045185	maintenance of protein location	42	3	0.29	0.00291
GO:0051651	maintenance of location in cell	42	3	0.29	0.00291
GO:0044089	positive regulation of cellular component biogenesis	145	5	0.99	0.00313

GO:0050789	regulation of biological process	6805	61	46.33	0.00329
GO:0030855	epithelial cell differentiation	289	7	1.97	0.00362
GO:0002521	leukocyte differentiation	152	5	1.03	0.00383
GO:0007016	cytoskeletal anchoring at plasma membrane	14	2	0.1	0.00396
GO:0045446	endothelial cell differentiation	47	3	0.32	0.004
GO:0060429	epithelium development	927	14	6.31	0.00411
GO:0000902	cell morphogenesis	836	13	5.69	0.00445
GO:0035476	angioblast cell migration	15	2	0.1	0.00455
GO:0032989	cellular component morphogenesis	938	14	6.39	0.00456
GO:0022008	neurogenesis	1145	16	7.79	0.00461
GO:0010646	regulation of cell communication	1361	18	9.27	0.00472
GO:0009966	regulation of signal transduction	1155	16	7.86	0.00501
GO:0044763	single-organism cellular process	9616	79	65.46	0.00512
GO:0065007	biological regulation	7215	63	49.12	0.00529
GO:0007154	cell communication	4308	42	29.33	0.00539
GO:0030031	cell projection assembly	165	5	1.12	0.00541
GO:0009888	tissue development	1616	20	11	0.00602
GO:0048699	generation of neurons	1073	15	7.3	0.00606
GO:0014033	neural crest cell differentiation	108	4	0.74	0.00634
GO:0030182	neuron differentiation	978	14	6.66	0.00654
GO:0031346	positive regulation of cell projection organization	110	4	0.75	0.00676
GO:0006335	DNA replication-dependent nucleosome assembly	1	1	0.01	0.00681
GO:0046959	habituation	1	1	0.01	0.00681
GO:0050894	determination of affect	1	1	0.01	0.00681
GO:0060365	coronal suture morphogenesis	1	1	0.01	0.00681
GO:0070978	voltage-gated calcium channel complex assembly	1	1	0.01	0.00681
GO:0003158	endothelium development	57	3	0.39	0.00688
GO:0071840	cellular component organization or biogenesis	3332	34	22.68	0.00709

GO:0006383	transcription from RNA polymerase III promoter	19	2	0.13	0.00728
GO:0035176	social behavior	19	2	0.13	0.00728
GO:0051703	intraspecies interaction between organisms	19	2	0.13	0.00728
GO:0048513	organ development	2482	27	16.9	0.00786
GO:0048858	cell projection morphogenesis	603	10	4.11	0.00806
GO:0007165	signal transduction	3877	38	26.39	0.00811
GO:0044700	single organism signaling	4154	40	28.28	0.00864
GO:0044087	regulation of cellular component biogenesis	260	6	1.77	0.00873
GO:0023052	signaling	4159	40	28.31	0.00882
GO:0001755	neural crest cell migration	63	3	0.43	0.00907
GO:0030217	T cell differentiation	63	3	0.43	0.00907
GO:0016043	cellular component organization	3260	33	22.19	0.00916
GO:0032990	cell part morphogenesis	617	10	4.2	0.00939
GO:0023051	regulation of signaling	1349	17	9.18	0.00962
GO:0046649	lymphocyte activation	190	5	1.29	0.00967

Shaded terms indicate relation to the nervous system.

Supplementary Table 4. Enrichment of GO-terms associated with genes that contain DMRs and are common in testis of farmed fish (FLT vs WF) and fish reared at high temperature during early development (FHT vs FLT)

GO ID	GO term	Annotated	Significant	Expected	Fisher's *p*-value
GO:0001755	neural crest cell migration	63	4	0.27	0.00016
GO:0008582	regulation of synaptic growth at neuromuscular junction	5	2	0.02	0.00019
GO:1904396	regulation of neuromuscular junction development	5	2	0.02	0.00019
GO:0030500	regulation of bone mineralization	36	3	0.16	0.00051
GO:0051124	synaptic growth at neuromuscular junction	8	2	0.03	0.00051
GO:0070167	regulation of biomineral tissue development	37	3	0.16	0.00055
GO:0051963	regulation of synapse assembly	38	3	0.17	0.0006
GO:0014032	neural crest cell development	90	4	0.39	0.00064
GO:0007267	cell-cell signaling	567	9	2.47	0.00076
GO:0032927	positive regulation of activin receptor signaling pathway	11	2	0.05	0.001
GO:0014033	neural crest cell differentiation	108	4	0.47	0.00126
GO:0030282	bone mineralization	52	3	0.23	0.0015
GO:0050807	regulation of synapse organization	54	3	0.24	0.00167
GO:0031214	biomineral tissue development	63	3	0.27	0.00261
GO:0030501	positive regulation of bone mineralization	19	2	0.08	0.00305
GO:0007194	negative regulation of adenylate cyclase activity	20	2	0.09	0.00338
GO:0031280	negative regulation of cyclase activity	20	2	0.09	0.00338
GO:0070169	positive regulation of biomineral tissue development	20	2	0.09	0.00338
GO:0007416	synapse assembly	71	3	0.31	0.00366
GO:0051350	negative regulation of lyase activity	21	2	0.09	0.00372
GO:0030643	cellular phosphate ion homeostasis	1	1	0	0.00435
GO:0072502	cellular trivalent inorganic anion homeostasis	1	1	0	0.00435
GO:0014031	mesenchymal cell development	162	4	0.71	0.00544
GO:0048864	stem cell development	163	4	0.71	0.00556

GO:0001667	ameboidal-type cell migration	263	5	1.15	0.00576
GO:0051965	positive regulation of synapse assembly	27	2	0.12	0.00612
GO:0032925	regulation of activin receptor signaling pathway	28	2	0.12	0.00657
GO:0048259	regulation of receptor-mediated endocytosis	29	2	0.13	0.00704
GO:0048762	mesenchymal cell differentiation	180	4	0.78	0.00786
GO:0030803	negative regulation of cyclic nucleotide biosynthetic process	32	2	0.14	0.00852
GO:0030818	negative regulation of cAMP biosynthetic process	32	2	0.14	0.00852
GO:0010727	negative regulation of hydrogen peroxide metabolic process	2	1	0.01	0.00869
GO:0010728	regulation of hydrogen peroxide biosynthetic process	2	1	0.01	0.00869
GO:0010730	negative regulation of hydrogen peroxide biosynthetic process	2	1	0.01	0.00869
GO:0015942	formate metabolic process	2	1	0.01	0.00869
GO:0019556	histidine catabolic process to glutamate and formamide	2	1	0.01	0.00869
GO:0019557	histidine catabolic process to glutamate and formate	2	1	0.01	0.00869
GO:0035513	oxidative RNA demethylation	2	1	0.01	0.00869
GO:0035553	oxidative single-stranded RNA demethylation	2	1	0.01	0.00869
GO:0043606	formamide metabolic process	2	1	0.01	0.00869
GO:0071340	skeletal muscle acetylcholine-gated channel clustering	2	1	0.01	0.00869
GO:0072501	cellular divalent inorganic anion homeostasis	2	1	0.01	0.00869
GO:2000539	regulation of protein geranylgeranylation	2	1	0.01	0.00869
GO:2000541	positive regulation of protein geranylgeranylation	2	1	0.01	0.00869
GO:0045669	positive regulation of osteoblast differentiation	33	2	0.14	0.00905
GO:0030278	regulation of ossification	99	3	0.43	0.00921
GO:0030809	negative regulation of nucleotide biosynthetic process	34	2	0.15	0.00959
GO:0032924	activin receptor signaling pathway	34	2	0.15	0.00959
GO:0043113	receptor clustering	34	2	0.15	0.00959
GO:1900372	negative regulation of purine nucleotide biosynthetic process	34	2	0.15	0.00959

Shaded terms indicate relation to the nervous system.

Supplementary Table 5. Enrichment of GO-terms associated with differentially expressed genes between wild and farmed fish (FLT vs WF)

Tissue	GO term ID	GO-term	Annotated	Significant	Expected	Fisher p-value
Muscle	GO:0009066	aspartate family amino acid metabolic process	26	11	1.49	7.2E-08
	GO:1901605	alpha-amino acid metabolic process	175	28	10.02	7.4E-07
	GO:0001757	somite specification	27	9	1.55	1.2E-05
	GO:1901607	alpha-amino acid biosynthetic process	62	13	3.55	4.0E-05
	GO:0000302	response to reactive oxygen species	47	11	2.69	5.3E-05
	GO:0007379	segment specification	35	9	2	1.2E-04
	GO:0009086	methionine biosynthetic process	10	5	0.57	1.2E-04
	GO:0043648	dicarboxylic acid metabolic process	60	12	3.44	1.3E-04
	GO:0008652	cellular amino acid biosynthetic process	71	13	4.07	1.7E-04
	GO:0006555	methionine metabolic process	11	5	0.63	2.1E-04
	GO:0016447	somatic recombination of immunoglobulin gene segments	17	6	0.97	2.5E-04
	GO:0000097	sulfur amino acid biosynthetic process	12	5	0.69	3.4E-04
	GO:0009067	aspartate family amino acid biosynthetic process	12	5	0.69	3.4E-04
	GO:0043401	steroid hormone mediated signaling pathway	86	14	4.93	3.6E-04
	GO:0000096	sulfur amino acid metabolic process	25	7	1.43	3.8E-04
	GO:0006520	cellular amino acid metabolic process	282	31	16.15	3.8E-04
	GO:0048545	response to steroid hormone	97	15	5.56	3.9E-04
	GO:0071383	cellular response to steroid hormone stimulus	97	15	5.56	3.9E-04
	GO:0009628	response to abiotic stimulus	466	45	26.69	4.2E-04
	GO:0006979	response to oxidative stress	119	17	6.82	4.3E-04
	GO:0002204	somatic recombination of immunoglobulin genes involved in immune response	13	5	0.74	5.3E-04
	GO:0002208	somatic diversification of immunoglobulins involved in	13	5	0.74	5.3E-04

	immune response				
GO:0002381	immunoglobulin production involved in immunoglobulin mediated immune response	13	5	0.74	5.3E-04
GO:0045190	isotype switching	13	5	0.74	5.3E-04
GO:0009068	aspartate family amino acid catabolic process	8	4	0.46	6.2E-04
GO:0006950	response to stress	1669	126	95.61	6.7E-04
GO:0042113	B cell activation	72	12	4.12	7.4E-04
GO:0002377	immunoglobulin production	28	7	1.6	8.1E-04
GO:0042542	response to hydrogen peroxide	28	7	1.6	8.1E-04
GO:0016445	somatic diversification of immunoglobulins	21	6	1.2	9.0E-04
GO:0071396	cellular response to lipid	163	20	9.34	1.1E-03
GO:0010035	response to inorganic substance	164	20	9.39	1.1E-03
GO:0002562	somatic diversification of immune receptors via germline recombination within a single locus	22	6	1.26	1.2E-03
GO:0019724	B cell mediated immunity	22	6	1.26	1.2E-03
GO:0050871	positive regulation of B cell activation	22	6	1.26	1.2E-03
GO:0044283	small molecule biosynthetic process	289	30	16.55	1.2E-03
GO:0007275	multicellular organismal development	3957	266	226.67	1.3E-03
GO:0019752	carboxylic acid metabolic process	535	48	30.65	1.4E-03
GO:0002449	lymphocyte mediated immunity	48	9	2.75	1.4E-03
GO:0002312	B cell activation involved in immune response	16	5	0.92	1.6E-03
GO:0002440	production of molecular mediator of immune response	40	8	2.29	1.7E-03
GO:0019885	antigen processing and presentation of endogenous peptide antigen via MHC class I	10	4	0.57	1.7E-03
GO:0061298	retina vasculature development in camera-type eye	10	4	0.57	1.7E-03
GO:0009649	entrainment of circadian clock	5	3	0.29	1.7E-03
GO:1901606	alpha-amino acid catabolic process	59	10	3.38	1.8E-03
GO:0009755	hormone-mediated signaling pathway	101	14	5.79	1.8E-03

GO:0009069	serine family amino acid metabolic process	32	7	1.83	1.9E-03
GO:0016444	somatic cell DNA recombination	24	6	1.37	1.9E-03
GO:0009063	cellular amino acid catabolic process	70	11	4.01	2.0E-03
GO:0007155	cell adhesion	659	56	37.75	2.0E-03
GO:0022610	biological adhesion	659	56	37.75	2.0E-03
GO:0009084	glutamine family amino acid biosynthetic process	25	6	1.43	2.4E-03
GO:0016053	organic acid biosynthetic process	175	20	10.02	2.5E-03
GO:0046394	carboxylic acid biosynthetic process	175	20	10.02	2.5E-03
GO:0016054	organic acid catabolic process	116	15	6.64	2.6E-03
GO:0046395	carboxylic acid catabolic process	116	15	6.64	2.6E-03
GO:0002483	antigen processing and presentation of endogenous peptide antigen	11	4	0.63	2.6E-03
GO:0019883	antigen processing and presentation of endogenous antigen	11	4	0.63	2.6E-03
GO:0007271	synaptic transmission, cholinergic	18	5	1.03	2.8E-03
GO:0002200	somatic diversification of immune receptors	26	6	1.49	3.0E-03
GO:0001558	regulation of cell growth	178	20	10.2	3.0E-03
GO:0002368	B cell cytokine production	2	2	0.11	3.3E-03
GO:0003428	chondrocyte intercalation involved in growth plate cartilage morphogenesis	2	2	0.11	3.3E-03
GO:0006104	succinyl-CoA metabolic process	2	2	0.11	3.3E-03
GO:0006566	threonine metabolic process	2	2	0.11	3.3E-03
GO:0006567	threonine catabolic process	2	2	0.11	3.3E-03
GO:0018277	protein deamination	2	2	0.11	3.3E-03
GO:0019284	L-methionine biosynthetic process from S-adenosylmethionine	2	2	0.11	3.3E-03
GO:0021747	cochlear nucleus development	2	2	0.11	3.3E-03
GO:0043102	amino acid salvage	2	2	0.11	3.3E-03

GO:0051045	negative regulation of membrane protein ectodomain proteolysis	2	2	0.11	3.3E-03
GO:0071265	L-methionine biosynthetic process	2	2	0.11	3.3E-03
GO:0071267	L-methionine salvage	2	2	0.11	3.3E-03
GO:0071947	protein deubiquitination involved in ubiquitin-dependent protein catabolic process	2	2	0.11	3.3E-03
GO:0002639	positive regulation of immunoglobulin production	6	3	0.34	3.3E-03
GO:0006103	2-oxoglutarate metabolic process	6	3	0.34	3.3E-03
GO:0032729	positive regulation of interferon-gamma production	6	3	0.34	3.3E-03
GO:0044767	single-organism developmental process	4389	288	251.42	3.4E-03
GO:0002285	lymphocyte activation involved in immune response	27	6	1.55	3.6E-03
GO:0032502	developmental process	4465	292	255.77	3.8E-03
GO:0006941	striated muscle contraction	98	13	5.61	3.8E-03
GO:0043436	oxoacid metabolic process	580	49	33.22	4.1E-03
GO:0006730	one-carbon metabolic process	37	7	2.12	4.5E-03
GO:0009064	glutamine family amino acid metabolic process	57	9	3.27	4.8E-03
GO:0002460	adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	47	8	2.69	4.8E-03
GO:0071407	cellular response to organic cyclic compound	161	18	9.22	5.0E-03
GO:0006082	organic acid metabolic process	586	49	33.57	5.0E-03
GO:0009725	response to hormone	239	24	13.69	5.4E-03
GO:0032870	cellular response to hormone stimulus	239	24	13.69	5.4E-03
GO:0032649	regulation of interferon-gamma production	7	3	0.4	5.5E-03
GO:0016064	immunoglobulin mediated immune response	21	5	1.2	5.8E-03
GO:0051251	positive regulation of lymphocyte activation	59	9	3.38	6.1E-03
GO:0051289	protein homotetramerization	49	8	2.81	6.3E-03
GO:0050867	positive regulation of cell activation	70	10	4.01	6.3E-03

GO:0071214	cellular response to abiotic stimulus	70	10	4.01	6.3E-03
GO:0050864	regulation of B cell activation	30	6	1.72	6.3E-03
GO:0046649	lymphocyte activation	190	20	10.88	6.3E-03
GO:0033993	response to lipid	216	22	12.37	6.4E-03
GO:0045862	positive regulation of proteolysis	104	13	5.96	6.4E-03
GO:0009119	ribonucleoside metabolic process	128	15	7.33	6.5E-03
GO:0035914	skeletal muscle cell differentiation	14	4	0.8	6.7E-03
GO:0009880	embryonic pattern specification	105	13	6.01	6.9E-03
GO:0002250	adaptive immune response	50	8	2.86	7.1E-03
GO:0042100	B cell proliferation	22	5	1.26	7.1E-03
GO:0014070	response to organic cyclic compound	261	25	14.95	8.3E-03
GO:0006576	cellular biogenic amine metabolic process	62	9	3.55	8.4E-03
GO:0003254	regulation of membrane depolarization	8	3	0.46	8.4E-03
GO:0019695	choline metabolic process	8	3	0.46	8.4E-03
GO:0032609	interferon-gamma production	8	3	0.46	8.4E-03
GO:0061299	retina vasculature morphogenesis in camera-type eye	8	3	0.46	8.4E-03
GO:0030890	positive regulation of B cell proliferation	15	4	0.86	8.8E-03
GO:0009314	response to radiation	209	21	11.97	8.8E-03
GO:0044282	small molecule catabolic process	145	16	8.31	8.9E-03
GO:0009719	response to endogenous stimulus	649	52	37.18	9.0E-03
GO:0016311	dephosphorylation	304	28	17.41	9.1E-03
GO:0002696	positive regulation of leukocyte activation	63	9	3.61	9.3E-03
GO:0002347	response to tumor cell	3	2	0.17	9.5E-03
GO:0003422	growth plate cartilage morphogenesis	3	2	0.17	9.5E-03
GO:0006579	amino-acid betaine catabolic process	3	2	0.17	9.5E-03
GO:0006657	CDP-choline pathway	3	2	0.17	9.5E-03
GO:0030299	intestinal cholesterol absorption	3	2	0.17	9.5E-03
GO:0033137	negative regulation of peptidyl-serine phosphorylation	3	2	0.17	9.5E-03

	GO:0033353	S-adenosylmethionine cycle	3	2	0.17	9.5E-03
	GO:0036295	cellular response to increased oxygen levels	3	2	0.17	9.5E-03
	GO:0061034	olfactory bulb mitral cell layer development	3	2	0.17	9.5E-03
	GO:0071455	cellular response to hyperoxia	3	2	0.17	9.5E-03
Testis	GO:0002376	immune system process	1175	24	7.06	7.9E-08
	GO:0019882	antigen processing and presentation	124	6	0.75	1.0E-04
	GO:0006955	immune response	561	12	3.37	1.3E-04
	GO:0006542	glutamine biosynthetic process	4	2	0.02	2.1E-04
	GO:0030097	hemopoiesis	429	10	2.58	2.5E-04
	GO:0009064	glutamine family amino acid metabolic process	57	4	0.34	3.8E-04
	GO:0009084	glutamine family amino acid biosynthetic process	25	3	0.15	4.4E-04
	GO:0048534	hematopoietic or lymphoid organ development	461	10	2.77	4.5E-04
	GO:0002520	immune system development	481	10	2.89	6.3E-04
	GO:0050873	brown fat cell differentiation	7	2	0.04	7.4E-04
	GO:0008652	cellular amino acid biosynthetic process	71	4	0.43	8.8E-04
	GO:0010001	glial cell differentiation	124	5	0.75	9.1E-04
	GO:0042063	gliogenesis	145	5	0.87	1.8E-03
	GO:0045321	leukocyte activation	228	6	1.37	2.5E-03
	GO:2000106	regulation of leukocyte apoptotic process	15	2	0.09	3.6E-03
	GO:0043207	response to external biotic stimulus	245	6	1.47	3.6E-03
	GO:0051707	response to other organism	245	6	1.47	3.6E-03
	GO:0016053	organic acid biosynthetic process	175	5	1.05	4.1E-03
	GO:0046394	carboxylic acid biosynthetic process	175	5	1.05	4.1E-03
	GO:0048709	oligodendrocyte differentiation	60	3	0.36	5.6E-03
	GO:0009607	response to biotic stimulus	270	6	1.62	5.8E-03
	GO:0010940	positive regulation of necrotic cell death	1	1	0.01	6.0E-03
	GO:0032740	positive regulation of interleukin-17 production	1	1	0.01	6.0E-03

