UNIVERSITY OF ICELAND

Ph.D. Thesis
in Biology

Diversity of DNA methylation signals in teleosts with focus on the sympatric Arctic charr (*Salvelinus alpinus*) morphs of Thingvallavatn

Sébastien Léon Tadeus Matlosz

June 2024

FACULTY OF LIFE AND ENVIRONMENTAL SCIENCES

Diversity of DNA methylation signals in teleosts with focus on the sympatric Arctic charr (*Salvelinus alpinus*) morphs of Thingvallavatn

Sébastien Léon Tadeus Matlosz

Dissertation submitted in partial fulfillment of a
Philosophiae Doctor degree in Biology

Ph.D. Committee
Zophonías Oddur Jónsson (Supervisor)
Sigríður Rut Franzdóttir (Co-supervisor)
Arnar Pálsson (Co-supervior)
Robert Hindges

Opponents
Professor Irene Adrian-Kalchhauser
Professor Kristinn Pétur Magnússon

Faculty of Life and Environmental Sciences
School of Engineering and Natural Sciences
University of Iceland
Reykjavik, June 2024

Diversity of DNA methylation signals in teleosts with focus on the sympatric Arctic charr (*Salvelinus alpinus*) morphs of Thingvallavatn
Dissertation submitted in partial fulfillment of a Philosophiae Doctor degree in Biology

Faculty of Life and Environmental Sciences
School of Engineering and Natural Sciences
University of Iceland
Sturlugata 7 (Askja)
102, Reykjavik
Iceland

Telephone: 525 4000

Bibliographic information:
Sébastien Matlosz, 2024, *Diversity of DNA methylation signals in teleosts with focus on the sympatric Arctic charr (Salvelinus alpinus) morphs of Thingvallavatn*, PhD dissertation, Faculty of Life and Environmental Sciences, University of Iceland, 185 pp.

Author ORCID: 0000-0002-4995-6682
ISBN: 978-9935-9769-8-7

Abstract

For their sensitivity to environmental variables, their impact on early developmental processes and their importance to the establishment of multiple phenotypes, epigenetic mechanisms have been increasingly studied during the last decades. Like other molecular processes at play in building and shaping adaptive diversity, they can be studied over a wide range of species at different taxonomic levels. Broader levels, such as within vertebrate groups, allow for the identification of signals that have over time become assimilated while narrower levels, such as within species morphotypes, allow for the identification of molecular signals in the first stages of speciation, which might get lost over longer evolutionary periods. The four Arctic charr morphs of Lake Thingvallavatn are an example of morphotypes within species, and are a useful tool to assess molecular differences between recently diverged groups. In this thesis, I investigated DNA methylation and gene expression differences between these four morphs during early development (**Paper I**), I identified correlations between the divergence of genome, transcriptome, and methylome in three of these morphs (**Paper II**) and I reviewed the state of knowledge about DNA methylation reprogramming in teleosts (**Paper III**). This jointly adds an important perspective to the field of evolutionary and developmental epigenetics.

Útdráttur

Þegar svipgerðareinkenni erfast án þess að vera beinlínis skráð í DNA er talað um umframerfðir. Það hversu næmir ferlarnir að baki umframerfðum eru fyrir umhverfisþáttum, áhrif þeirra á snemmþroskun, og þar af leiðandi mikilvægi þeirra fyrir myndun svipgerðar hefur á síðustu áratugum vakið athygli þróunarfræðinga, sem hafa í auknum mæli leitast við að samþætta þessi ferli inn í viðteknar kenningar þróunarfræðinnar. Til að mögulegt sé að greina og skilja hin ýmsu lífefnaferli sem tengjast tilurð aðlaganna er nauðsynlegt að rannsaka margar bæði fjarskylda og náskylda hópa lífvera. Breiðari flokkunareiningar, eins og t.d. innan hryggdýra, gera það kleift að greina ferli sem orðin eru hluti af sameiginlegum þroskunarferlum þessara lífveruhópa. Þrengri samanburður, t.d. á afbrigðum innan tegunda, gefur færi á að greina sameindamerki á fyrstu stigum aðskilnaðar, sérstaklega umframerfðamerki sem gætu tapast vegna náttúrulegs vals á þróunarfræðilegum tímaskala. Bleikjuafbrigðin fjögur í Þingvallvatni eru einstakt dæmi um svipgerðarfjölbreytileika sem skapar kjörið tækifæri til að rannsaka sameindagrunn breytileika á milli einstaklinga af erfðafræðilega náskyldum hópum. Í þessari doktorsritgerð rannsaka ég í fyrsta lagi DNA methylmerkingar og mismun á genatjáningu allra fjögura afbrigðanna snemma í fósturþroskun (**Grein I**). Í öðru lagi greini ég fylgni milli methylmerkinga, erfðabreytileika og breytilegrar genatjáningar í þremur afbrigðum (**Grein II**) og í þriðja lagi tek ég saman yfirlit yfir rannsóknir á erfðum og endurstillingu DNA methylmerkja í beinfiskum (**Grein III**). Niðurstöðurnar gefa með því mikilvæga innsýn í stöðu þekkingar á sviði þróunar- og þroskunalegrar umframerfðafræði.

To my parents.

Table of Contents

List of Figures

List of Tables

List of Publications

This thesis is a collection of three scientific papers in varied states of publication:

Paper I:
Matlosz, Sébastien, Benjamín Sigurgeirsson, Sigrídur Rut Franzdóttir, Arnar Pálsson, and Zophonías O. Jónsson. 2022. 'DNA Methylation Differences during Development Distinguish Sympatric Morphs of Arctic Charr (*Salvelinus Alpinus*)'. *Molecular Ecology* 31 (18): 4739–61. https://doi.org/10.1111/mec.16620.

Paper II:
S. Matlosz, L. Jerman-Plesec, H. Xiao, A. Guðjónsson, S.R. Franzdóttir, A. Pálsson, Z.O. Jónsson, 'Patterns of genetic, methylomic, and expression divergence among three sympatric resource morphs of Arctic charr' (*Manuscript*).

Paper III:
Matlosz, Sébastien, Sigríður R. Franzdóttir, Arnar Pálsson, and Zophonías O. Jónsson. 2024. 'DNA Methylation Reprogramming in Teleosts'. Evolution & Development: e12486. https://doi.org/10.1111/ede.12486.

Contribution of the author.

Paper I: S. Matlosz participated in sampling, performed the experiments, analyzed the data and wrote the manuscript.

Paper II: S. Matlosz performed comparative analyses between the previously generated datasets and wrote the first version of the manuscript.

Paper III: S. Matlosz wrote the manuscript.

Abbreviations

5-CaC: 5-carboxylcytosine

5-fC: 5-formylcytosine

5.8S rRNA: 5.8S ribosomal RNA

5hmC: 5-hydroxymethylcytosine

5mC: 5-methylcytosine

6mA: 6-methyl adenine

CpG: Cytosine followed by guanine (dinucleotide)

Cs4R: Carp specific whole genome duplication event

ddRADseq: Double digest restriction site associated DNA sequencing

DNA: Deoxyribonucleic acid

Dnmt: DNA methyltransferase (enzyme)

DMR: Differentially methylated region

F_{ST}: Fixation index

glm: Generalized linear model

GO: Gene ontology

LB: Large benthic (morph)

LncRNA: Long non-coding RNA

MeHg: Methylmercury

miRNA: MicroRNA

mya: Million years ago

PCA: Principal component analysis

PGC: Primordial germ cells

PI: Piscivorous (morph)

piRNA: PIWI interacting RNA

PIWI: type of regulatory protein

PL: Planktivorous (morph)

qPCR: Quantitative polymerase chain reaction

RNA: Ribonucleic acid

RRBS: Reduced representation bisulfite sequencing

SB: Small benthic (morph)

SNP: Single nucleotide polymorphism

Ss4R: Salmonid specific whole genome duplication event

TE: Transposable element

TET: Ten-eleven translocation (enzyme)

Tol2: transposon closely related to the hAT family of non-autonomous elements

Ts3R: Teleost specific whole genome duplication event

tRNA: Transfer RNA

tsRNA: Transfer-RNA derived small RNAs

WGD: Whole genome duplication (event)

WGS: Whole genome sequencing

Acknowledgements

I would like to begin by thanking RANNÍS and the University of Iceland for funding this research.

In addition, this PhD project would have been naught but for the involvement of, and help from, Zophonías Oddur Jónsson, Sigríður Rut Franzdóttir and Arnar Pálsson, whom I deeply thank for their advice and support, both professional and personal. I would also like to thank Robert Hindges, external committee member, for his advice and comments on the thesis.

Science advances through the ideas, not of an individual but of an entire community. Thus, everyone I met at work or conferences, who discussed with me their own research and/or commented on mine deserves a thank you for the completion of this project. I also thank Josh Kinman, then staff scientist at Perkin Elmer, who in 2018 remotely helped me with troubleshooting the RRBS experiment when nothing was working.

No less important than ideas and lab work, placing orders, repairing machines, keeping common areas clean, providing fresh coffee to staff, managing grant money and generally doing any type of administrative work is central to the completion of such research. Thus, I would like to thank Sigrún Reynisdóttir and Snorri Páll Davíðsson as well as Þórhallur Aðalsteinsson, Karl Arnason and Hlynur Sigurgislason for helping with these matters and sharing daily smiles and Icelandic vocabulary with me. I would also like to thank Þuríður Sigurjónsdóttir for her continuous support.

More than the scientific upheavals, doctoral studies are a human adventure, and I would like to thank anyone who has crossed paths with me during these years and offered me a little bit of their life vision which led me to be who I am today.

Therefore, somewhat chronological thanks to:

Grétar for instantly introducing me to the hiking group and sharing some Icelandic knowledge with me from day one. Ölvir, for the plant knowledge and letting me keep Pétur the fern for 3 years. Pepe, for being insane. Sólrún and Inga for the laughs, Icelandic lessons and introducing me to Eurovision. Alejandro, Baldur, Charles, Dagný, Edwin, Friðrik, Gotje, Hildur, Jóhannes, Kate, Kera, Julie, Ragnhildur and Rob for sharing many moments in my early PhD days. Kalina, for the talks about life, shoe-pellets and introducing me to Terry Pratchett. The many sweet technicians, students and PIs from Hólar who provided either a bed, a meal or a nice chat in the beer club whenever we came to work North. Samia, Leifur and Samantha for some language-nerd talk, IT help, and making a snow Pepe, respectively. All of the summer interns in our research group, as well as all the master students I had the pleasure to meet during my time in Askja. Scout, for all the laughs and parties, as well as Martina C. and Martina K. for the sweet coffee-time moments and friendship over the years. Klara, for being herself and inspiring me to do the same. Lea and Denis for being craycray and the best office buddies. Xindan for the chinese lessons and for the plush dog. Ása for teaching us kettlebells with so much energy. Will for the laughs and that one race across the graveyard. Revathy, for the drawings and music/mediation talks. Gabe, Joaquin and the whole traditional Irish music gang for letting me play Celtic music with you. Henna for teaching me about quarks. Hannah, for being an

amazing running partner and friend, and for helping me these months in winter 2018 when I was in a bad place. Quentin, for being an amazing roommate, colleague and surf companion (and general adventure buddy). Petra, Alana, Caro and Tjorben need a surfing shoutout as well, catching waves with you was a blast. Marta and Jake for being great friends to this day and for, amongst all memories, that one night-hiking to the volcano. Marina for the Cádiz shenanigans. Heather and Matthew for listening to my existential thoughts and being the best car rental provider. All the sweet people from Háskolakorinn for the epic singing. Lucía, for the hikes, shared meals and helping me with important life decisions. Diego and Thejus, for always motivating me to run with you when I was down. Renée for the whale discussions and orca watching. Alix for the Icelandic grammar tips and getting me a barista job. All the artsy colleagues from Eymundsson. Han, for our art-related discussions. Juan for the scientific talks and hot pots. Vyshnav for the physics and video-game talks. Cecilie and Maiken for many nights northern light watching and being great photographers. Ivan, for discussions about the state of the world. Matthis and Méline, for letting me stay at your place when I was homeless. Ayça, Anna, Axel, Bastien, Emily, Fia, Guðbjorg, Ian, Ingeborg, Ingvild, Joren, Julia, Lieke, Maite, Marcos, Noémie, Silvia, Samoantha, Suzanne and Theresa for the pool trips, post-work ice-creams, hugs, hikes, emotional support, life discussions, writing motivation and overall sharing many amazing moments in my late PhD days.

This is a non-exhaustive list, and I apologize to anyone whom I have not mentioned. I wish all of you the best for whatever comes your way.

Last, I would like to thank my family.

Mom, Dad, Loïc and Julie (and cat, Donut) for always supporting me through this long and sometimes exhausting journey. I could not have done this without your presence in the good and bad moments. I am lucky to have you by my side. All uncles, aunts and cousins I thank for the same reasons.

1 Introduction

The integration of Mendelian genetics with the Darwinian theory of evolution through natural selection, referred to as "the new synthesis", not only revolutionized biology but the life-sciences in general (Barton, 2022). Since its conceptualization in the 20th century, scientists have worked hard to test the principles and implications of evolutionary biology. The importance of genetic mutations in the establishment of new traits has been proven repeatedly (Hamilton, 2021), for example through gene expression changes caused by a now-modified transcript, or through the loss or gain of a specific binding motif necessary for the binding of the proper factors to DNA at the right location. A large number of traits are also caused by mutations in pleiotropic genes which affect poorly understood polygenic pathways. However, it is becoming increasingly clear that alterations of the DNA sequence itself are not the only actors contributing heritable variation for selection to act on and that non-genetic molecular pathways, sometimes driven by the environment itself, can shape individuals in specific ways by contributing to phenotypic plasticity and divergence (Bonduriansky and Day, 2018; Lind and Spagopoulou, 2018). These mechanisms revolving around, and interacting with, the DNA sequence are at the core of the field of epigenetics, which Waddington (1942) originally described as "the branch of biology which studies the causal interactions between genes and their products, which bring the phenotype into being", and by adding an extra layer to gene expression regulation these mechanisms are of rapidly increasing interest to developmental and evolutionary biologists. Research has now shown examples of the establishment of epigenetic marks through environmental cues, as well as the transmission (direct or indirect) of said marks to the offspring (e.g. Major et al., 2020; Pierron et al., 2021).

It could therefore be argued that the Lamarckian proposition of "the inheritance of acquired traits" has some material foundation (Lamarck, 1802). Some argue this prompts the establishment of a revised view of evolution (Skinner, 2015). However, revising long-standing evolutionary concepts to incorporate the epigenetic layer is no easy task, and the study of model organisms such as mouse, zebrafish or stickleback (Jones et al., 2012; Metzger and Schulte, 2018) will only get us so far in the understanding of evolution's long and winding roads, for which "non-model" species also need to be studied in molecular detail. Indeed, the more varied the organismal representation is, the more clades and lineages investigated, the better we will be able to link molecular changes to specific evolutionary histories. Moreover, being able to investigate epigenetic signals in groups of organisms affected by multiple environmental variables or closely related populations or species occupying varied ecological niches is essential to decipher the role the environment plays in adaptation and divergence.

The phenotypic diversity of teleosts, the largest and most diverse group of vertebrates with 30,000 described species (Nelson, Grande and Wilson, 2016), makes them an excellent research playground for the investigation of evolutionary processes (Volff, 2005). As ectothermic aquatic organisms, water temperature and chemical content are amongst many environmental variables they have had to adapt to, which makes them good models for the study of these important ecological factors on development and evolution. Along the same lines, they have been used extensively to study speciation and phenotypic plasticity and are now being increasingly investigated with regards to the impact of the environment on epigenetically-mediated evolutionary and developmental processes (Best et al., 2018). Out

of all teleosts, species showcasing sympatric polymorphism (that is, species for which different morphotypes are found in the same geographic location) are of great interest, as they allow for the observation of more recent evolutionary processes, therefore allowing for the identification of signals that might have possibly been lost between sister lineages which diverged a longer time ago. The Arctic charr morphs of Lake Thingvallavatn form such a system: following the end of the last ice age, an anadromous population of *S. alpinus* colonized the lake and eventually became trapped there, due to volcanic activity and isostatic rebound (Kapralova et al., 2011). Four morphs can now be found in the lake, small benthic (SB), large benthic (LB), planktivorous (PL) and piscivorous (PI), diverging along the benthic limnetic axis and showing some phenotypic differences (Sandlund et al., 1992), despite some still ongoing gene flow (Xiao et al., unpublished). As the morphs spawn in different geographic locations in the lake (Skúlason, Noakes, & Snorrason, 1989; Skúlason et al., 1989), some of which colocate with inflows of glacial springs, it is likely that the environment contributes to different extent to their development and to the setup of epigenetic marks (whether through their induction, or because some marks are environmentally selected). If these marks are transferred to the next generation, they could be responsible for the phenotypic differences observed in the wild in this species.

Out of the different epigenetic marks, DNA methylation is highly studied because it has the capacity to respond to environmental changes. In principle, specific changes in DNA methylation at a locus, leading to an increase in fitness, can be selected for and become stable in populations, potentially leading to reproductive isolation and speciation. This potential has been suggested as a way for rapid plastic responses in populations in response to the environment, that can consequently enable later fixation of these phenotypes by genetic assimilation (West-Eberhard, 2005; Pal & Miklos 1999). On the other hand, de Jong (2005) and Greenspoon et al. (2022) argue that epigenetic mechanisms can also slow down divergence if fast acting epigenetic adaptation hides beneficial alleles from selection. Whether or not methylation changes help or hinder divergence is an open question in evolutionary biology, and the answer may be context dependent. Studies showing differences of methylation patterns between teleosts are numerous and span from broad to narrow taxonomic contexts. For instance, Varriale and Bernardi (2006) identified important global methylation differences between fish species belonging to multiple orders of teleosts and we showcased methylation differences between the aforementioned recently diverged sympatric morphs of Arctic charr (Matlosz et al., 2022) (see **Paper I**), highlighting the potential for methylation marks to fuel the evolutionary diversity of fishes, even in the early stages of speciation.

This diversity may have also been fueled in part by multiple whole genome duplication (WGD) events (see Ravi & Venkatesh, 2018; Volff, 2005). In addition to the two basal vertebrate WGD events (Dehal and Boore, 2005; Simakov et al., 2020), the ancestor of modern teleosts experienced a WGD event, termed ‘Ts3R’, about 350 million years ago (Hoegg et al., 2004; Meyer and Van de Peer, 2005). Following this, some teleost lineages have experienced additional WGDs, with the *Salmonidae* family undergoing the ‘Ss4R’ event about 95 million years ago (Macqueen and Johnston, 2014) and the carp specific ‘Cs4R’ event happening more recently, around 12 million years ago (Ma et al., 2014). As genetic and epigenetic mechanisms go hand in hand, differences in genomic composition should be taken into consideration when studying epigenetic pathways. Indeed, major genomic events such as WGDs could have had an impact on epigenetic mechanisms in cases where duplications lead to neo- or sub-functionalization of genes responsible for the epigenetic machinery.

In this introduction, I aim to provide insights into the epigenetic molecular mechanisms that have contributed to or accompanied teleost evolution, with a focus on DNA methylation mechanisms.

First, I provide a brief overview of different types of epigenetic mechanisms in eukaryotes.

Second, studies of the possible impact of different environmental factors on the methylome of various fish species will be summarized, with a special focus on those showing inter- or transgenerational inheritance of methylation patterns. This inheritance is of great importance, as new epi-alleles being shared between individuals of a population provide the raw material for adaptive changes.

Third, I explore the interplay between methylation signals and genomic events during teleost evolution. This includes the impact that WGD events have had on the methylation machinery as well as the role of methylation in silencing transposable elements, rebalancing gene dosage and accompanying newly duplicated genes through neo- or sub-functionalization.

Last, I summarize the state of the art on the sympatric morphs of Arctic charr from lake Thingvallavatn, the research model at the core of this PhD research.

1.1 A Brief Overview of Epigenetic Mechanisms

1.1.1 Histone Modifications

Histone modifications are arguably the most studied epigenetic mechanisms, playing a role in multiple aspects of gene- and genome regulation. Highly conserved in eukaryotes, four core histones (H2A, H2B, H3 and H4) associate to form octamers which, by interacting with genomic DNA, create units called nucleosomes and subsequently allow for the compaction of a higher order structure of DNA arrangement called chromatin (Kornberg, 1974). A fifth type of histone called H1, known as linker histone, links the DNA between nucleosomes together, further increasing compaction. The N-terminal tails of the core histone proteins, which protrude from the nucleosome complex, are subject to a whole range of chemical modifications. These modifications can, through modifying inter-nucleosomal interactions, stabilize or destabilize chromatin and therefore impact accessibility of genes and thus their expression (Bannister & Kouzarides, 2011). The most studied histone modifications are methylation (mono-, di- or tri-), acetylation and phosphorylation, but many other modifications also take place, such as ubiquitination, sumoylation, glycosylation and biotinylation, all of which target specific amino acids on the histone tails. Some of these modifications induce chromatin changes by recruiting auxiliary proteins responsible for chromatin remodelling (Wysocka et al., 2006), while others such as acetylation can also induce the destabilization of chromatin by neutralizing the positive charge of their targeted residue, weakening the interaction between DNA and histone proteins (Verdone et al., 2006). The picture is more complicated however, as these modifications can have different biological roles depending on the residue they are attached to, and even the number of methyl groups involved changes the consequence of a histone methylation mark. For example, in human, while additions of a single methyl group to lysine 9, 27 or 79 of Histone H3 (H3K9Me, H3K27Me, H3K79Me) are all linked with increased transcription, tri-methylation at the same positions correlates with gene repression (Barski et al., 2007). These numerous chemical histone modifications are

established and removed by diverse enzymes, such as histone acetyl transferases (HATs) and deacetylases, or histone methyltransferases and demethylases. For more detailed review of the highly diverse and intricate histone modification molecular machinery in teleosts, see (Best et al., 2018; Fellous and Shama, 2019).

1.1.2 DNA Modifications

Unlike histones, for which the types of chemical modifications are numerous, the eukaryotic DNA itself is almost exclusively modified by methylation. This process does not impact all DNA bases as it happens predominantly on cytosines in a CpG setting, that is cytosines positioned next to guanines (Cooper, 1983). Specifically, a methyl group is added to the fifth carbon of the cytosine, hereafter denoted as 5mC. Interestingly, this epigenetic modification can affect chromatin structure, similarly to histone modifications, and thus can impact different nuclear mechanisms, from transcriptional activity (Ng and Bird, 1999) and transposable element (TE) silencing (Karimi et al., 2011) to X chromosome inactivation in mammals (Lee, 2003). In most eukaryotes, the DNA methyltransferase enzyme Dnmt1 is responsible for maintaining methylation marks through rounds of DNA replication (Bestor et al., 1988) while Dnmt3A and Dnmt3B (Okano, Xie, and Li 1998) are responsible for the establishment of new marks *de novo*. The 'Ts3R' and/or 'Ss4R' WGD events affected all three genes, resulting in various numbers of paralogs in different fish species (Best et al., 2018). As will be discussed later, gene duplication, followed by neo-functionalization and subfunctionalization (Force et al., 1999) can lead to changes in molecular pathways, whether that be through partitioning of ancestral functions or the acquisition of novel ones. Thus, the duplication of this methylation machinery could have had major impacts on how DNA methylation is established and regulated in fish.

The impact of DNA methylation on gene expression depends on the location of methylated cytosines in relation to genomic features, with DNA methylation at promoters and first introns generally correlating with a decrease in gene expression (Esteller, 2002; Anastasiadi, Esteve-Codina, and Piferrer, 2018) while hypermethylation in the gene body (exons and introns) of oncogenes correlates with their overexpression (Arechederra et al., 2018). Methylation marks can be passively lost during DNA replication if they are not maintained. Additionally, the process of active demethylation involves conversion of 5mC residues by TET enzymes (reviewed in (Rasmussen and Helin, 2016) and (Wu and Zhang, 2017)) into different modified nucleobases (hydroxymethylcytosine (5hmC), formylcytosine (5fC) and carboxycytosine (5CaC)) before the enzyme thymine DNA glycosylase (TDG) removes them, initiating the repair of the resulting abasic site by base excision repair (Maiti et al. 2013). Interestingly, 5hmC can be more than a transient intermediate and has been found to be in some cases stably present in genomic DNA in mammals (Bachman et al., 2014). This seems to be the case in specific cell types such as neurons where 5hmC might play a role in neurodevelopment, aging and neurological disorders (Szulwach et al., 2011; Mellén et al., 2012). While TET proteins usually catalyze the conversion of 5mC into 5hmC and subsequent bases, recent work has shown that the subsequent oxidation of 5hmC into 5fC (and then further to 5caC) does not always happen, which could lead to an accumulation of 5hmC in the genome (Rasmussen and Helin, 2016). In teleosts, the biological impact of 5hmC is starting to be investigated, and differential DNA hydroxymethylation between muscle, liver and pituitary was found at genes responsible for somatotropic growth in Nile tilapia (Konstantinidis et al., 2021).

1.1.3 Non-Coding RNAs

Non-coding RNAs play various important roles in the eukaryotic cell. With regard to gene expression, they are best known for participating in its post-transcriptional modulation, via what is commonly known as RNA interference pathways, guiding RNA interference silencing complexes (RISCs) to mRNAs and inducing their degradation or translational repression (Pratt and MacRae, 2009). However, they can also impact chromatin structure and induce gene silencing that can persist through several generations in the germline of *C. elegans*, possibly through the transmission of RNAs bound to argonaute complexes (Volpe et al., 2002; Duempelmann, Skribbe, and Bühler, 2020). As such, they provide a clear example of an epigenetic mechanism acting both during development and with potential to transfer between generations. In this subsection, we will discuss various types of non-coding RNAs such as miRNA, lncRNA, piRNA and tsRNA.

Micro RNAs (miRNAs) are an extensively studied group of small non-coding RNAs that have clear influence on gene regulation. Discovered in the 1990s for their role in targeting complementary mRNA sequences and regulating gene function (Lee, Feinbaum, and Ambros, 1993), they sparked the general interest of molecular biologists (Pasquinelli et al., 2000; Iorio et al., 2005; Kozomara and Griffiths-Jones, 2011). A large proportion of miRNA genes are organized in clusters (Lau et al., 2001), but in some cases these non-coding miRNA sequences are nested in UTRs or intronic sequences of protein coding genes (Lin, Miller, and Ying, 2006). On top of their post-transcriptional gene silencing role, miRNAs can also impact chromatin structure and play a role at the transcriptional level by targeting nascent RNA transcripts and driving heterochromatin formation and silencing at the neighboring genes during overlapping transcription (Gonzalez, Pisano, and Serrano, 2008). As in other eukaryotes, miRNAs play a role in a wide range of biological processes in teleost fish (Bizuayehu and Babiak, 2014).

PIWI-interacting RNAs (piRNAs) are small non-coding RNAs, that are part of the most important defenses against transposable elements. By associating with PIWI- and several other types of proteins to form an induced silencing complex called piRISC, those RNAs bind to transposable elements in the genome, induce *de novo* methylation and change chromatin structure, and thus repress their transcription and transposition (Siomi et al., 2011). These RNAs are especially enriched in cells of the germ line, and piRNA clusters were recently found to be differentially expressed between mature gonads of the cichlid *Astatotilapia latifasciata*, based on sex (Oliveira et al., 2022).

Lastly, long non-coding RNAs (lncRNAs) longer than 200 nucleotides (Kapranov et al., 2007) have been shown to be regulated by other epigenetic signals such as DNA methylation and histone modification marks (Wu, Kallin, and Zhang, 2010) and exhibit cell-type specific and developmental stage specific expression (Dinger et al., 2008).

On top of their ability to change chromatin states, non-coding RNAs are present in the egg and sperm, and are transmitted to the zygote during fertilization, providing an extra layer of epigenetic inheritance (Chen et al., 2016) (McJunkin, 2018). For example, tsRNAs (transfer-RNA derived small RNAs) located in the sperm (Peng et al., 2012) have been shown to drive the inheritance of metabolic disorders in mice (Chen et al., 2016), and the expression of several miRNAs has been shown to be altered, following unpredictable maternal stress (MSUS) in mouse, in F1 sperm and unexposed F2 serum and hippocampus (Gapp et al., 2014). Moreover, in this last example, the injection of RNAs purified from the sperm of males having experienced MSUS into wild-type fertilized mouse oocytes induced comparable behavioral, metabolic and molecular effects to the direct exposure to MSUS

during early post-natal life, suggesting a causal link between the sperm RNAs and the observed effects over generations (Gapp et al., 2014). However, it is interesting to note that the relative miRNA levels in the F3 generation were back to normal, despite the behavioral symptoms still being present, suggesting that the changes in miRNAs were possibly transferred to other epigenetic marks such as DNA or histone modifications for further transmission (Gapp et al., 2014).

1.2 Inheritance of Environmentally-Induced Epi-Alleles in Fish

Environmentally induced molecular variation transmitted over generations has the potential to generate diversity within populations by creating specific phenotypes that may vary in adaptive value, helping with niche colonization and adaptation (Herman and Sultan, 2016). A central question is therefore how much of the variance in phenotypic traits, subject to natural selection, is caused by epigenetic variation inherited over the generations? Furthermore, for how many generations are epi-alleles stable and how penetrant are they with respect to the phenotype in question?

The inheritance of acquired traits or environmentally induced traits has been hard to measure until now, but the field of epigenetics provides us with identified and potentially measurable molecular signals that can both be impacted by environmental variables and transmitted to the offspring, such as DNA or histone modifications. In methylomics, epi-alleles refer to copies of a gene differing by their methylation signal (as opposed to SNPs, indels or larger structural variations which vary at the DNA level), and they have been identified in various organisms following exposure to different environmental variables (see below). Similar to "regular alleles", even if these epi-alleles do not directly increase an organism's fitness in its current environment, they could provide bet hedging strategies in cases where they confer a survival advantage over a longer period of time, in changing or stressful environments.

The realization that there are drastic differences in methylation reprogramming mechanisms between mammals and fish (see **Paper III**) generated excitement in the field of teleost epigenetics. For example, the fact that the zebrafish paternal methylome undergoes very little reprogramming post-fertilization offers the opportunity for inheritance of epi-alleles. In this section, I will first summarize several studies on the influence of the environment on the fish methylome, and then review other work showcasing the potential inheritance of these methylation marks. The impact of these methylation changes on the phenotype will be discussed when applicable.

1.2.1 Different Factors Can Impact Methylation

The scientific work investigating environmentally-induced methylation changes in teleosts can be largely divided into studies of the impact of human derived or natural factors. First, anthropogenic factors, such as the repercussions of pollutants (van der Ven et al., 2017; Hu et al., 2021) and rearing or farming activities (Podgorniak et al., 2022) can affect fish methylomes. These are important contemporary problems that threaten natural populations and biodiversity. Moreover, this highlights how methylation, and potentially also phenotypes, can be impacted by drastic environmental changes. Second, many natural environmental factors have been found to affect DNA methylation, including temperature (Metzger and Schulte, 2017; Anastasiadi et al., 2021; Sävilammi et al., 2021), salinity (Li

et al., 2017; Si et al., 2021), hypoxia (Beemelmanns et al., 2021), pathogens/parasites (Berbel-Filho et al., 2019) (Sagonas et al., 2020), diet (Marandel et al., 2022) and darkness or other unknown environmental factors tied with cave dwelling (Gore et al., 2018). The list of studies investigating environmentally-induced methylation changes in fish (without investigating the inheritance of said marks) is quite long and the studies cited above only comprise some of the more recent work. For a more extensive list of this rich literature, I refer to review by Best et al (2018), which highlights most different types of environmental variables known to impact methylation in fish.

The field of research on environmentally-induced methylation marks contains recurring themes. Perhaps most importantly, the impact of a specific environmental variable appears to be context dependent, differs between species and tissues, and varies by genomic regions. For instance, changes in water temperature (both reduction and increase) increased global methylation in stickleback muscle, while at a locus level both hyper- and hypomethylated loci were observed (Metzger and Schulte, 2017). Interestingly, hyper- and hypomethylation at specific loci following changes in temperature were also observed in European sea bass, and the ratios differed between tissues, with the muscle and liver exhibiting more hypomethylated genes compared to the testis and brain (Anastasiadi et al., 2021).

Ultimately, the most important epi-alleles, from an evolutionary perspective, are the ones that lead to differences in fitness between individuals. Thus, changes in methylation leading to altered phenotypes are of interest as, although not all phenotypic differences do so, they have the potential to alter fitness. Furthemore, they are important to developmental biologists trying to understand the molecular forces driving changes in phenotypic traits, whether they operate in an ecological, evolutionary or medical field.

Several studies have addressed phenotypic changes driven by environmentally-induced methylation differences. A striking example is that DNA methylation was shown to play a role in environmental sex determination in several fish species through changes in the expression of genes important for gonadal development (Ortega-Recalde et al., 2020). Negative correlations between methylation and expression of the genes *cyp19a1a* (an aromatase converting testosterone into estrogen (Kitano et al., 1999)) and *dmrt1* (a widely studied sexual developmental factor (Herpin and Schartl, 2011)) have indeed been documented in teleosts (Piferrer et al., 2019). As with changes in other types of phenotypic variables, environmentally induced epigenetic regulation of sex ratio may have an evolutionary advantage when survival is biased towards a specific sex.

Last, it is important to note that many environmental variables can, and do, act in concert to affect methylation patterns and gene expression, leading to altered phenotypes. For example, Guo et al., (2018) identified interaction of two environmental variables (heat shock and metal exposure) on methylation in zebrafish. They focused on four CpG dinucleotides inside an HSF (heat shock factor) binding motif in the promoter of the heat shock protein HSP70 and found that heat shock reduced methylation at these sites while cadmium exposure both reduced and increased methylation levels depending on the CpG. However, combined exposure to cadmium and high heat led to hypermethylation of the whole motif.

In an evolutionary context, methylation changes impacting phenotype have no potential to become assimilated in populations if they are not inherited, and we will now focus on studies that have revealed inheritance of environmentally induced DNA methylation marks.

1.2.2 Inheritance of DNA Methylation in Fish

Before going further, it is important to distinguish between intergenerational and transgenerational epigenetic inheritance (Perez and Lehner, 2019). In fish, intergenerational inheritance refers to signals being transmitted from parents (F0) to their offspring (F1), as the environment experienced by the parents affects not only their own methylome but also the methylome of their germ cells, which will give rise to the next generation (F1). In contrast, transgenerational inheritance refers to signals transmitted from the parental generation (F0) to the F2 or subsequent generations. Note the differences to mammals, where the F2 generation (germ cells in the foetus) are still exposed to the original environmental cues experienced by the pregnant mothers (F0), requiring that such signals in mammals span an extra generation to be classified as true transgenerational inheritance.

Importantly, DNA methylation signals seem to be reprogrammed during early embryonic development in mammals, limiting the possibility of the inheritance of methylation signals.

Indeed, a first wave of demethylation is observed in mammals following fertilization, when both parental methylomes undergo reprogramming (Santos et al., 2002), although the paternal genome is believed to be actively demethylated through the activity of TET enzymes (Oswald et al. 2000; Seisenberger et al. 2013) while the maternal genome gets passively demethylated by dilution (Zeng and Chen 2019). Following implantation, the embryonic methylation levels then rise (Wang et al. 2014), completing the first demethylation/remethylation reprogramming event. The second wave of demethylation is observed in PGCs during their proliferation and migration to the genital ridge (Smallwood and Kelsey 2012). It is during this event that genomic imprints, which have escaped the first round of reprogramming, are reset (Yamaguchi et al. 2013), before being reestablished later during the formation of the gametes.

Reprogramming of DNA methylation in fish is reviewed in detail in **Paper III**, but to summarize, the zebrafish zygote doesn't seem to undergo any reprogramming, and although a small demethylation event is observed in zebrafish PGCs, its scale is greatly reduced compared to what happens in mouse PGCs. The state of reprogramming in medaka is debated, as methods used to assess global methylation in the zygote of this species were biased by mitochondrial levels in the oocyte and early cleavage stage. However, the expression of TET enzymes during early development in medaka suggests that demethylation is present post-fertilization in this species, although the extent of it is probably greatly reduced compared to mammals, as by the 32-cell stage embryonic methylation levels are measured at 80% (similar to gametic levels). In medaka PGCs, a reprogramming event was measured, although the scale of demethylation is not as important as in mammals.

To summarize, in species with DNA methylation reprogramming (such as mammals and possibly medaka), methylation marks induced by the environment need to escape two rounds of reprogramming for each generation, in order to persist in the germline and have a chance to impact somatic cell development and the phenotype of the offspring. Nevertheless, instances of methylation-reprogramming bypass have been observed in mammals: imprinted genes (Barlow and Bartolomei, 2014) as well as repeat elements such as Intracisternal A-particles (IAP) (Lane et al., 2003) have been shown to escape reprogramming (Guibert, Forné, and Weber, 2012). Considering that the methylome of mammals is more dramatically demethylated and reprogrammed than that of zebrafish,

medaka and likely other fishes also, it can be argued that teleosts are more likely to have the capacity to transmit environmentally induced methylation marks between generations.

Importantly, several recent studies have indeed documented such inheritance in fish, at specific loci (see **Table 1**). Most of these studies document inheritance of methylation marks, but do not provide assessment of the molecular mechanisms that are responsible for this. Thus, it is often unclear whether specific epi-alleles are transmitted directly (i.e escaping methylation reprogramming) or whether other genetic or epigenetic mechanisms lead to re-establishment of methylation marks in the offspring (termed indirect transmission). Nevertheless, five studies have reported the inheritance of epi-alleles over two generations in a wide range of species and contexts, and two over three generations (see **Table 1**). Moreover, inherited methylation marks were found in multiple tissues, from the gonads to the liver or blood cells, and understanding why some epi-alleles appear in the offspring in specific tissues but not others is of major importance to the understanding of the role of heritable methylation differences in evolution.

Table 1: Summary of studies investigating the inheritance of environmentally-induced methylation changes in fish. This inheritance has been demonstrated in multiple species, following exposure to different types of environmental variables, and can sometimes be detected up to the third generation.

Multigenerational studies of methylation inheritance				
Variable	**Species**	**Tissue**	**Changes until**	**Study**
Maternal effects	*Oncorhynchus tshawytscha*	Whole embryo/ Whole body	F1	Venney et al., 2020
Temperature	*Salvelinus fontinalis*	Liver	F1	Venney et al., 2022
Endocrine disruptor exposure	*Menidia beryllina*	Whole body	F2	Major et al., 2020
Methylmercury exposure	*Danio rerio*	Sperm	F2	Carvan et al., 2017
Hypoxia	*Oryzias melastigma*	Ovaries	F2	Lai et al., 2019
Sulfidic water environment	*Poecilia mexicana*	Red blood cells	F2	Kelley et al., 2021
Salinity	*Gasterosteus aculeatus*	Gill	F2	Heckwolf et al., 2020
Ionizing radiation	*Danio rerio*	Whole embryo	F3	Kamstra et al., 2018
Cadmium exposure	*Danio rerio*	Gonads	F3	Pierron et al., 2021

Worryingly, methylation changes following exposure to anthropogenic environmental variables have now been shown to be inherited in fish. First, inland silverside (*Menidia beryllina*) exposed to different concentrations of endocrine disruptors, such as the insecticide bifenthrin or the androgen trenbolone, see their methylation levels change at several gene loci (e.g. *17β-hsd* and *Dnmt3a*), with some methylation differences observed in unexposed F1 and F2 generation (Major et al., 2020). Importantly, in some treatments/generations, DNA methylation changes in genes involved in chondrocyte differentiation, cartilage condensation and other molecular pathways were observed, and correlated with observed craniofacial and/or skeletal deformity phenotypes (Major et al.,

2020). Secondly, cadmium exposure in zebrafish induced hypomethylation of *foxl2a*, a gene involved in sexual development, in females and their female offspring up to the third generation, inducing a progressive feminization of the population (Pierron et al., 2021). Third, F0 zebrafish exposed to methylmercury (MeHg) during early development see their sperm methylation change and exhibit behavioural and vision-related defects, with sperm DMRs and phenotypic changes still observed in the unexposed F2 generation (Carvan et al., 2017). However, the DMRs identified in the F0 generation following exposure were different from the DMRs identified in the F2 generation, suggesting that the induced methylation changes observed for F0 were reprogrammed in the F1 generation, while a different set of loci underwent methylation changes in the F2 following ancestral exposure to MeHg (Carvan et al., 2017). These patterns highlight a gap in knowledge about the mechanisms responsible as well as the stability of inheritance of these methylation signals. Fourth, in zebrafish, persistent effects of ionizing radiation on DNA methylation at specific loci can be found up to the F3 unexposed generation (Kamstra et al., 2018). In a world increasingly altered by humans, results from these four studies are of great concern for the conservation of marine and lacustrine ecological systems.

Studies have also revealed the inheritance of methylation marks driven by what we classify as more natural environmental factors. Hypoxia impacts the ovarian methylome of the ricefish *Oryzias melastigma*, leading to methylation and expression changes in two gene clusters associated with apoptosis, and impaired oocyte development in the F2 generation (Lai et al., 2019). A study of *Poecilia mexicana* from sulfidic waters grown in a non-sulfidic laboratory environment for multiple generations revealed a propagation of methylation changes identified between sulfidic versus non-sulfidic environments until at least the F2 (Kelley et al., 2021). Last, while not directly mentioning patterns of inheritance, Heckwolf et al (2020) showed that threespine stickleback, acclimatized to different salinity over two consecutive generations (F2), showed progressively (over generations) more similarity to the methylation patterns observed in wild populations living at the target salinity level (Heckwolf et al., 2020). This is particularly important as it implicates that the same environmental change can induce predictable and stable methylation states over multiple loci, that may then be transmitted onwards. On another note, through the generation of multiple crosses of Chinook salmon (*Oncorhynchus tshawytscha*) using different sire and dam pairs, Venney et al., (2020) showed that DNA methylation at individual genes in the offspring at the eyed egg and alevin stages correlated with the dam, but not the sire. This suggests an influence of maternal factors on DNA methylation, possibly through escaping of methylation reprogramming for specific loci.

A number of studies have explored the potential for intergenerational inheritance of environmentally-induced methylation marks through the modification of the sperm methylome (**Table 2**). The sperm methylome of rainbow trout (*Oncorhynchus mykiss*) reared for their first year in hatchery differs from that of genetically related fish caught straight from the wild (Gavery et al., 2018; Nilsson et al., 2021), and a similar pattern was found in the sperm of reared atlantic salmon (*Salmo salar*) compared to its wild counterpart (Rodriguez Barreto et al., 2019). Interestingly, rearing-induced sperm methylation marks in coho salmon (*Oncorhynchus kisutch*) persisted in their germ cells even after 1.5 years in the ocean following release, suggesting that the methylation changes were induced early in the development of spermatogonia, possibly imprinting sperm-stem cells and providing individuals with sperm specific epigenetic memory (Leitwein et al., 2021). On the other hand, a study on variation in adult rearing showed that rearing-induced sperm marks can also be acquired later in life in atlantic salmon (Wellband et al., 2021). It would be interesting to compare, in a given species, the genomic locations of rearing-induced methylation marks arising in early development with those induced later in life.

Are the same regions impacted? Are early-life- or late-life methylation changes more likely to be transmitted to the offspring?

Table 2: Summary of studies investigating environmentally induced methylation changes in the gametes of different fish species.

Potential for methylation mark inheritance through gamete methylome changes		
Variable	**Species**	**Study**
Rearing	*Oncorhynchus mykiss*	Gavery et al., 2018
Rearing	*Oncorhynchus mykiss*	Nilsson et al., 2021
Rearing	*Oncorhynchus mykiss*	Leitwein et al., 2021
Rearing	*Salmo salar*	Rodriguez Barreto et al., 2019
Rearing	*Salmo salar*	Wellband et al., 2021

As an epigenetic mechanism sensitive to environmental variables, DNA methylation has the ability to induce short term acclimation (Venney et al., 2023), which is not possible through the slower Darwinian evolution of genetic variants. However, methylation also has the ability to impact phenotypic evolution when the methylomic changes persist long enough to be inherited by the offspring. Moreover, some epi-alleles might not be environmentally induced, but environmentally selected. In this case, epi-alleles act very much like regular alleles, on which selection can act (Shea, Pen, and Uller, 2011, Greenspoon, Spencer, and M'Gonigle, 2022). Interestingly, epialleles are in theory more plastic than genetic variants and could return to their original condition or keep on changing with the environment if it continues to fluctuate.

However, the inheritance of environmentally-induced methylation marks means that the developing offspring may have to deal with two sets of (possibly contradicting) environmental cues, the environmental signals imprinted by its parents and influences received from its current surroundings. In cases of contradiction, which cue takes precedent and why? A recent experiment by Venney et al. (2022) on brook charr (*Salvelinus fontinalis*) addresses this directly. They used bisulfite sequencing of DNA extracted from male liver tissues, aimed to disentangle the impact of temperature during parental sexual maturation and the impact of temperature during rearing on the offspring's methylome. The strongest clustering of samples was based on adult rearing temperature, suggesting that the thermal regime experienced by the parents had the highest effect on the offspring's methylome. This variable was then followed by family effects, whereas the offspring's thermal regime during development had less impact. Future experiments of this sort will be vital to understand the interplay between epigenetic signals transmitted to the offspring from their parents or past generations, and their own interpretation of the present environmental conditions.

Methylation differences induced by environmental changes may act through the activity of key pathways and enzymes responsible for the maintenance and establishment of said marks. Thus, it is of interest to examine the impact that environmental variables have on the expression and function of these genes. Fellous et al (2022) investigated the impact of rearing temperature on the expression of chromatin-modifying enzymes, including genes coding for Dnmts and histone modification enzymes, in the gametes of marine stickleback (*Gasterosteus aculeatus*). Interestingly, while histone modification factors were present in both oocyte and sperm, maintenance and *de novo* Dnmts were only expressed in the oocyte. This suggests that the remethylation phase observed at the morula stage is likely under maternal control. For parents reared at higher temperature, the expression of factors involved in the establishment or inheritance of epigenetic marks (either Dnmts or histone

modification factors) increased in oocytes and decreased in sperm, and embryos from high-temperature crosses showed different global methylation reprogramming dynamics after the morula stage compared to embryos reared at an ambient temperature (Fellous et al., 2022). The influence of different environmental factors and animal condition on expression of the methylation machinery in gametes, and related reprogramming events, is something that could be further explored and included in future investigations on transgenerational inheritance of methylation marks. This is also important in terms of evolutionary studies as, as we will see in the next section, genes coding for the enzymes part of the methylation/demethylation machinery have diverged substantially in teleosts.

1.2.3 Influence of Genetic Variation on DNA Methylation

While epigenetic processes are usually defined as those inducing changes in gene expression without involving changes in the genetic code (Hamilton, 2011), it is undeniable that epigenetic and genetic mechanisms are constantly intertwined (Cavalli and Heard, 2019). Several recent studies have assessed the relationship or contribution of genetic variation to epigenetic differences in fish populations. Studies in both Atlantic salmon (Burgerhout et al., 2017) and Rainbow trout (Lallias et al., 2021) have shown that the effects of incubation temperature on the DNA methylome depends to an extent on genetic background. Similarly, a plant-based diet affects the global hepatic methylome in Rainbow trout to different extents by genetic background (Marandel et al., 2022). Comparisons between two sympatric freshwater fish species (*Gobio occitaniae* and *Phoxinus phoxinus*), indicate that the majority of epigenetic differences between these species are caused by genetic variation (Fargeot et al., 2021). Only a third of differentially methylated sites between stickleback populations locally adapted to different salinities were inducible by salinity manipulation, which suggests that the remaining 62% of pop-DMS are either regulated by *cis*- or *trans*- genomic factors (Heckwolf et al., 2020). The studies above compared distinct genetic groups or lineages but the heritability and inheritance of DNA methylation can only be explored with crosses. Furthemore, an experiment that involved experimental hybrids of two ecotypes of threespine stickleback, (Hu et al., 2021) showed that 24-35% of the global variation in methylation was explained by additive genetic variance.

Together these studies indicate that genetic background of species or populations may influence how organisms can respond to environmental factors, by altering patterns of DNA methylation in their cells, and is an important reminder that genetic variation must be controlled for when investigating the impact of the environment on the methylome.

1.3 Duplication Events and Methylation Changes

Duplication of genes or whole genomes can fuel evolution (Ohno, 1970) as when duplicated genes diverge functionally, through sub- or neo-functionalization (Force et al., 1999). After whole genome duplication, rediploidization eventually occurs, although this can take a long time and can vary across genomes (Wolfe, 2001; Gundappa et al., 2022). This new genetic material has the potential to contribute to rapid divergence and speciation, even though rapid speciation (which does not necessarily need these duplications) sometimes occurs long after whole genome duplication events. For instance, the teleost clades Ostariophysi and Percomorpha appeared ~100-130 Ma, or about ~220-250 million years after the teleost specific WGD (Santini et al., 2009). Robertson et al (2017) proposed a model where speciation happens prior to rediploidization (at least

partially), thus allowing unique ohnolog (paralogs resulting from WGD) divergence in sister lineages, possibly due to specific selective pressures. Studying the Ss4R they found one quarter of each salmonid genome possesses ohnologs that have evolved independently multiple times, and that these ohnologs are enriched in specific GO terms possibly impacting smoltification compared to ohnologs that diverged quickly in the salmonid ancestor (Robertson et al., 2017).

With teleosts as the focal group, the extent to which 1) genes from the methylation machinery were impacted by these WGD events, 2) methylation signals accompanied duplicated genes and 3) methylation signals interact with TEs to generate new functions will be discussed in this section.

1.3.1 Duplication of Methylation Machinery Genes

The duplication of genes of the methylation machinery (including demethylation) could have impacted the potential diversity of methylation marks produced during the embryo- or PGC- reprogramming and the fine-tuning of methylation signals during development, ultimately leading to phenotypic differences and speciation.

As mentioned in the first section of this introduction, in eukaryotes, Dnmts can be separated into two classes: maintenance Dnmts which play a role in conserving methylation states through cellular divisions and *de novo* Dnmts which target previously non-methylated regions. The human genome encodes 5 different Dnmts (Lyko, 2018). Dnmt1 plays the maintenance role and is recruited to replication forks by the multi-functional multidomain protein UHRF1 which has affinity to DNA containing hemi-methylated CpG (Bostick et al., 2007; Bronner et al., 2019). At the replication fork Dnmt1 interacts with several other proteins, most notably the replication coordinator proliferating cell nuclear antigen (PCNA) (Iida et al., 2002; Jimenji et al., 2019). Dnmt2 possesses a high structural similarity with other Dnmts, but acts on a different substrate as it is predominantly a tRNA methyltransferase (Goll et al., 2006; Jeltsch et al., 2017; Lyko, 2018) and, as such, will not be discussed further. Dnmt3A and Dnmt3B are both *de novo* methyltransferases (Okano et al., 1999), and their *de novo* activity is assisted by the last type of Dnmt, Dnmt3L, which does not possess catalytic Dnmt activity but interacts with Dnmt3 proteins through its C-terminal domain (Suetake et al., 2004; Z.-X. Chen et al., 2005) and increases their affinity for DNA by recognizing and binding to unmethylated H3K4 (the 4th lysine on the histone protein H3) (Ooi et al., 2007; Jia et al., 2007).

While less studied than in mammals, recent studies have investigated the methylation machinery in teleosts, highlighting differences both between these vertebrate groups and among fishes. Similarly to mammals, *Dnmt1* is only found as a single copy in the zebrafish genome, but two paralogs exist for trout and salmon, likely resulting from the salmonid specific WGD event (Best et al., 2018). Similarly, the *Salvelinus sp*. genome suggests two paralogs of *Dnmt1* in Arctic charr. The *de novo* methyltransferase genes are more numerous in fish, as two paralogs of *Dnmt3a* and four paralogs of *Dnmt3b* are found in zebrafish alone, with even more orthologs of these genes found in trout and salmon (Best et al., 2018). Despite being close relatives, rainbow trout and Atlantic salmon also exhibit differences, with 3 paralogs of *Dnmt3a* in trout and 4 in salmon (Best et al., 2018). The *Salvelinus sp.* genome suggests three paralogs of *Dnmt3a* (one *Dnmt3aa* gene and two paralogs of *Dnmt3ab*) as well as five paralogs of *Dnmt3b* (two *Dnmt3ba*, two *Dnmt3bb.1*, and one *Dnmt3bb.2*) for Arctic charr. While *Dnmt3aa* and *Dnmt3ab* isoforms had similar and additive functions in modulating thermal plasticity in zebrafish (Loughland, Little, and Seebacher, 2021), *Dnmt3b* paralogs exhibit different spatial expression during early

development in this species, likely reflecting sub-functionalization of these genes (Campos, Valente, and Fernandes, 2012). Similarly, Nile tilapia (*Oreochromis niloticus*) possess multiple *Dnmt3* genes showing different expression levels across tissues (F.-L. Wang et al., 2018) with, for example, only one of the *Dnmt3a* paralogs (*Dnmt3aa*) playing a role in maintaining gametogenesis (F. Wang et al., 2021). The differential expression of *Dnmt3b* paralogous genes between oogenesis and spermatogenesis in rainbow trout is another example of neo- or sub-functionalization of Dnmt genes following WGD in teleosts (J. Liu et al., 2020). Teleosts do not seem to possess any homologs of *Dnmt3L* (Best et al., 2018) suggesting possible mechanistic differences in the way their Dnmt3 *de novo* methylation enzymes are recruited to chromatin.

While passive DNA demethylation occurs through cell division via an absence (or lack) of maintenance Dnmt1 activity, active demethylation also exists. Many molecular pathways have been hypothesized to play a role in active DNA demethylation, for example AICDA/APOBEC (Nabel et al., 2012), or even the Dnmts themselves (C.-C. Chen, Wang, and Shen, 2013; Chatterjee et al., 2018). The best understood players are the TET-family enzymes which drive hydroxylation/oxidation of 5mC into 5-hydroxymethylcytosine, 5-formylcytosine (5fC) or 5-carboxycytosine (5caC) (Xu and Wong, 2015), eventually leading to base-excision repair by the thymine DNA glycosylase (TDG) and the insertion of an unmethylated cytosine to replace the previously methylated one (Kohli and Zhang, 2013).

Similar to methylation genes, paralogs of demethylation genes also differ between teleost species. For example, while *TET1*, *TET2* and *TET3* are found in single copy in the zebrafish genome, two paralogs of *TET1* and *TET3*, as well as 3 paralogs of *TET2*, are found in trout and salmon. Similarly, *TDG* genes were observed in two copies for zebrafish and 4 copies for trout and salmon (Best et al., 2018). This most likely impacts the way demethylation events such as the ones required to reset the methylome in the zygote or PGCs are orchestrated in salmonids. It would be interesting to explore the expression of these genes across ontogeny, in different tissues and the germ cell lineages and primordia.

The functional relevance of duplications and divergence of genes related to the methylation/demethylation machinery needs to be further studied in teleosts. Did they have a role in shaping development and phenotypic differences throughout the taxonomic group, or do they mainly reflect subdivisions of ancestral functions?

1.3.2 Methylation and Functional Divergence Following Gene Duplication

It was suggested that DNA methylation may play a role in the subfunctionalization of newly duplicated genes (Rodin and Riggs, 2003), but only in the past decade have technological advances allowed the investigation of such concepts. Recent studies in a range of different species have now found interesting positive correlations between recently duplicated genes and methylation. Whether the changes in methylation follow subfunctionalization, or preempt it, is an interesting question. In humans, evolutionarily recent paralogs are often hypermethylated, and differences discernable between the paralogs. Moreover, different promoter methylation levels between paralogs correlate with their different expression levels (Keller and Yi, 2014). Similar results were observed in stickleback: the promoters of recently duplicated genes were found to have higher methylation levels than older ones, as did the genes found in multiple copies compared to non-CNVs (Huang and Chain, 2021). Overall, transcription levels of these highly

methylated genes were negatively correlated with their methylation levels, an indication of silencing to compensate for extra gene copies.

Sub-functionalization is most readily observable when it impacts the spatial expression of duplicated genes, and whether changes in methylation levels associate with this process is of interest. In humans, the tissue-specificity of promoter methylation levels between duplicate partners was shown to increase with evolutionary time (Keller and Yi, 2014), which most likely reflects neutral divergence. It would be of interest to evaluate whether this correlation stands true in teleosts. In stickleback, younger genes were recently found to show more tissue-specific expression patterns (Schmitz, Chain, and Bornberg-Bauer, 2020). However, Schmitz et al (2020) were investigating novel transcripts as a whole, not distinguishing between the recent genes that arose from duplication events and the ones that arose *de novo*. Thus, more work is needed to compare these trends between human and teleost.

These results suggest that increased methylation follows gene duplication in teleosts (based on the stickleback data cited above), potentially to reduce or suppress the expression of new genes in order to rebalance gene dosage, as in mammals (Chang and Liao, 2012). Specific examples of this process exist in teleosts; the global methylation levels increased for *Sox9a*, *Amh* and *Cyp19a1a* while their global expression reduced after autotetraploidization of *Carassius auratus* (X. Huang et al., 2020), allotetraploid hybrids of *Carassius auratus* and *Cyprinus carpio* show hypermethylation of genes related to metabolism or cell cycle regulation compared to their diploid parents (Xiao et al., 2013) and methylation-mediated gene dosage compensation was observed in allotriploid fish compared to their diploid parents (Ren et al., 2022). Then, as time passes, duplicate genes will gradually lose methylation and diverge either functionally, spatially (in different tissues) or both (not necessarily in that order).

The processes of sub- or neo-functionalisation, and their interaction with methylation changes, have likely happened in different ways for distinct genes across the teleost group. Therefore, there is plenty of room for exploration of methylation differences between duplicate partners by tissue and taxa, and studies in this field are needed to determine if changes in methylation preceded or accompanied the divergence of newly duplicated genes in teleosts.

1.3.3 Methylation and Transposable Elements

Transposable elements (TEs) can be divided into two groups based on their strategies for transposition: Class I TEs (retrotransposons) require reverse transcription in order to transpose, while class II TEs (DNA transposons) do not (Wicker et al., 2007). Furthermore, TEs can be autonomous or non-autonomous, depending on whether they encode the proteins required for their own transposition (Wicker et al., 2007). TEs constitute a large proportion of eukaryotic genomes, for instance around 40% of the human genome (Venter et al., 2001), although only a minority of those TEs are active. Their abundance and relative proportions in teleost genomes vary between species, with zebrafish possessing both a high abundance (55%) and rich diversity of TE compared to tilapia (24%), stickleback (14%), fugu (7%) and tetraodon (6%), for which transposon diversity is not as rich (Chalopin et al., 2015).