GO:0035712	T-helper 2 cell activation	1	1	0.01	6.0E-03
GO:0070233	negative regulation of T cell apoptotic process	1	1	0.01	6.0E-03
GO:2000316	regulation of T-helper 17 type immune response	1	1	0.01	6.0E-03
GO:2000318	positive regulation of T-helper 17 type immune response	1	1	0.01	6.0E-03
GO:2000569	regulation of T-helper 2 cell activation	1	1	0.01	6.0E-03
GO:2000570	positive regulation of T-helper 2 cell activation	1	1	0.01	6.0E-03
GO:1901607	alpha-amino acid biosynthetic process	62	3	0.37	6.2E-03
GO:0030217	T cell differentiation	63	3	0.38	6.4E-03
GO:0042110	T cell activation	125	4	0.75	6.8E-03
GO:0070489	T cell aggregation	125	4	0.75	6.8E-03
GO:0071887	leukocyte apoptotic process	21	2	0.13	7.0E-03
GO:0071593	lymphocyte aggregation	126	4	0.76	7.0E-03
GO:0045444	fat cell differentiation	65	3	0.39	7.0E-03
GO:0070486	leukocyte aggregation	130	4	0.78	7.8E-03
GO:0002688	regulation of leukocyte chemotaxis	23	2	0.14	8.3E-03
GO:0006541	glutamine metabolic process	23	2	0.14	8.3E-03
GO:0030099	myeloid cell differentiation	213	5	1.28	9.2E-03

Supplementary Table 6. Enrichment of GO-terms associated with differentially expressed genes between farmed fish reared at natural and high temperature during early development (FHT vs FLT)

Tissue	GO term ID	GO-term	Annotated	Significant	Expected	Fisher p-value
Muscle	GO:0030890	positive regulation of B cell proliferation	15	3	0.04	0.0000
	GO:0030888	regulation of B cell proliferation	18	3	0.05	0.0000
	GO:0042100	B cell proliferation	22	3	0.06	0.0000
	GO:0050871	positive regulation of B cell activation	22	3	0.06	0.0000
	GO:0006542	glutamine biosynthetic process	4	2	0.01	0.0000
	GO:0032946	positive regulation of mononuclear cell proliferation	30	3	0.09	0.0001
	GO:0050671	positive regulation of lymphocyte proliferation	30	3	0.09	0.0001
	GO:0050864	regulation of B cell activation	30	3	0.09	0.0001
	GO:0070665	positive regulation of leukocyte proliferation	31	3	0.09	0.0001
	GO:0032944	regulation of mononuclear cell proliferation	42	3	0.12	0.0002
	GO:0050670	regulation of lymphocyte proliferation	42	3	0.12	0.0002
	GO:0070663	regulation of leukocyte proliferation	43	3	0.12	0.0003
	GO:0043433	negative regulation of sequence-specific DNA binding transcription factor activity	48	3	0.14	0.0004
	GO:0032943	mononuclear cell proliferation	56	3	0.16	0.0006
	GO:0046651	lymphocyte proliferation	56	3	0.16	0.0006
	GO:0070661	leukocyte proliferation	57	3	0.16	0.0006
	GO:0006950	response to stress	1669	13	4.81	0.0007
	GO:0051251	positive regulation of lymphocyte activation	59	3	0.17	0.0007
	GO:0002696	positive regulation of leukocyte activation	63	3	0.18	0.0008
	GO:0019752	carboxylic acid metabolic process	535	7	1.54	0.0008
	GO:0050867	positive regulation of cell activation	70	3	0.2	0.0011
	GO:1901700	response to oxygen-containing compound	411	6	1.18	0.0011
	GO:0042113	B cell activation	72	3	0.21	0.0012

GO:0050868	negative regulation of T cell activation	18	2	0.05	0.0012
GO:1903038	negative regulation of leukocyte cell-cell adhesion	18	2	0.05	0.0012
GO:0043436	oxoacid metabolic process	580	7	1.67	0.0013
GO:0006082	organic acid metabolic process	586	7	1.69	0.0014
GO:0051090	regulation of sequence-specific DNA binding transcription factor activity	168	4	0.48	0.0014
GO:0051249	regulation of lymphocyte activation	85	3	0.25	0.0019
GO:0006541	glutamine metabolic process	23	2	0.07	0.0020
GO:0051250	negative regulation of lymphocyte activation	23	2	0.07	0.0020
GO:0034111	negative regulation of homotypic cell-cell adhesion	24	2	0.07	0.0022
GO:0006091	generation of precursor metabolites and energy	193	4	0.56	0.0023
GO:0009084	glutamine family amino acid biosynthetic process	25	2	0.07	0.0023
GO:0002695	negative regulation of leukocyte activation	26	2	0.07	0.0025
GO:0002694	regulation of leukocyte activation	97	3	0.28	0.0027
GO:0002638	negative regulation of immunoglobulin production	1	1	0	0.0029
GO:0032763	regulation of mast cell cytokine production	1	1	0	0.0029
GO:0032764	negative regulation of mast cell cytokine production	1	1	0	0.0029
GO:0043379	memory T cell differentiation	1	1	0	0.0029
GO:0043380	regulation of memory T cell differentiation	1	1	0	0.0029
GO:0044324	regulation of transcription involved in anterior/posterior axis specification	1	1	0	0.0029
GO:0045829	negative regulation of isotype switching	1	1	0	0.0029
GO:0048294	negative regulation of isotype switching to IgE isotypes	1	1	0	0.0029
GO:0050866	negative regulation of cell activation	31	2	0.09	0.0036
GO:0050865	regulation of cell activation	116	3	0.33	0.0045
GO:0022408	negative regulation of cell-cell adhesion	39	2	0.11	0.0056
GO:0002713	negative regulation of B cell mediated immunity	2	1	0.01	0.0058
GO:0002829	negative regulation of type 2 immune response	2	1	0.01	0.0058

GO:0002890	negative regulation of immunoglobulin mediated immune response	2	1	0.01	0.0058
GO:0006431	methionyl-tRNA aminoacylation	2	1	0.01	0.0058
GO:0032762	mast cell cytokine production	2	1	0.01	0.0058
GO:0043371	negative regulation of CD4-positive, alpha-beta T cell differentiation	2	1	0.01	0.0058
GO:0045623	negative regulation of T-helper cell differentiation	2	1	0.01	0.0058
GO:0045629	negative regulation of T-helper 2 cell differentiation	2	1	0.01	0.0058
GO:0046639	negative regulation of alpha-beta T cell differentiation	2	1	0.01	0.0058
GO:0048289	isotype switching to IgE isotypes	2	1	0.01	0.0058
GO:0048293	regulation of isotype switching to IgE isotypes	2	1	0.01	0.0058
GO:0051791	medium-chain fatty acid metabolic process	2	1	0.01	0.0058
GO:2000016	negative regulation of determination of dorsal identity	2	1	0.01	0.0058
GO:2000515	negative regulation of CD4-positive, alpha-beta T cell activation	2	1	0.01	0.0058
GO:0031324	negative regulation of cellular metabolic process	792	7	2.28	0.0072
GO:0001817	regulation of cytokine production	137	3	0.39	0.0072
GO:0043549	regulation of kinase activity	427	5	1.23	0.0075
GO:1901701	cellular response to oxygen-containing compound	270	4	0.78	0.0075
GO:0071216	cellular response to biotic stimulus	46	2	0.13	0.0078
GO:0023057	negative regulation of signaling	431	5	1.24	0.0078
GO:0010629	negative regulation of gene expression	432	5	1.25	0.0078
GO:0000302	response to reactive oxygen species	47	2	0.14	0.0081
GO:0010648	negative regulation of cell communication	436	5	1.26	0.0081
GO:0005977	glycogen metabolic process	48	2	0.14	0.0084
GO:0002903	negative regulation of B cell apoptotic process	3	1	0.01	0.0086
GO:0009299	mRNA transcription	3	1	0.01	0.0086
GO:0031665	negative regulation of lipopolysaccharide-mediated	3	1	0.01	0.0086

		signaling pathway				
	GO:0034123	positive regulation of toll-like receptor signaling pathway	3	1	0.01	0.0086
	GO:0034139	regulation of toll-like receptor 3 signaling pathway	3	1	0.01	0.0086
	GO:0034141	positive regulation of toll-like receptor 3 signaling pathway	3	1	0.01	0.0086
	GO:0034145	positive regulation of toll-like receptor 4 signaling pathway	3	1	0.01	0.0086
	GO:0042789	mRNA transcription from RNA polymerase II promoter	3	1	0.01	0.0086
	GO:0045628	regulation of T-helper 2 cell differentiation	3	1	0.01	0.0086
	GO:0045662	negative regulation of myoblast differentiation	3	1	0.01	0.0086
	GO:0060397	JAK-STAT cascade involved in growth hormone signaling pathway	3	1	0.01	0.0086
	GO:0006073	cellular glucan metabolic process	49	2	0.14	0.0088
	GO:0044042	glucan metabolic process	49	2	0.14	0.0088
	GO:0042326	negative regulation of phosphorylation	148	3	0.43	0.0089
	GO:0051338	regulation of transferase activity	456	5	1.31	0.0098
	GO:0006986	response to unfolded protein	52	2	0.15	0.0098
Testis	GO:0007175	negative regulation of epidermal growth factor-activated receptor activity	4	1	0	0.0010
	GO:2000272	negative regulation of receptor activity	6	1	0	0.0015
	GO:0061099	negative regulation of protein tyrosine kinase activity	7	1	0	0.0017
	GO:0007176	regulation of epidermal growth factor-activated receptor activity	16	1	0	0.0039
	GO:0050732	negative regulation of peptidyl-tyrosine phosphorylation	16	1	0	0.0039
	GO:0042059	negative regulation of epidermal growth factor receptor signaling pathway	17	1	0	0.0042

GO:1901185	negative regulation of ERBB signaling pathway	17	1	0	0.0042
GO:0042058	regulation of epidermal growth factor receptor signaling pathway	32	1	0.01	0.0078
GO:0061097	regulation of protein tyrosine kinase activity	32	1	0.01	0.0078
GO:1901184	regulation of ERBB signaling pathway	32	1	0.01	0.0078

General discussion

General discussion and future directions

The overall aim of this thesis was to study intrinsic and environmental influences on the DNA methylation and gene expression patterns in a fish that ultimately contribute to its phenotype. To achieve this goal, a variety of methods was used at different levels and resolutions, allowing us to draw specific conclusions on the effects of ageing (Chapter 1), sex (Chapter 3), tissue specificity (Chapter 1 and 4), genetic background (Chapter 3), developmental temperature increases (Chapters 2, 3 and 5) and farming (Chapter 5), as well as disentangle part of the relationship of gene feature-methylation with gene expression (Chapter 4). Each chapter addressed a set of specific influences discussed in detail in each corresponding section. The results obtained could serve as a guide for stimulating and focusing future research. However, there are some common methodological considerations among the data presented that deserve to be discussed globally. In addition, we were aware that other epigenetic mechanisms are likely to be involved in the regulation of the phenotype. Thus, although not considered in this thesis, experiments were conducted to study the potential role of histone modifications in the long-lasting phenotypic consequences of temperature. Also, ncRNAs likely participate in this regulation. However, not to save efforts, but simply because DNA methylation is, on its own right, a big prey, we aimed our shots to this type of epigenetic modification.

1. Methodological considerations

Among the variety of available methods to assess DNA methylation, we chose three: 1) MSAP, which is enzyme-based, low-resolution and provides genome-wide coverage in Chapter 2, 2) RRBS, which is enzyme- and bisulifte-based, provides high-resolution and genome-wide coverage in Chapters 4 and 5, and 3) designed our own bisulifte-based high-resolution low-coverage method applied in Chapters 1 and 3, named Multiplex Bisulfite Sequencing (MBS).

MBS was discussed in detail in its corresponding section in Chapter 1. Nevertheless, it is worth mentioning that the majority of methods available for assessing DNA methylation are either genome-wide/whole genome or target a small number of genes (Fig. 1). MBS is suitable for targeting a higher number of genes than traditional bisulfite conversion followed by cloning with a significanlty higher coverage.

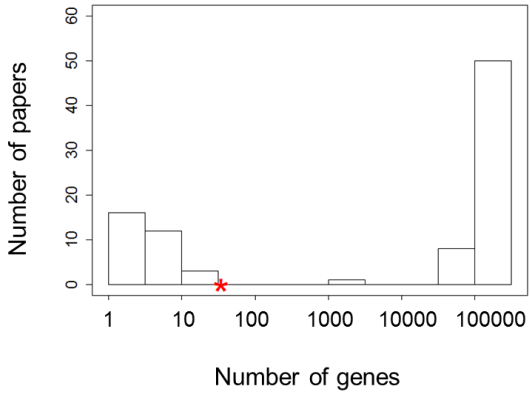


Figure 1. Number of genes interrogated in DNA methylation studies. The figure originates from a literature survey of publications between 2013-2015 containing "DNA methylation" in the title in the journal "Epigenetics". The number of genes is shown in a log scale. The red asterisk indicates the number of genes MBS targeted in Chapter 1.

The MSAP analysis is a measure of random CpG sites in the order of hundreds which are thought to be representative of the genome, as to conclude on "global" DNA methylation and epigenetic "differentiation", like the AFLP markers have long been used to conclude on genetic differentiation (Alonso et al., 2016; Foust et al., 2016; Hegarty et al., 2011; Smith and Meissner, 2013; Wenzel and Piertney, 2014). However, one limitation of MSAP is the low resolution, implying that the exact location of methylation changes remains unknown, unless specific polymorphic loci of interest are identified by gel electrophoresis, extracted and sequenced by Sanger. We deliberately preferred to automatize the fragment analysis by using automated gel electrophoresis on a Genetic

Analyzer and use appropriate software to avoid the manual scoring of bands as it is susceptible to introduce errors, therefore polymorphic loci were not sequenced by Sanger. In addition, MSAP allows for the detection only of loci with variably methylated status between samples and can be used only for comparisons between groups. A further limitation is the definition of methylation states based on the sensitivities of the enzymes used for digestion. Hemi-methylation of the external cytosine, as deduced from *Hpa*II cuts in the 5'-CCGG-3' recognition site, therefore, cannot be directly compared to non-CpG methylation levels reported by NGS-based methods, since this is distinguished only if these sites are hemi-methylated in at least only one sample, while the unmethylated cases are not distinguished. *In silico* digestion of the genome, if its sequence is available, may be used to approximately estimate the number of expected bands. However, it is impossible to predict the status of DNA methylation of the outer and inner cytosines of the restriction sites using only the genomic sequence, neither to predict whether these two cytosines will be methylated in the forward or reverse strand, or both. Thus, *in silico* MSAP digestion is only valid to obtain the maximum number of fragments expected. On the other hand, MSAP allows for the simultaneous comparison of DNA methylation in many samples, a fact that permitted us to process 11 experimental groups with 12-23 replicates each. Since our question in Chapter 2 was whether there were any effects on global DNA methylation, rather than the details of these changes, and pools of samples had to be used due to the small size of fish larvae, using the MSAP approach was the best compromise that we could settle for.

On the contrary, in Chapters 4 and 5, the DNA methylation levels of exact genomic locations were of interest and, accordingly, an NGS-based method was used. RRBS enriches for parts of the genome which are relevant to DNA methylation, i.e., CpG-rich regions. The data analysis, in this case, is most complete when a genome assembly and annotation are available. The European sea bass genome was published in 2014 and is one of the most high-quality fish genomes (Tine et al., 2014). However, there are obvious limitations when compared to model species such as the human, the mouse or even the zebrafish genome. The gene annotation was performed in the majority of cases *in silico*, since functional studies in sea bass are targeted to few genes of interest. This means that

defining the promoters of the sea bass genome was based on predicted transcription start sites (TSSes). In addition, if functional promoter annotation is not present, the definition of promoter length has to be arbitrary. In our case, we tried various sizes, varying from 500 to 4000 bp of length before ending up with a 1000 bp upstream of the TSS definition of the promoter for subsequent analysis. The definition of first exon/intron followed the published gene annotation of the sea bass genome; however, in some genes, alternative transcription start sites may be present and unknown and, therefore, could not be taken into account.

In parallel, the definition of differentially methylated loci (DMC) or regions (DMR) is necessarily arbitrary to some degree. Various statistical models have been suggested to identify differential methylation (Robinson et al., 2014), each one holding limitations. In our case, with the vast majority of softwares configured only for genomes of model organisms, we had limited choices and several analysis had to be performed by combining standardized utilities such as methylKit (Akalin et al., 2012) with custom-made packages such as the BSgenome package of the sea bass genome. In addition, differential methylation, in whole-genome or genome-wide studies, is usually defined by thresholds of 10-25% differences in methylation, but these thresholds are loose. To illustrate, differences of less than 5% have been reported to associate with phenotypic changes in the most famous case of epigenetic inheritance (Heijmans et al., 2008). The definition of DMR is also subject to arbitrary decisions on how many DMCs are required to be present, how much percent of differences is used to define a DMC and how much percent of differences is used to define a DMR, in addition to the different approaches of statistical modeling testing (Li et al., 2013; Zhang et al., 2011). In this study, we used a 15% cutoff for DMCs and 10% cutoff for DMRs in methylation differences, which, after exploratory analyses with variable thresholds, we found as a good compromise between robust differential methylation and retention of potentially interesting loci.

2. Other epigenetic mechanisms

In this thesis, we have focused on DNA methylation, as explained above. Nevertheless, the importance of other epigenetic mechanisms, namely histone modifications and variants and ncRNA may not be neglected. Therefore, we planned to investigate the potential long-lasting influences of early developmental temperature on histone modifications and one histone variant because there was one study showing histone variant changes during temperature acclimation in carp (Simonet et al., 2013) and temperature is the main environmental factor tested in this thesis. The study of histone modifications and variants may be complicated by the lack of previous knowledge, since there are a variety of them.

Testis and muscle samples from the farmed fish of Chapter 5 subjected to low (FLT) or high temperature (FHT) were used to perform Chromatin Immunoprecipitation (ChIP) followed by qPCR on target regions. In order to obtain candidate target genes, we employed three criteria: 1) genes differentially expressed based on the RNA-seq data, 2) genes with DMCs based on RRBS data, and 3) genes known to present long-lasting changes in their expression due to early temperature in sea bass (Díaz and Piferrer, 2015). We, then, measured the expression of 5-10 candidate genes also by qPCR. Two histone modifications and one histone variant were selected for the ChIP: H3K4me3, which is typically associated with transcriptional activation; H3K9me3, which is associated with gene silencing (Rose and Klose, 2014), and the H2A.z histone variant, previously found affected by seasonal temperature variations in other fish (Simonet et al., 2013).