The multiplication of TEs in genomes is suppressed by various mechanisms, DNA methylation being of particular importance (Slotkin and Martienssen, 2007) (Karimi et al., 2011), especially in the context of this thesis. The demethylation events that take place

during early development and in germ cell development offers an opportunity for active TEs to jump and, most importantly in the case of PGCs, to be transmitted to the next generation in a new location (to multiply). Interestingly, alternate mechanisms for silencing TEs in gametes, in the absence of DNA methylation control, can be found such as piRNAs (Carmell et al., 2007).

Jumping TEs often have deleterious effects for their hosts: new insertions of elements in coding sequences and other functional genetic elements can lead to gene inactivation or changes in gene expression (Elbarbary, Lucas, and Maquat, 2016). These new alleles can also lead to non-lethal, even adaptively favourable phenotypic changes. One such example has been observed in medaka where spontaneous excision of a Tol2 element inserted in a tyrosinase pigmentation gene gives rise to different phenotypes (Koga et al., 2006). TE silencing by DNA methylation (or other mechanisms), subsequent inactivation of the TE and particular adaptive genetic changes can eventually lead to a process coined molecular domestication in which TE sequences acquire new functions (Sinzelle, Izsvák, and Ivics, 2009). For example, TEs are responsible for maintaining telomeres in Drosophila (Pardue and DeBaryshe, 2011), and a notable proportion of human protein coding genes contain fragments of TE sequences (M. Wu, Li, and Sun, 2007). TEs thus have the potential to inactivate genes and generate new functions, and thus have likely contributed to evolutionary divergence and the phenotypic differences observed in teleost today (Casacuberta and González, 2013; Monsen, 2022). As an example, the salmonid sex determining SdY gene is flanked by transposons and has been, throughout the evolutionary history of salmonids, transposed or translocated on different chromosomes (Bertho et al., 2021).

Large scale events such as WGD (Marburger et al., 2018) and hybridization (Laporte et al., 2019) (S. Liu et al., 2022) may lead to TE expansion, if general or specific TE suppression is impaired. In salmonids rediploidization post-Ss4R happened in concert with large genomic rearrangements, and it has been suggested that this WGD event coincided with an expansion of transposable elements in the different salmonid genomes (Lien et al., 2016). Thus, both WGD and episodes of TE expansion may have generated diversity for adaptive divergence. Furthermore, DNA methylation mechanisms involved in silencing newly inserted elements might impact the expression of nearby loci. Indeed, the silencing of TE elements by DNA methylation could have collateral effect in situations when TEs landed close to promoters or enhancers, or when they interact with gene sequences located sterically close to them in terms of DNA quaternary structure. The most famous of these examples is the retrotransposon-mediated silencing of the agouti gene in mice (Morgan et al., 1999) but a similar example is known in fish. In zebrafish, differential methylation of a transposable element located upstream of the *mef2ca* gene correlated with the gene's expression, leading to increased variance in development of several skeletal structures (Nichols et al., 2016). In this example, stochasticity in DNA methylated suppression (both genotypes had the TE close to the gene) is the reason for highly variable expressivity of a particular allele.

While transposable elements can escape methylation reprogramming during periods of genomic stress such as duplication events, exposure to different environmental variables can also have this effect. Indeed, lower DNA methylation levels of the long interspersed nuclear element (LINE) UnaL2 correlated with the up-regulation of retrotransposons in the liver of European eels (*Anguilla anguilla*) when exposed to high water contamination levels (Pierron et al., 2019). This illustrates that in addition to duplication events, environmental variables impacting DNA methylation are also likely to have contributed to the differential TE content observed in teleosts.

The potential influence of TE insertion site variation on DNA methylation (and variation in DNA methylation on TEs), on the function, mean or variability of expression of specific genes could be investigated in fish, particularly in salmonid species due to the relatively recent WGD event that both generated additional paralogous genes but also may have led to increased TE mobility and structural rearrangements in their genomes.

1.4 The Arctic Charr Morphs of Thingvallavatn

While comparative studies of sister lineages or related species can offer some insights into the evolutionary mechanisms that led to their ancestral divergence (Hamilton, 2021), considering the more recent evolutionary processes is also important in order to identify potential signals which might have been lost to time in broader taxonomic contexts. Examples allowing for the observation of molecular signals during the early stages of speciation or divergence are rare, although present (van Schooten et al., 2020). Indeed, sympatric populations of a given species, gradually experiencing a reduction in gene flow between individuals due to reproductive (pre- or post-zygotic, temporal or spatial) barriers, are perfect candidates to study the interplay between genetic, epigenetic and transcriptomic changes, reproductive isolation and eventually speciation. Such models can be seen in finches (Huber et al., 2007), but are quite common in fish, with cichlids (Kautt et al., 2020) and salmonids (Salisbury and Ruzzante, 2022) being prime examples. As was mentioned above, salmonids underwent a fourth WGD event some 85 mya, which may have contributed to diversity in this group. Among the salmonids, the *Salvelinus* genus is particularly interesting due to its high level of within species polymorphism, found both in allopatry and sympatry, which shows indication of parallel phenotypic evolution (Salisbury and Ruzzante, 2022), and makes it particulary suitable for addressing questions about the influence of genomic and epigenomic variation on adaptive divergence.

Out of the different species in this genus, the Arctic charr (*Salvelinus alpinus*), exhibits drastic cases of sympatric phenotypic polymorphism in different locations around the globe. This species invaded multiple freshwater systems after the last glacial retreat and is now found in many Arctic and sub-Arctic lakes where it has in some cases lost its anadromy. For instance, sympatric charr morphs are observed in Russia (Markevich, Esin, and Anisimova, 2018), Norway (Hindar and Jonsson, 1982) and Canada (Arbour, Hardie, and Hutchings, 2011).

Out of the different Arctic charr systems displaying polymorphic sympatry, few have been as intensely studied as the Arctic charr populations of Lake Thingvallavatn (Jonsson et al., 1988; Sandlund et al., 1992), one of Iceland's biggest lakes. The invasion of *S. alpinus* in Thingvallavatn is thought to have happened ~10 000 years ago (Kapralova et al., 2011), which equals ~3000/3500 generations for this species in which sexual maturity is reached after 3 years. In this system, four different morphs coexist: two benthic morphs Small Benthic (SB) and Large Benthic (LB), and two limnetic morphs Planktivorous (PL) and Piscivorous (PI). As suggested by their names, these morphs differ by their habitat use (Sandlund et al., 1987), diet (Malmquist et al., 1992) and size. Their reproductive behaviours also differ, with morphs spawning not only at different sites in the lake but also at slightly different times of the year, with in particular the LB charr spawning earlier than the other three morphs (late July - mid August) around cold inflowing springs on the northern/eastern bank of the lake (Skúlason, Noakes and Snorrason, 1989; Skúlason et al., 1989). Perhaps most importantly, their diets differ, with some morphs preferring to eat snails in the lava and rocky benthic areas (LB and SB) while others feed on plankton in the

water column (PL), or capture threespine stickleback in the nitella belt (PI), and it is of little surprise that their trophic apparatus also displays contrasting structures. Externally, limnetic morphs (PL and PI) display a terminal mouth and pointed snout while benthic morphs (SB and LB) possess a subterminal mouth and blunt snout (Sandlund et al., 1992). Interestingly, hatching seemed to correlate with an increase in allometric shape changes in these morphs, suggesting that molecular events around the hatching period may be partly responsible for the contrasting phenotypes observed in nature (Kapralova et al., 2015).

These external characteristics have provided easy identification of these morphs in the field for more than 30 years. Recent geometric morphometric analyses to assess variation in shape and size of six upper and lower jaw bones in this system have unveiled the anatomical foundations behind these external differences. Indeed, four bones (dentary, articular-angular, premaxilla and maxilla) display bone shape divergence between the morphs, along a benthic-limnetic axis, with for instance benthic morphs possessing a taller and more compact dentary with a shorter dentary palate (Jónsdóttir et al., 2023).

While SB and PL morphs can hybridize in common garden (Horta-Lacueva et al., 2021), showing intermediate, transgressive and parental-like phenotypes (de la Cámara et al., 2023), gene flow between morphs in this population is reduced, most likely due to both the pre-zygotic spatial and temporal reproductive barriers mentioned above, as well as post-zygotic barriers such as fertilization failure or early-development mortality (Horta-Lacueva, 2022). Importantly, genetic differences between the morphs have been observed. Single nucleotide polymorphism (SNP) analysis on RNA-seq data from three of the four morphs (PL, LB and SB) showed a clear separation between the three of them (Guðbrandsson et al., 2019), with markers separating the morphs located accross the genome (Guðbrandsson et al., 2019; Brachmann et al., 2022). Interestingly, SNP derived from ddRAD- and WGS data from all four morphs showed the PI to be more genetically ambiguous, with most individuals genetically close to PL while others showed varying levels of admixture with LB (Xiao et al., unpublished data, and **Paper II**). Matching the results from studies of genetic variation, transcriptomic studies had previously identified more than a thousand differentially expressed genes between PL, SB and LB morphs during early development (Guðbrandsson et al., 2018). In more targeted studies, qPCR on candidate genes (identified by surveying 2018 data and an earlier transcriptome, see Guðbrandsson et al., 2016) revealed some genes involved in bone development to be differently expressed between the morphs. Two dramatic examples were the genes *Mmp2* and *Sparc*, both associated with matrix remodeling, which differed in expression between benthic and limnetic individuals (Ahi et al., 2013). Whether differential expression of these hundreds of genes is caused by sequence divergence or because they are located downstream of some regulator being modulated is an interesting question, which I address in **Paper II**. Regarding epigenetic studies, miRNAs were found to be differentially expressed between the SB morph and aquaculture charr (displaying a limnetic craniofacial phenotype) during embryonic development (Gudbrandsson et al., 2016; Kapralova et al., 2014). However, no histone- or DNA modification studies have so far been performed in this system, creating somewhat of a gap between genetic and environmental variables. The work presented in this thesis (mainly **Paper I**) provides the first look into epigenetic differences between these sympatric Arctic charr morphs.

2 Aims of The Thesis

DNA methylation is amongst one of the most studied epigenetic mechanisms. Through its impact on chromatin accessibility and gene expression, it helps orchestrate early embryonic development and, through its sensitivity to environmental variables, is an intermediate between the environment and the genome. However, the impact DNA methylation (and its machinery) has had on evolutionary processes is still poorly understood. To investigate such questions, teleost fishes seem like the perfect candidates, due to their extensive species richness, phenotypic diversity and due to the WGD events they have undergone. And amongst them, closely related sympatric populations, such as the Arctic charr morphs from Lake Thingvallavatn, allow for the observation of recent evolutionary processes. This thesis brings a new point of view to the numerous studies on the Thingvallavatn Arctic charr system, as it is the first to investigate DNA methylation patterns in these sympatric morphs.

Its aims are to:

- I: Identify methylome differences between Arctic charr morphs from Lake Thingvallavatn, at different early embryonic stages.
- II: Assess the relationship between methylation and expression of genes of interest.
- III: Identify overlaps – or lack thereof – between peaks of methylation differentiation and peaks of genetic differentiation between the morphs.
- IV: Identify overlaps – or lack thereof – between differently expressed genes and peaks of genetic differentiation between the morphs.

Additionally:

- V: Provide a literature review summarizing the extent of DNA methylation reprogramming in teleosts, both in the zygote and PGCs.

3 Methods

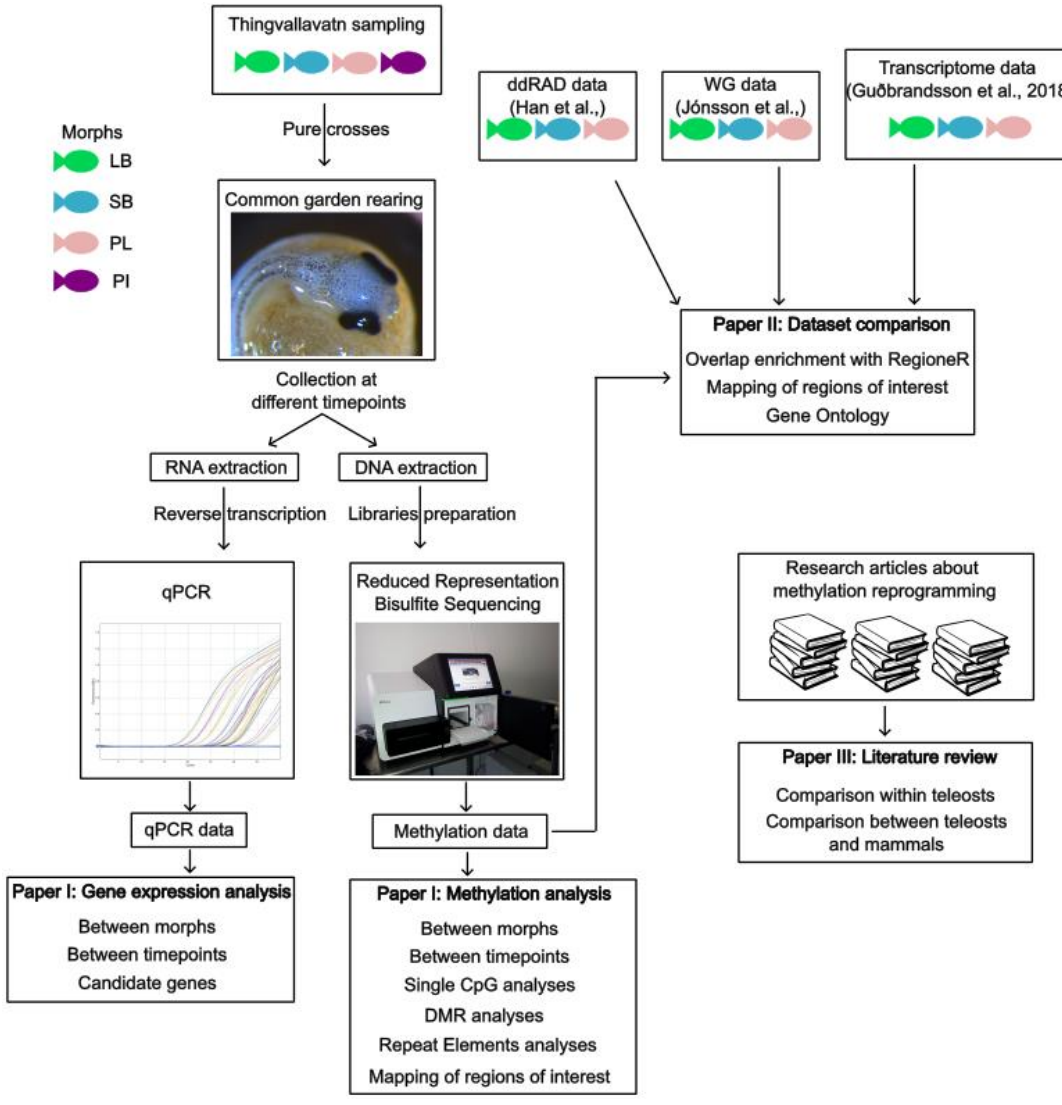


Figure 1: Description of the work performed for each part of the thesis (Papers I to III)

3.1 Methylation, and Related Expression, Differences Between Arctic Charr Morphs

Methods are described in detail in **Paper I**.

Briefly, we set out to investigate methylomic differences between Arctic charr morphs from lake Thingvallavatn. Sexually mature adults from each of the four morphs were sampled during spawning period and milt and eggs used to create pure morph crosses, which were reared in common garden under the same conditions. This common garden rearing was important, as we wanted to identify methylation differences being either directly inherited or genetically induced, and not methylation differences caused by differences in environment. The different charr morphs spawn at different sites in the lake, with some of these sites being fed by glacial streams, creating different environments which could potentially alter the methylome of the developing embryos. Investigating direct environmentally induced methylation differences between morphs is also of great interest but was not the aim of this paper. Reduced representation bisulfite sequencing was used to obtain methylation data for 48 individuals, from four morphs at four different timepoints, spanning stages from gastrulation to late organogenesis, which was analyzed using several different methods. Last, as methylation can impact gene expression, we then set out to investigate the expression of 14 candidate genes located close to DMRs, using qPCR.

3.2 Overlaps Between Regions of Genetic, Transcriptomic and Methylomic Differentiation

Methods are described in detail in **Paper II**.

Briefly, we used previously generated datasets: ddRAD data, WGS data, transcriptome data and RRBS data (see **Paper I**) to identify possible congruence between signals at different biological levels (genomic, transcriptomic, epigenomic). Because each of these datasets was originally generated for specific purposes, and not with the overall aim of comparing them all together, this task was arduous. For instance, many of the datasets had low genomic coverage, leading to overlap comparisons being impossible in some cases. Overlap analyzes were performed with the help of the RegioneR R package and follow up analyzes and plotting were made with custom scripts.

3.3 Literature Review: Differences in Methylation Reprogramming Patterns Amongst Teleosts and Comparison to Mammals

This literature review (**Paper III**) did not require any lab work or computational analysis.

However, a number of developmental stages key to the methylation reprogramming processes, for example the beginning and end of migration period for the PGCs, needed to

be inferred. Indeed, research groups evaluating methylation reprogramming patterns often describe developmental stages relative to the time of fecundation (i.e. 5 days post fecundation - 5dpf, 3 hours post fecundation - 3hpf, etc..) and, as developmental stages vary in timing and duration between species, it can be difficult to determine the precise correspondence from one species to the next. Thus, research articles describing the precise embryonic development of each of the investigated species were used in conjunction with methylation reprogramming articles to provide the most accurate comparison between species. Moreover, the methodologies used by the different research groups to assess global CpG methylation levels differ, and comparisons between the different technologies at play needed to be established in order to disentangle whether differences in results between studies stemmed from technological biases.

4 Summary of the Papers

In **Paper I**, we identified differences in methylation signals between developmental stages for each of the morphs. This was expected, as it is well established that methylation is dynamic during early embryonic development. More specifically, we identified a decrease in overall methylation over time for all morphs, between gastrulation and late organogenesis. It is important to note that this is not representative of the overall state of the methylome as RRBS mainly focuses on CpG islands, and more global methylation analyses at these (and earlier) timepoints are needed to determine whether this reflects a wave of methylation reprogramming. A less expected result was the strong methylomic difference observed between the morphs, which was most pronounced between benthic and limnetic individuals. These differences were observed both in intergenic regions and close to gene bodies such as *GLI3-like*, *NFIX* and *MEGF9*. In fact, DMRs between both morphs and timepoints were present all over the genome. Interestingly, these DMRs were enriched in RNA sequences such as tRNAs and 5.8S rRNA and, although underrepresented in DMRs, a number of transposons were also found to be differentially methylated between morphs or timepoints. Multiple histone genes, located in histone gene clusters, also showed differences in methylation between morphs and timepoints.

We then used qPCR to assess the expression levels of 14 genes lying close to, or overlapping with, DMRs. While these genes of interest showed differences in both methylation and expression, these differences did not correlate in most cases. This might be due in part to our analysis in whole embryos, which blends methylation and expression levels from different cell types and tissues. However, for four genes a correlation was observed. It has been reported that the impact of DNA methylation on gene expression depends on the location of the methylated site with regards to gene features. For instance, methylation in promoters tends to repress transcription while methylation in gene bodies often enhances it. Our results agreed with this statement as *MEGF9* and *MPP3*, for which the DMR was located in promoter region, showed a negative correlation between methylation and expression, while *ARHGEF37-like* and *H3-like*, for which the DMRs spanned exon features, displayed a positive correlation.

In **Paper II**, we investigated possible overlaps between genomic, transcriptomic and methylomic signals. Centrally, the somewhat limited WGS data available was found to be highly congruent with ddRAD-seq data, encouraging us to use this WGS data for comparisons with other datasets. Regions of methylation and genetic differentiation were then found to be overlapping slightly more than expected, although this signal needs to be confirmed by further studies. DE genes did very strongly overlap with regions of genomic differentiation identified through WGS although this was not the case with regions of differentiation identified through ddRAD-seq. Surprisingly, only SB-specific DE genes were found to be enriched in regions of SB-specific genetic differentiation. The same did not apply for the PL- and LB-specific genes, suggesting a higher proportion of transcriptomic changes related to *cis*-genetic variation in the SB morph, while *trans* acting factors might lie high in the developmental hierarchy differentiating the LB and PL morphs. Of the DE genes showing an SB- specific pattern and located in SB- specific regions of differentiation, *brorin* is of upmost interest, as it is known to interact with bone morphogenetic protein 2 (BMP2) and could be responsible for some of the observed skeletal differences between the SB and the other morphs. Moreover, three genes (*ALP1-*

like, *PTCHD3-like* and *YME1L*) were, at once, differently expressed, located in regions of genetic differentiation and close to differentially methylated CpGs and their biological importance could be further investigated.

An interesting methodological note was revealed in the ddRAD-seq data, which contained the most reliable genetic information (because of a large number of samples). Markers from this set proved to be biased towards a certain subset of CpGs, most likely because of the methylation sensitivity of one of the enzymes (*Ape*KI) used during the library preparation. Also, as both datasets are reduced representations of the genome, the ddRAD data could not be directly compared with methylation data.

All in all, while overlaps between transcriptomic, genomic and methylomic datasets were not always enriched compared to random distributions, overlaps are still present and genes located in the regions spanned by these different biological signals need to be further investigated.

In **Paper III**, we described the known state of methylation reprogramming in different species of teleosts, and compared the observed patterns with mammalian trends. This was done both for the first wave of reprogramming happening at fecundation and for the second wave of reprogramming in PGCs. The comparison revealed some differences within the teleost group. For instance, while zebrafish PGCs showed very little reprogramming, this was not the case for medaka PGCs, for which the extent of demethylation and remethylation was more important. However, whether this demethylation in medaka PGCs happened during migration of the PGCs to the genital ridge, or following their arrival at this location is still to be established. Importantly, previously identified differences in zygotic DNA methylation between teleost species need to be reassessed: the methods used by the research teams investigating methylation reprogramming post-fertilization in the zygotes of medaka, stickleback and mangrove rivulus suffer from a bias caused by their inability to distinguish genomic from mitochondrial DNA (which is mainly unmethylated) in the oocyte and early cleavage stage. Because of this bias, it is unclear whether the non-reprogramming of DNA methylation in the zebrafish zygote is, as has sometimes been previously suggested, an exception, or whether all teleosts differ in this regard from mammals, for which important reprogramming happens post-fertilization. The presence of TET enzymes expression in the medaka zygote is an indication that some demethylation probably happens in the medaka zygote, separating it from zebrafish, although to which extent this demethylation event happens is unknown.

Last, it is interesting to note that gametic methylation levels vary between species, although sperm methylation levels are always higher than those of the oocytes, and that maternal and paternal methylomes do not always undergo similar reprogramming. For instance, in zebrafish, the maternal methylome gets modified (demethylated or remethylated depending on locus) to match paternal patterns, a process that was recently understood to involve, at least in part, H2AFV placeholder nucleosomes.

5 Conclusion, future perspectives and discussion

Our methylation studies in Arctic charr highlight amazing differences of methylation signals both between developmental timepoints and between morphs, showcasing a great example of methylome divergence in closely related individuals. The fact that methylation differences between morphs overlap more than expected by chance with regions of genetic differentiation also hints once again to the interplay between the evolution of genetic and epigenetic signals.

Future directions to follow up on this project are numerous:

First, the improvement in methylation sequencing technology (i.e. Nanopore sequencing) and single- cell -omics could be embraced to investigate methylation differences between Arctic charr morphs at the tissue- or cell-level. Indeed, our current study was made on whole embryos, and while the presence of strong methylome differences is indicative of general differences in epigenomics between these morphs, more subtle tissue-specific methylation differences could not be detected. This information would be of great interest to help explain variations in craniofacial development between the morphs.

Second, performing WGS with a higher number of samples and at higher depth could be a way to follow up on the overlap analyses shown in **Paper II**. While we were able to show some enrichment of DMRs and DE genes in regions of genetic differentiation, these regions only had a resolution of 100kb and it is therefore impossible to know whether the SNPs themselves were in *cis*- or *trans*- compared to these epigenomic/transcriptomic signals. Moreover, adding more types of data to this comparison study would be of interest. Investigating possible overlaps between different types of epigenetic data such as histone modifications, or other less common DNA modified bases (5hmC, 5fC, 5CaC) could enhance the results obtained in **Paper I**. The setup of an Arctic charr cell line has, over the years, proven hard to establish, limiting single-cell level analysis, but performing ChIP sequencing from trypsinized tissue to investigate histone modifications could be an alternative.

Third, investigating duplication of methylation machinery genes, as well as their expression, in salmonids (post Ss4R) including charr would be a great addition to the field.

Fourth, transgenesis studies could be done to further investigate the role of specific candidate genes highlighted in **Paper I** and **Paper II**. For instance, the use of Crispr I or Crispr A to inhibit or activate gene expression, as well as other knockout or knockdown methods, could help identify whether these genes are key actors in the establishment of phenotypic traits during development, or simply part of a co-expression network. While this might be difficult to setup in Arctic charr due to their long generation time, investigating the impact of candidate gene expression changes in zebrafish, which has been extensively used for transgenesis studies, could prove useful.

Last, the literature comparison we established in **Paper III** highlights some differences in methylation reprogramming mechanisms between teleosts, and summarizes technological

biases due to the use of different methods of measuring global DNA methylation, which are an important issue. Unfortunately, no research groups have so far investigated methylation reprogramming in salmonids, let alone charr. Doing so would provide valuable information to the field. Originally set-up as a side project, I attempted to measure global methylation levels in very early embryonic stages in charr (the first hours/days post fertilization) using an ELISA assay (which I now realize would have suffered from the mitochondrial bias described in **Paper III**), to try to detect possible waves of demethylation. However, technical issues (possibly linked to the yolk content, sensitivity of the spectrophotometer, or both) prevented me from getting usable data.

References

Ahi, Ehsan Pashay, Jóhannes Guðbrandsson, Kalina H. Kapralova, Sigríður R. Franzdóttir, Sigurður S. Snorrason, Valerie H. Maier, and Zophonías O. Jónsson. 2013. 'Validation of Reference Genes for Expression Studies during Craniofacial Development in Arctic Charr'. PloS One 8 (6): e66389. https://doi.org/10.1371/journal.pone.0066389.

Anastasiadi, Dafni, Anna Esteve-Codina, and Francesc Piferrer. 2018. 'Consistent Inverse Correlation between DNA Methylation of the First Intron and Gene Expression across Tissues and Species'. Epigenetics & Chromatin 11 (1): 37. https://doi.org/10.1186/s13072-018-0205-1.

Anastasiadi, Dafni, Changwei Shao, Songlin Chen, and Francesc Piferrer. 2021. 'Footprints of Global Change in Marine Life: Inferring Past Environment Based on DNA Methylation and Gene Expression Marks'. Molecular Ecology 30 (3): 747–60. https://doi.org/10.1111/mec.15764.

Arbour, Jessica, David Hardie, and Jeffrey Hutchings. 2011. 'Morphometric and Genetic Analyses of Two Sympatric Morphs of Arctic Char (Salvelinus Alpinus) in the Canadian High Arctic'. Canadian Journal of Zoology 89 (January): 19–30. https://doi.org/10.1139/Z10-100.

Arechederra, Maria, Fabrice Daian, Annie Yim, Sehrish K. Bazai, Sylvie Richelme, Rosanna Dono, Andrew J. Saurin, Bianca H. Habermann, and Flavio Maina. 2018. 'Hypermethylation of Gene Body CpG Islands Predicts High Dosage of Functional Oncogenes in Liver Cancer'. Nature Communications 9 (1): 3164. https://doi.org/10.1038/s41467-018-05550-5.

Bachman, Martin, Santiago Uribe-Lewis, Xiaoping Yang, Michael Williams, Adele Murrell, and Shankar Balasubramanian. 2014. '5-Hydroxymethylcytosine Is a Predominantly Stable DNA Modification'. Nature Chemistry 6 (12): 1049–55. https://doi.org/10.1038/nchem.2064.

Bannister, Andrew J., and Tony Kouzarides. 2011. 'Regulation of Chromatin by Histone Modifications'. Cell Research 21 (3): 381–95. https://doi.org/10.1038/cr.2011.22.

Barlow, Denise P., and Marisa S. Bartolomei. 2014. 'Genomic Imprinting in Mammals'. Cold Spring Harbor Perspectives in Biology 6 (2): a018382. https://doi.org/10.1101/cshperspect.a018382.

Barski, Artem, Suresh Cuddapah, Kairong Cui, Tae-Young Roh, Dustin E. Schones, Zhibin Wang, Gang Wei, Iouri Chepelev, and Keji Zhao. 2007. 'High-Resolution Profiling of Histone Methylations in the Human Genome'. Cell 129 (4): 823–37. https://doi.org/10.1016/j.cell.2007.05.009.

Barton, Nicholas H. 2022. 'The "New Synthesis"'. Proceedings of the National Academy of Sciences 119 (30): e2122147119. https://doi.org/10.1073/pnas.2122147119.

Beemelmanns, Anne, Laia Ribas, Dafni Anastasiadi, Javier Moraleda-Prados, Fábio S. Zanuzzo, Matthew L. Rise, and A. Kurt Gamperl. 2021. 'DNA Methylation Dynamics in Atlantic Salmon (Salmo Salar) Challenged With High Temperature and Moderate Hypoxia'. Frontiers in Marine Science 7. https://www.frontiersin.org/articles/10.3389/fmars.2020.604878.

Berbel-Filho, Waldir M., Carlos Garcia de Leaniz, Paloma Morán, Joanne Cable, Sergio M. Q. Lima, and Sofia Consuegra. 2019. 'Local Parasite Pressures and Host Genotype Modulate Epigenetic Diversity in a Mixed-Mating Fish'. Ecology and Evolution 9 (15): 8736–48. https://doi.org/10.1002/ece3.5426.

Bertho, Sylvain, Amaury Herpin, Manfred Schartl, and Yann Guiguen. 2021. 'Lessons from an Unusual Vertebrate Sex-Determining Gene'. Philosophical Transactions of the Royal Society B: Biological Sciences 376 (1832): 20200092. https://doi.org/10.1098/rstb.2020.0092.

Best, Carol, Heather Ikert, Daniel J. Kostyniuk, Paul M. Craig, Laia Navarro-Martin, Lucie Marandel, and Jan A. Mennigen. 2018. 'Epigenetics in Teleost Fish: From Molecular Mechanisms to Physiological Phenotypes'. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology, 50 Years of Comparative Biochemistry: The Legacy of Peter Hochachka, 224 (October): 210–44. https://doi.org/10.1016/j.cbpb.2018.01.006.

Bestor, Timothy, Andrew Laudano, Robert Mattaliano, and Vernon Ingram. 1988. 'Cloning and Sequencing of a cDNA Encoding DNA Methyltransferase of Mouse Cells: The Carboxyl-Terminal Domain of the Mammalian Enzymes Is Related to Bacterial Restriction Methyltransferases'. Journal of Molecular Biology 203 (4): 971–83. https://doi.org/10.1016/0022-2836(88)90122-2.

Bizuayehu, Teshome Tilahun, and Igor Babiak. 2014. 'MicroRNA in Teleost Fish'. Genome Biology and Evolution 6 (8): 1911–37. https://doi.org/10.1093/gbe/evu151.

Bonduriansky, Russell, and Troy Day. 2018. Extended Heredity: A New Understanding of Inheritance and Evolution. Princeton University Press. https://doi.org/10.2307/j.ctvc77mbg.

Bostick, Magnolia, Jong Kyong Kim, Pierre-Olivier Estève, Amander Clark, Sriharsa Pradhan, and Steven E. Jacobsen. 2007. 'UHRF1 Plays a Role in Maintaining DNA Methylation in Mammalian Cells'. Science (New York, N.Y.) 317 (5845): 1760–64. https://doi.org/10.1126/science.1147939.

Brachmann, Matthew K., Kevin Parsons, Skúli Skúlason, Oscar Gaggiotti, and Moira Ferguson. 2022. 'Variation in the Genomic Basis of Parallel Phenotypic and Ecological Divergence in Benthic and Pelagic Morphs of Icelandic Arctic Charr (Salvelinus Alpinus)'. Molecular Ecology 31 (18): 4688–4706. https://doi.org/10.1111/mec.16625.

Bronner, Christian, Mahmoud Alhosin, Ali Hamiche, and Marc Mousli. 2019. 'Coordinated Dialogue between UHRF1 and DNMT1 to Ensure Faithful Inheritance of Methylated DNA Patterns'. Genes 10 (1): 65. https://doi.org/10.3390/genes10010065.

Burgerhout, Erik, Maren Mommens, Hanne Johnsen, Arnfinn Aunsmo, Nina Santi, and Øivind Andersen. 2017. 'Genetic Background and Embryonic Temperature Affect DNA Methylation and Expression of Myogenin and Muscle Development in Atlantic Salmon

(Salmo Salar)'. PLOS ONE 12 (6): e0179918. https://doi.org/10.1371/journal.pone.0179918.

Cámara, Marina de la, Lieke Ponsioen, Quentin J. B. Horta-Lacueva, and Kalina H. Kapralova. 2023. 'The Dynamic Ontogenetic Shape Patterns of Adaptive Divergence and Sexual Dimorphism'. Evolutionary Biology 50 (2): 170–80. https://doi.org/10.1007/s11692-022-09592-y.

Campos, Catarina, Luisa M. P. Valente, and Jorge M. O. Fernandes. 2012. 'Molecular Evolution of Zebrafish Dnmt3 Genes and Thermal Plasticity of Their Expression during Embryonic Development'. Gene 500 (1): 93–100. https://doi.org/10.1016/j.gene.2012.03.041.

Carmell, Michelle A., Angélique Girard, Henk J. G. van de Kant, Deborah Bourc'his, Timothy H. Bestor, Dirk G. de Rooij, and Gregory J. Hannon. 2007. 'MIWI2 Is Essential for Spermatogenesis and Repression of Transposons in the Mouse Male Germline'. Developmental Cell 12 (4): 503–14. https://doi.org/10.1016/j.devcel.2007.03.001.

Carvan, Michael J., Thomas A. Kalluvila, Rebekah H. Klingler, Jeremy K. Larson, Matthew Pickens, Francisco X. Mora-Zamorano, Victoria P. Connaughton, Ingrid Sadler-Riggleman, Daniel Beck, and Michael K. Skinner. 2017. 'Mercury-Induced Epigenetic Transgenerational Inheritance of Abnormal Neurobehavior Is Correlated with Sperm Epimutations in Zebrafish'. PLOS ONE 12 (5): e0176155. https://doi.org/10.1371/journal.pone.0176155.

Casacuberta, Elena, and Josefa González. 2013. 'The Impact of Transposable Elements in Environmental Adaptation'. Molecular Ecology 22 (6): 1503–17. https://doi.org/10.1111/mec.12170.

Cavalli, Giacomo, and Edith Heard. 2019. 'Advances in Epigenetics Link Genetics to the Environment and Disease'. Nature 571 (7766): 489–99. https://doi.org/10.1038/s41586-019-1411-0.

Chalopin, Domitille, Magali Naville, Floriane Plard, Delphine Galiana, and Jean-Nicolas Volff. 2015. 'Comparative Analysis of Transposable Elements Highlights Mobilome Diversity and Evolution in Vertebrates'. Genome Biology and Evolution 7 (2): 567–80. https://doi.org/10.1093/gbe/evv005.

Chang, Andrew Ying-Fei, and Ben-Yang Liao. 2012. 'DNA Methylation Rebalances Gene Dosage after Mammalian Gene Duplications'. Molecular Biology and Evolution 29 (1): 133–44. https://doi.org/10.1093/molbev/msr174.

Chatterjee, Biswanath, Miao-Hsia Lin, Chun-Chang Chen, Kai-Lin Peng, Mu-Sheng Wu, Mei-Chun Tseng, Yu-Ju Chen, and Che-Kun James Shen. 2018. 'DNA Demethylation by DNMT3A and DNMT3B in Vitro and of Methylated Episomal DNA in Transiently Transfected Cells'. Biochimica Et Biophysica Acta. Gene Regulatory Mechanisms 1861 (11): 1048–61. https://doi.org/10.1016/j.bbagrm.2018.09.009.

Chen, Chun-Chang, Keh-Yang Wang, and Che-Kun James Shen. 2013. 'DNA 5-Methylcytosine Demethylation Activities of the Mammalian DNA Methyltransferases'. The Journal of Biological Chemistry 288 (13): 9084–91. https://doi.org/10.1074/jbc.M112.445585.

Chen, Qi, Menghong Yan, Zhonghong Cao, Xin Li, Yunfang Zhang, Junchao Shi, Gui-hai Feng, et al. 2016. 'Sperm tsRNAs Contribute to Intergenerational Inheritance of an Acquired Metabolic Disorder'. Science 351 (6271): 397–400. https://doi.org/10.1126/science.aad7977.

Chen, Zhao-Xia, Jeffrey R. Mann, Chih-Lin Hsieh, Arthur D. Riggs, and Frédéric Chédin. 2005. 'Physical and Functional Interactions between the Human DNMT3L Protein and Members of the de Novo Methyltransferase Family'. Journal of Cellular Biochemistry 95 (5): 902–17. https://doi.org/10.1002/jcb.20447.

Cooper, D. N. 1983. 'Eukaryotic DNA Methylation'. Human Genetics 64 (4): 315–33. https://doi.org/10.1007/BF00292363.

Dehal, Paramvir, and Jeffrey L. Boore. 2005. 'Two Rounds of Whole Genome Duplication in the Ancestral Vertebrate'. PLoS Biology 3 (10): e314. https://doi.org/10.1371/journal.pbio.0030314.

Dinger, Marcel E., Paulo P. Amaral, Tim R. Mercer, Ken C. Pang, Stephen J. Bruce, Brooke B. Gardiner, Marjan E. Askarian-Amiri, et al. 2008. 'Long Noncoding RNAs in Mouse Embryonic Stem Cell Pluripotency and Differentiation'. Genome Research 18 (9): 1433–45. https://doi.org/10.1101/gr.078378.108.

Duempelmann, Lea, Merle Skribbe, and Marc Bühler. 2020. 'Small RNAs in the Transgenerational Inheritance of Epigenetic Information'. Trends in Genetics 36 (3): 203–14. https://doi.org/10.1016/j.tig.2019.12.001.

Elbarbary, Reyad A., Bronwyn A. Lucas, and Lynne E. Maquat. 2016. 'Retrotransposons as Regulators of Gene Expression'. Science (New York, N.Y.) 351 (6274): aac7247. https://doi.org/10.1126/science.aac7247.

Esteller, Manel. 2002. 'CpG Island Hypermethylation and Tumor Suppressor Genes: A Booming Present, a Brighter Future'. Oncogene 21 (35): 5427–40. https://doi.org/10.1038/sj.onc.1205600.

Fargeot, Laura, Géraldine Loot, Jérôme G. Prunier, Olivier Rey, Charlotte Veyssière, and Simon Blanchet. 2021. 'Patterns of Epigenetic Diversity in Two Sympatric Fish Species: Genetic vs. Environmental Determinants'. Genes 12 (1): 107. https://doi.org/10.3390/genes12010107.

Fellous, Alexandre, and Lisa N. S. Shama. 2019. 'Genome Survey of Chromatin-Modifying Enzymes in Threespine Stickleback: A Crucial Epigenetic Toolkit for Adaptation?' Frontiers in Marine Science 6. https://www.frontiersin.org/articles/10.3389/fmars.2019.00721.

Fellous, Alexandre, K. Mathias Wegner, Uwe John, Felix C. Mark, and Lisa N. S. Shama. 2022. 'Windows of Opportunity: Ocean Warming Shapes Temperature-Sensitive Epigenetic Reprogramming and Gene Expression across Gametogenesis and Embryogenesis in Marine Stickleback'. Global Change Biology 28 (1): 54–71. https://doi.org/10.1111/gcb.15942.

Force, A., M. Lynch, F. B. Pickett, A. Amores, Y. L. Yan, and J. Postlethwait. 1999. 'Preservation of Duplicate Genes by Complementary, Degenerative Mutations'. Genetics 151 (4): 1531–45. https://doi.org/10.1093/genetics/151.4.1531.

Gapp, Katharina, Ali Jawaid, Peter Sarkies, Johannes Bohacek, Pawel Pelczar, Julien Prados, Laurent Farinelli, Eric Miska, and Isabelle M. Mansuy. 2014. 'Implication of Sperm RNAs in Transgenerational Inheritance of the Effects of Early Trauma in Mice'. Nature Neuroscience 17 (5): 667–69. https://doi.org/10.1038/nn.3695.

Gavery, Mackenzie R., Krista M. Nichols, Giles W. Goetz, Mollie A. Middleton, and Penny Swanson. 2018. 'Characterization of Genetic and Epigenetic Variation in Sperm and Red Blood Cells from Adult Hatchery and Natural-Origin Steelhead, Oncorhynchus Mykiss'. G3 (Bethesda, Md.) 8 (11): 3723–36. https://doi.org/10.1534/g3.118.200458.

Goll, Mary Grace, Finn Kirpekar, Keith A. Maggert, Jeffrey A. Yoder, Chih-Lin Hsieh, Xiaoyu Zhang, Kent G. Golic, Steven E. Jacobsen, and Timothy H. Bestor. 2006. 'Methylation of tRNAAsp by the DNA Methyltransferase Homolog Dnmt2'. Science (New York, N.Y.) 311 (5759): 395–98. https://doi.org/10.1126/science.1120976.

Gonzalez, Susana, David G. Pisano, and Manuel Serrano. 2008. 'Mechanistic Principles of Chromatin Remodeling Guided by siRNAs and miRNAs'. Cell Cycle (Georgetown, Tex.) 7 (16): 2601–8. https://doi.org/10.4161/cc.7.16.6541.

Gore, Aniket V., Kelly A. Tomins, James Iben, Li Ma, Daniel Castranova, Andrew E. Davis, Amy Parkhurst, William R. Jeffery, and Brant M. Weinstein. 2018. 'An Epigenetic Mechanism for Cavefish Eye Degeneration'. Nature Ecology & Evolution 2 (7): 1155–60. https://doi.org/10.1038/s41559-018-0569-4.

Greenspoon, Philip B., Hamish G. Spencer, and Leithen K. M'Gonigle. 2022. 'Epigenetic Induction May Speed up or Slow down Speciation with Gene Flow'. Evolution; International Journal of Organic Evolution 76 (6): 1170–82. https://doi.org/10.1111/evo.14494.

Gudbrandsson, Johannes, Ehsan P. Ahi, Sigridur R. Franzdottir, Kalina H. Kapralova, Bjarni K. Kristjansson, S. Sophie Steinhaeuser, Valerie H. Maier, et al. 2016. 'The Developmental Transcriptome of Contrasting Arctic Charr (Salvelinus Alpinus) Morphs'. F1000Research 4 (December): 136. https://doi.org/10.12688/f1000research.6402.3.

Guðbrandsson, Jóhannes, Sigríður Rut Franzdóttir, Bjarni Kristófer Kristjánsson, Ehsan Pashay Ahi, Valerie Helene Maier, Kalina Hristova Kapralova, Sigurður Sveinn Snorrason, Zophonías Oddur Jónsson, and Arnar Pálsson. 2018. 'Differential Gene Expression during Early Development in Recently Evolved and Sympatric Arctic Charr Morphs'. PeerJ 6: e4345. https://doi.org/10.7717/peerj.4345.

Guðbrandsson, Jóhannes, Kalina H. Kapralova, Sigríður R. Franzdóttir, Þóra Margrét Bergsveinsdóttir, Völundur Hafstað, Zophonías O. Jónsson, Sigurður S. Snorrason, and Arnar Pálsson. 2019. 'Extensive Genetic Differentiation between Recently Evolved Sympatric Arctic Charr Morphs'. Ecology and Evolution 9 (19): 10964–83. https://doi.org/10.1002/ece3.5516.

Guibert, Sylvain, Thierry Forné, and Michael Weber. 2012. 'Global Profiling of DNA Methylation Erasure in Mouse Primordial Germ Cells'. Genome Research 22 (4): 633–41. https://doi.org/10.1101/gr.130997.111.

Gundappa, Manu Kumar, Thu-Hien To, Lars Grønvold, Samuel A. M. Martin, Sigbjørn Lien, Juergen Geist, David Hazlerigg, Simen R. Sandve, and Daniel J. Macqueen. 2022. 'Genome-Wide Reconstruction of Rediploidization Following Autopolyploidization across

One Hundred Million Years of Salmonid Evolution'. Molecular Biology and Evolution 39 (1): msab310. https://doi.org/10.1093/molbev/msab310.

Guo, Sai-Nan, Jia-Lang Zheng, Shuang-Shuang Yuan, and Qing-Ling Zhu. 2018. 'Effects of Heat and Cadmium Exposure on Stress-Related Responses in the Liver of Female Zebrafish: Heat Increases Cadmium Toxicity'. The Science of the Total Environment 618 (March): 1363–70. https://doi.org/10.1016/j.scitotenv.2017.09.264.

Hamilton, Matthew B. 2021. 'Population Genetics, 2nd Edition | Wiley'. Wiley.Com. 2021. https://www.wiley.com/en-us/Population+Genetics%2C+2nd+Edition-p-9781118436943.

Heckwolf, Melanie J., Britta S. Meyer, Robert Häsler, Marc P. Höppner, Christophe Eizaguirre, and Thorsten B. H. Reusch. 2020. 'Two Different Epigenetic Information Channels in Wild Three-Spined Sticklebacks Are Involved in Salinity Adaptation'. Science Advances 6 (12): eaaz1138. https://doi.org/10.1126/sciadv.aaz1138.

Herman, Jacob J., and Sonia E. Sultan. 2016. 'DNA Methylation Mediates Genetic Variation for Adaptive Transgenerational Plasticity'. Proceedings of the Royal Society B: Biological Sciences 283 (1838): 20160988. https://doi.org/10.1098/rspb.2016.0988.

Herpin, Amaury, and Manfred Schartl. 2011. 'Dmrt1 Genes at the Crossroads: A Widespread and Central Class of Sexual Development Factors in Fish'. The FEBS Journal 278 (7): 1010–19. https://doi.org/10.1111/j.1742-4658.2011.08030.x.

Hindar, Kjetil, and Bror Jonsson. 1982. 'Habitat and Food Segregation of Dwarf and Normal Arctic Charr (Salvelinus Alpinus) from Vangsvatnet Lake, Western Norway'. Canadian Journal of Fisheries and Aquatic Sciences, July. https://doi.org/10.1139/f82-138.

Hoegg, Simone, Henner Brinkmann, John S. Taylor, and Axel Meyer. 2004. 'Phylogenetic Timing of the Fish-Specific Genome Duplication Correlates with the Diversification of Teleost Fish'. Journal of Molecular Evolution 59 (2): 190–203. https://doi.org/10.1007/s00239-004-2613-z.

Horta-Lacueva, Quentin. 2022. 'Evolution of Reproductive Isolation in Sympatric Arctic Charr Morphs (Salvelinus Alpinus)'.

Horta-Lacueva, Quentin J.-B., Sigurður S. Snorrason, Michael B. Morrissey, Camille A.-L. Leblanc, and Kalina H. Kapralova. 2021. 'Multivariate Analysis of Morphology, Behaviour, Growth and Developmental Timing in Hybrids Brings New Insights into the Divergence of Sympatric Arctic Charr Morphs'. BMC Ecology and Evolution 21 (1): 170. https://doi.org/10.1186/s12862-021-01904-8.

Hu, Junjie, Jinhuan Liu, Jinyun Li, Xiaomei Lv, Lili Yu, Kangming Wu, and Yan Yang. 2021. 'Metal Contamination, Bioaccumulation, ROS Generation, and Epigenotoxicity Influences on Zebrafish Exposed to River Water Polluted by Mining Activities'. Journal of Hazardous Materials 405 (March): 124150. https://doi.org/10.1016/j.jhazmat.2020.124150.

Hu, Juntao, Sara J. S. Wuitchik, Tegan N. Barry, Heather A. Jamniczky, Sean M. Rogers, and Rowan D. H. Barrett. 2021. 'Heritability of DNA Methylation in Threespine Stickleback (Gasterosteus Aculeatus)'. Genetics 217 (1): 1–15. https://doi.org/10.1093/genetics/iyab001.

Huang, Katherine M., and Frédéric J. J. Chain. 2021. ‘Copy Number Variations and Young Duplicate Genes Have High Methylation Levels in Sticklebacks’. Evolution; International Journal of Organic Evolution 75 (3): 706–18. https://doi.org/10.1111/evo.14184.

Huang, Xu, Qinbo Qin, Kaijun Gong, Chang Wu, Yuwei Zhou, Qian Chen, Wenjing Feng, et al. 2020. ‘Comparative Analyses of the Sox9a-Amh-Cyp19a1a Regulatory Cascade in Autotetraploid Fish and Its Diploid Parent’. BMC Genetics 21 (1): 35. https://doi.org/10.1186/s12863-020-00840-8.

Huber, Sarah K., Luis Fernando De León, Andrew P. Hendry, Eldredge Bermingham, and Jeffrey Podos. 2007. ‘Reproductive Isolation of Sympatric Morphs in a Population of Darwin’s Finches’. Proceedings. Biological Sciences 274 (1619): 1709–14. https://doi.org/10.1098/rspb.2007.0224.

Iida, Tetsuo, Isao Suetake, Shoji Tajima, Hiroshi Morioka, Satoshi Ohta, Chikashi Obuse, and Toshiki Tsurimoto. 2002. ‘PCNA Clamp Facilitates Action of DNA Cytosine Methyltransferase 1 on Hemimethylated DNA’. Genes to Cells: Devoted to Molecular & Cellular Mechanisms 7 (10): 997–1007. https://doi.org/10.1046/j.1365-2443.2002.00584.x.

Iorio, Marilena V., Manuela Ferracin, Chang-Gong Liu, Angelo Veronese, Riccardo Spizzo, Silvia Sabbioni, Eros Magri, et al. 2005. ‘MicroRNA Gene Expression Deregulation in Human Breast Cancer’. Cancer Research 65 (16): 7065–70. https://doi.org/10.1158/0008-5472.CAN-05-1783.

Jeltsch, Albert, Ann Ehrenhofer-Murray, Tomasz P. Jurkowski, Frank Lyko, Gunter Reuter, Serge Ankri, Wolfgang Nellen, Matthias Schaefer, and Mark Helm. 2017. ‘Mechanism and Biological Role of Dnmt2 in Nucleic Acid Methylation’. RNA Biology 14 (9): 1108–23. https://doi.org/10.1080/15476286.2016.1191737.

Jia, Da, Renata Z. Jurkowska, Xing Zhang, Albert Jeltsch, and Xiaodong Cheng. 2007. ‘Structure of Dnmt3a Bound to Dnmt3L Suggests a Model for de Novo DNA Methylation’. Nature 449 (7159): 248–51. https://doi.org/10.1038/nature06146.

Jimenji, Tomohiro, Rumie Matsumura, Satomi Kori, and Kyohei Arita. 2019. ‘Structure of PCNA in Complex with DNMT1 PIP Box Reveals the Basis for the Molecular Mechanism of the Interaction’. Biochemical and Biophysical Research Communications 516 (2): 578–83. https://doi.org/10.1016/j.bbrc.2019.06.060.

Jones, Felicity C, Manfred G Grabherr, Yingguang Frank Chan, Pamela Russell, Evan Mauceli, Jeremy Johnson, Ross Swofford, et al. 2012. ‘The Genomic Basis of Adaptive Evolution in Threespine Sticklebacks’. Nature 484 (7392): 55–61. https://doi.org/10.1038/nature10944.

Jong, Gerdien de. 2005. ‘Evolution of Phenotypic Plasticity: Patterns of Plasticity and the Emergence of Ecotypes’. The New Phytologist 166 (1): 101–17. https://doi.org/10.1111/j.1469-8137.2005.01322.x.

Jónsdóttir, Guðbjörg Ósk, Laura-Marie von Elm, Finnur Ingimarsson, Samuel Tersigni, Sigurður Sveinn Snorrason, Arnar Pálsson, and Sarah Elisabeth Steele. 2023. ‘Diversity in the Internal Functional Feeding Elements of Sympatric Morphs of Arctic Charr (Salvelinus Alpinus)’. bioRxiv. https://doi.org/10.1101/2023.02.17.528955.

Jonsson, B., S. Skúlason, S. S. Snorrason, O. T. Sandlund, H. J. Malmquist, P. M. Jónasson, R. Cydemo, and T. Lindem. 1988. 'Life History Variation of Polymorphic Arctic Charr (Salvelinus Alpinus) in Thingvallavatn, Iceland'. Canadian Journal of Fisheries and Aquatic Sciences 45 (9): 1537–47. https://doi.org/10.1139/f88-182.

Kamstra, Jorke H., Selma Hurem, Leonardo Martin Martin, Leif C. Lindeman, Juliette Legler, Deborah Oughton, Brit Salbu, Dag Anders Brede, Jan Ludvig Lyche, and Peter Aleström. 2018. 'Ionizing Radiation Induces Transgenerational Effects of DNA Methylation in Zebrafish'. Scientific Reports 8 (1): 15373. https://doi.org/10.1038/s41598-018-33817-w.

Kapralova, K. H., M. B. Morrissey, B. K. Kristjánsson, G. Á Ólafsdóttir, S. S. Snorrason, and M. M. Ferguson. 2011. 'Evolution of Adaptive Diversity and Genetic Connectivity in Arctic Charr (Salvelinus Alpinus) in Iceland'. Heredity 106 (3): 472–87. https://doi.org/10.1038/hdy.2010.161.

Kapralova, Kalina H., Sigrídur Rut Franzdóttir, Hákon Jónsson, Sigurður S. Snorrason, and Zophonías O. Jónsson. 2014. 'Patterns of miRNA Expression in Arctic Charr Development'. PloS One 9 (8): e106084. https://doi.org/10.1371/journal.pone.0106084.

Kapralova, Kalina H., Zophonías O. Jónsson, Arnar Palsson, Sigrídur Rut Franzdóttir, Soizic le Deuff, Bjarni Kristófer Kristjánsson, and Sigurður S. Snorrason. 2015. 'Bones in Motion: Ontogeny of Craniofacial Development in Sympatric Arctic Charr Morphs'. Developmental Dynamics 244 (9): 1168–78. https://doi.org/10.1002/dvdy.24302.

Kapranov, Philipp, Jill Cheng, Sujit Dike, David A. Nix, Radharani Duttagupta, Aarron T. Willingham, Peter F. Stadler, et al. 2007. 'RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription'. Science (New York, N.Y.) 316 (5830): 1484–88. https://doi.org/10.1126/science.1138341.

Karimi, Mohammad M., Preeti Goyal, Irina A. Maksakova, Misha Bilenky, Danny Leung, Jie Xin Tang, Yoichi Shinkai, et al. 2011. 'DNA Methylation and SETDB1/H3K9me3 Regulate Predominantly Distinct Sets of Genes, Retroelements and Chimaeric Transcripts in Mouse ES Cells'. Cell Stem Cell 8 (6): 676–87. https://doi.org/10.1016/j.stem.2011.04.004.

Kautt, Andreas F., Claudius F. Kratochwil, Alexander Nater, Gonzalo Machado-Schiaffino, Melisa Olave, Frederico Henning, Julián Torres-Dowdall, et al. 2020. 'Contrasting Signatures of Genomic Divergence during Sympatric Speciation'. Nature 588 (7836): 106–11. https://doi.org/10.1038/s41586-020-2845-0.

Keller, Thomas E., and Soojin V. Yi. 2014. 'DNA Methylation and Evolution of Duplicate Genes'. Proceedings of the National Academy of Sciences 111 (16): 5932–37. https://doi.org/10.1073/pnas.1321420111.

Kelley, Joanna L., Michael Tobler, Daniel Beck, Ingrid Sadler-Riggleman, Corey R. Quackenbush, Lenin Arias Rodriguez, and Michael K. Skinner. 2021. 'Epigenetic Inheritance of DNA Methylation Changes in Fish Living in Hydrogen Sulfide-Rich Springs'. Proceedings of the National Academy of Sciences of the United States of America 118 (26): e2014929118. https://doi.org/10.1073/pnas.2014929118.

Kitano, T., K. Takamune, T. Kobayashi, Y. Nagahama, and S. I. Abe. 1999. 'Suppression of P450 Aromatase Gene Expression in Sex-Reversed Males Produced by Rearing

Genetically Female Larvae at a High Water Temperature during a Period of Sex Differentiation in the Japanese Flounder (Paralichthys Olivaceus)'. Journal of Molecular Endocrinology 23 (2): 167–76. https://doi.org/10.1677/jme.0.0230167.

Koga, Akihiko, Atsuo Iida, Hiroshi Hori, Atsuko Shimada, and Akihiro Shima. 2006. 'Vertebrate DNA Transposon as a Natural Mutator: The Medaka Fish Tol2 Element Contributes to Genetic Variation without Recognizable Traces'. Molecular Biology and Evolution 23 (7): 1414–19. https://doi.org/10.1093/molbev/msl003.

Kohli, Rahul M., and Yi Zhang. 2013. 'TET Enzymes, TDG and the Dynamics of DNA Demethylation'. Nature 502 (7472): 472–79. https://doi.org/10.1038/nature12750.

Konstantinidis, Ioannis, Dafni Anastasiadi, Pål Sætrom, Artem V. Nedoluzhko, Robin Mjelle, Tomasz Podgorniak, Francesc Piferrer, and Jorge M. O. Fernandes. 2021. 'Epigenetic Mapping of the Somatotropic Axis in Nile Tilapia Reveals Differential DNA Hydroxymethylation Marks Associated with Growth'. Genomics 113 (5): 2953–64. https://doi.org/10.1016/j.ygeno.2021.06.037.

Kornberg, R. D. 1974. 'Chromatin Structure: A Repeating Unit of Histones and DNA'. Science (New York, N.Y.) 184 (4139): 868–71. https://doi.org/10.1126/science.184.4139.868.

Kozomara, Ana, and Sam Griffiths-Jones. 2011. 'miRBase: Integrating microRNA Annotation and Deep-Sequencing Data'. Nucleic Acids Research 39 (Database issue): D152-157. https://doi.org/10.1093/nar/gkq1027.

Lai, Keng Po, Simon Yuan Wang, Jing Woei Li, Yin Tong, Ting Fung Chan, Nana Jin, Anna Tse, et al. 2019. 'Hypoxia Causes Transgenerational Impairment of Ovarian Development and Hatching Success in Fish'. Environmental Science & Technology 53 (7): 3917–28. https://doi.org/10.1021/acs.est.8b07250.

Lallias, Delphine, Maria Bernard, Céline Ciobotaru, Nicolas Dechamp, Laurent Labbé, Lionel Goardon, Jean-Michel Le Calvez, et al. 2021. 'Sources of Variation of DNA Methylation in Rainbow Trout: Combined Effects of Temperature and Genetic Background'. Epigenetics 16 (9): 1031–52. https://doi.org/10.1080/15592294.2020.1834924.