Unfortunately, tissue size was small, frozen at -80ºC, while most parts of the testis and muscle samples from the FLT and FHT had been already used for RRBS and RNA-seq analysis. The sample conditions constituted a limitation for ChIP, since ~1µg of chromatin is necessary per IP, a minimum of four IPs per sample had to be performed and freezing affects the fragmentation procedure (Schoppee Bortz and Wamhoff, 2011). Optimizations were performed at every step of the procedure, from nuclei extraction to the IP, and ChIPs were carried out several times. In all cases, pooling of samples was avoided, but only 3 biological

replicates could be used. Therefore, we obtained large biological variation, which prevented us from confidently drawing conclusions.

However, some indications were acquired when targeting the following genomic regions of *myog*: ~2000 bp upstream of the TSS, around the TSS and ~2000 bp downstream of the TSS. In the muscle, *myog* was up-regulated in European sea bass exposed to high temperature, as it had been observed in the Senegalese sole, *Solea senegalensis* (Campos et al., 2013a) . There was enrichment of the activation marks, H3K4me3 and H2A.z, in high temperature fish and enrichment of the repression mark H3K9me3 in low temperature fish (Fig. 2). These results indicate the potential presence of long-lasting effects of early developmental temperature on histone modifications and variants as well.

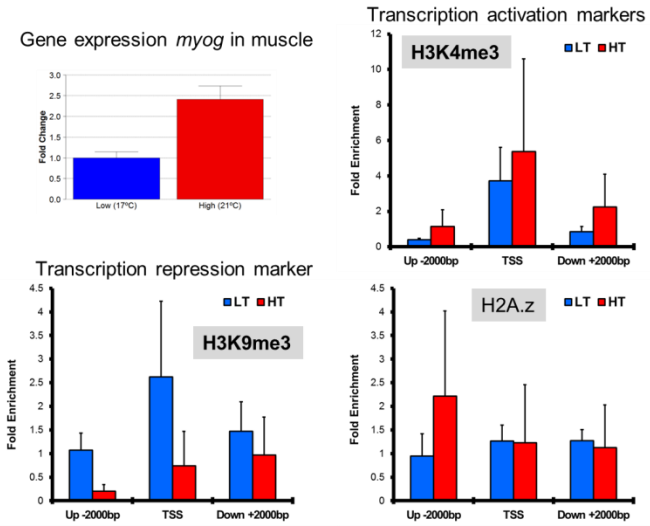


Fig. 2. Gene expression and enrichment of histone modifications and an histone variant in *myog* in the muscle of fish reared at low (LT; blue) or high (HT; red) temperature during early development. Fold enrichment of histones is shown for ~2000 bp upstream of the transcription start site (TSS), around the TSS and ~2000 bp downstream of the TSS.

The importance of ncRNAs cannot be neglected, although not addressed in this study, and especially miRNAs for which there is increasing knowledge in teleost fish. In teleost fish, miRNAs have a significant role in the regulation of early development and formation of various tissues, including the eye, the heart, the gonads, the brain and the muscle (Bizuayehu and Babiak, 2014). miRNAs are present in the gonads of teleosts and exhibit sexually dimorphic patterns of expression (Bizuayehu et al., 2012, 2013; Jing et al., 2014). Temperature is able to influence the expression of miRNAs during early development in the Senegalese sole (Campos et al., 2014). In addition, miRNAs although themselves versatile, their expression is altered at the long-term due to early developmental temperature increases in the Atlantic cod possibly through other epigenetic regulatory mechanisms (Bizuayehu et al., 2015).

3. Future directions

The results presented in this thesis can serve as a guide for future research. In Chapter 1, tissue-independent CpG loci that change methylation with age unidirectionally were identified that could be used as a starting point for the development of a piscine epigenetic clock after validation and calibration. Perhaps the best known molecular clock of ageing is the telomeric clock, based on the observation that telomere length is negatively correlated with age, with a correlation coefficient between 0.5 and 0.6 (Gibss, 2014). In comparison, the epigenetic clock is much more accurate, with a correlation coefficient of >0.95. Thus, based on the extremely accurate epigenetic clock in humans, confirmed in chimpanzees and whales (Horvath, 2013; Polanowski et al., 2014), the potential of age estimation in fish based on a DNA methylation clock is great. The CpG loci that we suggested can serve for focusing the efforts of future research. Firstly, European sea bass of more age classes along its natural lifespan should be included for measuring DNA methylation of candidate CpG loci, as well as more tissues and especially, easily accessible tissues like fins, gills or scales. Then, these CpGs should be tested in other fish species of distinct growth rates and longevities in order to calibrate the epigenetic clock. In parallel, the epigenetic clock has to be validated using tissues and

fish from different origins and age classes. This could be achieved at a reasonable cost using the MBS method developed in this thesis.

In Chapter 4, we identified a clear negative correlation between the DNA methylation of the first intron and gene expression in muscle and in testis. However, in each expression decile there was variation of DNA methylation, although the median was negatively correlated. We expect that the genes with DNA methylation around the median should share some common characteristics, either structural or functional. In addition, in the testis there was higher variation of DNA methylation of the promoter and first exon in the 20% of the lowest expressed genes when compared to the muscle. The genes that disobey the rule of low methylation in the promoter and first exon in testis could also share structural or functional characteristics that distinguish them from the rest. These relationships will be investigated in the sea bass in the near future. However, since this genome-wide relationship of clear negative correlation is firstly identified here in any species, these results should be confirmed in other vertebrates. Furthermore, genome-wide DNA methylation studies should include the first intron as such in the analysis.

Another major finding of this thesis is that our data obtained comparing wild vs. farmed sea bass strongly suggest an epigenetic link between the first steps to domestication and the NCC hypothesis to explain Darwin's domestication syndrome. However, these results more than a definite proof should serve to stimulate further research. Importantly, the localization of DNA methylation changes in genes related to NCC should be evaluated in other tissues known to be direct derivatives of NCC, ideally in homogeneous cell populations of unique lineage. This should also include analysis of extensive gene expression and phenotypic traits associated to the domestication syndrome. Moreover, these DNA methylation alterations should be confirmed in other vertebrate species, including fish species at the lowest levels of domestication. Although in mammalian domesticates genetic differences are already in place, non-domesticated research model species could be used to evaluate whether there is a phylogenetic conservation of the epigenetic responses to the farming environment. Also, the link between initial epigenetic response and genetic integration of

domestication signatures should be investigated and for this, fish such as the European sea bass, for which selective breeding programs exist at various stages of selection, are ideal models.

Lastly, we described a series of DNA methylation changes associated with early developmental temperature. These were global changes in fish larvae, but not juveniles as defined in Chapter 2. Additional time-course experiments at various time intervals of the first 15 dpf using moderately elevated temperature can be performed to further narrow down, if possible, the sensitive developmental period and, subsequently, identify the exact position of these changes to deduce the processes affected. In Chapter 3, we showed gene-specific changes in association with the genetic component of the sex determination system, which deserve further investigation, as well as the probable epigenetic inheritance of the DNA methylation patterns from father to daughter. Potential associations between differential methylation of specific CpG loci in the genes identified as responsive to temperature and *cis* acting SNPs should be evaluated, in search of potentially methylation quantitative trait loci (meQTL). Also, the DNA methylation and gene expression patterns, together with the SNPs and meQTLs of other genes involved in gonad differentiation, will help to elucidate the complex gene network regulating sex and the genetically-conditioned response to temperature.

Conclusions

1. A low-cost method was developed to interrogate the DNA methylation status of genomic regions up to ~550 bp at single nucleotide resolution in many samples. This method is available for 22 key genes of sea bass while it is easily customizable through new primers design for additional genes and/or other species.

2. We have identified a series of CpG loci that exhibit age-related decreases or increases in methylation in somatic and gonadal tissues. These loci constitute promising candidates for the development of a piscine epigenetic clock, with potential uses to estimate age in fish population and fisheries science studies.

3. Temperature increases within the range predicted by the latest global warming models result in stage-dependent changes in global DNA methylation and expression of ecologically important genes.

4. The inhibitory role of developmental temperature on the expression of gonadal aromatase is non-linearly mediated by DNA methylation, which in turn is influenced by the male parent. Thus, in a species with polygenic sex determination with mixed genetic and environmental influences, the genetic make-up influences the epigenetic response to temperature.

5. Developmental temperature increases are associated with higher methylation in genes related to female development and lower methylation in genes related to male development. Therefore, other genes in addition to c*yp19a1a* related to gonad formation participate in the epigenetic response to temperature.

6. A clear, genome-wide, quasi-linear, inverse relationship between DNA methylation of the first intron and gene expression is present in tissues of different cellular heterogeneity. These results highlight an overlooked gene feature for the regulation of gene expression by DNA methylation, which deserves testing in other vertebrate genomes.

7. Part of the tissue-specific differences in DNA methylation of the first intron is positively associated with gene expression. This association confirms the tissue-specific mechanisms of gene expression regulation and includes the first intron as a member of this regulation.

8. The first steps of domestication in the absence of yet major genetic differences encompass dynamic changes of DNA methylation in genes of somatic and gonadal tissues related to the nervous system, including neural crest cell development, differentiation and migration. This is, to our knowledge, the first empirical demonstration incorporating epigenetic mechanisms in support of the neural crest cell deficit hypothesis to explain Darwin's domestication syndrome.

9. Because of the conserved nature of fundamental regulatory mechanisms, the findings of this thesis concerning the effects of the environment, the age-related changes and the intricacies of gene features, constitute a solid starting point for further research—not only in fishes but in other vertebrates, including mammals—to advance our understanding of the contribution of epigenetics to the emergence of the phenotype.

References

A

Acar, M., Becskei, A., and van Oudenaarden, A. (2005). Enhancement of cellular memory by reducing stochastic transitions. Nature *435*, 228–232.

Adam, M., Murali, B., Glenn, N.O., and Potter, S.S. (2008). Epigenetic inheritance based evolution of antibiotic resistance in bacteria. BMC Evol. Biol. *8*, 52.

Adler, D. (2005). vioplot: Violin plot.

Agrawal, V., Brown, B.N., Beattie, A.J., Gilbert, T.W., and Badylak, S.F. (2009). Evidence of innervation following extracellular matrix scaffold-mediated remodelling of muscular tissues. J. Tissue Eng. Regen. Med. *3*, 590–600.

Aguilera, O., Fernandez, A.F., Munoz, A., and Fraga, M.F. (2010). Epigenetics and environment: a complex relationship. J. Appl. Physiol. *109*, 243–251.

Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F.E., Figueroa, M.E., Melnick, A., Mason, C.E., and others (2012). methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. Genome Biol *13*, R87.

Alexa, A., and Rahnenfuhrer, J. (2016). topGO: Enrichment Analysis for Gene Ontology. R Package Version 2240.

Alonso, C., Pérez, R., Bazaga, P., Medrano, M., and Herrera, C.M. (2016). MSAP markers and global cytosine methylation in plants: a literature survey and comparative analysis for a wild-growing species. Mol. Ecol. Resour. *16*, 80–90.

Arechavala-Lopez, P., Sanchez-Jerez, P., Bayle-Sempere, J.T., Sfakianakis, D.G., and Somarakis, S. (2012). Morphological differences between wild and farmed Mediterranean fish. Hydrobiologia *679*, 217–231.

Arechavala-Lopez, P., Fernandez-Jover, D., Black, K.D., Ladoukakis, E., Bayle-Sempere, J.T., Sanchez-Jerez, P., and

Dempster, T. (2013). Differentiating the wild or farmed origin of Mediterranean fish: a review of tools for sea bream and sea bass. Rev. Aquac. *5*, 137–157.

Arefev, V.A. (1989). Cytogenetic analysis of the common bass, *Dicentrarchus labrax*, and the structure of the nucleolar organizer areas in fish. Vopr. Ikhtiologii 812–822.

Arrigo, N., Tuszynski, J.W., Ehrich, D., Gerdes, T., and Alvarez, N. (2009). Evaluating the impact of scoring parameters on the structure of intra-specific genetic variation using RawGeno, an R package for automating AFLP scoring. BMC Bioinformatics *10*, 33.

Ayala, M.D., López-Albors, O., Gil, F., Latorre, R., Vázquez, J.M., García-Alcázar, A., Abellán, E., Ramírez, G., and Moreno, F. (2000). Temperature Effect on Muscle Growth of the Axial Musculature of the Sea Bass (*Dicentrarchus labrax* L.). Anat. Histol. Embryol. *29*, 235–242.

B

Baert, Y., Stukenborg, J.-B., Landreh, M., De Kock, J., Jornvall, H., Soder, O., and Goossens, E. (2015). Derivation and characterization of a cytocompatible scaffold from human testis. Hum. Reprod. *30*, 256–267.

Baerwald, M.R., Meek, M.H., Stephens, M.R., Nagarajan, R.P., Goodbla, A.M., Tomalty, K.M.H., Thorgaard, G.H., May, B., and Nichols, K.M. (2016). Migration-related phenotypic divergence is associated with epigenetic modifications in rainbow trout. Mol. Ecol. *25*, 1785–1800.

Bagni, M. (2005). Cultured Aquatic Species Information Programme. *Dicentrarchus labrax*. Cultured Aquatic Species Information Programme. In FAO Fisheries and Aquaculture Department, (Rome) http://www.fao.org/fishery/culturedspecies/Dicentrarchus_labrax/en

Bahri-Sfar, L., Lemaire, C., Ben Hassine, O.K., and Bonhomme, F. (2000). Fragmentation of sea bass populations in the western and eastern Mediterranean as revealed by microsatellite polymorphism. Proc. Biol. Sci. *267*, 929–935.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME Suite: tools for motif discovery and searching. Nucleic Acids Res. *37*, W202–W208.

Balaban, N.Q., Merrin, J., Chait, R., Kowalik, L., and Leibler, S. (2004). Bacterial Persistence as a Phenotypic Switch. Science *305*, 1622–1625.

Ball, M.P., Li, J.B., Gao, Y., Lee, J.-H., LeProust, E.M., Park, I.-H., Xie, B., Daley, G.Q., and Church, G.M. (2009). Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat. Biotechnol. *27*, 361–368.

Balon, E.K. (2004). About the oldest domesticates among fishes. J. Fish Biol. *65*, 1–27.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res. *21*, 381–395.

Banovich, N.E., Lan, X., McVicker, G., van de Geijn, B., Degner, J.F., Blischak, J.D., Roux, J., Pritchard, J.K., and Gilad, Y. (2014). Methylation QTLs are associated with coordinated changes in transcription factor binding, histone modifications, and gene expression levels. PLoS Genet. *10*, e1004663.

Barnabé, G. (1980). Exposé synoptique des données biologiques sur le loup du bar (*Dicentrarchus labrax* (Linné, 1758)). (Rome: Organisation des nations unies pour l'alimentation et l'agriculture).

Barnabé, G. (1989). L'élevage du loup et de la daurade. In Aquaculture, (Paris: Lavoisier Technique et Documentation), pp. 675–720.

Bauchot, M.-L. (1987). Poisson osseux. In Fiches FAO D'identification Pour Les Besoins de La Pêche. (Rev. 1).

Méditerranée et Mer Noire. Zone de Pêche 37, (Rome: Commission des Communautés Européennes and FAO), pp. 891–1421.

Beaman, J.E., White, C.R., and Seebacher, F. (2016). Evolution of Plasticity: Mechanistic Link between Development and Reversible Acclimation. Trends Ecol. Evol. *31*, 237–249.

Bell, J.T., Pai, A.A., Pickrell, J.K., Gaffney, D.J., Pique-Regi, R., Degner, J.F., Gilad, Y., and Pritchard, J.K. (2011). DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. Genome Biol. *12*, R10.

Benayoun, B.A., Pollina, E.A., and Brunet, A. (2015). Epigenetic regulation of ageing: linking environmental inputs to genomic stability. Nat. Rev. Mol. Cell Biol. *16*, 593–610.

Bender, J., and Fink, G.R. (1995). Epigenetic control of an endogenous gene family is revealed by a novel blue fluorescent mutant of Arabidopsis. Cell *83*, 725–734.

Benhaïm, D., Péan, S., Lucas, G., Blanc, N., Chatain, B., and Bégout, M.-L. (2012). Early life behavioural differences in wild caught and domesticated sea bass (*Dicentrarchus labrax*). Appl. Anim. Behav. Sci. *141*, 79–90.

Benhaïm, D., Bégout, M.-L., Lucas, G., and Chatain, B. (2013). First Insight into Exploration and Cognition in Wild Caught and Domesticated Sea Bass (*Dicentrarchus labrax*) in a Maze. PLOS ONE *8*, e65872.

Berdyshev, G.D., Korotaev, G.K., Boiarskikh, G.V., and Vaniushin, B.F. (1967). Nucleotide composition of DNA and RNA from somatic tissues of humpback and its changes during spawning. Biokhimiia Mosc. Russ. *32*, 988–993.

Berger, S.L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. (2009). An operational definition of epigenetics. Genes Dev. *23*, 781–783.

Bernstein, D.L., Kameswaran, V., Le Lay, J.E., Sheaffer, K.L., and Kaestner, K.H. (2015). The BisPCR2 method for targeted bisulfite sequencing. Epigenetics Chromatin *8*, 27.

Bicskei, B., Taggart, J.B., Glover, K.A., and 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. GSE *48*.

Bilio, M. (2007). Controlled reproduction and domestication in aquaculture – the current state of the art, Part II. Aquac. Eur. 5–23.

Bird, A.P. (1980). DNA methylation and the frequency of CpG in animal DNA. Nucleic Acids Res. *8*, 1499–1504.

Bizuayehu, T.T., and Babiak, I. (2014). MicroRNA in teleost fish. Genome Biol. Evol. *6*, 1911–1937.

Bizuayehu, T.T., Babiak, J., Norberg, B., Fernandes, J.M.O., Johansen, S.D., and Babiak, I. (2012). Sex-biased miRNA expression in Atlantic halibut (*Hippoglossus hippoglossus*) brain and gonads. Sex. Dev. Genet. Mol. Biol. Evol. Endocrinol. Embryol. Pathol. Sex Determ. Differ. *6*, 257–266.

Bizuayehu, T.T., Fernandes, J.M.O., Johansen, S.D., and Babiak, I. (2013). Characterization of Novel Precursor miRNAs Using Next Generation Sequencing and Prediction of miRNA Targets in Atlantic Halibut. PLOS ONE *8*, e61378.

Bizuayehu, T.T., Johansen, S.D., Puvanendran, V., Toften, H., and Babiak, I. (2015). Temperature during early development has long-term effects on microRNA expression in Atlantic cod. BMC Genomics *16*, 305.

Blattler, A., Yao, L., Witt, H., Guo, Y., Nicolet, C.M., Berman, B.P., and Farnham, P.J. (2014). Global loss of DNA methylation uncovers intronic enhancers in genes showing expression changes. Genome Biol. *15*, 1.

Blázquez, M., Carrillo, M., Zanuy, S., and Piferrer, F. (1999). Sex ratios in offspring of sex-reversed sea bass and the relationship between growth and phenotypic sex differentiation. J. Fish Biol. *55*, 916–930.

Blázquez, M., González, A., Papadaki, M., Mylonas, C., and Piferrer, F. (2008). Sex-related changes in estrogen receptors and aromatase gene expression and enzymatic activity during early development and sex differentiation in the European sea bass (*Dicentrarchus labrax*). Gen. Comp. Endocrinol. *158*, 95–101.

Blázquez, M., González, A., Mylonas, C.C., and Piferrer, F. (2011). Cloning and sequence analysis of a vasa homolog in the European sea bass (*Dicentrarchus labrax*): Tissue distribution and mRNA expression levels during early development and sex differentiation. Gen. Comp. Endocrinol. *170*, 322–333.

Bock, C. (2012). Analysing and interpreting DNA methylation data. Nat. Rev. Genet. *13*, 705–719.

Bock, C., Tomazou, E.M., Brinkman, A.B., Müller, F., Simmer, F., Gu, H., Jäger, N., Gnirke, A., Stunnenberg, H.G., and Meissner, A. (2010). Quantitative comparison of genome-wide DNA methylation mapping technologies. Nat. Biotechnol. *28*, 1106–1114.