Lamarck, Jean-Baptiste de Monet de (1744-1829) Auteur du texte. 1802. Recherches Sur l'organisation Des Corps Vivans. https://gallica.bnf.fr/ark:/12148/bpt6k4226016j.

Lane, Natasha, Wendy Dean, Sylvia Erhardt, Petra Hajkova, Azim Surani, Jörn Walter, and Wolf Reik. 2003. 'Resistance of IAPs to Methylation Reprogramming May Provide a Mechanism for Epigenetic Inheritance in the Mouse'. Genesis (New York, N.Y.: 2000) 35 (2): 88–93. https://doi.org/10.1002/gene.10168.

Laporte, M., J. Le Luyer, C. Rougeux, A.-M. Dion-Côté, M. Krick, and L. Bernatchez. 2019. 'DNA Methylation Reprogramming, TE Derepression, and Postzygotic Isolation of Nascent Animal Species'. Science Advances 5 (10): eaaw1644. https://doi.org/10.1126/sciadv.aaw1644.

Lau, Nelson C., Lee P. Lim, Earl G. Weinstein, and David P. Bartel. 2001. 'An Abundant Class of Tiny RNAs with Probable Regulatory Roles in Caenorhabditis Elegans'. Science 294 (5543): 858–62. https://doi.org/10.1126/science.1065062.

Lee, Jeannie T. 2003. ‘Molecular Links between X-Inactivation and Autosomal Imprinting: X-Inactivation as a Driving Force for the Evolution of Imprinting?’ Current Biology: CB 13 (6): R242-254. https://doi.org/10.1016/s0960-9822(03)00162-3.

Lee, R. C., R. L. Feinbaum, and V. Ambros. 1993. ‘The C. Elegans Heterochronic Gene Lin-4 Encodes Small RNAs with Antisense Complementarity to Lin-14’. Cell 75 (5): 843–54. https://doi.org/10.1016/0092-8674(93)90529-y.

Leitwein, Maeva, Martin Laporte, Jeremy Le Luyer, Kayla Mohns, Eric Normandeau, Ruth Withler, and Louis Bernatchez. 2021. ‘Epigenomic Modifications Induced by Hatchery Rearing Persist in Germ Line Cells of Adult Salmon after Their Oceanic Migration’. Evolutionary Applications 14 (10): 2402–13. https://doi.org/10.1111/eva.13235.

Li, Siping, Feng He, Haishen Wen, Jifang Li, Yufeng Si, Mingyuan Liu, Yajuan Huang, and Lingcai Meng. 2017. ‘Low Salinity Affects Cellularity, DNA Methylation, and mRNA Expression of Igf1 in the Liver of Half Smooth Tongue Sole (Cynoglossus Semilaevis)’. Fish Physiology and Biochemistry 43 (6): 1587–1602. https://doi.org/10.1007/s10695-017-0395-7.

Lien, Sigbjørn, Ben F. Koop, Simen R. Sandve, Jason R. Miller, Matthew P. Kent, Torfinn Nome, Torgeir R. Hvidsten, et al. 2016. ‘The Atlantic Salmon Genome Provides Insights into Rediploidization’. Nature 533 (7602): 200–205. https://doi.org/10.1038/nature17164.

Lin, Shi-Lung, Joseph D. Miller, and Shao-Yao Ying. 2006. ‘Intronic microRNA (miRNA)’. Journal of Biomedicine & Biotechnology 2006 (4): 26818. https://doi.org/10.1155/JBB/2006/26818.

Lind, Martin I., and Foteini Spagopoulou. 2018. ‘Evolutionary Consequences of Epigenetic Inheritance’. Heredity 121 (3): 205–9. https://doi.org/10.1038/s41437-018-0113-y.

Liu, Jingwei, Huihua Hu, Stéphane Panserat, and Lucie Marandel. 2020. ‘Evolutionary History of DNA Methylation Related Genes in Chordates: New Insights from Multiple Whole Genome Duplications’. Scientific Reports 10 (1): 970. https://doi.org/10.1038/s41598-020-57753-w.

Liu, Shenglin, Aja Noersgaard Buur Tengstedt, Magnus W. Jacobsen, Jose Martin Pujolar, Bjarni Jónsson, Javier Lobón-Cervià, Louis Bernatchez, and Michael M. Hansen. 2022. ‘Genome-Wide Methylation in the Panmictic European Eel (Anguilla Anguilla)’. Molecular Ecology 31 (16): 4286–4306. https://doi.org/10.1111/mec.16586.

Loughland, Isabella, Alexander Little, and Frank Seebacher. 2021. ‘DNA Methyltransferase 3a Mediates Developmental Thermal Plasticity’. BMC Biology 19 (1): 11. https://doi.org/10.1186/s12915-020-00942-w.

Lyko, Frank. 2018. ‘The DNA Methyltransferase Family: A Versatile Toolkit for Epigenetic Regulation’. Nature Reviews. Genetics 19 (2): 81–92. https://doi.org/10.1038/nrg.2017.80.

Ma, W., Z.-H. Zhu, X.-Y. Bi, R. W. Murphy, S.-Y. Wang, Y. Gao, H. Xiao, Y.-P. Zhang, and J. Luo. 2014. ‘Allopolyploidization Is Not so Simple: Evidence from the Origin of the

Tribe Cyprinini (Teleostei: Cypriniformes)'. Current Molecular Medicine 14 (10): 1331–38. https://doi.org/10.2174/1566524014666141203101543.

Macqueen, Daniel, and Ian Johnston. 2014. 'A Well-Constrained Estimate for the Timing of the Salmonid Whole Genome Duplication Reveals Major Decoupling from Species Diversification'. Proceedings. Biological Sciences / The Royal Society 281 (January): 20132881. https://doi.org/10.1098/rspb.2013.2881.

Maiti, Atanu, Anna Zhachkina Michelson, Cherece J. Armwood, Jeehiun K. Lee, and Alexander C. Drohat. 2013. 'Divergent Mechanisms for Enzymatic Excision of 5-Formylcytosine and 5-Carboxylcytosine from DNA'. Journal of the American Chemical Society 135 (42): 15813–22. https://doi.org/10.1021/ja406444x.

Major, Kaley M., Bethany M. DeCourten, Jie Li, Monica Britton, Matthew L. Settles, Alvine C. Mehinto, Richard E. Connon, and Susanne M. Brander. 2020. 'Early Life Exposure to Environmentally Relevant Levels of Endocrine Disruptors Drive Multigenerational and Transgenerational Epigenetic Changes in a Fish Model'. Frontiers in Marine Science 7. https://www.frontiersin.org/articles/10.3389/fmars.2020.00471.

Malmquist, H. J., S. S. Snorrason, S. Skúlason, B. Jonsson, O. T. Sandlund, and P. M. Jonasson. 1992. 'Diet Differentiation in Polymorphic Arctic Charr in Thingvallavatn, Iceland'. Journal of Animal Ecology 61 (1): 21–35. https://doi.org/10.2307/5505.

Marandel, L, Heraud C, Véron V, Laithier J, Marchand M, Quillet E, Callet T, Dupont-Nivet M, and Médale F. 2022. 'A Plant-Based Diet Differentially Affects the Global Hepatic Methylome in Rainbow Trout Depending on Genetic Background'. Epigenetics, March, 1–12. https://doi.org/10.1080/15592294.2022.2058226.

Marburger, Sarah, Markos A. Alexandrou, John B. Taggart, Simon Creer, Gary Carvalho, Claudio Oliveira, and Martin I. Taylor. 2018. 'Whole Genome Duplication and Transposable Element Proliferation Drive Genome Expansion in Corydoradinae Catfishes'. Proceedings of the Royal Society B: Biological Sciences 285 (1872): 20172732. https://doi.org/10.1098/rspb.2017.2732.

Markevich, Grigorii, Evgeny Esin, and Liudmila Anisimova. 2018. 'Basic Description and Some Notes on the Evolution of Seven Sympatric Morphs of Dolly Varden Salvelinus Malma from the Lake Kronotskoe Basin'. Ecology and Evolution 8 (5): 2554–67. https://doi.org/10.1002/ece3.3806.

Matlosz, Sébastien, Benjamín Sigurgeirsson, Sigrídur Rut Franzdóttir, Arnar Pálsson, and Zophonías O. Jónsson. 2022. 'DNA Methylation Differences during Development Distinguish Sympatric Morphs of Arctic Charr (Salvelinus Alpinus)'. Molecular Ecology 31 (18): 4739–61. https://doi.org/10.1111/mec.16620.

McJunkin, Katherine. 2018. 'Maternal Effects of microRNAs in Early Embryogenesis'. RNA Biology 15 (2): 165–69. https://doi.org/10.1080/15476286.2017.1402999.

Mellén, Marian, Pinar Ayata, Scott Dewell, Skirmantas Kriaucionis, and Nathaniel Heintz. 2012. 'MeCP2 Binds to 5hmC Enriched within Active Genes and Accessible Chromatin in the Nervous System'. Cell 151 (7): 1417–30. https://doi.org/10.1016/j.cell.2012.11.022.

Metzger, David C. H., and Patricia M. Schulte. 2017. 'Persistent and Plastic Effects of Temperature on DNA Methylation across the Genome of Threespine Stickleback

(Gasterosteus Aculeatus)'. Proceedings. Biological Sciences 284 (1864): 20171667. https://doi.org/10.1098/rspb.2017.1667.

Metzger, David C H, and Patricia M Schulte. 2018. 'The DNA Methylation Landscape of Stickleback Reveals Patterns of Sex Chromosome Evolution and Effects of Environmental Salinity'. Genome Biology and Evolution 10 (3): 775–85. https://doi.org/10.1093/gbe/evy034.

Meyer, Axel, and Yves Van de Peer. 2005. 'From 2R to 3R: Evidence for a Fish-Specific Genome Duplication (FSGD)'. BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology 27 (9): 937–45. https://doi.org/10.1002/bies.20293.

Monsen, Øystein. 2022. 'The Contribution of Repetitive Elements to Salmonid Genome Evolution'. Doctoral thesis, Norwegian University of Life Sciences, Ås. https://nmbu.brage.unit.no/nmbu-xmlui/handle/11250/3047268.

Morgan, Hugh D., Heidi G. E. Sutherland, David I. K. Martin, and Emma Whitelaw. 1999. 'Epigenetic Inheritance at the Agouti Locus in the Mouse'. Nature Genetics 23 (3): 314–18. https://doi.org/10.1038/15490.

Nabel, Christopher S., Huijue Jia, Yu Ye, Li Shen, Hana L. Goldschmidt, James T. Stivers, Yi Zhang, and Rahul M. Kohli. 2012. 'AID/APOBEC Deaminases Disfavor Modified Cytosines Implicated in DNA Demethylation'. Nature Chemical Biology 8 (9): 751–58. https://doi.org/10.1038/nchembio.1042.

Nelson, Joseph, Terry Grande, and Mark Wilson. 2016. Fishes of the World, Fifth Edition. Fishes of the World: Fifth Edition. https://doi.org/10.1002/9781119174844.

Ng, H. H., and A. Bird. 1999. 'DNA Methylation and Chromatin Modification'. Current Opinion in Genetics & Development 9 (2): 158–63. https://doi.org/10.1016/s0959-437x(99)80024-0.

Nichols, James T., Bernardo Blanco-Sánchez, Elliott P. Brooks, Raghuveer Parthasarathy, John Dowd, Arul Subramanian, Gregory Nachtrab, Kenneth D. Poss, Thomas F. Schilling, and Charles B. Kimmel. 2016. 'Ligament versus Bone Cell Identity in the Zebrafish Hyoid Skeleton Is Regulated by Mef2ca'. Development (Cambridge, England) 143 (23): 4430–40. https://doi.org/10.1242/dev.141036.

Nilsson, Eric, Ingrid Sadler-Riggleman, Daniel Beck, and Michael K. Skinner. 2021. 'Differential DNA Methylation in Somatic and Sperm Cells of Hatchery vs Wild (Natural-Origin) Steelhead Trout Populations'. Environmental Epigenetics 7 (1): dvab002. https://doi.org/10.1093/eep/dvab002.

Ohno, Susumu. 1970. Evolution by Gene Duplication. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-86659-3.

Okano, M., D. W. Bell, D. A. Haber, and E. Li. 1999. 'DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for de Novo Methylation and Mammalian Development'. Cell 99 (3): 247–57. https://doi.org/10.1016/s0092-8674(00)81656-6.

Okano, M., S. Xie, and E. Li. 1998. 'Cloning and Characterization of a Family of Novel Mammalian DNA (Cytosine-5) Methyltransferases'. Nature Genetics 19 (3): 219–20. https://doi.org/10.1038/890.

Oliveira, Jordana Inácio Nascimento, Adauto Lima Cardoso, Ivan Rodrigo Wolf, Rogério Antônio de Oliveira, and Cesar Martins. 2022. 'First Characterization of PIWI-Interacting RNA Clusters in a Cichlid Fish with a B Chromosome'. BMC Biology 20 (1): 204. https://doi.org/10.1186/s12915-022-01403-2.

Ooi, Steen K. T., Chen Qiu, Emily Bernstein, Keqin Li, Da Jia, Zhe Yang, Hediye Erdjument-Bromage, et al. 2007. 'DNMT3L Connects Unmethylated Lysine 4 of Histone H3 to de Novo Methylation of DNA'. Nature 448 (7154): 714–17. https://doi.org/10.1038/nature05987.

Ortega-Recalde, Oscar, Alexander Goikoetxea, Timothy A. Hore, Erica V. Todd, and Neil J. Gemmell. 2020. 'The Genetics and Epigenetics of Sex Change in Fish'. Annual Review of Animal Biosciences 8 (1): 47–69. https://doi.org/10.1146/annurev-animal-021419-083634.

Oswald, J, S Engemann, N Lane, W Mayer, A Olek, R Fundele, W Dean, W Reik, and J Walter. 2000. 'Active Demethylation of the Paternal Genome in the Mouse Zygote'. Current Biology 10 (8): 475–78. https://doi.org/10.1016/S0960-9822(00)00448-6.

Pál, C., and I. Miklós. 1999. 'Epigenetic Inheritance, Genetic Assimilation and Speciation'. Journal of Theoretical Biology 200 (1): 19–37. https://doi.org/10.1006/jtbi.1999.0974.

Pardue, Mary-Lou, and P. G. DeBaryshe. 2011. 'Retrotransposons That Maintain Chromosome Ends'. Proceedings of the National Academy of Sciences of the United States of America 108 (51): 20317–24. https://doi.org/10.1073/pnas.1100278108.

Pasquinelli, A. E., B. J. Reinhart, F. Slack, M. Q. Martindale, M. I. Kuroda, B. Maller, D. C. Hayward, et al. 2000. 'Conservation of the Sequence and Temporal Expression of Let-7 Heterochronic Regulatory RNA'. Nature 408 (6808): 86–89. https://doi.org/10.1038/35040556.

Peng, Hongying, Junchao Shi, Ying Zhang, He Zhang, Shangying Liao, Wei Li, Li Lei, et al. 2012. 'A Novel Class of tRNA-Derived Small RNAs Extremely Enriched in Mature Mouse Sperm'. Cell Research 22 (11): 1609–12. https://doi.org/10.1038/cr.2012.141.

Perez, Marcos Francisco, and Ben Lehner. 2019. 'Intergenerational and Transgenerational Epigenetic Inheritance in Animals'. Nature Cell Biology 21 (2): 143–51. https://doi.org/10.1038/s41556-018-0242-9.

Pierron, Fabien, Guillemine Daffe, Patrick Lambert, Patrice Couture, and Magalie Baudrimont. 2019. 'Retrotransposon Methylation and Activity in Wild Fish (A. Anguilla): A Matter of Size'. Environmental Pollution (Barking, Essex: 1987) 245 (February): 494–503. https://doi.org/10.1016/j.envpol.2018.11.014.

Pierron, Fabien, Sophie Lorioux, Débora Héroin, Guillemine Daffe, Bruno Etcheverria, Jérôme Cachot, Bénédicte Morin, Sylvie Dufour, and Patrice Gonzalez. 2021. 'Transgenerational Epigenetic Sex Determination: Environment Experienced by Female Fish Affects Offspring Sex Ratio'. Environmental Pollution (Barking, Essex: 1987) 277 (May): 116864. https://doi.org/10.1016/j.envpol.2021.116864.

Piferrer, Francesc, Dafni Anastasiadi, Alejandro Valdivieso, Núria Sánchez-Baizán, Javier Moraleda-Prados, and Laia Ribas. 2019. 'The Model of the Conserved Epigenetic

Regulation of Sex'. Frontiers in Genetics 10: 857. https://doi.org/10.3389/fgene.2019.00857.

Podgorniak, Tomasz, Anusha Dhanasiri, Xianquan Chen, Xu Ren, Pei-Fen Kuan, and Jorge Fernandes. 2022. 'Early Fish Domestication Affects Methylation of Key Genes Involved in the Rapid Onset of the Farmed Phenotype'. Epigenetics, January, 1–18. https://doi.org/10.1080/15592294.2021.2017554.

Pratt, Ashley J., and Ian J. MacRae. 2009. 'The RNA-Induced Silencing Complex: A Versatile Gene-Silencing Machine'. The Journal of Biological Chemistry 284 (27): 17897–901. https://doi.org/10.1074/jbc.R900012200.

Rasmussen, Kasper Dindler, and Kristian Helin. 2016. 'Role of TET Enzymes in DNA Methylation, Development, and Cancer'. Genes & Development 30 (7): 733–50. https://doi.org/10.1101/gad.276568.115.

Ravi, Vydianathan, and Byrappa Venkatesh. 2018. 'The Divergent Genomes of Teleosts'. Annual Review of Animal Biosciences 6 (February): 47–68. https://doi.org/10.1146/annurev-animal-030117-014821.

Ren, Li, Hong Zhang, Mengxue Luo, Xin Gao, Jialin Cui, Xueyin Zhang, and Shaojun Liu. 2022. 'Heterosis of Growth Trait Regulated by DNA Methylation and miRNA in Allotriploid Fish'. Epigenetics & Chromatin 15 (1): 19. https://doi.org/10.1186/s13072-022-00455-6.

Robertson, Fiona M., Manu Kumar Gundappa, Fabian Grammes, Torgeir R. Hvidsten, Anthony K. Redmond, Sigbjørn Lien, Samuel A. M. Martin, Peter W. H. Holland, Simen R. Sandve, and Daniel J. Macqueen. 2017. 'Lineage-Specific Rediploidization Is a Mechanism to Explain Time-Lags between Genome Duplication and Evolutionary Diversification'. Genome Biology 18 (1): 111. https://doi.org/10.1186/s13059-017-1241-z.

Rodin, Sergei N., and Arthur D. Riggs. 2003. 'Epigenetic Silencing May Aid Evolution by Gene Duplication'. Journal of Molecular Evolution 56 (6): 718–29. https://doi.org/10.1007/s00239-002-2446-6.

Rodriguez Barreto, Deiene, Carlos Garcia de Leaniz, Eric Verspoor, Halina Sobolewska, Mark Coulson, and Sofia Consuegra. 2019. 'DNA Methylation Changes in the Sperm of Captive-Reared Fish: A Route to Epigenetic Introgression in Wild Populations'. Molecular Biology and Evolution 36 (10): 2205–11. https://doi.org/10.1093/molbev/msz135.

Sagonas, Kostas, Britta S. Meyer, Joshka Kaufmann, Tobias L. Lenz, Robert Häsler, and Christophe Eizaguirre. 2020. 'Experimental Parasite Infection Causes Genome-Wide Changes in DNA Methylation'. Molecular Biology and Evolution 37 (8): 2287–99. https://doi.org/10.1093/molbev/msaa084.

Salisbury, S. J., and D. E. Ruzzante. 2022. 'Genetic Causes and Consequences of Sympatric Morph Divergence in Salmonidae: A Search for Mechanisms'. Annual Review of Animal Biosciences 10 (February): 81–106. https://doi.org/10.1146/annurev-animal-051021-080709.

Sandlund, Odd Ter, Bror Jonsson, Hilmar J. Malmquist, Rolf Gydemo, Torfinn Lindem, Skúli Skúlason, Sigurdur S. Snorrason, and Pétur M. Jónasson. 1987. 'Habitat Use of

Arctic charrSalvelinus Alpinus in Thingvallavatn, Iceland'. Environmental Biology of Fishes 20 (4): 263–74. https://doi.org/10.1007/BF00005297.

Sandlund, Odd Terje, Karl Gunnarsson, Pétur M. Jónasson, Bror Jonsson, Torfinn Lindem, Kristinn P. Magnússon, Hilmar J. Malmquist, Hrefna Sigurjónsdóttir, Skúli Skúlason, and Sigurdur S. Snorrason. 1992. 'The Arctic Charr Salvelinus Alpinus in Thingvallavatn'. Oikos 64 (1/2): 305–51. https://doi.org/10.2307/3545056.

Santini, Francesco, Luke J Harmon, Giorgio Carnevale, and Michael E Alfaro. 2009. 'Did Genome Duplication Drive the Origin of Teleosts? A Comparative Study of Diversification in Ray-Finned Fishes'. BMC Evolutionary Biology 9 (August): 194. https://doi.org/10.1186/1471-2148-9-194.

Santos, Fátima, Brian Hendrich, Wolf Reik, and Wendy Dean. 2002. 'Dynamic Reprogramming of DNA Methylation in the Early Mouse Embryo'. Developmental Biology 241 (1): 172–82. https://doi.org/10.1006/dbio.2001.0501.

Sävilammi, Tiina, Spiros Papakostas, Erica H. Leder, L. Asbjørn Vøllestad, Paul V. Debes, and Craig R. Primmer. 2021. 'Cytosine Methylation Patterns Suggest a Role of Methylation in Plastic and Adaptive Responses to Temperature in European Grayling (Thymallus Thymallus) Populations'. Epigenetics 16 (3): 271–88. https://doi.org/10.1080/15592294.2020.1795597.

Schmitz, Jonathan F., Frédéric J. J. Chain, and Erich Bornberg-Bauer. 2020. 'Evolution of Novel Genes in Three-Spined Stickleback Populations'. Heredity 125 (1–2): 50–59. https://doi.org/10.1038/s41437-020-0319-7.

Schooten, Bas van, Jesyka Meléndez-Rosa, Steven M. Van Belleghem, Chris D. Jiggins, John D. Tan, W. Owen McMillan, and Riccardo Papa. 2020. 'Divergence of Chemosensing during the Early Stages of Speciation'. Proceedings of the National Academy of Sciences of the United States of America 117 (28): 16438–47. https://doi.org/10.1073/pnas.1921318117.

Seisenberger, Stefanie, Julian R. Peat, Timothy A. Hore, Fátima Santos, Wendy Dean, and Wolf Reik. 2013. 'Reprogramming DNA Methylation in the Mammalian Life Cycle: Building and Breaking Epigenetic Barriers'. Philosophical Transactions of the Royal Society B: Biological Sciences 368 (1609): 20110330. https://doi.org/10.1098/rstb.2011.0330.

Shea, N., I. Pen, and T. Uller. 2011. 'Three Epigenetic Information Channels and Their Different Roles in Evolution'. Journal of Evolutionary Biology 24 (6): 1178–87. https://doi.org/10.1111/j.1420-9101.2011.02235.x.

Si, Yufeng, Feng He, Haishen Wen, Siping Li, and Huiwen He. 2021. 'Effects of Low Salinity on Epigenetic Changes of Growth Hormone and Growth Hormone Receptor in Half Smooth Tongue Sole (Cynoglossus Semilaevis)'. Reproduction and Breeding 1 (1): 11–21. https://doi.org/10.1016/j.repbre.2021.01.001.

Simakov, Oleg, Ferdinand Marlétaz, Jia-Xing Yue, Brendan O'Connell, Jerry Jenkins, Alexander Brandt, Robert Calef, et al. 2020. 'Deeply Conserved Synteny Resolves Early Events in Vertebrate Evolution'. Nature Ecology & Evolution 4 (6): 820–30. https://doi.org/10.1038/s41559-020-1156-z.

Sinzelle, L., Z. Izsvák, and Z. Ivics. 2009. ‘Molecular Domestication of Transposable Elements: From Detrimental Parasites to Useful Host Genes’. Cellular and Molecular Life Sciences: CMLS 66 (6): 1073–93. https://doi.org/10.1007/s00018-009-8376-3.

Siomi, Mikiko C., Kaoru Sato, Dubravka Pezic, and Alexei A. Aravin. 2011. ‘PIWI-Interacting Small RNAs: The Vanguard of Genome Defence’. Nature Reviews. Molecular Cell Biology 12 (4): 246–58. https://doi.org/10.1038/nrm3089.

Skinner, Michael K. 2015. ‘Environmental Epigenetics and a Unified Theory of the Molecular Aspects of Evolution: A Neo-Lamarckian Concept That Facilitates Neo-Darwinian Evolution’. Genome Biology and Evolution 7 (5): 1296–1302. https://doi.org/10.1093/gbe/evv073.

Skúlason, S., S. S. Snorrason, D. L. G. Noakes, M. M. Ferguson, and H. J. Malmquist. 1989. ‘Segregation in Spawning and Early Life History among Polymorphic Arctic Charr, Salvelinus Alpinus, in Thingvallavatn, Iceland’. Journal of Fish Biology 35 (sA): 225–32. https://doi.org/10.1111/j.1095-8649.1989.tb03065.x.

SKÚLASON, SKÚLI, DAVID L. G. NOAKES, and SIGURDUR S. SNORRASON. 1989. ‘Ontogeny of Trophic Morphology in Four Sympatric Morphs of Arctic Charr Salvelinus Alpinus in Thingvallavatn, Iceland’. Biological Journal of the Linnean Society 38 (3): 281–301. https://doi.org/10.1111/j.1095-8312.1989.tb01579.x.

Slotkin, R. Keith, and Robert Martienssen. 2007. ‘Transposable Elements and the Epigenetic Regulation of the Genome’. Nature Reviews Genetics 8 (4): 272–85. https://doi.org/10.1038/nrg2072.

Smallwood, Sébastien A., and Gavin Kelsey. 2012. ‘De Novo DNA Methylation: A Germ Cell Perspective’. Trends in Genetics 28 (1): 33–42. https://doi.org/10.1016/j.tig.2011.09.004.

Suetake, Isao, Fuminori Shinozaki, Junichi Miyagawa, Hideyuki Takeshima, and Shoji Tajima. 2004. ‘DNMT3L Stimulates the DNA Methylation Activity of Dnmt3a and Dnmt3b through a Direct Interaction’. The Journal of Biological Chemistry 279 (26): 27816–23. https://doi.org/10.1074/jbc.M400181200.

Szulwach, Keith E., Xuekun Li, Yujing Li, Chun-Xiao Song, Hao Wu, Qing Dai, Hasan Irier, et al. 2011. ‘5-hmC-Mediated Epigenetic Dynamics during Postnatal Neurodevelopment and Aging’. Nature Neuroscience 14 (12): 1607–16. https://doi.org/10.1038/nn.2959.

Varriale, Annalisa, and Giorgio Bernardi. 2006. ‘DNA Methylation and Body Temperature in Fishes’. Gene 385 (December): 111–21. https://doi.org/10.1016/j.gene.2006.05.031.

Ven, Leo T.M. van der, Jaroslav Jelinek, Hennie M. Hodemaekers, Edwin P. Zwart, Sander Ruiter, Evert-Jan van den Brandhof, Jean-Pierre J. Issa, Jeroen L.A. Pennings, and Mirjam Luijten. 2017. ‘An Adverse Outcome Pathway Analysis Employing DNA Methylation Effects in Arsenic-Exposed Zebrafish Embryos Supports a Role of Epigenetic Events in Arsenic-Induced Chronic Disease’. Applied In Vitro Toxicology 3 (4): 312–24. https://doi.org/10.1089/aivt.2017.0018.

Venney, Clare J, Dafni Anastasiadi, Maren Wellenreuther, and Louis Bernatchez. 2023. ‘The Evolutionary Complexities of DNA Methylation in Animals: From Plasticity to

Genetic Evolution'. Genome Biology and Evolution 15 (12): evad216. https://doi.org/10.1093/gbe/evad216.

Venney, Clare J., Kyle W. Wellband, Eric Normandeau, Carolyne Houle, Dany Garant, Céline Audet, and Louis Bernatchez. 2022. 'Thermal Regime during Parental Sexual Maturation, but Not during Offspring Rearing, Modulates DNA Methylation in Brook Charr (Salvelinus Fontinalis)'. bioRxiv. https://doi.org/10.1101/2022.02.25.481661.

Venney, Clare, Oliver Love, Ellen Drown, and Daniel Heath. 2020. 'DNA Methylation Profiles Suggest Intergenerational Transfer of Maternal Effects', February. https://doi.org/10.1093/molbev/msz244.

Venter, J. Craig, Mark D. Adams, Eugene W. Myers, Peter W. Li, Richard J. Mural, Granger G. Sutton, Hamilton O. Smith, et al. 2001. 'The Sequence of the Human Genome'. Science 291 (5507): 1304–51. https://doi.org/10.1126/science.1058040.

Verdone, Loredana, Eleonora Agricola, Micaela Caserta, and Ernesto Di Mauro. 2006. 'Histone Acetylation in Gene Regulation'. Briefings in Functional Genomics & Proteomics 5 (3): 209–21. https://doi.org/10.1093/bfgp/ell028.

Volff, J.-N. 2005. 'Genome Evolution and Biodiversity in Teleost Fish'. Heredity 94 (3): 280–94. https://doi.org/10.1038/sj.hdy.6800635.