Bogdanović, O., Smits, A.H., de la Calle Mustienes, E., Tena, J.J., Ford, E., Williams, R., Senanayake, U., Schultz, M.D., Hontelez, S., van Kruijsbergen, I., et al. (2016). Active DNA demethylation at enhancers during the vertebrate phylotypic period. Nat. Genet. *48*, 417–426.

Bonin, A., Bellemain, E., Bronken Eidesen, P., Pompanon, F., Brochmann, C., and Taberlet, P. (2004). How to track and assess genotyping errors in population genetics studies. Mol. Ecol. *13*, 3261–3273.

Bonin, A., Ehrich, D., and Manel, S. (2007). Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. Mol. Ecol. *16*, 3737–3758.

Bossdorf, O., and Zhang, Y. (2011). A truly ecological epigenetics study. Mol. Ecol. *20*, 1572–1574.

Bossdorf, O., Richards, C.L., and Pigliucci, M. (2008). Epigenetics for ecologists. Ecol. Lett. *11*, 106–115.

Brandon, R.N., and Nijhout, H.F. (2006). The Empirical Nonequivalence of Genic and Genotypic Models of Selection: A (Decisive) Refutation of Genic Selectionism and Pluralistic Genic Selectionism. Philos. Sci. *73*, 277–297.

Brenet, F., Moh, M., Funk, P., Feierstein, E., Viale, A.J., Socci, N.D., and Scandura, J.M. (2011). DNA Methylation of the First Exon Is Tightly Linked to Transcriptional Silencing. PLoS ONE *6*, e14524.

Brink, R.A. (1956). A Genetic Change Associated with the R Locus in Maize Which Is Directed and Potentially Reversible. Genetics *41*, 872–889.

Brock, H.W., and Fisher, C.L. (2005). Maintenance of gene expression patterns. Dev. Dyn. *232*, 633–655.

Bull, J.J. (1983). Evolution of sex determining mechanisms (Menlo Park: Benjamin/Cummings Pub. Co.).

Burdge, G.C., Slater-Jefferies, J., Torrens, C., Phillips, E.S., Hanson, M.A., and Lillycrop, K.A. (2007). Dietary protein restriction of pregnant rats in the F0 generation induces altered methylation of hepatic gene promoters in the adult male offspring in the F1 and F2 generations. Br. J. Nutr. *97*, 435–439.

C

Calvanese, V., Lara, E., Kahn, A., and Fraga, M.F. (2009). The role of epigenetics in aging and age-related diseases. Ageing Res. Rev. *8*, 268–276.

Campana, S.E. (2001). Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods. J. Fish Biol. *59*, 197–242.

Campos, C., Valente, L.M.P., and Fernandes, J.M.O. (2012). Molecular evolution of zebrafish *dnmt3* genes and thermal plasticity of their expression during embryonic development. Gene *500*, 93–100.

Campos, C., Valente, L., Conceição, L., Engrola, S., and Fernandes, J. (2013a). Temperature affects methylation of the *myogenin* putative promoter, its expression and muscle cellularity in Senegalese sole larvae. Epigenetics *8*, 389–397.

Campos, C., Valente, L.M.P., Conceição, L.E.C., Engrola, S., Sousa, V., Rocha, E., and Fernandes, J.M.O. (2013b). Incubation temperature induces changes in muscle cellularity and gene expression in Senegalese sole (*Solea senegalensis*). Gene *516*, 209–217.

Campos, C., Sundaram, A.Y., Valente, L.M., Conceição, L.E., Engrola, S., and Fernandes, J.M. (2014). Thermal plasticity of the miRNA transcriptome during Senegalese sole development. BMC Genomics *15*, 525.

Cao, J., Wei, C., Liu, D., Wang, H., Wu, M., Xie, Z., Capellini, T.D., Zhang, L., Zhao, F., Li, L., et al. (2015). DNA methylation Landscape of body size variation in sheep. Sci. Rep. *5*, 13950.

Carlberg, C., and Molnár, F. (2014a). Mechanisms of Gene Regulation (Dordrecht: Springer Netherlands).

Carlberg, C., and Molnár, F. (2014b). The Epigenome. In Mechanisms of Gene Regulation, (Springer Netherlands), pp. 155–167.

Carlson, M. GO.db: A set of annotation maps describing the entire Gene Ontology. R Package Version 322.

Castilho, R., and Ciftci, Y. (2005). Genetic differentiation between close eastern Mediterranean *Dicentrarchus labrax* (L.) populations. J. Fish Biol. *67*, 1746–1752.

Cedar, H., and Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. Nat. Rev. Genet. *10*, 295–304.

Chatain, B., and Chavanne, H. (2009). Genetics of European seabass (*Dicentrarchus labrax* L.). Cah. Agric. *18*, 249–255.

Chatziplis, D., Batargias, C., Tsigenopoulos, C.S., Magoulas, A., Kollias, S., Kotoulas, G., Volckaert, F.A.M., and Haley, C.S. (2007). Mapping quantitative trait loci in European sea bass (*Dicentrarchus labrax*): The BASSMAP pilot study. Aquaculture *272*, *Supplement 1*, S172–S182.

Chen, T., Ueda, Y., Dodge, J.E., Wang, Z., and Li, E. (2003). Establishment and Maintenance of Genomic Methylation Patterns in Mouse Embryonic Stem Cells by Dnmt3a and Dnmt3b. Mol. Cell. Biol. *23*, 5594–5605.

Chen, X., Wang, Z., Tang, S., Zhao, Y., and Zhao, J. (2016). Genome-wide mapping of DNA methylation in Nile Tilapia. Hydrobiologia 1–11.

Cheng, C.-H.C., and Chen, L. (1999). Evolution of an antifreeze glycoprotein. Nature *401*, 443–444.

Choi, M., Lee, J., Le, M.T., Nguyen, D.T., Park, S., Soundrarajan, N., Schachtschneider, K.M., Kim, J., Park, J.-K., Kim, J.-H., et al. (2015). Genome-wide analysis of DNA methylation in pigs using reduced representation bisulfite sequencing. DNA Res. *22*, 343–355.

Choudhuri, S., Cui, Y., and Klaassen, C.D. (2010). Molecular targets of epigenetic regulation and effectors of environmental influences. Toxicol. Appl. Pharmacol. *245*, 378–393.

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

Ci, W., and Liu, J. (2015). Programming and inheritance of parental DNA methylomes in vertebrates. Physiol. Bethesda Md *30*, 63–68.

Claireaux, G. (2006). Effect of temperature on maximum swimming speed and cost of transport in juvenile European sea bass (*Dicentrarchus labrax*). J. Exp. Biol. *209*, 3420–3428.

Claireaux, G., and Lagardère, J.-P. (1999). Influence of temperature, oxygen and salinity on the metabolism of the European sea bass. J. Sea Res. *42*, 157–168.

Colot, V., and Rossignol, J.-L. (1999). Eukaryotic DNA methylation as an evolutionary device. BioEssays *21*, 402–411.

Consuegra, S., and Rodríguez López, C.M. (2016). Epigenetic-induced alterations in sex-ratios in response to climate change: An epigenetic trap? BioEssays News Rev. Mol. Cell. Dev. Biol. doi: 10.1002/bies.201600058.

Coscia, I., and Mariani, S. (2011). Phylogeography and population structure of European sea bass in the north-east Atlantic. Biol. J. Linn. Soc. *104*, 364–377.

Coscia, I., Desmarais, E., Guinand, B., and Mariani, S. (2012). Phylogeography of European sea bass in the north-east Atlantic: a correction and reanalysis of the mitochondrial DNA data from Coscia & Mariani (2011). Biol. J. Linn. Soc. *106*, 455–458.

Costa, F.F. (2008). Non-coding RNAs, epigenetics and complexity. Gene *410*, 9–17.

Couldrey, C., Brauning, R., Bracegirdle, J., Maclean, P., Henderson, H.V., and McEwan, J.C. (2014). Genome-Wide DNA Methylation Patterns and Transcription Analysis in Sheep Muscle. PLoS ONE *9*, e101853.

Coulondre, C., Miller, J.H., Farabaugh, P.J., and Gilbert, W. (1978). Molecular basis of base substitution hotspots in *Escherichia coli*. Nature *274*, 775–780.

Crozier, L.G., and Hutchings, J.A. (2014). Plastic and evolutionary responses to climate change in fish. Evol. Appl. *7*, 68–87.

D

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158.

Darwin, C. (1868). The Variation of Animals and Plants Under Domestication (London: J. Murray).

Davidoff, M.S., Middendorff, R., Müller, D., and Holstein, A.F. (2009). The neuroendocrine Leydig cells and their stem cell progenitors, the pericytes. Adv. Anat. Embryol. Cell Biol. *205*, 1–107.

Daxinger, L., and Whitelaw, E. (2012). Understanding transgenerational epigenetic inheritance via the gametes in mammals. Nat. Rev. Genet. *13*, 153-62

Day, T., and Bonduriansky, R. (2011). A unified approach to the evolutionary consequences of genetic and nongenetic inheritance. Am. Nat. *178*, E18-36.

De Bustos, C., Ramos, E., Young, J.M., Tran, R.K., Menzel, U., Langford, C.F., Eichler, E.E., Hsu, L., Henikoff, S., Dumanski, J.P., et al. (2009). Tissue-specific variation in DNA methylation levels along human chromosome 1. Epigenetics Chromatin *2*, 7.

Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. Genes Dev. *25*, 1010–1022.

Deloffre, L.A.M., Martins, R.S.T., Mylonas, C.C., and Canario, A.V.M. (2009). Alternative transcripts of DMRT1 in the

European sea bass: Expression during gonadal differentiation. Aquaculture *293*, 89–99.

Derks, M.F.L., Schachtschneider, K.M., Madsen, O., Schijlen, E., Verhoeven, K.J.F., and van Oers, K. (2016). Gene and transposable element methylation in great tit (*Parus major*) brain and blood. BMC Genomics *17*, 332.

Devlin, R.H., and Nagahama, Y. (2002). Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. Aquaculture *208*, 191–364.

Diamond, J. (2002). Evolution, consequences and future of plant and animal domestication. Nature *418*, 700–707.

Díaz, N., and Piferrer, F. (2015). Lasting effects of early exposure to temperature on the gonadal transcriptome at the time of sex differentiation in the European sea bass, a fish with mixed genetic and environmental sex determination. BMC Genomics *16*.

Díaz, N., Ribas, L., and Piferrer, F. (2013). The relationship between growth and sex differentiation in the European sea bass (*Dicentrarchus labrax*). Aquaculture *408–409*, 191–202.

Donnelly, S.R., Hawkins, T.E., and Moss, S.E. (1999). A conserved nuclear element with a role in mammalian gene regulation. Hum. Mol. Genet. *8*, 1723–1728.

Dray, S., and Dufour, A.B. (2007). The ade4 package: implementing the duality diagram for ecologists. J. Stat. Softw. *22*, 1–20.

Dufour, V., Cantou, M., and Lecomte, F. (2009). Identification of sea bass (*Dicentrarchus labrax*) nursery areas in the north-western Mediterranean Sea. J. Mar. Biol. Assoc. U. K. *89*, 1367–1374.

E

Edgar, R., Tan, P.P.C., Portales-Casamar, E., and Pavlidis, P. (2014). Meta-analysis of human methylomes reveals stably methylated sequences surrounding CpG islands associated with high gene expression. Epigenetics Chromatin *7*, 28.

van Eijk, K.R., de Jong, S., Boks, M.P., Langeveld, T., Colas, F., Veldink, J.H., de Kovel, C.G., Janson, E., Strengman, E., Langfelder, P., et al. (2012). Genetic analysis of DNA methylation and gene expression levels in whole blood of healthy human subjects. BMC Genomics *13*, 636.

Elango, N., and Yi, S.V. (2008). DNA methylation and structural and functional bimodality of vertebrate promoters. Mol. Biol. Evol. *25*, 1602–1608.

Eschmeyer, W.N., Fricke, R., and van der Laan, R. (2016). Catalog of fishes: Genera, Species, References. http://www.calacademy.org/scientists/projects/catalog-of-fishes.

Excoffier, L., Smouse, P.E., and Quattro, J.M. (1992). Analysis of Molecular Variance Inferred from Metric Distances among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. Genetics *131*, 479–491.

F

Faulk, C., and Dolinoy, D.C. (2011). Timing is everything. Epigenetics *6*, 791–797.

Feidantsis, K., Pörtner, H.O., Lazou, A., Kostoglou, B., and Michaelidis, B. (2009). Metabolic and molecular stress responses of the gilthead seabream *Sparus aurata* during long-term exposure to increasing temperatures. Mar. Biol. *156*, 797–809.

Feil, R., and Fraga, M.F. (2012). Epigenetics and the environment: emerging patterns and implications. Nat. Rev. Genet. *13*, 97-109

Feinberg, A.P., Cui, H., and Ohlsson, R. (2002). DNA methylation and genomic imprinting: insights from cancer into epigenetic mechanisms. Semin. Cancer Biol. *12*, 389–398.

Fernandes, J.M.O., MacKenzie, M.G., Wright, P.A., Steele, S.L., Suzuki, Y., Kinghorn, J.R., and Johnston, I.A. (2006). Myogenin in model pufferfish species: Comparative genomic analysis and thermal plasticity of expression during early development. Comp. Biochem. Physiol. Part D Genomics Proteomics *1*, 35–45.

Fernandino, J.I., Hattori, R.S., Kishii, A., Strüssmann, C.A., and Somoza, G.M. (2012). The cortisol and androgen pathways cross talk in high temperature-induced masculinization: the 11β-hydroxysteroid dehydrogenase as a key enzyme. Endocrinology *153*, 6003–6011.

Fernandino, J.I., Hattori, R.S., Moreno Acosta, O.D., Strüssmann, C.A., and Somoza, G.M. (2013). Environmental stress-induced testis differentiation: androgen as a by-product of cortisol inactivation. Gen. Comp. Endocrinol. *192*, 36–44.

Ferraresso, S., Milan, M., Pellizzari, C., Vitulo, N., Reinhardt, R., Canario, A.V., Patarnello, T., and Bargelloni, L. (2010). Development of an oligo DNA microarray for the European sea bass and its application to expression profiling of jaw deformity. BMC Genomics *11*, 354.

Finn, R.N., and Kristoffersen, B.A. (2007). Vertebrate Vitellogenin Gene Duplication in Relation to the "3R Hypothesis": Correlation to the Pelagic Egg and the Oceanic Radiation of Teleosts. PLoS ONE *2*, e169.

Foust, C.M., Preite, V., Schrey, A.W., Alvarez, M., Robertson, M.H., Verhoeven, K.J.F., and Richards, C.L. (2016). Genetic and epigenetic differences associated with environmental gradients in replicate populations of two salt marsh perennials. Mol. Ecol. *25*, 1639–1652.

Freyhof, J., and Kottelat, M. (2008). *Dicentrarchus labrax*: Freyhof, J. & Kottelat, M.: The IUCN Red List of Threatened Species 2008: e.T135606A4159287.

Fuller, M.F. (2004). The Encyclopedia of Farm Animal Nutrition (CABI). doi. 10.1079/9780851993690.0000.

Furrow, R.E., and Feldman, M.W. (2014). Genetic Variation and the Evolution of Epigenetic Regulation. Evolution *68*, 673–683.

G

Gabriel, W., Luttbeg, B., Sih, A., and Tollrian, R. (2005). Environmental tolerance, heterogeneity, and the evolution of reversible plastic responses. Am. Nat. *166*, 339–353.

Galay-Burgos, M., Gealy, C., Navarro-Martín, L., Piferrer, F., Zanuy, S., and Sweeney, G.E. (2006). Cloning of the promoter from the gonadal aromatase gene of the European sea bass and identification of single nucleotide polymorphisms. Comp. Biochem. Physiol. A. Mol. Integr. Physiol. *145*, 47–53.

Gautier, M., Moazami-Goudarzi, K., Levéziel, H., Parinello, H., Grohs, C., Rialle, S., Kowalczyk, R., and Flori, L. (2016). Deciphering the Wisent Demographic and Adaptive Histories from Individual Whole-Genome Sequences. Mol. Biol. Evol. doi: 10.1093/molbev/msw144.

Gavery, M.R., and Roberts, S.B. (2010). DNA methylation patterns provide insight into epigenetic regulation in the Pacific oyster (*Crassostrea gigas*). BMC Genomics *11*, 483.

Gavery, M.R., and Roberts, S.B. (2013). Predominant intragenic methylation is associated with gene expression characteristics in a bivalve mollusc. PeerJ *1*, e215. doi 10.7717/peerj.215

Gavery, M.R., and Roberts, S.B. (2014). A context dependent role for DNA methylation in bivalves. Brief. Funct. Genomics *13*, 217–222.

Gavrilov, L.A., and Gavrilova, N.S. (2002). Evolutionary theories of aging and longevity. ScientificWorldJournal *2*, 339–356.

Geay, F., Ferraresso, S., Zambonino-Infante, J.L., Bargelloni, L., Quentel, C., Vandeputte, M., Kaushik, S., Cahu, C.L., and Mazurais, D. (2011). Effects of the total replacement of fish-based diet with plant-based diet on the hepatic transcriptome of two European sea bass (*Dicentrarchus labrax*) half-sibfamilies showing different growth rates with the plant-based diet. BMC Genomics *12*, 522.

Genicot, S., Rentier-Delrue, F., Edwards, D., VanBeeumen, J., and Gerday, C. (1996). Trypsin and trypsinogen from an Antarctic fish: molecular basis of cold adaptation. Biochim. Biophys. Acta *1298*, 45–57.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., et al. (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. *5*, R80.

Georgakopoulou, E., Sfakianakis, D.G., Kouttouki, S., Divanach, P., Kentouri, M., and Koumoundouros, G. (2007a). The influence of temperature during early life on phenotypic expression at later ontogenetic stages in sea bass. J. Fish Biol. *70*, 278–291.

Georgakopoulou, E., Angelopoulou, A., Kaspiris, P., Divanach, P., and Koumoundouros, G. (2007b). Temperature effects on cranial deformities in European sea bass, *Dicentrarchus labrax* (L.). J. Appl. Ichthyol. *23*, 99–103.

Ghosh, S., and Zhou, Z. (2015). Chapter 14 - Epigenetics of physiological and premature aging A2 - Huang, Suming. In Epigenetic Gene Expression and Regulation, M.D. Litt, and C.A. Blakey, eds. (Oxford: Academic Press), pp. 313–338.

Gibbs, W.W. (2014). Biomarkers and ageing: The clock-watcher. Nature *508*, 168–170.

Gibbs, J.R., van der Brug, M.P., Hernandez, D.G., Traynor, B.J., Nalls, M.A., Lai, S.-L., Arepalli, S., Dillman, A., Rafferty, I.P., Troncoso, J., et al. (2010). Abundant Quantitative Trait Loci

Exist for DNA Methylation and Gene Expression in Human Brain. PLoS Genet. *6*.

Gilbert, S.F., and Epel, D. (2008). Ecological Developmental Biology: Integrating Epigenetics, Medicine, and Evolution (Sunderland, MA: Sinauer Associates).

Goll, M.G., and Bestor, T.H. (2005). Eukaryotic cytosine methyltransferases. Annu. Rev. Biochem. *74*, 481–514.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics *27*, 1017–1018.

Greer, E.L., Maures, T.J., Ucar, D., Hauswirth, A.G., Mancini, E., Lim, J.P., Benayoun, B.A., Shi, Y., and Brunet, A. (2011). Transgenerational epigenetic inheritance of longevity in *Caenorhabditis elegans*. Nature *479*, 365–371.

Grosjean, H. (2013). Nucleic Acids Are Not Boring Long Polymers of Only Four Types of Nucleotides: A Guided Tour (Austin: Landes Bioscience).

Gross, M.R. (1998). One species with two biologies: Atlantic salmon (*Salmo salar*) in the wild and in aquaculture. Can. J. Fish. Aquat. Sci. *55*, 131–144.

GTEx Consortium (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science *348*, 648–660.

Gu, H., Smith, Z.D., Bock, C., Boyle, P., Gnirke, A., and Meissner, A. (2011). Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. Nat. Protoc. *6*, 468–481.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77–88.

Guiguen, Y., Fostier, A., Piferrer, F., and Chang, C.-F. (2010). Ovarian aromatase and estrogens: A pivotal role for gonadal sex

differentiation and sex change in fish. Gen. Comp. Endocrinol. *165*, 352–366.

Gutierrez de Paula, T., de Almeida, F.L.A., Carani, F.R., Vechetti-Júnior, I.J., Padovani, C.R., Salomão, R.A.S., Mareco, E.A., dos Santos, V.B., and Dal-Pai-Silva, M. (2014). Rearing temperature induces changes in muscle growth and gene expression in juvenile pacu (*Piaractus mesopotamicus*). Comp. Biochem. Physiol. B Biochem. Mol. Biol. *169*, 31–37.

Gutierrez-Arcelus, M., Lappalainen, T., Montgomery, S.B., Buil, A., Ongen, H., Yurovsky, A., Bryois, J., Giger, T., Romano, L., Planchon, A., et al. (2013). Passive and active DNA methylation and the interplay with genetic variation in gene regulation. Elife *2*, e00523.

H

Haller, R.G., Henriksson, K.G., Jorfeldt, L., Hultman, E., Wibom, R., Sahlin, K., Areskog, N.H., Gunder, M., Ayyad, K., and Blomqvist, C.G. (1991). Deficiency of skeletal muscle succinate dehydrogenase and aconitase. Pathophysiology of exercise in a novel human muscle oxidative defect. J. Clin. Invest. *88*, 1197–1206.

Hannon, E., Spiers, H., Viana, J., Pidsley, R., Burrage, J., Murphy, T.M., Troakes, C., Turecki, G., O'Donovan, M.C., Schalkwyk, L.C., et al. (2016). Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci. Nat. Neurosci. *19*, 48–54.

Hanson, M.A., and Skinner, M.K. (2016). Developmental origins of epigenetic transgenerational inheritance. Environ. Epigenetics *2*, doi: 10.1093/eep/dvw002.

Hashimoto, S., Ogoshi, K., Sasaki, A., Abe, J., Qu, W., Nakatani, Y., Ahsan, B., Oshima, K., Shand, F.H.W., Ametani, A., et al. (2013). Coordinated Changes in DNA Methylation in Antigen-Specific Memory CD4 T Cells. J. Immunol. *190*, 4076–4091.

Hayami, Y., Iida, S., Nakazawa, N., Hanamura, I., Kato, M., Komatsu, H., Miura, I., Dave, B.J., Sanger, W.G., Lim, B., et al. (2003). Inactivation of the *E3/LAPTm5* gene by chromosomal rearrangement and DNA methylation in human multiple myeloma. Leukemia *17*, 1650–1657.

Hegarty, M.J., Batstone, T., Barker, G.L., Edwards, K.J., Abbott, R.J., and Hiscock, S.J. (2011). Nonadditive changes to cytosine methylation as a consequence of hybridization and genome duplication in *Senecio* (Asteraceae). Mol. Ecol. *20*, 105–113.

Heijmans, B.T., Tobi, E.W., Stein, A.D., Putter, H., Blauw, G.J., Susser, E.S., Slagboom, P.E., and Lumey, L.H. (2008). Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc. Natl. Acad. Sci. *105*, 17046–17049.

Hemmer, H. (1990). Domestication: The Decline of Environmental Appreciation (Cambridge: Cambridge University Press).

Herman, J.J., Spencer, H.G., Donohue, K., and Sultan, S.E. (2014). How Stable "should" Epigenetic Modifications Be? Insights from Adaptive Plasticity and Bet Hedging. Evolution *68*, 632–643.

Hernando-Herraez, I., Prado-Martinez, J., Garg, P., Fernandez-Callejo, M., Heyn, H., Hvilsom, C., Navarro, A., Esteller, M., Sharp, A.J., and Marques-Bonet, T. (2013). Dynamics of DNA Methylation in Recent Human and Great Ape Evolution. PLOS Genet *9*, e1003763.

Hernando-Herraez, I., Heyn, H., Fernandez-Callejo, M., Vidal, E., Fernandez-Bellon, H., Prado-Martinez, J., Sharp, A.J., Esteller, M., and Marques-Bonet, T. (2015). The interplay between DNA methylation and sequence divergence in recent human evolution. Nucleic Acids Res. *43*, 8204–8214.

Herrera, C.M., and Bazaga, P. (2010). Epigenetic differentiation and relationship to adaptive genetic divergence in discrete populations of the violet *Viola cazorlensis*. New Phytol. *187*, 867–876.

Herrera, C.M., and Bazaga, P. (2013). Epigenetic correlates of plant phenotypic plasticity: DNA methylation differs between prickly and nonprickly leaves in heterophyllous *Ilex aquifolium* (Aquifoliaceae) trees. Bot. J. Linn. Soc. *171*, 441–452.

Heule, C., Salzburger, W., and Bohne, A. (2014). Genetics of Sexual Development: An Evolutionary Playground for Fish. Genetics *196*, 579–591.

Heyn, H., Li, N., Ferreira, H.J., Moran, S., Pisano, D.G., Gomez, A., Diez, J., Sanchez-Mut, J.V., Setien, F., Carmona, F.J., et al. (2012). Distinct DNA methylomes of newborns and centenarians. Proc. Natl. Acad. Sci. *109*, 10522–10527.

Hidalgo, F., Alliot, E., and Thebault, H. (1987). Influence of water temperature on food intake, food efficiency and gross composition of juvenile sea bass, *Dicentrarchus labrax*. Aquaculture *64*, 199–207.

Hillen, J., Coscia, I., and Volckaert, F. (2014). Aquatrace species leaflet. European sea bass (*Dicentrarchus labrax*). Available at https://aquatrace.eu/documents/80305/142567/european+sea+bass+leaflet.pdf. Accessed 14 Jul 2016.

Horowitz, M. (2014). Heat acclimation, epigenetics, and cytoprotection memory. Compr. Physiol. *4*, 199–230.

Horvath, S. (2013). DNA methylation age of human tissues and cell types. Genome Biol. *14*, R115.

Hosomichi, K., Mitsunaga, S., Nagasaki, H., and Inoue, I. (2014). A Bead-based Normalization for Uniform Sequencing depth (BeNUS) protocol for multi-samples sequencing exemplified by HLA-B. BMC Genomics *15*, 645.

Hu, S., Wan, J., Su, Y., Song, Q., Zeng, Y., Nguyen, H.N., Shin, J., Cox, E., Rho, H.S., Woodard, C., et al. (2013). DNA methylation presents distinct binding sites for human transcription factors. eLife 2:e00726.

Huber, W., Carey, V.J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B.S., Bravo, H.C., Davis, S., Gatto, L., Girke, T., et al.

(2015). Orchestrating high-throughput genomic analysis with Bioconductor. Nat. Methods *12*, 115–121.

Huidobro, C., Fernandez, A.F., and Fraga, M.F. (2013). Aging epigenetics: Causes and consequences. Mol. Aspects Med. *34*, 765–781.

Hunt, B.M., Hoefling, K., and Cheng, C.-H.C. (2003). Annual warming episodes in seawater temperatures in McMurdo Sound in relationship to endogenous ice in notothenioid fish. Antarct. Sci. *15*, 333–338.

Hurst, L.D., Sachenkova, O., Daub, C., Forrest, A.R., and Huminiecki, L. (2014). A simple metric of promoter architecture robustly predicts expression breadth of human genes suggesting that most transcription factors are positive regulators. Genome Biol. *15*, 413.

I

Ibeagha-Awemu, E.M., and Zhao, X. (2015). Epigenetic marks: regulators of livestock phenotypes and conceivable sources of missing variation in livestock improvement programs. Front. Genet. *6* doi: 10.3389/fgene.2015.00302.

Illingworth, R.S., and Bird, A.P. (2009). CpG islands – "A rough guide." FEBS Lett. *583*, 1713–1720.

IPCC (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (Geneva, Switzerland).

Irizarry, R.A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., Cui, H., Gabo, K., Rongione, M., Webster, M., et al. (2009). The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat. Genet. *41*, 178–186.

Issa, J.-P. (2003). Age-related epigenetic changes and the immune system. Clin. Immunol. Orlando Fla *109*, 103–108.

Issa, J.-P. (2014). Aging and epigenetic drift: a vicious cycle. J. Clin. Invest. *124*, 24–29.

Ivell, R., and Holstein, A.-F. (2012). The Fate of the Male Germ Cell (Springer Science & Business Media).

J

Jabbari, K., Cacciò, S., Païs de Barros, J.P., Desgrès, J., and Bernardi, G. (1997). Evolutionary changes in CpG and methylation levels in the genome of vertebrates. Gene *205*, 109–118.

Jablonka, E. (2005). Evolution in four dimensions: genetic, epigenetic, behavioral, and symbolic variation in the history of life (Cambridge, Mass: MIT Press).

Jablonka, E., and Raz, G. (2009). Transgenerational Epigenetic Inheritance: Prevalence, Mechanisms, and Implications for the Study of Heredity and Evolution. Q. Rev. Biol. *84*, 131–176.

Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat. Genet. *33 Suppl*, 245–254.

Janowitz Koch, I., Clark, M.M., Thompson, M.J., Deere-Machemer, K.A., Wang, J., Duarte, L., Gnanadesikan, G.E., McCoy, E.L., Rubbi, L., Stahler, D.R., et al. (2016). The concerted impact of domestication and transposon insertions on methylation patterns between dogs and grey wolves. Mol. Ecol. *25*, 1838–1855.

Janssen, K., Chavanne, H., Berentsen, P., and Komen, H. (2015). European seabass (*Dicentrarchus labrax*)–Current status of selective breeding in Europe.

Jensen, M.K., Madsen, S.S., and Kristiansen, K. (1998). Osmoregulation and salinity effects on the expression and

activity of Na+,K(+)-ATPase in the gills of European sea bass, *Dicentrarchus labrax* (L.). J. Exp. Zool. *282*, 290–300.

Jiang, L., Zhang, J., Wang, J.-J., Wang, L., Zhang, L., Li, G., Yang, X., Ma, X., Sun, X., Cai, J., et al. (2013). Sperm, but Not Oocyte, DNA Methylome Is Inherited by Zebrafish Early Embryos. Cell *153*, 773–784.

Jin, L., Jiang, Z., Xia, Y., Lou, P. 'er, Chen, L., Wang, H., Bai, L., Xie, Y., Liu, Y., Li, W., et al. (2014). Genome-wide DNA methylation changes in skeletal muscle between young and middle-aged pigs. BMC Genomics *15*, 653.

Jing, J., Wu, J., Liu, W., Xiong, S., Ma, W., Zhang, J., Wang, W., Gui, J.-F., and Mei, J. (2014). Sex-biased miRNAs in gonad and their potential roles for testis development in yellow catfish. PloS One *9*, e107946.

Johannes, F., Porcher, E., Teixeira, F.K., Saliba-Colombani, V., Simon, M., Agier, N., Bulski, A., Albuisson, J., Heredia, F., Audigier, P., et al. (2009). Assessing the impact of transgenerational epigenetic variation on complex traits. PLoS Genet. *5*, e1000530.

Johnston, I.A. (2006). Environment and plasticity of myogenesis in teleost fish. J. Exp. Biol. *209*, 2249–2264.

Johnston, I.A., Bower, N.I., and Macqueen, D.J. (2011). Growth and the regulation of myotomal muscle mass in teleost fish. J. Exp. Biol. *214*, 1617–1628.

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat. Rev. Genet. *13*, 484–492.

Jones, O.R., Scheuerlein, A., Salguero-Gómez, R., Camarda, C.G., Schaible, R., Casper, B.B., Dahlgren, J.P., Ehrlén, J., García, M.B., Menges, E.S., et al. (2013). Diversity of ageing across the tree of life. Nature *505*, 169–173.

Jonsson, B., and Jonsson, N. (2014). Early environment influences later performance in fishes: effects of early experiences. J. Fish Biol. *85*, 151–188.

Jung, M., and Pfeifer, G.P. (2015). Aging and DNA methylation. BMC Biol. *13*, 7.

K

Kalcheim, C. (2011). Regulation of Trunk Myogenesis by the Neural Crest: A New Facet of Neural Crest-Somite Interactions. Dev. Cell *21*, 187–188.

Kawashima, T., and Berger, F. (2014). Epigenetic reprogramming in plant sexual reproduction. Nat. Rev. Genet. *15*, 613–624.

Keller, C., and Bühler, M. (2013). Chromatin-associated ncRNA activities. Chromosome Res. *21*, 627–641.

Kelly, S.A., Panhuis, T.M., and Stoehr, A.M. (2011). Phenotypic Plasticity: Molecular Mechanisms and Adaptive Significance. In Comprehensive Physiology, (John Wiley & Sons, Inc.).

Khan, A.R., Enjalbert, J., Marsollier, A.-C., Rousselet, A., Goldringer, I., and Vitte, C. (2013). Vernalization treatment induces site-specific DNA hypermethylation at the VERNALIZATION-A1 (VRN-A1) locus in hexaploid winter wheat. BMC Plant Biol. *13*, 209.

Khavari, D.A., Sen, G.L., and Rinn, J.L. (2010). DNA methylation and epigenetic control of cellular differentiation. Cell Cycle *9*, 3880–3883.

Kiefer, J.C. (2007). Epigenetics in development. Dev. Dyn. *236*, 1144–1156.

Kikuchi, K., and Hamaguchi, S. (2013). Novel sex-determining genes in fish and sex chromosome evolution: Novel Sex-Determining Genes in Fish. Dev. Dyn. *242*, 339–353.

Kim, J., Bhattacharjee, R., Khalyfa, A., Kheirandish-Gozal, L., Capdevila, O.S., Wang, Y., and Gozal, D. (2012). DNA Methylation in Inflammatory Genes among Children with

Obstructive Sleep Apnea. Am. J. Respir. Crit. Care Med. *185*, 330–338.

Klughammer, J., Datlinger, P., Printz, D., Sheffield, N.C., Farlik, M., Hadler, J., Fritsch, G., and Bock, C. (2015). Differential DNA Methylation Analysis without a Reference Genome. Cell Rep. *13*, 2621–2633.

Korbie, D., Lin, E., Wall, D., Nair, S.S., Stirzaker, C., Clark, S.J., and Trau, M. (2015). Multiplex bisulfite PCR resequencing of clinical FFPE DNA. Clin. Epigenetics *7*, 28.

Kottelat, M., and Freyhof, J. (2007). Handbook of European freshwater fishes (Publications Kottelat).

Koumoundouros, G., Divanach, P., Anezaki, L., and Kentouri, M. (2001). Temperature-induced ontogenetic plasticity in sea bass (*Dicentrarchus labrax*). Mar. Biol. *139*, 817–830.

Koumoundouros, G., Pavlidis, M., Anezaki, L., Kokkari, C., Sterioti, A., Divanach, P., and Kentouri, M. (2002). Temperature sex determination in the European sea bass, *Dicentrarchus labrax* (L., 1758) (Teleostei, Perciformes, Moronidae): Critical sensitive ontogenetic phase. J. Exp. Zool. *292*, 573–579.

Koumoundouros, G., Ashton, C., Sfakianakis, D.G., Divanach, P., Kentouri, M., Anthwal, N., and Stickland, N.C. (2009). Thermally induced phenotypic plasticity of swimming performance in European sea bass *Dicentrarchus labrax* juveniles. J. Fish Biol. *74*, 1309–1322.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693–705.

Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics *27*, 1571–1572.

Kubista, M., Sindelka, R., Tichopad, A., Bergkvist, A., Lindh, D., and Forooran, A. (2007). The prime technique—real-time PCR data analysis. GIT Lab. J. *9*, 33–35.

Kuhl, H., Tine, M., Beck, A., Timmermann, B., Kodira, C., and Reinhardt, R. (2011a). Directed sequencing and annotation of three *Dicentrarchus labrax* L. chromosomes by applying Sanger- and pyrosequencing technologies on pooled DNA of comparatively mapped BAC clones. Genomics *98*, 202–212.

Kuhl, H., Tine, M., Hecht, J., Knaust, F., and Reinhardt, R. (2011b). Analysis of single nucleotide polymorphisms in three chromosomes of European sea bass *Dicentrarchus labrax*. Comp. Biochem. Physiol. Part D Genomics Proteomics *6*, 70–75.

Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M.J., et al. (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317–330.

Kyprianou, T.-D., Pörtner, H.O., Anestis, A., Kostoglou, B., Feidantsis, K., and Michaelidis, B. (2010). Metabolic and molecular stress responses of gilthead seam bream *Sparus aurata* during exposure to low ambient temperature: an analysis of mechanisms underlying the winter syndrome. J. Comp. Physiol. B *180*, 1005–1018.

L

Labbé, C., Robles, V., and Herraez, M.P. (2016). Epigenetics in fish gametes and early embryo. Aquaculture. http://dx.doi.org/10.1016/j.aquaculture.2016.07.026

Lance, V.A. (2009). Is regulation of aromatase expression in reptiles the key to understanding temperature-dependent sex determination? J. Exp. Zool. Part Ecol. Genet. Physiol. *311A*, 314–322.

Larsen, F., Gundersen, G., Lopez, R., and Prydz, H. (1992). CpG islands as gene markers in the human genome. Genomics *13*, 1095–1107.

Larson, G., and Fuller, D.Q. (2014). The Evolution of Animal Domestication. Annu. Rev. Ecol. Evol. Syst. *45*, 115–136.

Laurent, L., Wong, E., Li, G., Huynh, T., Tsirigos, A., Ong, C.T., Low, H.M., Kin Sung, K.W., Rigoutsos, I., Loring, J., et al. (2010). Dynamic changes in the human methylome during differentiation. Genome Res. *20*, 320–331.

Law, J.A., and Jacobsen, S.E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat. Rev. Genet. *11*, 204–220.

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for Computing and Annotating Genomic Ranges. PLoS Comput. Biol. *9*, e1003118.

Lea, A.J., Altmann, J., Alberts, S.C., and Tung, J. (2016). Resource base influences genome-wide DNA methylation levels in wild baboons (*Papio cynocephalus*). Mol. Ecol. *25*, 1681–1696.

Levy, S.F., Ziv, N., and Siegal, M.L. (2012). Bet hedging in yeast by heterogeneous, age-correlated expression of a stress protectant. PLoS Biol. *10*, e1001325.

Li, L.-C., and Dahiya, R. (2002). MethPrimer: designing primers for methylation PCRs. Bioinformatics *18*, 1427–1431.

Li, M., and Leatherland, J.F. (2013). The implications for aquaculture practice of epigenomic programming of components of the endocrine system of teleostean embryos: lessons learned from mammalian studies. Fish Fish. *14*, 528–553.

Li, C.G., Wang, H., Chen, H.J., Zhao, Y., Fu, P.S., and Ji, X.S. (2014). Differential expression analysis of genes involved in high-temperature induced sex differentiation in Nile tilapia. Comp. Biochem. Physiol. B Biochem. Mol. Biol. *177–178*, 36–45.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence

Alignment/Map format and SAMtools. Bioinforma. Oxf. Engl. *25*, 2078–2079.

Li, H., Chen, D., and Zhang, J. (2012). Analysis of Intron Sequence Features Associated with Transcriptional Regulation in Human Genes. PLOS ONE *7*, e46784.