Volpe, Thomas A., Catherine Kidner, Ira M. Hall, Grace Teng, Shiv I. S. Grewal, and Robert A. Martienssen. 2002. 'Regulation of Heterochromatic Silencing and Histone H3 Lysine-9 Methylation by RNAi'. Science 297 (5588): 1833–37. https://doi.org/10.1126/science.1074973.

Waddington, C. H. 1942. 'The Epigenotype'. Endeavour 1: 18–20.

Wang, Feilong, Zuliang Qin, Zhiqiang Li, Shuangyi Yang, Tian Gao, Lina Sun, and Deshou Wang. 2021. 'Dnmt3aa but Not Dnmt3ab Is Required for Maintenance of Gametogenesis in Nile Tilapia (Oreochromis Niloticus)'. International Journal of Molecular Sciences 22 (18): 10170. https://doi.org/10.3390/ijms221810170.

Wang, Fei-Long, Long-Xia Yan, Hong-Juan Shi, Xing-Yong Liu, Qiao-Yuan Zheng, Li-Na Sun, and De-Shou Wang. 2018. 'Genome-Wide Identification, Evolution of DNA Methyltransferases and Their Expression during Gonadal Development in Nile Tilapia'. Comparative Biochemistry and Physiology. Part B, Biochemistry & Molecular Biology 226 (December): 73–84. https://doi.org/10.1016/j.cbpb.2018.08.007.

Wang, Lu, Jun Zhang, Jialei Duan, Xinxing Gao, Wei Zhu, Xingyu Lu, Lu Yang, et al. 2014. 'Programming and Inheritance of Parental DNA Methylomes in Mammals'. Cell 157 (4): 979–91. https://doi.org/10.1016/j.cell.2014.04.017.

Wellband, Kyle, David Roth, Tommi Linnansaari, R. Allen Curry, and Louis Bernatchez. 2021. 'Environment-Driven Reprogramming of Gamete DNA Methylation Occurs during Maturation and Is Transmitted Intergenerationally in Atlantic Salmon'. G3 (Bethesda, Md.) 11 (12): jkab353. https://doi.org/10.1093/g3journal/jkab353.

West-Eberhard, Mary Jane. 2005. 'Developmental Plasticity and the Origin of Species Differences'. Proceedings of the National Academy of Sciences of the United States of America 102 Suppl 1 (Suppl 1): 6543–49. https://doi.org/10.1073/pnas.0501844102.

Wicker, Thomas, François Sabot, Aurélie Hua-Van, Jeffrey L. Bennetzen, Pierre Capy, Boulos Chalhoub, Andrew Flavell, et al. 2007. 'A Unified Classification System for Eukaryotic Transposable Elements'. Nature Reviews Genetics 8 (12): 973–82. https://doi.org/10.1038/nrg2165.

Wolfe, K. H. 2001. 'Yesterday's Polyploids and the Mystery of Diploidization'. Nature Reviews. Genetics 2 (5): 333–41. https://doi.org/10.1038/35072009.

Wu, Ming, Li Li, and Zhirong Sun. 2007. 'Transposable Element Fragments in Protein-Coding Regions and Their Contributions to Human Functional Proteins'. Gene 401 (1–2): 165–71. https://doi.org/10.1016/j.gene.2007.07.012.

Wu, Susan C., Eric M. Kallin, and Yi Zhang. 2010. 'Role of H3K27 Methylation in the Regulation of lncRNA Expression'. Cell Research 20 (10): 1109–16. https://doi.org/10.1038/cr.2010.114.

Wu, Xiaoji, and Yi Zhang. 2017. 'TET-Mediated Active DNA Demethylation: Mechanism, Function and Beyond'. Nature Reviews Genetics 18 (9): 517–34. https://doi.org/10.1038/nrg.2017.33.

Wysocka, Joanna, Tomek Swigut, Hua Xiao, Thomas A. Milne, So Yeon Kwon, Joe Landry, Monika Kauer, et al. 2006. 'A PHD Finger of NURF Couples Histone H3 Lysine 4 Trimethylation with Chromatin Remodelling'. Nature 442 (7098): 86–90. https://doi.org/10.1038/nature04815.

Xiao, Jun, Can Song, Shaojun Liu, Min Tao, Jie Hu, Jun Wang, Wei Liu, Ming Zeng, and Yun Liu. 2013. 'DNA Methylation Analysis of Allotetraploid Hybrids of Red Crucian Carp (Carassius Auratus Red Var.) and Common Carp (Cyprinus Carpio L.)'. PloS One 8 (2): e56409. https://doi.org/10.1371/journal.pone.0056409.

Xu, Guo-Liang, and Jiemin Wong. 2015. 'Oxidative DNA Demethylation Mediated by Tet Enzymes'. National Science Review 2 (3): 318–28. https://doi.org/10.1093/nsr/nwv029.

Yamaguchi, Shinpei, Li Shen, Yuting Liu, Damian Sendler, and Yi Zhang. 2013. 'Role of Tet1 in Erasure of Genomic Imprinting'. Nature 504 (7480): 460–64. https://doi.org/10.1038/nature12805.

Zeng, Yang, and Taiping Chen. 2019. 'DNA Methylation Reprogramming during Mammalian Development'. Genes 10 (4): 257. https://doi.org/10.3390/genes10040257.

Paper I

DNA Methylation Differences during Development Distinguish Sympatric Morphs of Arctic Charr (*Salvelinus Alpinus*)

Matlosz, Sébastien, Benjamín Sigurgeirsson, Sigrídur Rut Franzdóttir, Arnar Pálsson, and Zophonías O. Jónsson.

Molecular Ecology 2022, 31 (18): 4739–61.
https://doi.org/10.1111/mec.16620.

Author contribution: S. Matlosz participated in sampling, performed the experiments, analyzed the data and wrote the manuscript.

Received: 15 September 2021 | Revised: 13 June 2022 | Accepted: 4 July 2022

DOI: 10.1111/mec.16620

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

DNA methylation differences during development distinguish sympatric morphs of Arctic charr (*Salvelinus alpinus*)

Sébastien Matlosz | **Benjamín Sigurgeirsson** | **Sigrídur Rut Franzdóttir** | **Arnar Pálsson** | **Zophonías O. Jónsson**

Institute of Life and Environmental Sciences, University of Iceland, Reykjavík, Iceland

Correspondence
Sébastien Matlosz, Institute of Life and Environmental Sciences, University of Iceland, Sturlugata 7, Reykjavík, Iceland.
Email: slt2@hi.is

Funding information
Háskóli Íslands; Icelandic Centre for Research, Grant/Award Number: 163477-051

Handling Editor: Sean Rogers

Abstract

Changes in DNA methylation in specific coding or non-coding regions can influence development and potentially divergence in traits within species and groups. While the impact of epigenetic variation on developmental pathways associated with evolutionary divergence is the focus of intense investigation, few studies have looked at recently diverged systems. Phenotypic diversity between closely related populations of Arctic charr (*Salvelinus alpinus*), which diverged within the last 10,000 years, offers an interesting ecological model to address such effects. Using bisulphite sequencing, we studied general DNA methylation patterns during development in the four sympatric morphs of Arctic charr from Lake Thingvallavatn. The data revealed strong differences between developmental timepoints and between morphs (mainly along the benthic–limnetic axis), both at single CpG sites and in 1000 bp-regions. Genes located close to differentially methylated CpG sites were involved in nucleosome assembly, regulation of osteoclast differentiation, and cell-matrix adhesion. Differentially methylated regions were enriched in tRNA and rRNA sequences, and half of them were located close to transcription start sites. The expression of 14 genes showing methylation differences over time or between morphs was further investigated by qPCR and nine of these were found to be differentially expressed between morphs. Four genes (*ARHGEF37-like*, *H3-like*, *MPP3* and *MEGF9*) showed a correlation between methylation and expression. Lastly, histone gene clusters displayed interesting methylation differences between timepoints and morphs, as well as intragenic methylation variation. The results presented here provide a motivation for further studies on the contribution of epigenetic traits, such as DNA methylation, to phenotypic diversity and developmental mechanisms.

KEYWORDS
Bisulfite sequencing, DNA methylation, Ecological genetics, Molecular evolution, RRBS, Salmonid

1 | INTRODUCTION

Variation in development, driven by genetic differences, can provide raw material for adaptive evolution and the emergence of new traits (Richardson, 1999; Wilkins, 2001). By investigating differential regulation of molecular mechanisms between species or populations at the embryonic level, possible causes of adaptive divergence can be unravelled. Advances in high throughput sequencing technologies have enabled the study of the evolution of regulatory mechanisms in various systems (Brawand et al., 2014; Jones et al., 2012;

Palmer & Kronforst, 2015) and some studies have associated specific loci and mechanisms with examples of ecological adaptations (Han et al., 2017). Among the many levels of molecular regulation, the term epigenetics refers to molecular alterations that impact gene expression and phenotypes without directly changing the nucleotide sequence, such as DNA or histone chemical modifications. It has been hypothesized that epigenetic changes (conserved and inherited), could increase the rate of adaptive evolution (Klironomos et al., 2013; Schmitz et al., 2011; Skinner et al., 2014). At the individual level, epigenetic changes induced by environmental factors can also influence phenotypic plasticity (Geng et al., 2013; Putnam et al., 2016; Roberts & Gavery, 2012), the capacity of identical genomes to create different phenotypes. Despite significant advances in the field of evolutionary epigenetics many unresolved questions remain, concerning causes, consequences and interplay of epigenetic variation on evolution. Profiling epigenetic patterns during development can provide further understanding of how changes in the epigenome correlate with or contribute to phenotypic variation.

In vertebrates, DNA methylation occurs almost exclusively at CpG dinucleotides (cytosine bases located next to guanines). Regions of the genome containing large numbers of CpG dinucleotides are referred to as CpG islands and are primarily found in promoter regions. The complex interaction between methylation, chromatin state and gene expression has been a focus of epigenetic research for more than 20 years (Bird & Wolffe, 1999; Jones, 1999). The numerous studies on the matter make it clear that the impact of DNA methylation on gene expression is highly context dependent, with the position of DNA methylation marks in relation to different gene features playing an important role. Methylation in CpG islands located at promoters is associated with gene silencing, for example during abnormal gene silencing in cancers (Esteller, 2002). Similarly, a negative correlation between methylation of the first intron and gene expression was observed in different vertebrate species (Anastasiadi et al., 2018). However, hypermethylation of CpG islands at the 5′-UTR region or in the gene body of oncogenes has been correlated with their overexpression (Arechederra et al., 2018).

Several methods to quantify the state of cytosine methylation have been developed (Fraga & Esteller, 2002), the most widely used being bisulphite conversion (Frommer et al., 1992), a chemical reaction that converts unmethylated cytosine to uracil while leaving methylated cytosines unchanged. Sequencing of bisulphite converted DNA can be done at different scales: Whole genome bisulphite sequencing (WGBS) gives whole genome coverage whereas reduced representation bisulphite sequencing (RRBS) (Meissner et al., 2005) uses a CpG targeting enzyme (most commonly *Msp*I) to fragment the DNA at CpG rich regions and thus focus the sequencing on CpG islands, resulting in higher cytosine coverage at these sites than would be obtained through WGBS for similar sequencing costs. RRBS has been used to investigate DNA methylation differences between reared and wild fish populations (Gavery et al., 2018; Le Luyer et al., 2017). Most studies investigating methylation differences in fish have done so to evaluate the impact of the environment, such as temperature (Navarro-Martín et al., 2011; Varriale & Bernardi, 2006), salinity (Heckwolf et al., 2020) or other chemicals (Fuzinatto et al., 2015; Liu et al., 2016) on methylation. However, few studies have addressed how DNA methylation varies in diverging fish lineages (Cayuela et al., 2021) and rarely has this been done on within species variation (notable exceptions being; Hu et al., 2021; Vernaz et al., 2021).

Some questions remain unanswered regarding the role and causes of DNA methylation in evolutionary context. For instance: Does variation in DNA methylation patterns and/or the regulation of methylation play an active role in divergence? Are these patterns induced by genetic differences between groups or by environmental effects on individuals? And what is the interplay of expression and DNA methylation divergence during development (Lind & Spagopoulou, 2018; Verhoeven Koen et al., 2016)?

To inquire how DNA methylation during development has evolved in recently separated populations, we utilized sympatric ecomorphs (morphs hereafter) of Arctic charr (*Salvelinus alpinus*) found in Lake Thingvallavatn, Iceland. *S. alpinus* is a fish species present in Arctic and sub-Arctic lakes and coastal waters. It is well known for its high degree of resource and phenotypic polymorphism all over its range. Arctic charr invaded many freshwater systems after the last glacial retreat, in the case of lake Thingvallavatn about ~10,000 years (~3000 generations) ago (Kapralova et al., 2011). Cases of two or more sympatric charr morphs have been described in Iceland (Jonsson et al., 1988; Sandlund et al., 1992), Norway (Hindar & Jonsson, 1982; Tamayo et al., 2020), Scotland (Adams et al., 1998), and Canada (Arbour et al., 2010), while other *Salvelinus* species exhibit polymorphism in Russia (Markevich et al., 2018). The charr in Lake Thingvallavatn represents a well studied example, displaying prominent polymorphic diversity with four different morphs: Small Benthic (SB), Large Benthic (LB), Planktivorous (PL) and Piscivorous (PI). These morphs differ by their habitat use (Sandlund et al., 1987), diet (Malmquist et al., 1992) and reproductive behaviour (Skúlason, Noakes, & Snorrason, 1989; Skúlason, Snorrason, et al., 1989), but also by their size and the structure of their trophic apparatus. Specifically, benthic morphs (SB and LB) exhibit a subterminal mouth and blunt snout, while limnetic morphs (PL and PI) have a terminal mouth and pointed snout (Sandlund et al., 1992). Data indicate that a population of anadromous *S. alpinus* became landlocked in the Thingvallavatn watershed before diverging into the distinct morphs, as eruptions and tectonic forces isolated this population from the ocean (Kapralova et al., 2011). Analysis of single nucleotide polymorphism (SNP) in RNA-seq data revealed clear genetic separation between three of the morphs [PL, LB and SB (Guðbrandsson et al., 2019)]. More recent genome analysis using ddRADSeq and whole genome sequencing places most individuals of the PI morph genetically close to the PL morph, with others showing varying levels of admixture with LB (Han Xiao, Zophonias O. Jónsson et al., unpublished data). On the expression level, transcriptomic studies found more than a thousand genes to be differentially expressed between three of these morphs during early development (Guðbrandsson et al., 2017), and candidate gene analysis revealed gene expression differences strongly associated with this polymorphism. For instance, *Mmp2* and *Sparc*, two genes associated with matrix remodelling show different expression between

benthic and limnetic morphs (Ahi et al., 2013) and are part of a conserved network of coexpressed genes (possibly under the regulation of *Ets2* and *Ap1*) showing distinct transcriptional dynamics between the heads of benthic and limnetic embryos (Ahi et al., 2014). Similarly, divergence in the expression of genes associated with the mTOR pathway was associated with charr dwarfism in Icelandic charr (Macqueen et al., 2011) and significant differences in gene and miRNA expression were observed between Thingvallavatn SB charr and aquaculture charr (displaying a limnetic craniofacial phenotype) during embryonic development (Gudbrandsson et al., 2016; Kapralova et al., 2014). Together these studies demonstrate extensive variation in multiple molecular systems between the morphs. However, the molecular machinery underlying development is complex and multilayered. While the miRNA study showcases a type of epigenetic variation within Arctic charr, other epigenetic mechanisms impacting gene expression could be at play. In order to expand on the understanding of this system, we reared embryos from the four Arctic charr morphs of Lake Thingvallavatn under common garden conditions and profiled their DNA methylome at four timepoints prior to hatching, using RRBS. We also evaluated the expression of 14 differentially methylated genes with qPCR. It is important to note that the common garden approach taken in this study allows us to remove potential effects of environmentally induced phenotypic plasticity, leaving us with a focus on epigenetic marks transmitted directly or indirectly (via genetic differences) from the parents.

With this data, we first explore changes in methylation over development and potential differences in methylation patterns between the morphs. We hypothesize that differences in methylation could impact specific biological processes and perform gene ontology (GO) analysis of genes located in the vicinity of CpG sites showing methylation differences. Next, as some repeat elements (such as transposons) are known to be the targets of the methylation machinery (Karimi et al., 2011), we look at differentially methylated regions (DMRs) and determine their content in repeat element sequences to check whether specific repeats are differentially methylated between Arctic charr morphs or timepoints. Finally, we explore the correlations between methylation and nearby gene expression. By quantifying the expression of 14 genes showing varying methylation patterns, we investigate how methylation impacts gene expression in this model. If correlation is seen between gene expression and methylation, we hypothesize that the location of methylation sites with respect to specific gene features (e.g., exon vs. promoter) would be indicative of the direction of correlation.

2 | MATERIALS AND METHODS

2.1 | Sampling and rearing of embryos

Sexually mature fish belonging to all four Arctic charr morphs from Lake Thingvallavatn were caught in fall 2017 using gill nets. Fishing permissions were obtained from the Thingvellir National Park Commission and the Directorate of Fisheries. Fish were killed by a sharp blow to the head immediately before collecting eggs and milt. Crosses were generated by fertilizing eggs from individual females, each with milt from a single individual male (further details on families are in Table S1). Eggs were reared at approximately 5°C in a hatching tray under constant water flow and in complete darkness at the Keldur aquaculture facility in Reykjavík. The water temperature was monitored constantly (at 10 s intervals) and the daily averages were used to estimate the relative age of the embryos using tau-somite (τs) units defined as the time it takes for one somite pair to form at a given temperature (Gorodilov, 1996). Embryos were sampled into both 96% EtOH and RNA-later at four different timepoints after the onset of organogenesis and before hatching (50, 100, 150 and 200 τs, for pictures of 100–200 τs, see Gudbrandsson et al., 2016). The samples were stored overnight at 4°C. The next day, 96% EtOH was replaced using a sterile pipette. After 3–4 h, the tubes were stored at −20°C for further analysis.

2.2 | DNA extraction

For 100, 150 and 200 τs, DNA was extracted from single embryos. For 50 τs, each DNA sample comes from a pool of two embryos from the same cross to obtain sufficient yield for further bisulphite conversion and library preparation.

Each sample was digested overnight at 56°C in a microcentrifuge tube containing 10 µl of proteinase K at 800 U/ml, 10 µl of 1 M DTT and 480 µl of lysis buffer (25 mM Tris HCl pH 7.5, 100 mM EDTA pH 8, 1% SDS). Then, DNA was extracted using standard phenol–chloroform extraction with Isopropanol precipitation, and resuspended in 100 µl TE-low EDTA buffer (10 mM Tris, 0.1 mM EDTA). DNA concentrations were quantified using the Qubit dsDNA HS Assay Kit according to the manufacturer's instructions (Thermo Fisher Scientific).

2.3 | RRBS preparation

In total, 48 samples were prepared for sequencing: three individuals (or three pools of two individuals at 50 τs) per morph and developmental timepoint. To maximize biological variance and avoid parental effects, each of the three biological replicates came from a different cross, with each cross having unique parents (Table S1). For each sample, 3 µg of DNA was digested with 3 µl of type II restriction enzyme *Msp*I (New England Biolabs) (EC:3.1.21.4) in a total of 150 µl of 1× NEB buffer. The DNA was then cleaned up using NucleoMag NGS beads (Macherey-Nagel GmbH Co. KG) with a ratio of 1.8 µl beads for each µl of DNA. Elution was done with 50 µl of TE-low EDTA buffer (10 mM Tris pH 7.8, 0.001 mM EDTA). From this point, all the steps from the End Repair to the PCR amplification were done according to the NEXTflex bisulphite sequencing Kit (Bioscientific) protocol, except for the second step (Step B: Clean-Up) where Nucleomag NGS beads were again used instead of AMPure XP beads (Beckman-Coulter). NEXTflex Bisulphite-Seq Barcodes – 6 (Bioscientific) were used for barcoding.

2.4 | High throughput sequencing

For sequencing, 0.066 pmol of six different DNA samples (each with individualized barcodes) were combined (for a total of 0.4 pmol) and subsequently diluted in 10 nM Tris solution pH 8.5 (with 0.1% Tween), 0.2 M NaOH and HT1 buffer from the MiSeq reagent kits version 2 (Illumina, Inc.) to obtain 1 ml of 8 pM library. Because of the low GC percentage of the DNA samples due to the bisulphite conversion, and to avoid problems with the phasing of clusters on the Illumina machine, 100 µl of 12.5 pM PhiX control library was added to the 8 pM library and 600 µl of this final preparation were sequenced on an Illumina MiSeq 2×150 bp. In total, eight libraries were generated to sequence the 48 samples.

2.5 | Analysis of methylation-seq data

Given that C/T DNA polymorphisms in the study population can be falsely interpreted as epigenetic variation, data from whole genome sequencing of eight individuals (two per morph: one male, one female) were used to call C/T SNPs. The *Salvelinus* sp. genome (*Salvelinus* sp. IW2-2015, assembly ASM291031v2, https://www.ncbi.nlm.nih.gov/assembly/GCF_002910315.2/) was masked with 4,659,191 C/T SNPs using the bedtools maskfasta command (Quinlan & Hall, 2010). This masked genome was then indexed and prepared using the bismark genome preparation function (Bismark version 0.19.0) (Krueger & Andrews, 2011). Reads from the Illumina MiSeq sequencing were trimmed for quality ($q > 10$) and adapter and PhiX sequences were removed using bbduk.sh function (SourceForge). Paired reads were merged using bbmerge.sh function of the same package. The merged reads were aligned to the masked genome with Bismark aligner (Bismark version 0.19.0). The bismark_methylation_extractor function was used to get CpG methylation information for those merged reads.

Following this, methylation data was analysed by first examining single CpG sites, using only the 10,340 CpGs that had sufficient sequencing coverage (see below in Section 2.5.1) in each of the 48 samples. We then interrogated bigger regions (1000 bp) containing multiple CpGs. This second, more exploratory approach used pairwise comparisons in order to increase the number of CpG sites taken into account (see below in Section 2.5.2).

2.5.1 | Analysis of individual CpGs

The output files from the bismark_methylation_extractor function were filtered with the MethylKit R package (Akalin et al., 2012). For each sample, bases with less than 10× coverage or more than 99.9th percentile of coverage were discarded to account for potential PCR bias, and the normalize Coverage function was used to normalize coverage between samples. The unite function was then used to keep only bases covered in all 48 samples, leading to 10,340 CpG sites. For each of these sites, and for each sample, the methylation percentage was calculated by dividing the number of cytosines by the total number of bases sequenced at this position (cytosines + thymines).

The methylation percentage of these 10,340 CpGs was subject to principal component analysis (PCA) with the prcomp function in R. The contribution of each sample to each principal component (PC) was used to assess whether the variables (morph, timepoint, sex) explained any of the PCs:

Two linear models were used:

$$PCx = Morph + Timepoint + Sex$$

and

$$PCx = Morph \times Timepoint \times Sex$$

The model with the lowest AICc (or the fixed effect model when there were no significant differences between the models) was used for analysis of variance (ANOVA) testing. As most PCs follow a normal distribution, they were analysed as such. The models for the first four PCs were diagnosed with the DHARMa package (Hartig, 2022), and found to be satisfactory (Figure S1).

To explore the patterns observed with the PCA, we used two complementary approaches.

After identifying PCs explained by variables of interest, the loadings of each of the 10,340 sites for these PCs were studied, and CpGs with a loading lying outside two standard deviations of the distribution were used for GO analysis (see Section 2.8). Alternatively, we used a custom generalized mixed model to identify the CpGs most strongly influenced by the independent variables:

$$CpG = Morph + Timepoint + Sex + Morph \times Timepoint + Morph \times Sex + Timepoint \times Sex$$

This model (using glm in R) was run for each of the 10,340 positions, assuming Gaussian distributions. Corrections for multiple testing were done with the Bonferroni method and CpGs with a corrected p-value <.05 were used for GO analysis (see Section 2.8).

To compare the statistical approaches, for each CpG and independent variable (time and morph), the corrected p-value obtained in the linear model was plotted against the absolute value of the CpG's weight in the corresponding PC (1 and 4), see Figure S2.

2.5.2 | Differentially methylated regions

To identify methylation differences in regions containing three or more CpGs, we performed multiple pairwise comparisons [with the MethylKit R package (Akalin et al., 2012)], which reduced the number of samples taken into account for each analysis and, as a result, increased the number of CpG sites passing filtering criteria. For each sample, bases with less than 10× coverage or more than 99.9th percentile of coverage were discarded to account for potential PCR bias, and the normalizeCoverage function was

used to normalize coverage between samples (in each pairwise comparison).

The list of all 35 pairwise comparisons performed can be found in Table S2. For pairwise comparisons between morphs or between timepoints, developmental timepoints and morphs batch-effects were accounted for with the assocComp and removeComp functions in MethylKit (to detect components accounting for those variables and remove the contributing cytosines from the data). Methylation was summarized over tiling windows (1000 bp long regions with at least three CpGs) using the tileMethylCount function for further analysis. Logistic regression was then used to identify DMRs with the calculateDiffMeth function. Regions with q-value <.01 and at least 5% methylation difference between conditions in any of the comparisons were defined as DMRs. This threshold value was chosen empirically, to select a reasonable number of candidate regions for consideration, while avoiding an excessive number of sites likely to be false positives. However, it is important to note that candidate genes selected for qPCR based on this DMR approach were selected among the ones showing the highest methylation differences (i.e., up to 63% methylation difference for *SLC9A3R2-like*).

The genomic position of gene features (promoter, exon, intron, transcription start site [TSS]) were obtained from the annotated *Salvelinus* sp. genome (*Salvelinus* sp. IW2-2015, assembly ASM291031v2) using the readTranscriptFeatures function from the genomation package (Akalin et al., 2015).

2.6 | Calculating average methylation over a specific region

Methylation average was estimated for 16 regions of particular interest. Differences in RRBS coverage called for different ways of calculating average methylation: for the five regions with high coverage (regions close to *H1*, *H2B*, *H2A-like*, *H3-like* and *RASSF4-like*) only CpG sites covered in all 48 samples were used to calculate methylation means. This way, we reduced the influence of coverage and methylation heterogeneities over the region, and interindividual variation was taken into account with the use of every biological replicate. For the remaining 11 regions, methylation was first averaged over all cytosines for each sample and the mean calculated between biological replicates.

Two linear models were used:

$$\text{Average methylation} = \text{Morph} + \text{Timepoint}$$

and

$$\text{Average methylation} = \text{Morph} \times \text{Timepoint}$$

The model with the lowest AICc was used for ANOVA testing unless there were no significant differences between the models, in which case the fixed effect model was used. The data for *ARL16*, *MEGF9*, *MEIS1-like*, *NFIX*, *NKX23-like*, *ARHGEF37-like*, *SLC9A3R2-like* and *H2A-like* did not follow normal distribution and was Box-Cox transformed. ANOVA was then used to test for significant methylation differences between morphs or between timepoints and Tukey's HSD was used for pairwise tests between groups. Again the models were diagnosed with the DHARMa package, see Figure S3. For analysis of methylation differences between gene features (exon vs. promoter) for *H1* and *H2B*, a "feature" variable was added.

In parallel to average methylation graphs, more detailed methylation maps for these regions were made by plotting the methylation percentage of single CpGs (Figure S8).

2.7 | RNA extraction

A total of 36 embryos (three biological replicates per morph and developmental timepoints: 100, 150, 200 τs) were dechorionated in RNAlater (7 M $[NH_4]_2SO_4$, 16.7 mM $Na_3C_6H_5O_7$, 13.3 mM EDTA, pH = 5.2) and homogenized individually in TRI reagent solution (Thermo Fisher Scientific) using a Pellet Pestle Cordless Motor (Kimble Kontes). Phase separation was done by adding 1-bromo-3-chloropropane to the TRI reagent solution at a 1/5 ratio (v/v). After centrifugation, precipitation was done by adding Isopropanol to the upper phase at a 1/1 ratio (v/v). Standard EtOH washes were performed to clean the pellet. RNA was resuspended in RNase-free H_2O and stored at −150°C. Sufficient RNA could not be extracted from individual embryos at 50 τs.

2.8 | Analysis of candidate genes with qPCR

Reverse transcription was performed on the 36 RNA samples using high-capacity cDNA reverse transcription kit (Thermo Fisher Scientific).

qPCR reactions were performed with Maxima SYBR Green/ROX qPCR Master Mix (Thermo Fisher Scientific) in the QuantStudio3 machine (Thermo Fisher Scientific). Two technical replicates were used for each primer-sample combination. The reference genes were *Actb* and *Ub2l3*, which were previously demonstrated to be suitable for expression analysis in Arctic charr embryos (Ahi et al., 2013). Primer sequences and efficiencies are available in Table S3. To account for differences in primer efficiencies, normalized relative quantities were calculated using a model by Hellemans et al. (2007) which was adapted from Pfaffl's equation (Pfaffl, 2001) to take into account multiple reference genes. This equation requires a sample to be chosen as normalizer, and S80 (from the 17PI3_100 cross) was arbitrarily chosen as the normalizer.

The data for *ARL16*, *H2A-like*, *LMTK2*, *MPP3*, *NFIX* and *SLC9A3R2-like* did not follow normal distribution and was Box-Cox transformed. ANOVA was then used to test for significant expression differences between morphs or between timepoints and Tukey's HSD was used for pairwise tests between groups. Diagnostics of these linear models were made with the DHARMa package, see Figure S4.

2.9 | GO analysis

The genomation package (Akalin et al., 2015) and annotations available from the *Salvelinus* sp. IW2-2015 (assembly ASM291031v2) were used to annotate CpGs of interest and identify the genes located less than 10 kb away. These CpGs of interest were either the ones lying outside two standard deviations of the loading distribution in the PCA, or the ones showing a *p*-value <.05 in the glm analysis (Section 2.5.1).

Then, the GO term annotation of all proteins in the same assembly was downloaded from NCBI. This file was reformatted to match the required input format for the topGo R package (Alexa & Rahnenfuhrer, 2022) using custom scripts. A topGOdata object was created using this correctly formatted file, the list of genes of interest, and the following parameters: nodeSize = 5, Ontology = "BP" and annot = annFUN.gene2GO. Enrichment analysis was performed on this topGOdata object using the runTest function with the algorithm = "weight" parameter (designed to take into account GO hierarchy) and Fisher's exact test. Categories having a *p*-value <.01 with Fisher's test were considered significantly enriched.

2.10 | Repeat enrichment analysis

Repeat enrichment analysis was done using the online tool RepeatMasker (Smit et al., 2013) with rmblast as a search engine and the database Dfam3.0 (Hubley et al., 2016). DNA sequences in FASTA format were tested against the Dfam3.0 Danio rerio transposable elements database.

2.11 | Sex identification

PCR for a 177 base-long fragment of the sex determining gene *SdY* was performed on all RRBS samples at 100, 150 and 200 τs. This gene is present in males and absent in females. Mitochondrial DNA was also amplified as a positive control. Products were run on a 2% agarose gel at 100 mA for 40 min to check for the presence or absence of the *Sdy* band, allowing for gender identification of 100, 150 and 200 τs samples. We did not analyse the 50 τs samples (used for RRBS) as they were pools of two individuals. Primer sequences are available in Table S3.

3 | RESULTS

We studied methylation in embryos from four sympatric morphs (SB, LB, PL and PI) at four different timepoints (50, 100, 150 and 200 τs). RRBS yielded 971,126 unique CpG sites with at least 10× sequencing coverage in one of the 48 samples.

3.1 | Individual CpG sites show methylation differences between developmental timepoints and between morphs

As variation in methylation at single CpG sites can influence biological functions [i.e., transcription (Claus et al., 2012; Kitazawa & Kitazawa, 2007)], we first analysed individual CpGs. To identify the ones showing methylation differences during development and between the four morphs, we applied stringent filtering to the data, excluding sites with less than 10× coverage in any of the 48 samples, leaving 10,340 sites for analysis. The majority of these CpGs were heavily methylated (on average 80.3%) and showed a decrease in methylation over developmental time (Figure 1). Comparison of the youngest and oldest embryos showed a notable decrease, as 85.4% (8824/10,340) of the cytosines had less methylation at 200 τs than at 50 τs, while only 1503 sites showed an increase. However, the average decrease was rather small (3.6%) (Figure S5). While the methylation percentage distributions significantly shifted with time for each morph (Figure 1b), they were also significantly different between morphs, especially at 100 τs (Kruskal–Wallis *p*-value $<2.2^{-16}$, Table S4).

To better visualize these morph differences, we studied methylation changes at each interval separately (50–100 τs, 100–150 τs and 150–200 τs) looking at all CpGs, with an extra focus on CpGs showing more than 5% methylation differences (Figure 2). Most observed changes were between the earliest two timepoints, 50 and 100 τs (χ^2 *p*-value $<2.2^{-16}$). The number of sites showing more than 5% change in methylation between timepoints varied between morphs (χ^2 *p*-value $= 3.1^{-16}$). This was more pronounced in the limnetic morphs (PL and PI) than in benthic ones (LB and SB) with ~2300 CpGs showing decreased methylation between 50 and 100 τs in PI and PL morphs, while ~1400 showed a decrease in methylation in SB and LB morphs (χ^2 *p*-value $<2.2^{-16}$). A difference in the number of cytosines with increased methylation was also observed between the limnetic and benthic morphs. This was most notable between timepoints 100 and 150 τs when ~375 CpGs showed this increase for LB and SB morphs while ~875 were counted for PL and PI morphs (χ^2 *p*-value $<2.2^{-16}$). The results suggest differences in global methylation changes between benthic and limnetic morphs, mainly at early timepoints (50–150 τs).

Further validating these first results, PC decomposition and ANOVAs on them revealed significant differences in DNA methylation between morphs and developmental timepoints (Table 1). The first 23 PCs accounted for 60% of the variance within the data (Table S5). PC1 (8% of variance) separated the samples between timepoints while PC4 (2.9%) separated the benthic morphs (SB and LB) from the limnetic ones (PL and PI) (Figure 3). Furthermore, PC3 (3.2%) revealed time and morph separation, while PC2 (3.5%) was driven by technical effects (library differences). Analyses of variance revealed that some additional PCs were affected by either morph or developmental time, albeit more weakly (Table S6). No PC was

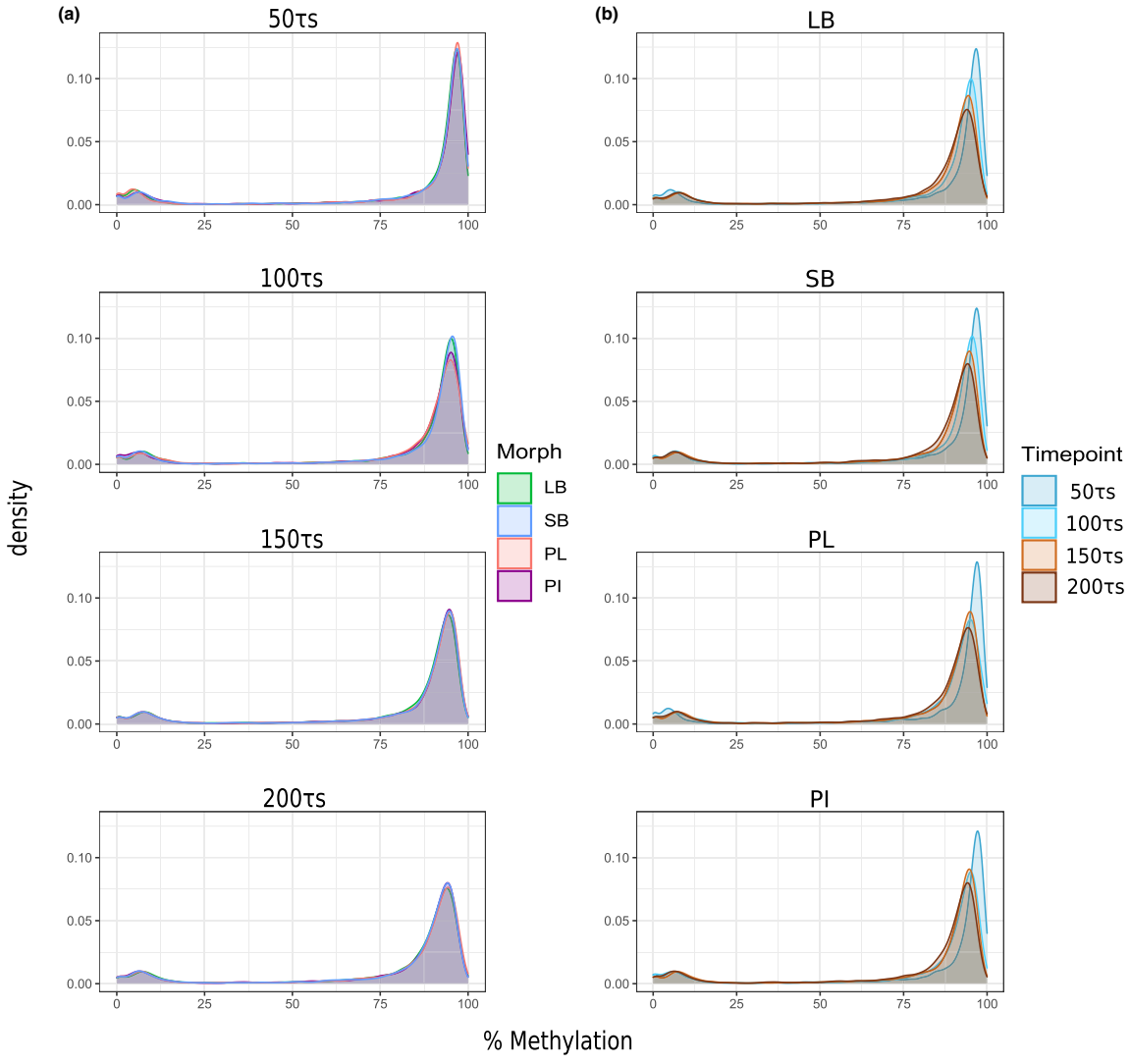


FIGURE 1 Distribution of methylation in 10,340 CpGs, over morphs and timepoints. The % methylation was calculated by dividing the number of cytosines sequenced at a given position by the total number of bases sequenced at this position (cytosines + thymines). (a) Difference in methylation distribution between morphs for each timepoint. (b) Difference in methylation distribution between timepoints for each morph. Significant differences were seen both between timepoints and between morphs (Table S4).

significantly influenced by sex. The contribution of each CpG to each PC allowed for identification of the sites driving this timepoint or morph variation, and the CpGs with the highest loading for each PC were kept for further analysis.

As a complementary approach, a generalized linear model was employed to identify cytosines most strongly influenced by timepoint, morph, and sex differences. Both analyses gave corresponding results for CpGs separating timepoints and morphs by their methylation (Figure S2), although three CpGs showed a difference in methylation between sexes with the glm. However, two of these cytosines were located on unplaced genomic scaffolds and these potentially sex-based methylation differences were not investigated further.

3.2 | Genes close to methylation sites separating timepoints and morphs are overrepresented in specific GO categories

To study the biological correlates of differential methylation, genes within 10 kb distance from these CpGs of interest were identified, and GO analysis for biological processes was performed. Many

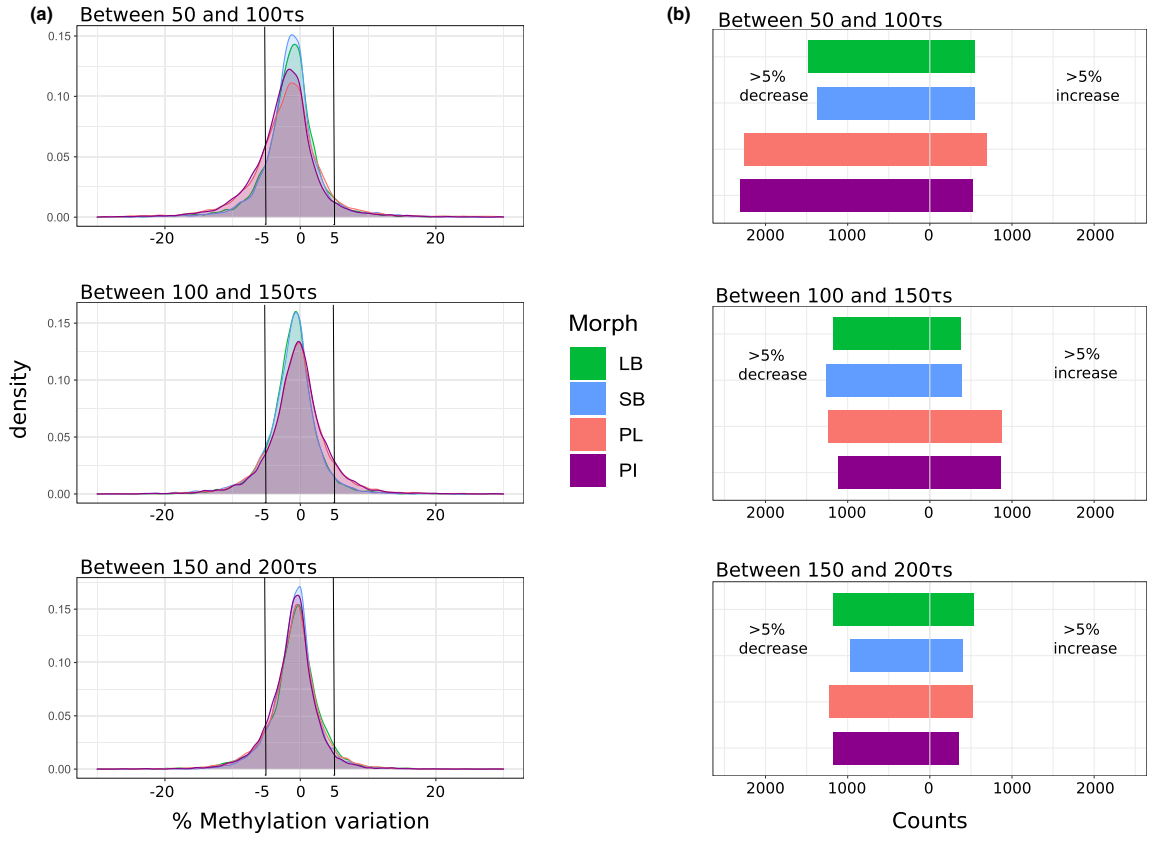


FIGURE 2 Variation in methylation of 10,340 CpGs between developmental timepoints, for each morph. (a) % methylation variation between each consecutive developmental timepoint. (b) Number of CpGs for which methylation varies by more than 5% (as an increase or a decrease) between consecutive developmental timepoints.

TABLE 1 Results of ANOVA on principal components derive from 10,340 CpGs in Arctic charr

PC	Morph	Time	Morph × time
PC1	1	1.06E-28	1
PC2	1	1.22E-07	1
PC3	0.052	0.003	0.369
PC4	9.80E-09	0.982	1

Note: *p*-values were corrected for multiple testing ($n = 47$) with Bonferroni.

Abbreviation: ANOVA, analysis of variance.

GO categories showed a significant enrichment with Fisher's test (Table 2; Table S7). For instance, genes close to methylation sites separating developmental timepoints were found to be significantly associated with nucleosome assembly (GO:0006334, $p = 1.1e\text{-}4$), IMP biosynthesis (GO:0006189, $p = 4.1e\text{-}4$), *TGF-β* regulation (GO:0030511, $p = 7.9e\text{-}3$) and pigment biosynthesis (GO:0046148, $p = 1.9e\text{-}3$) which is congruent with the emergence of pigmentation observed at 150 τs (Guðbrandsson et al., 2019). Genes in the vicinity of methylation differences between morphs were most significantly related to the nucleosome assembly category (GO:0006334, $p = 4.20e\text{-}6$), followed by defence response to bacterium (GO:0042742, $p = 8.20e\text{-}5$), the establishment of meiotic spindle localization (GO:0051295, $p = 2.2e\text{-}4$) and calcium dependent cell matrix adhesion (GO:0016340, $p = 4.3e\text{-}4$). Enrichment of negative regulation of osteoclast differentiation (GO:0045671, $p = 5.5e\text{-}3$) should also be mentioned because of the known craniofacial differences between morphs.

3.3 | Differentially methylated regions are distributed over the genome and are enriched in functional RNA sequences

To have a better representation of regions sequenced by RRBS, we probed the genomic distribution of these 10,340 CpG sites. Most of them were located in intergenic sequences (87.8%), 7.9% were in promoter regions, 3.7% in introns and 0.6% in exons.

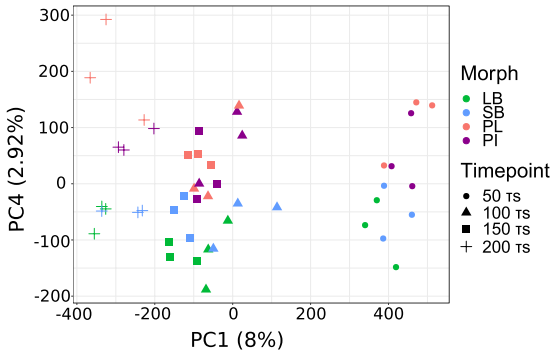


FIGURE 3 Developmental DNA methylation differs between benthic and limnetic morphs. PCA on 48 Arctic charr methylomes separates sympatric morphs (four in total: Two benthic: LB and SB and two limnetic: PL and PI) and the developmental timepoints (four timepoints, 50 to 200 τs). The first PC separates the samples based on their developmental stage while PC4 separates benthic from limnetic samples. The brackets indicate the proportion of variance explained. LB, large benthic; PCA, principal component analysis; PI, piscivorous; PL, planktivorous; SB, small benthic.

To study differential methylation in CpG islands and focus on important regions such as transcription start sites, we next studied regions containing multiple CpG sites as is commonly done (Irizarry et al., 2009). We focused on 1000 bp-long regions with at least three CpG sites (with >10× read coverage). These 18,469 regions were interrogated for differential methylation by morph, developmental time or sex. Similar to the single CpG results, more regions displayed methylation differences between timepoints (339) than between morphs (91). Interestingly, 49 DMRs were present in both categories, that is, showed a difference in methylation between morphs and between timepoints. The 478 DMRs (Table S8) were distributed throughout the whole genome, as mapping to a *Salvelinus* sp. genome revealed several DMRs on every linkage group (Table S9). Because of the RRBS methodology, the majority of sequenced regions were located close to promoters, and this trend was also observed for DMRs (Figure S6). However, it should be noted that multiple DMRs (between timepoints and between morphs) were located in intergenic regions further away from gene bodies. We also investigated whether DMRs were enriched or depleted in repeat element sequences. Generally, regions differentially methylated between morphs and timepoints contained significantly fewer repeat elements than non-DMRs (Figure 4a), with the most important difference observed for retroelements and DNA transposon sequences (Figure 4b). This was expected, as retroelements are known to be silenced by methylation (Karimi et al., 2011) and would therefore be expected to be highly methylated independently of morph or timepoint. Interestingly, DMRs were enriched in functional noncoding RNAs (Figure 4c). Although this category is admittedly small, many regions differentially methylated (both between timepoints and between morphs) are located close to functional RNA genes such as 5.8S rRNA, U1, U2 and U5 spliceosomal RNA, and multiple tRNAs (Figure S7).

3.4 | Further investigation of 14 differentially methylated regions

To better characterize the patterns of methylation at DMRs and to test for the relationship between differential methylation and gene expression, 14 DMRs were chosen for in depth analysis, based on the proximity of the DMRs to TSSs, methylation differences observed in pairwise comparisons, and the potential biological relevance of nearby genes (Table 3).

3.4.1 | Genes of interest show methylation and expression differences

The 14 genes associated with these regions represent examples of the differential methylation patterns observed in the data: changes in methylation over development, between morphs or both (Figure 5). Average methylation in each region, per morph and timepoint, was calculated to reveal trends and enable tests for correlation of methylation and expression (methylation maps of the 14 DMRs with individual CpGs are available in Figure S8). Some regions showed a substantial shift in methylation over time, including the regions close to *MEGF9* (ANOVA p-value = 5.5^{-16}), *H2A-like* (p-value = 3.6^{-21}), *H3-like* (p-value = 2.6^{-16}), *ARHGEF37-like* (p-value = 5.0^{-10}) and *MPP3* (p-value = 3.9^{-7}), with a decrease in methylation over time being most prominent. This is in accordance with the tendency towards reduced methylation over time observed for single CpG sites (above). For other regions, methylation remained stable over developmental timepoints but showed a considerable difference in methylation between morphs such as the regions close to *SLC9A3R2-like* (p-value = 7.4^{-5}), *NFIX* (p-value = 7.0^{-6}) and *ARL16* (p-value = 2.2^{-3}). The region close to *RASSF4-like* showed methylation differences both between morphs (p-value = 2.6^{-6}) and over development (p-value = 5.4^{-6}). The patterns of differential methylation by morph showed examples of clear benthic versus limnetic separation in regions close to *SLC9A3R2-like* and *RASSF4-like*. ANOVA and post hoc test results are shown in Table S10.

To study whether gene expression varied at these 14 loci, RNA was isolated from three embryos for each morph and timepoint (100, 150 and 200 τs) and analysed by qPCR (Figure 6). All genes except *GLI3-like* and *ARMC1* showed significant differences in expression between timepoints and *ARHGEF37-like*, *ARL16*, *ARMC1*, *GLI3-like*, *H2A-like*, *H3-like*, *MEGF9*, *NFIX* and *SLC9A3R2-like* showed significant differences in expression between morphs (ANOVA, see Table S11).

Next, we set out to test whether there was a correlation between methylation and expression. The analysis was performed on averages per morph and timepoint (whole embryos), because obtaining paired DNA and RNA samples from the same individuals was not technically feasible (Figure 7; Table S12). Ten genes showed no linear relationship between methylation and expression. Negative correlation with methylation levels was observed for *MPP3* and *MEGF9* expression, and *ARHGEF37-like* and *H3-like* expression was positively correlated with methylation. Similar to what has been observed in previous studies (Anastasiadi et al., 2018; Arechederra et al., 2018;

TABLE 2 Enrichment of gene ontology (GO) categories of genes located close to CpGs separating samples by their methylation patterns, whether in PCA or glm analysis

GO.ID	Term	Annotated	Observed	Expected	Weighted Fisher
Genes near CpGs separating timepoints (PC1)					
GO:0006334	Nucleosome assembly	152	5	0.46	1.10E-04
GO:0006189	De novo IMP biosynthetic process	10	2	0.03	4.10E-04
GO:0016559	Peroxisome fission	26	2	0.08	2.87E-03
GO:0045638	Negative regulation of myeloid cell differentiation	325	5	0.99	3.27E-03
GO:0007567	Parturition	98	3	0.3	3.42E-03
GO:0009113	Purine nucleobase biosynthetic process	29	2	0.09	3.56E-03
GO:0035584	Calcium-mediated signalling using intracellular calcium source	110	3	0.34	4.72E-03
GO:0090197	Positive regulation of chemokine secretion	34	2	0.1	4.87E-03
GO:0050847	Progesterone receptor signalling pathway	35	2	0.11	5.16E-03
GO:0030511	Positive regulation of transforming growth factor beta receptor signalling pathway	133	3	0.41	7.97E-03
Genes near CpGs separating timepoints (glm)					
GO:0006334	Nucleosome assembly	152	5	0.47	1.10E-04
GO:0042742	Defence response to bacterium	640	9	1.98	1.70E-04
GO:0006189	De novo IMP biosynthetic process	10	2	0.03	4.20E-04
GO:1900275	Negative regulation of phospholipase C activity	14	2	0.04	8.40E-04
GO:0046148	Pigment biosynthetic process	171	4	0.53	1.99E-03
GO:0015824	Proline transport	23	2	0.07	2.29E-03
GO:0051599	Response to hydrostatic pressure	28	2	0.09	3.39E-03
GO:0032328	Alanine transport	29	2	0.09	3.63E-03
GO:0051262	Protein tetramerization	491	6	1.52	4.24E-03
GO:0033209	Tumour necrosis factor-mediated signalling pathway	356	5	1.1	5.02E-03
Genes near CpGs separating morphs (PC4)					
GO:0006334	Nucleosome assembly	152	5	0.53	2.00E-04
GO:0016340	Calcium-dependent cell-matrix adhesion	9	2	0.03	4.30E-04
GO:0046548	Retinal rod cell development	51	3	0.18	7.70E-04
GO:0021963	Spinothalamic tract morphogenesis	12	2	0.04	7.80E-04
GO:0032780	Negative regulation of ATPase activity	53	3	0.19	8.60E-04
GO:0106028	Neuron projection retraction	14	2	0.05	1.07E-03
GO:0099557	Trans-synaptic signalling by trans-synaptic complex, modulating synaptic transmission	14	2	0.05	1.07E-03
GO:0055069	Zinc ion homeostasis	59	3	0.21	1.18E-03
GO:1904782	Negative regulation of NMDA glutamate receptor activity	16	2	0.06	1.41E-03
GO:0007216	G protein-coupled glutamate receptor signalling pathway	63	3	0.22	1.42E-03
Genes near CpGs separating morphs (glm)					
GO:0006334	Nucleosome assembly	152	4	0.11	4.20E-06
GO:0042742	Defence response to bacterium	640	5	0.46	8.20E-05
GO:0051295	Establishment of meiotic spindle localization	30	2	0.02	2.20E-04

TABLE 2 (Continued)

GO.ID	Term	Annotated	Observed	Expected	Weighted Fisher
GO:1901223	Negative regulation of NIK/NF-kappaB signalling	69	2	0.05	1.15E-03
GO:0030071	Regulation of mitotic metaphase/anaphase transition	118	2	0.09	3.30E-03
GO:0008205	Ecdysone metabolic process	5	1	0	3.62E-03
GO:0051673	Membrane disruption in other organism	5	1	0	3.62E-03
GO:0000964	Mitochondrial RNA 5′-end processing	5	1	0	3.62E-03
GO:0051663	Oocyte nucleus localization involved in oocyte dorsal/ventral axis specification	6	1	0	4.34E-03
GO:1902956	Regulation of mitochondrial electron transport, NADH to ubiquinone	6	1	0	4.34E-03

Note: See Table 1 for PCs. Only the first 10 categories are displayed. For a full list see Table S7.

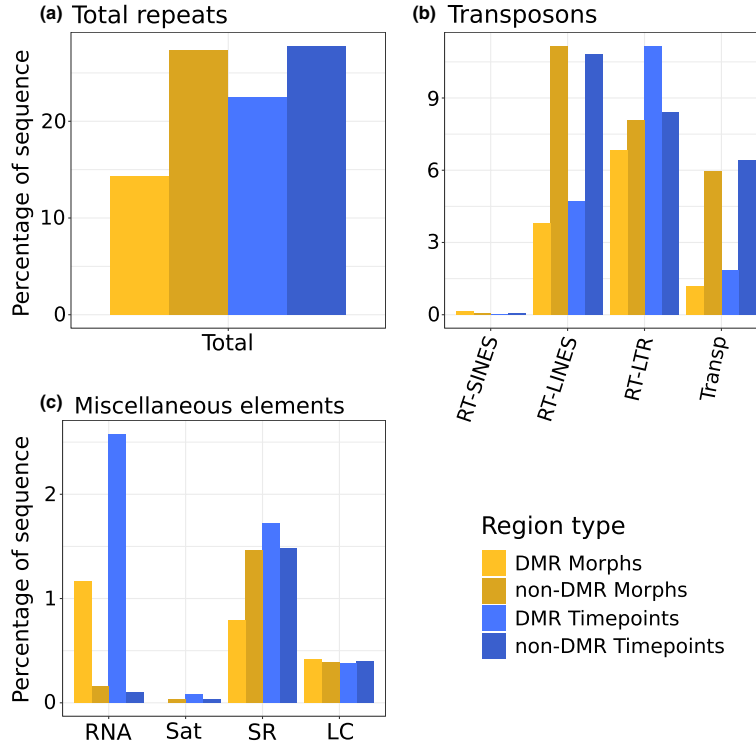


FIGURE 4 Differentially methylated regions are depleted of transposon sequences and enriched in functional RNAs. The total sequence percentage of the indicated DNA elements within all 1000 bp regions of each category (regions differentially methylated and nondifferentially methylated between either morph or time) is displayed. (a) All repeat elements. (b) Transposable elements. (c) Other repeat elements such as functional RNAs (RNA), satellites (sat), simple repeats (SR) and low complexity (LC) elements.

Esteller, 2002), the direction of correlation seems to reflect the position of DMRs with relation to gene features: the DMRs were located on the promoter regions of *MMP3* and *MEGF9*, but overlapped with exon sequences for *ARHGEF37-like* and *H3-like* (Table 3).

3.4.2 | Histone gene clusters showed differential methylation by morph and time

As the "nucleosome assembly" GO category (GO:0006334, Table 2) was most significantly enriched in genes close to CpGs differentially methylated between both morphs and timepoints, we investigated it further. All genes identified as part of this GO category were histone genes and both the PCA and glm on residues identified differentially methylated CpGs in a histone gene cluster (as did the DMR analyses). Histone genes encode the building blocks of nucleosomes and are an essential component of chromatin (Kornberg, 1974). In vertebrates, five histone genes are arranged in clusters present in multiple copies, with cluster genomic arrangements and the number of clusters varying between species. The precise genomic locations or numbers of those histone clusters are unknown in Arctic charr, and generally understudied in fish (Doenecke et al., 1997; Rooney et al., 2002). The

TABLE 3 Genes chosen for in depth analysis

Gene name	Abbreviation	Dist to DMR (bp)[a]	Locus-ID[b]	Feature overlap with DMR[c]
Lemur tyrosine kinase 2	LMTK2	0	lmtk2	Promoter—exon
Armadillo repeat-containing protein 1	ARMC1	5571	LOC111953658	Intergenic
Na(+)/H(+) exchange regulatory cofactor NHE-RF2-like	SLC9A3R2-like	2301	LOC112078129	Intergenic
Homeobox protein Nkx-2.3-like	NKX23-like	454	LOC112075116	Intergenic
Nuclear factor I X	NFIX	3009	nfix	Intergenic
Ras association domain-containing protein 4-like	RASSF4-like	175	LOC112080902	Promoter
ADP-ribosylation factor-like protein 16	ARL16	0	LOC112079300	Promoter
Homeobox protein Meis1-like	MEIS1-like	243	LOC112075762	Promoter
Histone H3-like	H3-like	0	LOC111971647	Promoter—exon
Histone H2A-like	H2A-like	0	LOC111971648	Promoter—exon
Transcriptional activator GLI3-like	GLI3-like	382	LOC111953336	Intergenic
Multiple epidermal growth factor-like domains protein 9	MEGF9	0	LOC112080620	Promoter
MAGUK p55 subfamily member 3	MPP3	0	LOC111981890	Promoter
Rho guanine nucleotide exchange factor 37-like	ARHGEF37-like	1989	LOC111965367	Exon

Abbreviations: DMR, differentially methylated region; TSS, transcription start site.