Li, S., Garrett-Bakelman, F.E., Akalin, A., Zumbo, P., Levine, R., To, B.L., Lewis, I.D., Brown, A.L., D'Andrea, R.J., Melnick, A., et al. (2013). An optimized algorithm for detecting and annotating regional differential methylation. BMC Bioinformatics *14*, S10.

Lillycrop, K.A., Phillips, E.S., Jackson, A.A., Hanson, M.A., and Burdge, G.C. (2005). Dietary protein restriction of pregnant rats induces and folic acid supplementation prevents epigenetic modification of hepatic gene expression in the offspring. J. Nutr. *135*, 1382–1386.

Lillycrop, K.A., Phillips, E.S., Torrens, C., Hanson, M.A., Jackson, A.A., and Burdge, G.C. (2008). Feeding pregnant rats a protein-restricted diet persistently alters the methylation of specific cytosines in the hepatic PPARα promoter of the offspring. Br. J. Nutr. *100*, 278–282.

Lim, H.N., and van Oudenaarden, A. (2007). A multistep epigenetic switch enables the stable inheritance of DNA methylation states. Nat. Genet. *39*, 269–275.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.-M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315–322.

Liu, Y., Siegmund, K.D., Laird, P.W., Berman, B.P., and others (2012). Bis-SNP: Combined DNA methylation and SNP calling for Bisulfite-seq data. Genome Biol *13*, R61.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods San Diego Calif *25*, 402–408.

Lloret, J., Shulman, G., and Love, R.M. (2013). Condition and Health Indicators of Exploited Marine Fishes (Chichester, West Sussex: John Wiley & Sons).

Lokk, K., Modhukur, V., Rajashekar, B., Märtens, K., Mägi, R., Kolde, R., Koltšina, M., Nilsson, T.K., Vilo, J., Salumets, A., et al. (2014). DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. Genome Biol. *15*, r54.

Long, H.K., Sims, D., Heger, A., Blackledge, N.P., Kutter, C., Wright, M.L., Grützner, F., Odom, D.T., Patient, R., Ponting, C.P., et al. (2013). Epigenetic conservation at gene regulatory elements revealed by non-methylated DNA profiling in seven vertebrates. Elife *2*, e00348.

López-Albors, O., Ayala, M.D., Gil, F., García-Alcázar, A., Abellán, E., Latorre, R., Ramírez-Zarzosa, G., and Vázquez, J.M. (2003). Early temperature effects on muscle growth dynamics and histochemical profile of muscle fibres of sea bass *Dicentrarchus labrax* L., during larval and juvenile stages. Aquaculture *220*, 385–406.

Lou, S., Lee, H.-M., Qin, H., Li, J.-W., Gao, Z., Liu, X., Chan, L.L., KL Lam, V., So, W.-Y., Wang, Y., et al. (2014). Whole-genome bisulfite sequencing of multiple individuals reveals complementary roles of promoter and gene body methylation in transcriptional regulation. Genome Biol. *15*, 408.

Loukovitis, D., Ioannidi, B., Chatziplis, D., Kotoulas, G., Magoulas, A., and Tsigenopoulos, C.S. (2015). Loss of genetic variation in Greek hatchery populations of the European sea bass (*Dicentrarchus labrax* L.) as revealed by microsatellite DNA analysis. Mediterr. Mar. Sci. *16*. doi: 10.12681/mms.1033

Louro, B., Passos, A.L.S., Souche, E.L., Tsigenopoulos, C., Beck, A., Lagnel, J., Bonhomme, F., Cancela, L., Cerdà, J., Clark, M.S., et al. (2010). Gilthead sea bream (*Sparus auratus*) and European sea bass (*Dicentrarchus labrax*) expressed sequence tags: Characterization, tissue-specific expression and gene markers. Mar. Genomics *3*, 179–191.

Louro, B., Power, D.M., and Canario, A.V.M. (2014). Advances in European sea bass genomics and future perspectives. Mar. Genomics *18*, 71–75.

Lowdon, R.F., Jang, H.S., and Wang, T. (2016). Evolution of Epigenetic Regulation in Vertebrate Genomes. Trends Genet. TIG *32*, 269–283.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251–260.

Lühmann, L.M., Knorr, C., Hörstgen-Schwark, G., and Wessels, S. (2012). First evidence for family-specific QTL for temperature-dependent sex reversal in Nile tilapia (*Oreochromis niloticus*). Sex. Dev. Genet. Mol. Biol. Evol. Endocrinol. Embryol. Pathol. Sex Determ. Differ. *6*, 247–256.

Lyko, F., Foret, S., Kucharski, R., Wolf, S., Falckenhayn, C., and Maleszka, R. (2010). The Honey Bee Epigenomes: Differential Methylation of Brain DNA in Queens and Workers. PLoS Biol. *8*, e1000506.

Lynch, M., and O'Hely, M. Captive breeding and the genetic fitness of natural populations. Conserv. Genet. *2*, 363–378.

M

Maamar, H., Raj, A., and Dubnau, D. (2007). Noise in gene expression determines cell fate in Bacillus subtilis. Science *317*, 526–529.

Macqueen, D.J., Robb, D.H.F., Olsen, T., Melstveit, L., Paxton, C.G.M., and Johnston, I.A. (2008). Temperature until the "eyed stage" of embryogenesis programmes the growth trajectory and muscle phenotype of adult Atlantic salmon. Biol. Lett. *4*, 294–298.

Magnanou, E., Klopp, C., Noirot, C., Besseau, L., and Falcón, J. (2014). Generation and characterization of the sea bass

Dicentrarchus labrax brain and liver transcriptomes. Gene *544*, 56–66.

Maiese, K., Li, F., Chong, Z.Z., and Shang, Y.C. (2008). The Wnt signaling pathway: aging gracefully as a protectionist? Pharmacol. Ther. *118*, 58–81.

Majewski, J., and Ott, J. (2002). Distribution and characterization of regulatory elements in the human genome. Genome Res. *12*, 1827–1836.

Maleszka, R., Mason, P.H., and Barron, A.B. (2014). Epigenomics and the concept of degeneracy in biological systems. Brief. Funct. Genomics *13*, 191–202.

Manchado, M., Infante, C., Asensio, E., Crespo, A., Zuasti, E., and Cañavate, J.P. (2008). Molecular characterization and gene expression of six trypsinogens in the flatfish Senegalese sole (*Solea senegalensis* Kaup) during larval development and in tissues. Comp. Biochem. Physiol. B Biochem. Mol. Biol. *149*, 334–344.

Marco-Sola, S., Sammeth, M., Guigó, R., and Ribeca, P. (2012). The GEM mapper: fast, accurate and versatile alignment by filtration. Nat. Methods *9*, 1185–1188.

Marcus, J., Dupree, J.L., and Popko, B. (2002). Myelin-associated glycoprotein and myelin galactolipids stabilize developing axo-glial interactions. J. Cell Biol. *156*, 567–577.

Marsh, A.G., and Pasqualone, A.A. (2014). DNA methylation and temperature stress in an Antarctic polychaete, *Spiophanes tcherniai*. Front. Physiol. *5*, 173.

Marttila, S., Kananen, L., Häyrynen, S., Jylhävä, J., Nevalainen, T., Hervonen, A., Jylhä, M., Nykter, M., and Hurme, M. (2015). Ageing-associated changes in the human DNA methylome: genomic locations and effects on gene expression. BMC Genomics *16*, 179.

Massault, C., Hellemans, B., Louro, B., Batargias, C., Van, H., Canario, A., Volckaert, F.A.M., Bovenhuis, H., Haley, C., and

De, K. (2010). QTL for body weight, morphometric traits and stress response in European sea bass *Dicentrarchus labrax*. Anim. Genet. *41*, 337–345.

Masser, D.R., Berg, A.S., and Freeman, W.M. (2013). Focused, high accuracy 5-methylcytosine quantitation with base resolution by benchtop next-generation sequencing. Epigenetics Chromatin *6*, 33.

Massicotte, R., Whitelaw, E., and Angers, B. (2011). DNA methylation: A source of random variation in natural populations. Epigenetics *6*, 421–427.

Mathelier, A., Zhao, X., Zhang, A.W., Parcy, F., Worsley-Hunt, R., Arenillas, D.J., Buchman, S., Chen, C., Chou, A., Ienasescu, H., et al. (2013). JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic Acids Res. gkt997. doi: 10.1093/nar/gkt997.

Matsumoto, Y., Buemio, A., Chu, R., Vafaee, M., and Crews, D. (2013). Epigenetic Control of Gonadal Aromatase (*cyp19a1*) in Temperature-Dependent Sex Determination of Red-Eared Slider Turtles. PLoS ONE *8*, e63599.

Mattick, J.S. (2007). A new paradigm for developmental biology. J. Exp. Biol. *210*, 1526–1547.

Mattick, J.S., and Makunin, I.V. (2006). Non-coding RNA. Hum. Mol. Genet. *15 Spec No 1*, R17-29.

McCully, K.K., Argov, Z., Boden, B.P., Brown, R.L., Bank, W.J., and Chance, B. (1988). Detection of muscle injury in humans with 31-P magnetic resonance spectroscopy. Muscle Nerve *11*, 212–216.

Medrano, M., Herrera, C.M., and Bazaga, P. (2014). Epigenetic variation predicts regional and local intraspecific functional diversity in a perennial herb. Mol. Ecol. *23*, 4926–4938.

Meehl, G., Stocker, T., Collins, W.D., Friedlingstein, P., Gaye, A.T., Gregory, J.M., Kitoh, A., Knutti, R., Murphy, J.M., Noda, A., et al. (2007). Global Climate Projections. In Climate Change

2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, (Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press).

Melé, M., Ferreira, P.G., Reverter, F., DeLuca, D.S., Monlong, J., Sammeth, M., Young, T.R., Goldmann, J.M., Pervouchine, D.D., Sullivan, T.J., et al. (2015). Human genomics. The human transcriptome across tissues and individuals. Science *348*, 660–665.

Messerschmidt, D.M., Knowles, B.B., and Solter, D. (2014). DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. Genes Dev. *28*, 812–828.

Migaud, H., Bell, G., Cabrita, E., McAndrew, B., Davie, A., Bobe, J., Herráez, M. p., and Carrillo, M. (2013). Gamete quality and broodstock management in temperate fish. Rev. Aquac. *5*, S194–S223.

Millot, S., Péan, S., Chatain, B., and Bégout, M.-L. (2011). Self-feeding behavior changes induced by a first and a second generation of domestication or selection for growth in the European sea bass, *Dicentrarchus labrax*. Aquat. Living Resour. *24*, 53–61.

Mills, B.B., McBride, C.M., and Riddle, N.C. (2015). Chapter 8 - Epigenetic inheritance A2 - Huang, Suming. In Epigenetic Gene Expression and Regulation, M.D. Litt, and C.A. Blakey, eds. (Oxford: Academic Press), pp. 183–208.

Miner, B.G., Sultan, S.E., Morgan, S.G., Padilla, D.K., and Relyea, R.A. (2005). Ecological consequences of phenotypic plasticity. Trends Ecol. Evol. *20*, 685–692.

Mitter, K., Kotoulas, G., Magoulas, A., Mulero, V., Sepulcre, P., Figueras, A., Novoa, B., and Sarropoulou, E. (2009). Evaluation of candidate reference genes for QPCR during ontogenesis and of immune-relevant tissues of European seabass (*Dicentrarchus*

labrax). Comp. Biochem. Physiol. B Biochem. Mol. Biol. *153*, 340–347.

Moghadam, H., Mørkøre, T., and Robinson, N. (2015). Epigenetics—Potential for Programming Fish for Aquaculture? J. Mar. Sci. Eng. *3*, 175–192.

Monserrat, J.M., Martínez, P.E., Geracitano, L.A., Amado, L.L., Martins, C.M.G., Pinho, G.L.L., Chaves, I.S., Ferreira-Cravo, M., Ventura-Lima, J., and Bianchini, A. (2007). Pollution biomarkers in estuarine animals: critical review and new perspectives. Comp. Biochem. Physiol. Toxicol. Pharmacol. CBP *146*, 221–234.

Montague, M.J., Li, G., Gandolfi, B., Khan, R., Aken, B.L., Searle, S.M.J., Minx, P., Hillier, L.W., Koboldt, D.C., Davis, B.W., et al. (2014). Comparative analysis of the domestic cat genome reveals genetic signatures underlying feline biology and domestication. Proc. Natl. Acad. Sci. *111*, 17230–17235.

Moore, L.D., Le, T., and Fan, G. (2013). DNA methylation and its basic function. Neuropsychopharmacology *38*, 23–38.

Morán, P., and Pérez-Figueroa, A. (2011). Methylation changes associated with early maturation stages in the Atlantic salmon. BMC Genet. *12*, 86.

Moreira, I.S., Peres, H., Couto, A., Enes, P., and Oliva-Teles, A. (2008). Temperature and dietary carbohydrate level effects on performance and metabolic utilisation of diets in European sea bass (*Dicentrarchus labrax*) juveniles. Aquaculture *274*, 153–160.

Morgan, H.D., Sutherland, H.G., Martin, D.I., and Whitelaw, E. (1999). Epigenetic inheritance at the agouti locus in the mouse. Nat. Genet. *23*, 314–318.

Morimoto, R.I., and Milarski, K.L. (1990). Expression and Function of Vertebrate hsp70 Genes. Cold Spring Harb. Monogr. Arch. *19*, 323–359.

Morretti, A. (1999). Manual on Hatchery Production of Seabass and Gilthead Seabream (Food & Agriculture Org.).

Mugatroyd, C., Wu, Y., Bockmühl, Y., and Spengler, D. (2010). The Janus face of DNA methylation in aging. Aging *2*, 107–110.

Muller, F.L., Lustgarten, M.S., Jang, Y., Richardson, A., and Van Remmen, H. (2007). Trends in oxidative aging theories. Free Radic. Biol. Med. *43*, 477–503.

Muñoz-Najar, U., and Sedivy, J.M. (2011). Epigenetic control of aging. Antioxid. Redox Signal. *14*, 241–259.

Murray, H ., Perez-Casanova, J ., Gallant, J ., Johnson, S ., and Douglas, S. (2004). Trypsinogen expression during the development of the exocrine pancreas in winter flounder (*Pleuronectes americanus*). Comp. Biochem. Physiol. A. Mol. Integr. Physiol. *138*, 53–59.

N

Naciri, M., Lemaire, C., Borsa, P., and Bonhomme, F. (1999). Genetic study of the Atlantic/Mediterranean transition in sea bass (*Dicentrarchus labrax*). J. Hered. *90*, 591–596.

Nagai, T., Takehana, Y., Hamaguchi, S., and Sakaizumi, M. (2008). Identification of the sex-determining locus in the Thai medaka, *Oryzias minutillus*. Cytogenet. Genome Res. *121*, 137–142.

Nätt, D., Rubin, C.-J., Wright, D., Johnsson, M., Beltéky, J., Andersson, L., and Jensen, P. (2012). Heritable genome-wide variation of gene expression and promoter methylation between wild and domesticated chickens. BMC Genomics *13*, 59.

Navarro-Martín, L., Blázquez, M., Viñas, J., Joly, S., and Piferrer, F. (2009). Balancing the effects of rearing at low temperature during early development on sex ratios, growth and maturation in the European sea bass (*Dicentrarchus labrax*). Aquaculture *296*, 347–358.

Navarro-Martín, L., Viñas, J., Ribas, L., Díaz, N., Gutiérrez, A., Di Croce, L., and Piferrer, F. (2011). DNA Methylation of the Gonadal Aromatase (*cyp19a*) Promoter Is Involved in Temperature-Dependent Sex Ratio Shifts in the European Sea Bass. PLoS Genet *7*, e1002447.

Ndlovu, 'Matladi N., Denis, H., and Fuks, F. (2011). Exposing the DNA methylome iceberg. Trends Biochem. Sci *36*, 381-387.

Nicol, J.W., Helt, G.A., Blanchard, S.G., Raja, A., and Loraine, A.E. (2009). The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. Bioinformatics *25*, 2730–2731.

Noble, D. (2015). Conrad Waddington and the origin of epigenetics. J. Exp. Biol. *218*, 816–818.

Novel, P., Porta, J.M., Porta, J., Béjar, J., and Alvarez, M.C. (2010). PCR multiplex tool with 10 microsatellites for the European seabass (*Dicentrarchus labrax*) — Applications in genetic differentiation of populations and parental assignment. Aquaculture *308, Supplement 1*, S34–S38.

Nutrition National Research Council (1984). Nutrient Requirements of Beef Cattle (Washington, DC: National Academies).

O

O'Dea, R.E., Noble, D.W.A., Johnson, S.L., Hesselson, D., and Nakagawa, S. (2016). The role of non-genetic inheritance in evolutionary rescue: epigenetic buffering, heritable bet hedging and epigenetic traps. Environ. Epigenetics *2*, dvv014.

O'Doherty, A.M., MacHugh, D.E., Spillane, C., and Magee, D.A. (2015). Genomic imprinting effects on complex traits in domesticated animal species. Livest. Genomics 156.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., et al. (2016). vegan: Community Ecology Package.

Oliveros, J.C. (2007). Venny. An interactive tool for comparing lists with Venn's diagrams.

Ong, C.-T., and Corces, V.G. (2011). Enhancer function: new insights into the regulation of tissue-specific gene expression. Nat. Rev. Genet. *12*, 283–293.

Ospina-Álvarez, N., and Piferrer, F. (2008). Temperature-Dependent Sex Determination in Fish Revisited: Prevalence, a Single Sex Ratio Response Pattern, and Possible Effects of Climate Change. PLoS ONE *3*, e2837.

P

Pagès, H. (2016). BSgenome: Infrastructure for Biostrings-based genome data packages and support for efficient SNP representation.

Palaiokostas, C., Bekaert, M., Taggart, J.B., Gharbi, K., McAndrew, B.J., Chatain, B., Penman, D.J., and Vandeputte, M. (2015). A new SNP-based vision of the genetics of sex determination in European sea bass (*Dicentrarchus labrax*). Genet. Sel. Evol. *47*, 68.

Pallavicini, A., Canapa, A., Barucca, M., Alföldi, J., Biscotti, M.A., Buonocore, F., De Moro, G., Di Palma, F., Fausto, A.M., Forconi, M., et al. (2013). Analysis of the transcriptome of the Indonesian coelacanth *Latimeria menadoensis*. BMC Genomics *14*, 538.

Pan, J.-B., Hu, S.-C., Shi, D., Cai, M.-C., Li, Y.-B., Zou, Q., and Ji, Z.-L. (2013). PaGenBase: A Pattern Gene Database for the Global and Dynamic Understanding of Gene Function. PLOS ONE *8*, e80747.

Pannetier, M., Fabre, S., Batista, F., Kocer, A., Renault, L., Jolivet, G., Mandon-Pépin, B., Cotinot, C., Veitia, R., and Pailhoux, E. (2006). FOXL2 activates *P450 aromatase* gene transcription: towards a better characterization of the early steps of mammalian ovarian development. J. Mol. Endocrinol. *36*, 399–413.

Park, J., Xu, K., Park, T., and Yi, S.V. (2011). What are the determinants of gene expression levels and breadths in the human genome? Hum. Mol. Genet. doi:10.1093/hmg/ddr436.

Parrott, B.B., Kohno, S., Cloy-McCoy, J.A., and Guillette, L.J. (2014). Differential incubation temperatures result in dimorphic DNA methylation patterning of the *SOX9* and aromatase promoters in gonads of alligator (*Alligator mississippiensis*) embryos. Biol. Reprod. *90*, 2.

Paun, O., Bateman, R.M., Fay, M.F., Hedren, M., Civeyrel, L., and Chase, M.W. (2010). Stable Epigenetic Effects Impact Adaptation in Allopolyploid Orchids (*Dactylorhiza*: Orchidaceae). Mol. Biol. Evol. *27*, 2465–2473.

Pavlidis, M., Koumoundouros, G., Sterioti, A., Somarakis, S., Divanach, P., Kentouri, M., and others (2000). Evidence of temperature-dependent sex determination in the European sea bass (*Dicentrarchus labrax* L.). J. Exp. Zool. *287*, 225–232.

Penman, D.J., and Piferrer, F. (2008). Fish Gonadogenesis. Part I: Genetic and Environmental Mechanisms of Sex Determination. Rev. Fish. Sci. *16*, 16–34.

Perez-Casanova, J.C., Murray, H.M., Gallant, J.W., Ross, N.W., Douglas, S.E., and Johnson, S.C. (2006). Development of the digestive capacity in larvae of haddock (*Melanogrammus aeglefinus*) and Atlantic cod (*Gadus morhua*). Aquaculture *251*, 377–401.

Pérez-Figueroa, A. (2013). *msap* : a tool for the statistical analysis of methylation-sensitive amplified polymorphism data. Mol. Ecol. Resour. *13*, 522–527.

Periago, M.J., Ayala, M.D., López-Albors, O., Abdel, I., Martínez, C., García-Alcázar, A., Ros, G., and Gil, F. (2005). Muscle cellularity and flesh quality of wild and farmed sea bass, *Dicentrarchus labrax* L. Aquaculture *249*, 175–188.

Perkins, J.R., Dawes, J.M., McMahon, S.B., Bennett, D.L., Orengo, C., and Kohl, M. (2012). ReadqPCR and NormqPCR: R packages for the reading, quality checking and normalisation of

RT-qPCR quantification cycle (Cq) data. BMC Genomics *13*, 296.

Person-Le Ruyet, J., Mahé, K., Le Bayon, N., and Le Delliou, H. (2004). Effects of temperature on growth and metabolism in a Mediterranean population of European sea bass, *Dicentrarchus labrax*. Aquaculture *237*, 269–280.

Pfeifer, G.P. (2016). Epigenetics: An elusive DNA base in mammals. Nature *532*, 319–320.

Pickett, G.D., and Pawson, M.G. (1994). Sea Bass: Biology (Springer Science & Business Media).

Piferrer, F. (2013). Epigenetics of sex determination and gonadogenesis: Epigenetics and Sex Determination. Dev. Dyn. *242*, 360–370.

Piferrer, F. (2016). Altered sex ratios in response to climate change—Who will fall into the (epigenetic) trap? BioEssays doi: 10.1002/bies.201600172.

Piferrer, F., and Blázquez, M. (2005). Aromatase distribution and regulation in fish. Fish Physiol. Biochem. *31*, 215-226.

Piferrer, F., and Guiguen, Y. (2008). Fish Gonadogenesis. Part II: Molecular Biology and Genomics of Sex Differentiation. Rev. Fish. Sci. *16*(51), 35–55.

Piferrer, F., Blázquez, M., Navarro, L., and González, A. (2005). Genetic, endocrine, and environmental components of sex determination and differentiation in the European sea bass (*Dicentrarchus labrax* L.). Gen. Comp. Endocrinol. *142*, 102–110.

Piferrer, F., Ribas, L., and Díaz, N. (2012). Genomic Approaches to Study Genetic and Environmental Influences on Fish Sex Determination and Differentiation. Mar. Biotechnol. *14*, 591–604.

Pittman, K., Yúfera, M., Pavlidis, M., Geffen, A.J., Koven, W., Ribeiro, L., Zambonino-Infante, J.L., and Tandler, A. (2013).

Fantastically plastic: fish larvae equipped for a new world. Rev. Aquac. *5*, S224–S267.

Polanowski, A.M., Robbins, J., Chandler, D., and Jarman, S.N. (2014). Epigenetic estimation of age in humpback whales. Mol. Ecol. Resour. *14*, 976–987.

Poonlaphdecha, S., Pepey, E., Canonne, M., de Verdal, H., Baroiller, J.-F., and D'Cotta, H. (2013). Temperature induced-masculinisation in the Nile tilapia causes rapid up-regulation of both *dmrt1* and *amh* expressions. Gen. Comp. Endocrinol. *193*, 234–242.

Pope, E.C., Ellis, R.P., Scolamacchia, M., Scolding, J.W.S., Keay, A., Chingombe, P., Shields, R.J., Wilcox, R., Speirs, D.C., Wilson, R.W., et al. (2014). European sea bass, *Dicentrarchus labrax*, in a changing ocean. Biogeosciences *11*, 2519–2530.

Pörtner, H.O., and Peck, M.A. (2010). Climate change effects on fishes and fisheries: towards a cause-and-effect understanding. J. Fish Biol. *77*, 1745–1779.

Potok, M.E., Nix, D.A., Parnell, T.J., and Cairns, B.R. (2013). Reprogramming the Maternal Zebrafish Genome after Fertilization to Match the Paternal Methylation Pattern. Cell *153*, 759–772.

Probst, A.V., Dunleavy, E., and Almouzni, G. (2009). Epigenetic inheritance during the cell cycle. Nat. Rev. Mol. Cell Biol. *10*, 192–206.

Q

Quéré, N., Guinand, B., Kuhl, H., Reinhardt, R., Bonhomme, F., and Desmarais, E. (2010). Genomic sequences and genetic differentiation at associated tandem repeat markers in growth hormone, somatolactin and insulin-like growth factor-1 genes of the sea bass, *Dicentrarchus labrax*. Aquat. Living Resour. *23*, 285–296.

Quéré, N., Desmarais, E., Tsigenopoulos, C.S., Belkhir, K., Bonhomme, F., and Guinand, B. (2012). Gene flow at major transitional areas in sea bass (*Dicentrarchus labrax*) and the possible emergence of a hybrid swarm. Ecol. Evol. *2*, 3061–3078.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

R

R Core Team (2015). R: A Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing).

Ramsköld, D., Wang, E.T., Burge, C.B., and Sandberg, R. (2009). An Abundance of Ubiquitously Expressed Genes Revealed by Tissue Transcriptome Sequence Data. PLOS Comput Biol *5*, e1000598.

Rassoulzadegan, M., Grandjean, V., Gounon, P., Vincent, S., Gillot, I., and Cuzin, F. (2006). RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. Nature *441*, 469–474.

Ratel, D., Ravanat, J.-L., Berger, F., and Wion, D. (2006). N6-methyladenine: the other methylated base of DNA. Bioessays *28*, 309–315.

Rechache, N.S., Wang, Y., Stevenson, H.S., Killian, J.K., Edelman, D.C., Merino, M., Zhang, L., Nilubol, N., Stratakis, C.A., Meltzer, P.S., et al. (2012). DNA Methylation Profiling Identifies Global Methylation Differences and Markers of Adrenocortical Tumors. J. Clin. Endocrinol. Metab. *97*, E1004–E1013.

Reinders, J., Wulff, B.B.H., Mirouze, M., Marí-Ordóñez, A., Dapp, M., Rozhon, W., Bucher, E., Theiler, G., and Paszkowski, J. (2009). Compromised stability of DNA methylation and transposon immobilization in mosaic *Arabidopsis* epigenomes. Genes Dev. *23*, 939–950.

Reyna-López, G.E., Simpson, J., and Ruíz-Herrera, J. (1997). Differences in DNA methylation patterns are detectable during the dimorphic transition of fungi by amplification of restriction polymorphisms. Mol. Gen. Genet. MGG *253*, 703–710.

Reznick, D., Ghalambor, C., and Nunney, L. (2002). The evolution of senescence in fish. Mech. Ageing Dev. *123*, 773–789.

Roberge, C., Einum, S., Guderley, H., and Bernatchez, L. (2005). Rapid parallel evolutionary changes of gene transcription profiles in farmed Atlantic salmon. Mol. Ecol. *15*, 9–20.

Robinson, M.D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. *11*, R25.

Robinson, M.D., Kahraman, A., Law, C.W., Lindsay, H., Nowicka, M., Weber, L.M., and Zhou, X. (2014). Statistical methods for detecting differentially methylated loci and regions. Bioinforma. Comput. Biol. *5*, 324.

Rocha, A., Gómez, A., Zanuy, S., Cerdá-Reverter, J.M., and Carrillo, M. (2007). Molecular characterization of two sea bass gonadotropin receptors: cDNA cloning, expression analysis, and functional activity. Mol. Cell. Endocrinol. *272*, 63–76.

Rocha, A., Zanuy, S., Carrillo, M., and Gómez, A. (2009). Seasonal changes in gonadal expression of gonadotropin receptors, steroidogenic acute regulatory protein and steroidogenic enzymes in the European sea bass. Gen. Comp. Endocrinol. *162*, 265–275.

Rocha, A., Zanuy, S., and Gómez, A. (2016). Conserved Anti-Müllerian Hormone: Anti-Müllerian Hormone Type-2 Receptor Specific Interaction and Intracellular Signaling in Teleosts. Biol. Reprod. *94*, 141.

Rohland, N., and Reich, D. (2012). Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Res. *22*, 939–946.

Romanov, G.A., and Vanyushin, B.F. (1981). Methylation of reiterated sequences in mammalian DNAs. Effects of the tissue type, age, malignancy and hormonal induction. Biochim. Biophys. Acta *653*, 204–218.

Rose, N.R., and Klose, R.J. (2014). Understanding the relationship between DNA methylation and histone lysine methylation. Biochim. Biophys. Acta BBA - Gene Regul. Mech. *1839*, 1362–1372.

Rountree, M.R., and Selker, E.U. (1997). DNA methylation inhibits elongation but not initiation of transcription in *Neurospora crassa*. Genes Dev. *11*, 2383–2395.

RStudio Team (2015). RStudio: Integrated Development Environment for R (Boston, MA: RStudio, Inc.).

Russell, N.R., Fish, J.D., and Wootton, R.J. (1996). Feeding and growth of juvenile sea bass: the effect of ration and temperature on growth rate and efficiency. J. Fish Biol. *49*, 206–220.

Russo, V.E., Martienssen, R.A., Riggs, A.D., and others (1996). Epigenetic mechanisms of gene regulation. (Cold Spring Harbor Laboratory Press).

S

Saillant, E., Fostier, A., Menu, B., Haffray, P., and Chatain, B. (2001). Sexual growth dimorphism in sea bass *Dicentrarchus labrax*. Aquaculture *202*, 371–387.

Saillant, E., Fostier, A., Haffray, P., Menu, B., Thimonier, J., and Chatain, B. (2002). Temperature effects and genotype-temperature interactions on sex determination in the European sea bass (*Dicentrarchus labrax* L.). J. Exp. Zool. *292*, 494–505.

Saillant, E., Chatain, B., Menu, B., Fauvel, C., Vidal, M.O., and Fostier, A. (2003). Sexual differentiation and juvenile intersexuality in the European sea bass (*Dicentrarchus labrax*). J. Zool. *260*, 53–63.

Sánchez-Villagra, M.R., Geiger, M., and Schneider, R.A. (2016). The taming of the neural crest: a developmental perspective on the origins of morphological covariation in domesticated mammals. Open Sci. *3*, 160107.

Sandoval, J., Heyn, H., Moran, S., Serra-Musach, J., Pujana, M.A., Bibikova, M., and Esteller, M. (2011). Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. Epigenetics *6*, 692–702.

Sarropoulou, E., Galindo-Villegas, J., García-Alcázar, A., Kasapidis, P., and Mulero, V. (2012). Characterization of European Sea Bass Transcripts by RNA SEQ After Oral Vaccine Against V. anguillarum. Mar. Biotechnol. *14*, 634–642.

Saxonov, S., Berg, P., and Brutlag, D.L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc. Natl. Acad. Sci. U. S. A. *103*, 1412–1417.

Schlichting, C.D., and Smith, H. (2002). Phenotypic plasticity: linking molecular mechanisms with evolutionary outcomes. Evol. Ecol. *16*, 189–211.

Schlichting, C.D., and Wund, M.A. (2014). Phenotypic Plasticity and Epigenetic Marking: An Assessment of Evidence for Genetic Accommodation. Evolution *68*, 656–672.

Schoppee Bortz, P.D., and Wamhoff, B.R. (2011). Chromatin immunoprecipitation (ChIP): revisiting the efficacy of sample preparation, sonication, quantification of sheared DNA, and analysis via PCR. PloS One *6*, e26015.

Schübeler, D. (2015). Function and information content of DNA methylation. Nature *517*, 321–326.

Schug, J., Schuller, W.-P., Kappen, C., Salbaum, J.M., Bucan, M., and Stoeckert, C.J. (2005). Promoter features related to tissue specificity as measured by Shannon entropy. Genome Biol. *6*, R33.

Schulz, B., Eckstein, R.L., and Durka, W. (2013). Scoring and analysis of methylation-sensitive amplification polymorphisms for epigenetic population studies. Mol. Ecol. Resour. *13*, 642–653.

Schulz, B., Eckstein, R.L., and Durka, W. (2014). Epigenetic variation reflects dynamic habitat conditions in a rare floodplain herb. Mol. Ecol. *23*, 3523–3537.

Seebacher, F., Beaman, J., and Little, A.G. (2014). Regulation of thermal acclimation varies between generations of the short-lived mosquitofish that developed in different environmental conditions. Funct. Ecol. *28*, 137–148.

Ser, J.R., Roberts, R.B., and Kocher, T.D. (2010). Multiple interacting loci control sex determination in lake Malawi cichlid fish. Evol. Int. J. Org. Evol. *64*, 486–501.

Serrana, D.G. de la, Vieira, V.L.A., Andree, K.B., Darias, M., Estévez, A., Gisbert, E., and Johnston, I.A. (2012). Development Temperature Has Persistent Effects on Muscle Growth Responses in Gilthead Sea Bream. PLOS ONE *7*, e51884.

Shao, C., Li, Q., Chen, S., Zhang, P., Lian, J., Hu, Q., Sun, B., Jin, L., Liu, S., Wang, Z., et al. (2014). Epigenetic modification and inheritance in sexual reversal of fish. Genome Res. *24*, 604–615.

Shi, Y.Y., Yan, W.Y., Huang, Z.Y., Wang, Z.L., Wu, X.B., and Zeng, Z.J. (2013). Genomewide analysis indicates that queen larvae have lower methylation levels in the honey bee (*Apis mellifera*). Naturwissenschaften *100*, 193–197.

Shimoda, N., Izawa, T., Yoshizawa, A., Yokoi, H., Kikuchi, Y., and Hashimoto, N. (2014). Decrease in cytosine methylation at CpG island shores and increase in DNA fragmentation during zebrafish aging. AGE *36*, 103–115.

Shoemaker, R., Deng, J., Wang, W., and Zhang, K. (2010). Allele-specific methylation is prevalent and is contributed by CpG-SNPs in the human genome. Genome Res. *20*, 883–889.

Simonet, N.G., Reyes, M., Nardocci, G., Molina, A., and Alvarez, M. (2013). Epigenetic regulation of the ribosomal cistron seasonally modulates enrichment of H2A. Z and H2A. Zub in response to different environmental inputs in carp (*Cyprinus carpio*). Epigenetics Chromatin *6*, 22.

Skjærven, K.H., Hamre, K., Penglase, S., Finn, R.N., and Olsvik, P.A. (2014). Thermal stress alters expression of genes involved in one carbon and DNA methylation pathways in Atlantic cod embryos. Comp. Biochem. Physiol. A. Mol. Integr. Physiol. *173*, 17–27.

Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. Nat. Rev. Genet. *14*, 204–220.

Smith, T.A., Martin, M.D., Nguyen, M., and Mendelson, T.C. (2016). Epigenetic divergence as a potential first step in darter speciation. Mol. Ecol. *25*, 1883–1894.

Souche, E.L., Hellemans, B., Babbucci, M., MacAoidh, E., Guinand, B., Bargelloni, L., Chistiakov, D.A., Patarnello, T., Bonhomme, F., Martinsohn, J.T., et al. (2015). Range-wide population structure of European sea bass *Dicentrarchus labrax*. Biol. J. Linn. Soc. *116*, 86–105.

Stadler, M.B., Murr, R., Burger, L., Ivanek, R., Lienert, F., Schöler, A., Erik van Nimwegen, Wirbelauer, C., Oakeley, E.J., Gaidatzis, D., et al. (2011). DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature *480*, 490–495.

Stanley, C.E., and Kulathinal, R.J. (2016). Genomic signatures of domestication on neurogenetic genes in *Drosophila melanogaster*. BMC Evol. Biol. *16*, 6.

Stone, A., Zotenko, E., Locke, W.J., Korbie, D., Millar, E.K.A., Pidsley, R., Stirzaker, C., Graham, P., Trau, M., Musgrove, E.A., et al. (2015). DNA methylation of oestrogen-regulated enhancers defines endocrine sensitivity in breast cancer. Nat. Commun. *6*, 7758.

Storey, J.D., Bass, A., Dabney, A., and Robinson, D. (2015). qvalue: Q-value estimation for false discovery rate control. R Package Version 222.

Straussman, R., Nejman, D., Roberts, D., Steinfeld, I., Blum, B., Benvenisty, N., Simon, I., Yakhini, Z., and Cedar, H. (2009). Developmental programming of CpG island methylation profiles in the human genome. Nat. Struct. Mol. Biol. *16*, 564–571.

Stumpf, M.P.H., Laidlaw, Z., and Jansen, V.A.A. (2002). Herpes viruses hedge their bets. Proc. Natl. Acad. Sci. U. S. A. *99*, 15234–15237.

Su, J., Shao, X., Liu, H., Liu, S., Wu, Q., and Zhang, Y. (2012). Genome-wide dynamic changes of DNA methylation of repetitive elements in human embryonic stem cells and fetal fibroblasts. Genomics *99*, 10–17.

Sultan, S.E. (1995). Phenotypic plasticity and plant adaptation. Acta Bot. Neerlandica *44*, 363–383.

Sultan, S.E. (2007). Development in context: the timely emergence of eco-devo. Trends Ecol. Evol. *22*, 575–582.

Sun, L.-X., Wang, Y.-Y., Zhao, Y., Wang, H., Li, N., and Ji, X.S. (2016). Global DNA Methylation Changes in Nile Tilapia Gonads during High Temperature-Induced Masculinization. PLOS ONE *11*, e0158483.

Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. PLOS ONE *6*, e21800.

T

Takehana, Y., Matsuda, M., Myosho, T., Suster, M.L., Kawakami, K., Shin-I, T., Kohara, Y., Kuroki, Y., Toyoda, A., Fujiyama, A., et al. (2014). Co-option of Sox3 as the male-determining factor on the Y chromosome in the fish *Oryzias dancena*. Nat. Commun. *5*, 4157.

Tanaka, H., and Baba, T. (2005). Gene expression in spermiogenesis. Cell. Mol. Life Sci. CMLS *62*, 344–354.

Tazi, J., and Bird, A. (1990). Alternative chromatin structure at CpG islands. Cell *60*, 909–920.

Teh, A.L., Pan, H., Chen, L., Ong, M.-L., Dogra, S., Wong, J., MacIsaac, J.L., Mah, S.M., McEwen, L.M., Saw, S.-M., et al. (2014). The effect of genotype and in utero environment on interindividual variation in neonate DNA methylomes. Genome Res. *24*, 1064–1074.

Teletchea, F. (2015). Domestication of Marine Fish Species: Update and Perspectives. J. Mar. Sci. Eng. *3*, 1227–1243.

Teletchea, F., and Fontaine, P. (2014). Levels of domestication in fish: implications for the sustainable future of aquaculture. Fish Fish. *15*, 181–195.

Tine, M., Kuhl, H., Gagnaire, P.-A., Louro, B., Desmarais, E., Martins, R.S.T., Hecht, J., Knaust, F., Belkhir, K., Klages, S., et al. (2014). European sea bass genome and its variation provide insights into adaptation to euryhalinity and speciation. Nat. Commun. *5*, 5770.

Tomazou, E.M., Sheffield, N.C., Schmidl, C., Schuster, M., Schönegger, A., Datlinger, P., Kubicek, S., Bock, C., and Kovar, H. (2015). Epigenome mapping reveals distinct modes of gene regulation and widespread enhancer reprogramming by the oncogenic fusion protein EWS-FLI1. Cell Rep. *10*, 1082–1095.

Toraño, E.G., García, M.G., Fernández-Morera, J.L., Niño-García, P., and Fernández, A.F. (2016). The Impact of External Factors on the Epigenome: In Utero and over Lifetime. BioMed Res. Int. BioMed Res. Int. *2016*, *2016*, e2568635.

Triantaphyllopoulos, K.A., Ikonomopoulos, I., and Bannister, A.J. (2016). Epigenetics and inheritance of phenotype variation in livestock. Epigenetics Chromatin *9*, 31.

Trucchi, E., Mazzarella, A.B., Gilfillan, G.D., Lorenzo, M.T., Schönswetter, P., and Paun, O. (2016). BsRADseq: screening

DNA methylation in natural populations of non-model species. Mol. Ecol. *25*, 1697–1713.

Turner, B.M. (2005). Reading signals on the nucleosome with a new nomenclature for modified histones. Nat. Struct. Mol. Biol. *12*, 110–112.

Turner, B.M. (2007). Defining an epigenetic code. Nat. Cell Biol. *9*, 2–6.

Turner, B.M. (2009). Epigenetic responses to environmental change and their evolutionary implications. Philos. Trans. R. Soc. B Biol. Sci. *364*, 3403–3418.

Turner, B.M. (2011). Environmental sensing by chromatin: An epigenetic contribution to evolutionary change. FEBS Lett. *585*, 2032–2040.

U

Unoki, M., and Nakamura, Y. (2003). Methylation at CpG islands in intron 1 of *EGR2* confers enhancer-like activity. FEBS Lett. *554*, 67–72.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012). Primer3—new capabilities and interfaces. Nucleic Acids Res. *40*, e115.

V

Valente, L.M.P., Moutou, K.A., Conceição, L.E.C., Engrola, S., Fernandes, J.M.O., and Johnston, I.A. (2013). What determines growth potential and juvenile quality of farmed fish species? Rev. Aquac. *5*, S168–S193.

Valenzuela, N. (2008). Sexual development and the evolution of sex determination. Sex. Dev. Genet. Mol. Biol. Evol. Endocrinol. Embryol. Pathol. Sex Determ. Differ. *2*, 64–72.

Valenzuela, N., and Lance, V. (2004). Temperature-dependent sex determination in vertebrates (Smithsonian Books Washington, DC).

Van Ho, A.T., Hayashi, S., Bröhl, D., Auradé, F., Rattenbach, R., and Relaix, F. (2011). Neural Crest Cell Lineage Restricts Skeletal Muscle Progenitor Cell Differentiation through Neuregulin1-ErbB3 Signaling. Dev. Cell *21*, 273–287.

Vandeputte, M., Mauger, S., and Dupont-Nivet, M. (2006). An evaluation of allowing for mismatches as a way to manage genotyping errors in parentage assignment by exclusion. Mol. Ecol. Notes *6*, 265–267.

Vandeputte, M., Dupont-Nivet, M., Chavanne, H., and Chatain, B. (2007). A polygenic hypothesis for sex determination in the European sea bass *Dicentrarchus labrax*. Genetics *176*, 1049–1057.

Vandeputte, M., Quillet, E., and Chatain, B. (2012). Are sex ratios in wild European sea bass (*Dicentrarchus labrax*) populations biased? Aquat. Living Resour. *25*, 77–81.

Vandeputte, M., Garouste, R., Dupont-Nivet, M., Haffray, P., Vergnet, A., Chavanne, H., Laureau, S., Ron, T.B., Pagelson, G., Mazorra, C., et al. (2014). Multi-site evaluation of the rearing performances of 5 wild populations of European sea bass (*Dicentrarchus labrax*). Aquaculture *424–425*, 239–248.

Varriale, A. (2014). DNA Methylation, Epigenetics, and Evolution in Vertebrates: Facts and Challenges. Int. J. Evol. Biol. *2014*, e475981.

Varriale, A., and Bernardi, G. (2006). DNA methylation and body temperature in fishes. Gene *385*, 111–121.

Veilleux, H.D., Ryu, T., Donelson, J.M., van Herwerden, L., Seridi, L., Ghosheh, Y., Berumen, M.L., Leggat, W., Ravasi, T., and Munday, P.L. (2015). Molecular processes of transgenerational acclimation to a warming ocean. Nat. Clim. Change *5*, 1074–1078.

Venney, C.J., Johansson, M.L., and Heath, D.D. (2016). Inbreeding effects on gene-specific DNA methylation among tissues of Chinook salmon. Mol. Ecol. *18*, 4521-4533.

Verhoeven, K.J.F., Jansen, J.J., Van Dijk, P.J., and Biere, A. (2010). Stress-induced DNA methylation changes and their heritability in asexual dandelions. New Phytol. *185*, 1108–1118.

Vickaryous, M.K., and Hall, B.K. (2006). Human cell type diversity, evolution, development, and classification with special reference to cells derived from the neural crest. Biol. Rev. Camb. Philos. Soc. *81*, 425–455.

Vinagre, C., Ferreira, T., Matos, L., Costa, M.J., and Cabral, H.N. (2009). Latitudinal gradients in growth and spawning of sea bass, *Dicentrarchus labrax*, and their relationship with temperature and photoperiod. Estuar. Coast. Shelf Sci. *81*, 375–380.

Viñas, J., and Piferrer, F. (2008). Stage-Specific Gene Expression During Fish Spermatogenesis as Determined by Laser-Capture Microdissection and Quantitative-PCR in Sea Bass (*Dicentrarchus labrax*) Gonads. Biol. Reprod. *79*, 738–747.

Vladutiu, G.D., and Heffner, R.R. (2000). Succinate dehydrogenase deficiency. Arch. Pathol. Lab. Med. *124*, 1755–1758.

W

Waddington, C.H. (1956). Genetic Assimilation of the Bithorax Phenotype. Evolution *10*, 1–13.

Wan, J., Oliver, V.F., Wang, G., Zhu, H., Zack, D.J., Merbs, S.L., and Qian, J. (2015). Characterization of tissue-specific differential DNA methylation suggests distinct modes of positive and negative gene expression regulation. BMC Genomics *16*, 49.

Wang, D.-S., Kobayashi, T., Zhou, L.-Y., Paul-Prasanth, B., Ijiri, S., Sakai, F., Okubo, K., Morohashi, K., and Nagahama, Y. (2007). Foxl2 up-regulates aromatase gene transcription in a female-specific manner by binding to the promoter as well as

interacting with ad4 binding protein/steroidogenic factor 1. Mol. Endocrinol. Baltim. Md *21*, 712–725.

Waples, R.S. (1998). Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. J. Hered. *89*, 438–450.

Warne, R.W., Proudfoot, G.A., and Crespi, E.J. (2015). Biomarkers of animal health: integrating nutritional ecology, endocrine ecophysiology, ecoimmunology, and geospatial ecology. Ecol. Evol. *5*, 557–566.

Waterland, R.A., and Jirtle, R.L. (2003). Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol. Cell. Biol. *23*, 5293–5300.

Weaver, I.C.G., Cervoni, N., Champagne, F.A., D'Alessio, A.C., Sharma, S., Seckl, J.R., Dymov, S., Szyf, M., and Meaney, M.J. (2004). Epigenetic programming by maternal behavior. Nat. Neurosci. *7*, 847–854.

Weaver, I.C.G., D'Alessio, A.C., Brown, S.E., Hellstrom, I.C., Dymov, S., Sharma, S., Szyf, M., and Meaney, M.J. (2007). The Transcription Factor Nerve Growth Factor-Inducible Protein A Mediates Epigenetic Programming: Altering Epigenetic Marks by Immediate-Early Genes. J. Neurosci. *27*, 1756–1768.

Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Pääbo, S., Rebhan, M., and Schübeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat. Genet. *39*, 457–466.

Weidner, C.I., and Wagner, W. (2014). The epigenetic tracks of aging. Biol. Chem. *395*, 1307–1314.

Weidner, C.I., Lin, Q., Koch, C.M., Eisele, L., Beier, F., Ziegler, P., Bauerschlag, D.O., Jöckel, K.-H., Erbel, R., Mühleisen, T.W., et al. (2014). Aging of blood can be tracked by DNA methylation changes at just three CpG sites. Genome Biol. *15*, R24.

Weir, B.S., and Cockerham, C.C. (1984). Estimating F-Statistics for the Analysis of Population Structure. Evolution *38*, 1358–1370.

Wen, A.Y., You, F., Sun, P., Li, J., Xu, D.D., Wu, Z.H., Ma, D.Y., and Zhang, P.J. (2014). CpG methylation of dmrt1 and cyp19a promoters in relation to their sexual dimorphic expression in the Japanese flounder *Paralichthys olivaceus*. J. Fish Biol. *84*, 193–205.

Wenzel, M.A., and Piertney, S.B. (2014). Fine-scale population epigenetic structure in relation to gastrointestinal parasite load in red grouse (*Lagopus lagopus scotica*). Mol. Ecol. *23*, 4256–4273.

Weyrich, A., Lenz, D., Jeschek, M., Chung, T.H., Rübensam, K., Göritz, F., Jewgenow, K., and Fickel, J. (2015). Paternal intergenerational epigenetic response to heat exposure in male Wild guinea pigs. Mol. Ecol. *25*, 1729-1740.

Wilkes, D., Xie, S.Q., Stickland, N.C., Alami-Durante, H., Kentouri, M., Sterioti, A., Koumoundouros, G., Fauconneau, B., and Goldspink, G. (2001). Temperature and Myogenic Factor Transcript Levels During Early Development Determines Muscle Growth Potential in Rainbow Trout (*Oncorhynchus mykiss*) and Sea Bass (*Dicentrarchus Labrax*). J. Exp. Biol. *204*, 2763–2771.

Wilkins, A.S., Wrangham, R.W., and Fitch, W.T. (2014). The "Domestication Syndrome" in Mammals: A Unified Explanation Based on Neural Crest Cell Behavior and Genetics. Genetics *197*, 795–808.

Wright, D. (2015). The Genetic Architecture of Domestication in Animals. Bioinforma. Biol. Insights *9*, 11–20.

Wright, P.A., and Turko, A.J. (2016). Amphibious fishes: evolution and phenotypic plasticity. J. Exp. Biol. *219*, 2245–2259.

Wu, J., Xiong, S., Jing, J., Chen, X., Wang, W., Gui, J.-F., and Mei, J. (2015). Comparative Transcriptome Analysis of Differentially Expressed Genes and Signaling Pathways between XY and YY Testis in Yellow Catfish. PLOS ONE *10*, e0134626.

X

Xi, Y., and Li, W. (2009). BSMAP: whole genome bisulfite sequence MAPping program. BMC Bioinformatics *10*, 232.

Xu, M., Li, X., and Korban, S.S. (2000). AFLP-based detection of DNA methylation. Plant Mol. Biol. Report. *18*, 361–368.

Y

Yamaguchi, T., Yamaguchi, S., Hirai, T., and Kitano, T. (2007). Follicle-stimulating hormone signaling and Foxl2 are involved in transcriptional regulation of aromatase gene during gonadal sex differentiation in Japanese flounder, *Paralichthys olivaceus*. Biochem. Biophys. Res. Commun. *359*, 935–940.

Yamano, K., and Miwa, S. (1998). Differential gene expression of thyroid hormone receptor alpha and beta in fish development. Gen. Comp. Endocrinol. *109*, 75–85.

Yan, W. (2014). Potential roles of noncoding RNAs in environmental epigenetic transgenerational inheritance. Mol. Cell. Endocrinol. *398*, 24–30.

Yang, H., Lu, P., Wang, Y., and Ma, H. (2011). The transcriptome landscape of Arabidopsis male meiocytes from high-throughput sequencing: the complexity and evolution of the meiotic process. Plant J. Cell Mol. Biol. *65*, 503–516.

Z

Zemach, A., McDaniel, I.E., Silva, P., and Zilberman, D. (2010). Genome-Wide Evolutionary Analysis of Eukaryotic DNA Methylation. Science *328*, 916–919.

Zhang, Y., Liu, H., Lv, J., Xiao, X., Zhu, J., Liu, X., Su, J., Li, X., Wu, Q., Wang, F., et al. (2011). QDMR: a quantitative method

for identification of differentially methylated regions by entropy. Nucleic Acids Res. *39*, e58.

Zhang, Y., Zhang, S., Liu, Z., Zhang, L., and Zhang, W. (2013). Epigenetic Modifications During Sex Change Repress Gonadotropin Stimulation of *Cyp19a1a* in a Teleost Ricefield Eel (*Monopterus albus*). Endocrinology *154*, 2881–2890.

Zhong, X. (2016). Comparative epigenomics: a powerful tool to understand the evolution of DNA methylation. New Phytol. *210*, 76–80.

Zhou, X., Lindsay, H., and Robinson, M.D. (2014). Robustly detecting differential expression in RNA sequencing data using observation weights. Nucleic Acids Res. *42*, e91–e91.

Zhu, J., He, F., Hu, S., and Yu, J. (2008). On the nature of human housekeeping genes. Trends Genet. TIG *24*, 481–484.

Zykovich, A., Hubbard, A., Flynn, J.M., Tarnopolsky, M., Fraga, M.F., Kerksick, C., Ogborn, D., MacNeil, L., Mooney, S.D., and Melov, S. (2014). Genome-wide DNA methylation changes with age in disease-free human skeletal muscle. Aging Cell *13*, 360–366.

List of scientific computational tools

A list of softwares and tools used in this thesis is presented below, divided in categories of purpose of use, a website link and a short description. All NGS data processings were performed under Ubuntu OS (>14.04), implying that unix and perl functions were also used when needed. All web links were working when accessed on 20/09/2016.

1) Next generation sequencing data processing

Quality controls:

FastQC: quality control checks on raw sequence data coming from high throughput sequencing pipelines.
http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Filtering and trimming:

Trim Galore: a wrapper tool to automate quality and adapter trimming as well as quality control.
http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/

Alignments and extraction of information:

Bismark: map bisulfite converted sequence reads and determine cytosine methylation states.
http://www.bioinformatics.babraham.ac.uk/projects/bismark/

BSMAP: short reads mapping program for bisulfite sequencing.
https://github.com/genome-vendor/bsmap

Bis-SNP: framework for genotyping and accurate DNA methylation calling in bisulfite treated massively parallel sequencing.
http://people.csail.mit.edu/dnaase/bissnp2011/

GEMtools: map raw reads to both a genome and a transcriptome.
http://gemtools.github.io/

Differential analysis:

methylKit: R package for DNA methylation analysis and annotation from high-throughput bisulfite sequencing. Not restricted to specific genomes.
https://github.com/al2na/methylKit

edgeR: Differential expression analysis of RNA-seq expression profiles with biological replication.
https://bioconductor.org/packages/release/bioc/html/edgeR.html

edmr: Optimized DMR analysis based on bimodal normal distribution model and cost function for regional methylation analysis.
https://github.com/ShengLi/edmr

NGS-associated files handling:

VCFtools: a program package designed for working with VCF files (Variant call format-complex genetic variation data).
https://vcftools.github.io/index.html

Samtools: a suite of programs for interacting with high-throughput sequencing data.
http://www.htslib.org/

BEDtools: a swiss-army knife of tools for a wide-range of genomics analysis tasks.
http://bedtools.readthedocs.io/en/latest/

BSgenome: Infrastructure for Biostrings-based genome data packages. Storing of full genome sequences.
https://bioconductor.org/packages/release/bioc/html/BSgenome.html

GenomicRanges: Representation and manipulation of genomic intervals and variables defined along a genome.
https://bioconductor.org/packages/release/bioc/html/GenomicRanges.html

2) Functional analysis

GO.db: A set of annotation maps describing the entire Gene Ontology.
https://bioconductor.org/packages/release/data/annotation/html/GO.db.html

topGO: Enrichment Analysis for Gene Ontology.
http://bioconductor.org/packages/release/bioc/html/topGO.html

REViGO: Reduce and visualize Gene Ontology.
http://revigo.irb.hr/

PaGenBase: database that provides information of pattern genes (specific genes, selective genes, housekeeping genes and repressed genes).
http://bioinf.xmu.edu.cn/PaGenBase/

MEME Suite: Motif-based sequence analysis tools.
http://meme-suite.org/index.html

3) Data handling in R

plyr: Tools for Splitting, Applying and Combining Data.
https://cran.r-project.org/web/packages/plyr/index.html

tidyr: Easily Tidy Data with 'spread()' and 'gather()' Functions.
https://cran.r-project.org/web/packages/tidyr/index.html

reshape: restructure and aggregate data.
https://cran.r-project.org/web/packages/reshape/index.html

data.table: aggregation of large data, joins, add/modify/delete of columns by group using no copies at all, list columns and a fast file reader.
https://cran.r-project.org/web/packages/data.table/index.html

4) Visualizations (except of R base)

biovizBase: Basic graphic utilities for visualization of genomic data.
http://bioconductor.org/packages/release/bioc/html/biovizBase.html

vioplot (Ch. 4, Fig. 4): combination of a box plot and a kernel density plot.
https://cran.r-project.org/web/packages/vioplot/index.html

hexbin (Ch. 4, Fig. 1A; Fig. 2A): Binning and plotting functions for hexagonal bins.
https://cran.r-project.org/web/packages/hexbin/index.html

PerformanceAnalytics (Ch. 3; Fig. S5): Econometric tools for performance and risk analysis. Includes the "chart.correlation" function for visualizing correlations.
https://cran.r-project.org/web/packages/PerformanceAnalytics/index.html

Ideogram (Ch. 4, Fig. S2; Ch. 5, Fig. S4): Plotting differentially methylated bases on an ideogram.
http://zvfak.blogspot.com.es/2012/06/plotting-differentially-methylated.html

Beeswarm (Ch. 3, Fig. 8 and Fig. S6): one-dimensional scatter plot like "stripchart", but with closely-packed, non-overlapping points.
https://cran.r-project.org/web/packages/beeswarm/index.html

Integrated Genome Browser (Ch. 4, Fig. 9; Ch. 5, Fig. 8): interactive desktop genome browser for exploring and understanding large-scale data sets from genomics.
https://bitbucket.org/lorainelab/integrated-genome-browser

Venny: Venn diagrams.
http://bioinfogp.cnb.csic.es/tools/venny/

5) Statistics (except of R base)

car: Companion to Applied Regression. Includes the Levene's test for homogeneity of variance across groups.
https://cran.r-project.org/web/packages/car/index.html

vegan: Ordination methods, diversity analysis and other functions for community and vegetation ecologists. Includes the "adonis" function for Permutational Multivariate Analysis of Variance Using Distance Matrices.
https://cran.r-project.org/web/packages/vegan/index.html

ade4: Multivariate data analysis and graphical display.
https://cran.r-project.org/web/packages/ade4/index.html

6) MSAP data processing

msap: Statistical Analyses of Methylation-sensitive Amplification Polymorphism (MSAP) assays.
https://cran.r-project.org/web/packages/msap/index.html

RawGeno: automatization of scoring of AFLP datasets.
https://sourceforge.net/projects/rawgeno/

7) qPCR analysis

ReadqPCR: read raw RT-qPCR data.
http://www.bioconductor.org/packages/release/bioc/html/ReadqPCR.html

NormqPCR: selection of optimal reference genes and the normalisation of real-time quantitative PCR data.
https://www.bioconductor.org/packages/release/bioc/html/NormqPCR.html

qPCR differential analysis: A spreadsheet with the correct formulas for qPCR normalization with two reference genes and correct standard error of propagation calculation.
http://www.mcbryan.co.uk/blog/PCR.xls