[a]Dist to DMR: distance of gene to the closest DMR. 0 means the DMR overlaps with the gene's TSS.

[b]Locus-ID: ID of gene in *Salvelinus* sp. IW2-2015 (assembly ASM291031v2).

[c]Feature overlap with DMR: indicates whether the DMR overlaps with specific features on the respective gene of interest (exon, intron or promoter), or lies outside the gene body (intergenic).

four core histone genes (*H2A*, *H2B*, *H3* and *H4*) and the linker histone *H1* are arranged in a cluster in the *Salvelinus* sp. assembly. Data from this genome suggests clustering of those five histone genes, but it does not confirm that every copy of the cluster is composed of all five genes. Because of their multicopy status, these histone gene clusters had a high combined RRBS coverage (286× on average) leading to a larger number of CpGs with sufficient coverage allowing for robust analysis. Notably, the PL morph had lower sequencing coverage over *H4* than other morphs (Figure S9) and therefore we decided not to analyse the methylation patterns in this gene. It is tempting to speculate that this reflects a lower copy number in the PL morph although other explanation such as technical issues cannot be ruled out.

ANOVA revealed that the other four genes in the cluster were differentially methylated between timepoints (*H1* p-value = 2.83^{-6}, *H2A-like* p-value = 3.63^{-21}, *H2B* p-value = 5.51^{-4}, *H3-like* p-value = 2.6^{-16}), and two of them were differentially methylated between morphs (*H1* p-value = 2.87^{-4}, *H2B* p-value = .027). These morph differences were driven by a benthic-limnetic separation for *H1*, and by the PI morph for *H2B* (Figure 8a; Figure S8).

The high sequencing coverage in these regions also allowed us to determine intra-DMR patterns, that is, the distribution of methylation sites over gene features (promoter vs. gene body). Methylation patterns were compared between gene features for *H1* and *H2B*, and ANOVA revealed a significant difference between the methylation levels in the promoter of *H1* compared to its gene body (p-value = 3.7^{-11}) (Figure 8b).

These results show clear methylation differences between morphs and over development in the histone clusters, and provide an example of the intricacy of methylation patterns at the individual gene level.

4 | DISCUSSION

In this study we set out to investigate the developmental methylome of phenotypically different Arctic charr ecomorphs. Previous studies have addressed the differences of these morphs in adult morphology, niche utilization as well as genetic and gene expression differences (Guðbrandsson et al., 2017, 2019; Skúlason, Noakes, & Snorrason, 1989; Skúlason, Snorrason, et al., 1989), but to date no study has addressed epigenetic differences between the morphs. We applied RRBS, a reduced representation method that enables the targeting of CpG-rich regions and the screening of a targeted part of the genome for methylation marks. This method is particularly useful to obtain cost-efficient methylation sequencing in species with a genome of large size such as Salmonids (which have undergone the teleost-specific whole-genome duplication [WGD] event ~350 million years ago (Meyer & Van de Peer, 2005; Hoegg

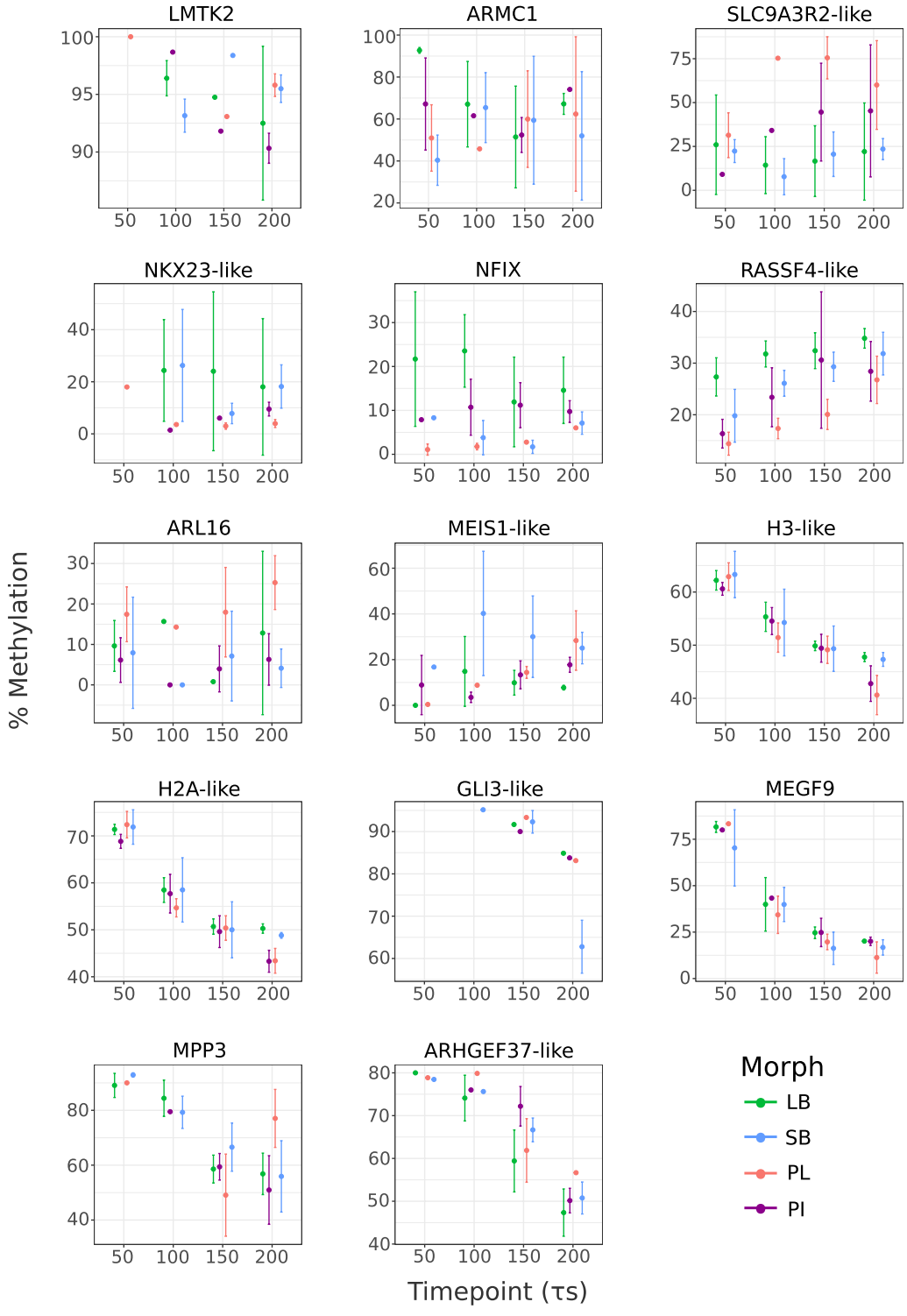


FIGURE 5 Average methylation of 14 DMRs of interest over developmental timepoints in each of the four morphs. The coverage varied by morph and timepoints. In some cases estimates are missing because of low coverage. Standard deviation is indicated when applicable. For significance of terms from ANOVA, see Table S10. ANOVA, analysis of variance; DMR, differentially methylated region.

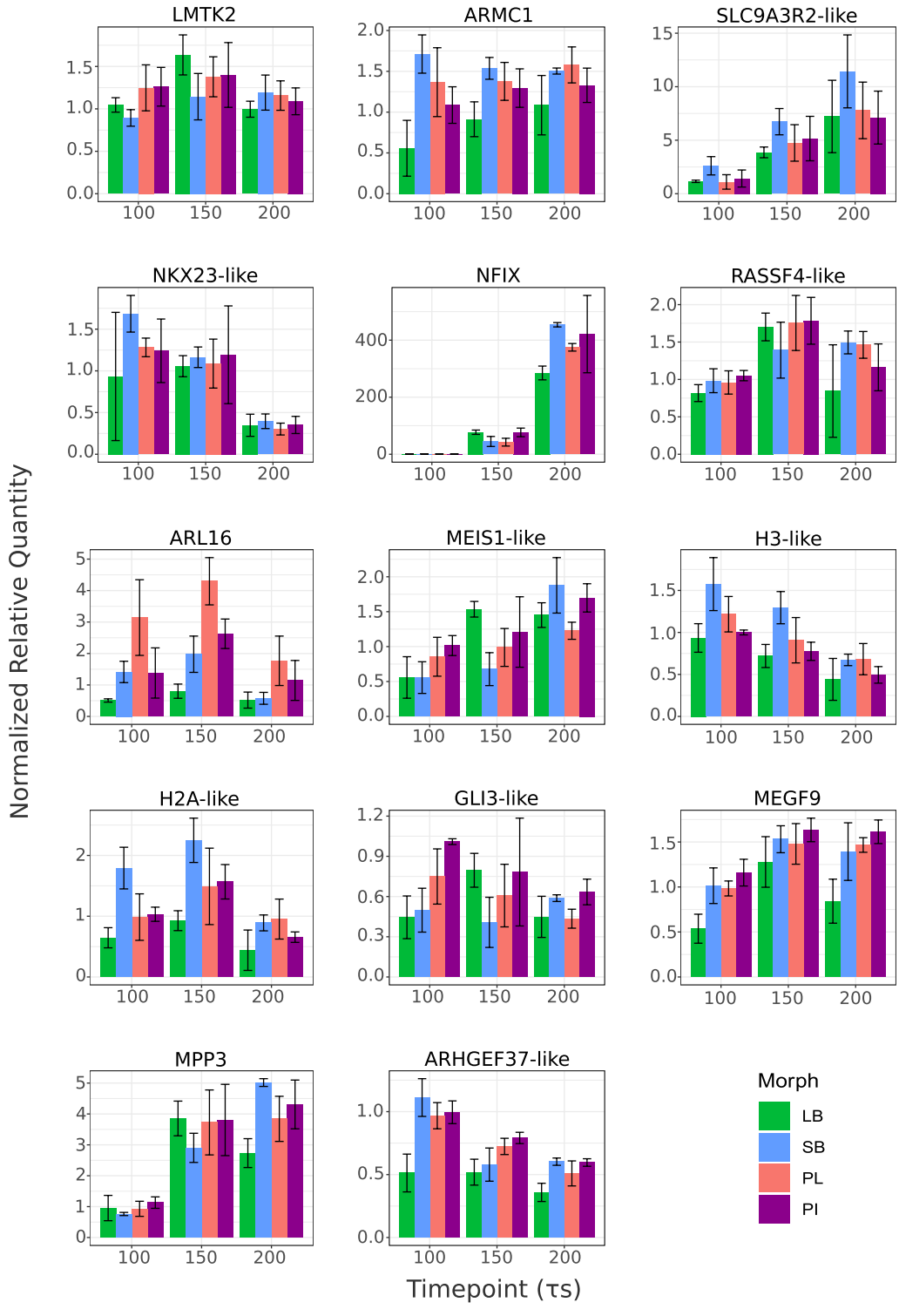


FIGURE 6 Genes of interest displaying differences in methylation show differences in expression. Gene expression based on timepoint and morph. Normalized relative quantities were calculated using sample S80 as a control. For significance of terms from ANOVA, see Table S11. ANOVA, analysis of variance.

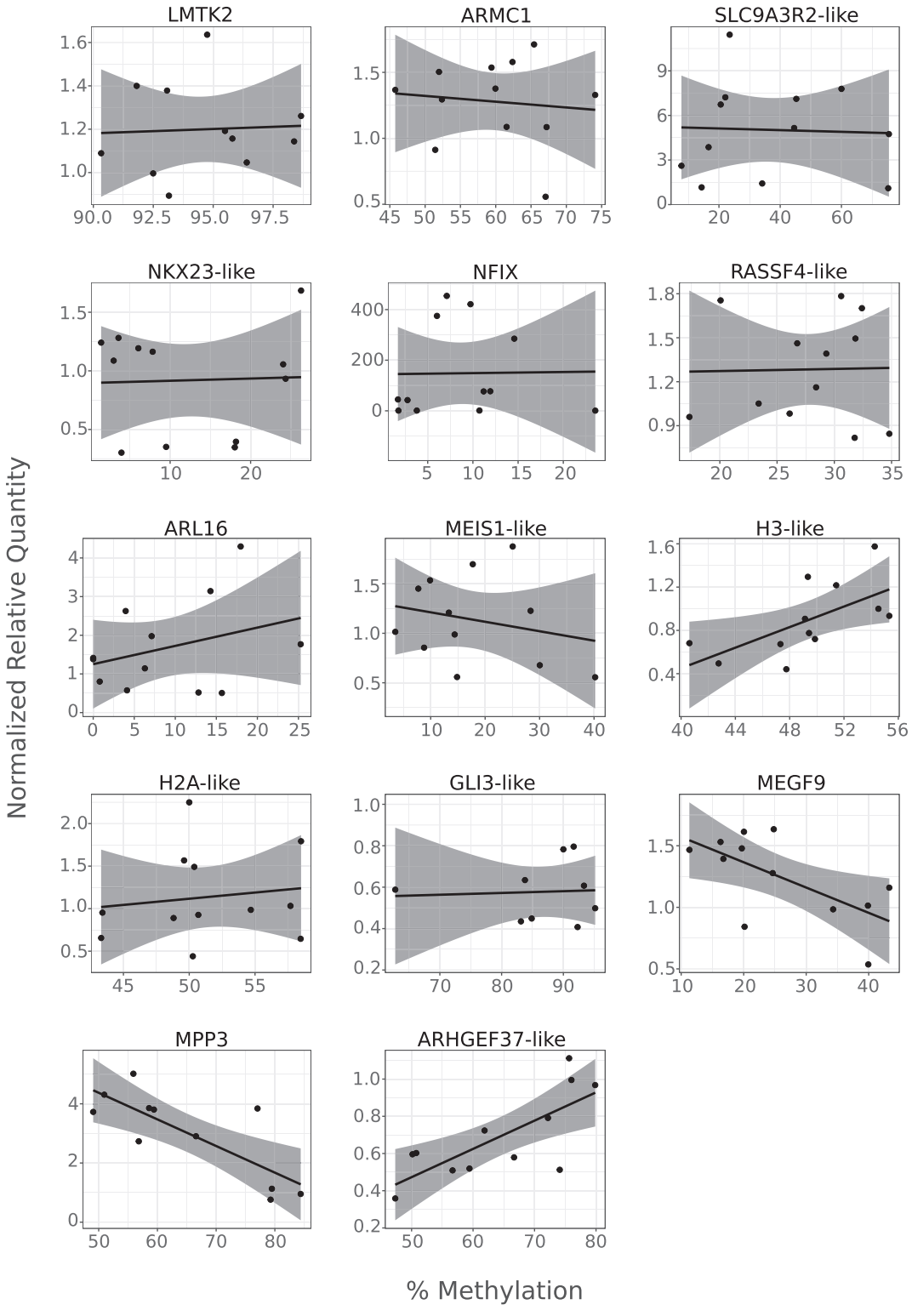


FIGURE 7 Legend on next page

FIGURE 7 Tests for correlation of the level of expression and methylation of nearby regions. Graphs showing the correlation (Pearson coefficient) between the average methylation and expression of each gene. For GLI3-like and NFIX for which the data was not parametric, the Kendall correlation and relevant test was used instead. The grey area represents confidence intervals (95%). ARHGEF37-like (p-value = .005, R = .75), MPP3 (p-value = .004, R = −.78), H3-like (p-value = .028, R = 0.63) and MEGF9 (p-value = .024, R = −.64) showed a correlation between both variables. For details see Table S12.

et al., 2004) as well as a more recent WGD ~80 million years ago (Macqueen & Johnston, 2014)).

4.1 | Reduction in methylation over time in all morphs

Our analyses revealed clear differences in the DNA methylome, mainly between developmental timepoints and to a lesser but notable extent between morphs, both at individual CpG sites and in the average methylation of larger regions containing multiple CpGs. It should be emphasized that we investigated methylome differences in whole embryos, which inevitably obscures subtle or tissue-specific patterns. Therefore, the results probably reflect large methylation differences (over development or between morphs) in specific cell types or tissues or, alternatively, more systemic changes in many tissues and cell types. Predictably, methylation changes are more pronounced during development than methylation differences between morphs, matching the results of transcriptomic studies in this model, showcasing a higher number of genes differentially expressed over time than between morphs in a similar setup (Beck et al., 2019; Gudbrandsson et al., 2016, 2017). More specifically, we observed broad reduction in methylation at single CpG sites between 50 and 200 τs (Figure 1), suggesting that overall, methylation in CpG islands (with respect to the reduced representation of the method) decreases over the course of development, from gastrulation through late organogenesis. Genome-wide gametic demethylation followed by a global remethylation shortly after fertilization has been observed in zebrafish (Andersen Ingrid et al., 2012; Jiang et al., 2013), but these studies focused on very early developmental stages, from fertilization to the mid-blastula transition, where zygotic genome activation occurs. The charr embryos used in our study are past this transition stage, and direct comparisons between our results and those zebrafish studies are therefore not meaningful. However, investigating global remethylation at very early stages in this system would be of great interest, as would mapping methylation and methylation differences in individual organ systems into adulthood.

4.2 | DNA methylation differences during development separate Arctic charr morphs

We were particularly keen to test for differences in methylation between morphs. Interestingly, they separate mainly along a benthic-limnetic axis: Principal component decomposition on single CpGs reveals clear separation on this axis in PC4 (Figure 3) while the other components separating morphs do so to a lesser extent. The generalized linear models also identified individual cytosines with differences in methylation between morphs, and corresponded with the PCA-based results (Figure S2). This is not entirely surprising given that population genomic data show that the limnetic morphs (PL and PI-charr) have a similar genetic constitution, while the two benthic morphs are distinct, but clearly more related to one another than the other two (Guðbrandsson et al., 2019). This is further corroborated by developmental transcriptome profiles, which show that PL, SB and LB-charr all differ at the gene expression level, with the PL morph being the most different from the other two. At the DMR level, niche-specific differences (in DMRs close to *RASSF4-like* and *SLC9A3R2-like*) and outlier morphs (individual morphs showing consistently higher methylation in regions close to *ARL16*, *NFIX* and *MEIS1-like*) were seen.

This study was done on embryos reared in the same experimental conditions, from multiple parents per morph caught in the wild. Because of this, and the fact that we established pure families and only one generation, we could not test for parental effects or disentangle the influence of genetic and environmental components, as was recently done by Venney et al. (2022). The observed methylation differences could thus be the result of multiple factors. First, genetic differences between the morphs could act as either *cis*- (genetic polymorphism influencing methylation of nearby CpGs) or *trans*-factors (for instance, through polymorphism in the machinery that alters methylation in specific locations, tissues, or developmental window). Second, parental effects (maternal or paternal, genetic or environmental) have been shown to be a source of methylation differences. Multiple epigenetic mechanisms acting across generations have been described, such as epi-alleles, maternal provisioning and metabolism, or sncRNAs in sperm [mainly tsRNAs (Peng et al., 2012) and miRNA (see Perez & Lehner, 2019)], all of which could possibly impact DNA methylation profiles of the offspring. In fish, specific cytosine methylation has been associated with maternal effects in Chinook salmon (Venney et al., 2020), parental maturation temperature corresponded with DNA methylation patterns in the offspring of Brook charr (Venney et al., 2022) and egg size has already been correlated with craniofacial shape at first feeding in this *S. alpinus* model (Beck et al., 2020). It would be very interesting to study parental effects and transgenerational epigenetic inheritance (see Baerwald et al., 2016; Heckwolf et al., 2020), with a multigeneration scheme and crosses (for instance, pure and reciprocal crosses of different morphs), for further inquiry into these potential causes. Yet, the 3–4 year generation time of Arctic charr makes these studies less than practical. An alternative approach would be to test whether differences in methylation between morphs associate with genomic peaks of genetic differentiation. This was, for instance,

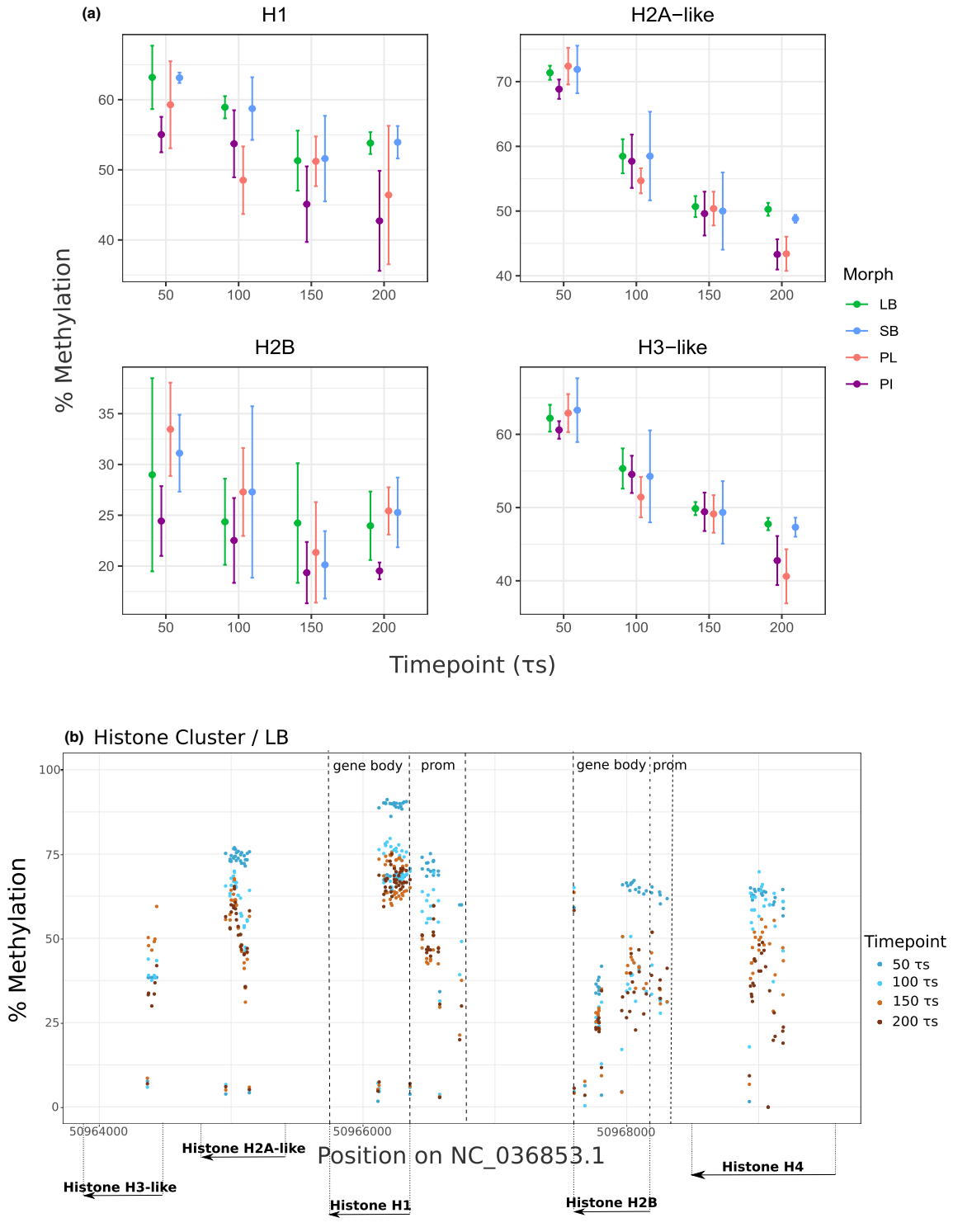


FIGURE 8 A cluster of histone genes is differentially methylated between morphs and developmental timepoints. (a) Average methylation of four genes in the histone clusters over developmental time in each of the four morphs. (b) Methylation map of the histone cluster for the LB morph at each developmental timepoint, with annotation of the five histone genes, as well the gene body and promoter regions of H1 and H2B. Methylation maps for this region for each morph can be found in Figure S8. LB, large benthic.

done in human monocytes which lead to the finding that DMRs in these cells had a methylation level that correlated with the genotype of nearby SNPs, suggesting a role for *cis*-acting genetic variation on the establishment of methylation marks (Schröder et al., 2017). Conversely, data of this sort could identify polymorphism in transcriptional or epigenetic regulators away from DMRs (*trans*-factors) that might be the drivers of the observed methylation differences between charr morphs. While our study shows clear DNA methylation differences between these morphs reared in common garden, it is quite likely that the methylation profiles of individuals developing in the wild will be different, and maybe reflect the response to their physical environment or microhabitat. LB charr, for instance, spawns in August, around very cold springs in the lake, while PL-charr spawns in October in shallows with less inflow of cold water (Skúlason, Snorrason, et al., 1989).

4.3 | DMRs are located throughout the whole genome and are enriched in RNA sequences

As DNA methylation is not distributed randomly in genomes (Bird & Wolffe, 1999), we set out to characterize the distribution of differentially methylated regions in Arctic charr development. Notably, DMRs appeared on every linkage group and were often seen close to promoters or gene bodies as these regions are being selected for during RRBS (Table S9; Figure S6). Repeat sequences were found to be underrepresented in DMRs. This was expected, as transposons are known to be actively silenced by methylation (Karimi et al., 2011) and would therefore be expected to be highly methylated independently of morph or timepoint. However, while underrepresented, some morph DMRs were in transposons. This can relate to morph differences in two ways. First, TE-derepression was shown to be a post-zygotic barrier in hybrids between morphs of another salmonid *Coregonus clupeaformis* (Laporte et al., 2019), and might also influence viability of hybrids between charr morphs. Secondly, differential methylation of a TE near the *mef2ca* gene in zebrafish influenced gene expression and bone development in the head (Nichols et al., 2016).

More interestingly, DMRs between both morph and timepoint showed an enrichment in functional RNA loci, particularly tRNAs and 5.8S rRNA, both of which can be dynamically expressed by tissues or during development (Figure 4; Figure S7). Maternal rRNAs such as 5.8S are, for instance, replaced by somatic equivalents during embryogenesis in zebrafish (Locati et al., 2017). The slight reduction in methylation of 5.8S rRNA loci observed between 50 and 200 τs in the charr could possibly reflect the activation of somatic ribosomal RNA to replace the maternal ones. Like rRNAs, tRNAs are dynamically expressed and changes in tRNA levels have been shown to influence cell and tissue growth (Grewal, 2015). Changes in methylation of the promoters of tRNA genes could potentially influence the pool of available tRNAs, and thus affect mRNA translation via codon bias (Hanson & Coller, 2018). Investigating the impact of methylation on those features could be of interest to further shed light on how methylation differences can contribute to phenotypic differences in this and other systems, and on the effect of methylation in genomic regions encoding functional RNA sequences on developmental pathways and evolutionary divergence.

4.4 | Differentially methylated genes show differences in expression

The most commonly regarded role of methylation is to influence the expression of nearby genes. The general notion is that increased methylation of a promoter will reduce transcription of a neighbouring unit (Esteller, 2002; Merlo et al., 1995). To test this, we studied the expression of 14 genes in the four morphs with qPCR. The data revealed differences in gene expression between morphs (*ARHGEF37-like*, *ARL16*, *ARMC1*, *GLI3-like*, *H2A-like*, *H3-like*, *MEGF9*, *NFIX*, and *SLC9A3R2-like*) and between developmental timepoints (all of the investigated genes except for *GLI3-like* and *ARMC1*). These patterns are of interest for the understanding of early development in salmonids. The genes differently expressed between morphs had various functions, from *H2A-like* and *H3-like* coding for core histone proteins, *SLC9A3R2-like* encoding the scaffold protein NHERF2 (Murthy et al., 1998; Oh et al., 2004), *ARHGEF37-like* coding for a putative RhoGEF protein which has been suggested to regulate endocytosis (Viplav et al., 2019), to the more relevant *GLI3-like*, *NFIX and MEGF9*. GLI3 is a zinc finger transcription factor that has been shown to be linked to tooth agenesis in humans (Liu et al., 2013) and hypothesized to play a role in skeletal malocclusion (Marañón-Vásquez et al., 2019), a defect associated with maxillary and mandibular development (Joshi et al., 2014). Similarly, different mutations of *NFIX* have been linked to the Marshall-Smith syndrome, the symptoms of which include facial dysmorphism (Martinez et al., 2015). MEGF9 is a putative transmembrane protein expressed in many tissues (Brandt-Bohne et al., 2007) which has been shown to affect the expression of EGFR and the matrix-remodelling factors MMP13 and ADAMTS-5 through PI3K/Akt signalling. Interestingly, misexpression of MEGF9 has been shown to induce cartilage degradation and is hypothesized to play a role in osteoarthritis (Jiang et al., 2021). As differences in the regulation of these pathways could contribute to the phenotypic differences observed between the morphs, investigating the tissue localization and activity of the products of these genes would be of interest. Local expression differences in bone primordia or surrounding tissues in the craniofacial region could, for instance, contribute to differences in the morphogenesis of the trophic apparatus of benthic and limnetic morphs. This study adds to previous studies (Ahi et al., 2014; Guðbrandsson et al., 2017; Kapralova et al., 2014) which have highlighted multiple molecular systems that may act differentially during development of these sympatric charr morphs. As sympatric charr morphs have evolved repeatedly in multiple lakes (Adams et al., 1998; Arbour et al., 2010; Hindar & Jonsson, 1982; Jonsson et al., 1988), investigating whether similar changes in methylation profiles have been influenced by natural selection to bring about

this parallel evolution is of considerable interest. Genome and transcriptome analyses of sympatric charr in Scotland has indeed suggested more parallelism at expression level than at the genetic level (Jacobs et al., 2020), while genome and methylome analyses between two lineages of capelin (which diverged 2.5 million years ago) suggest more parallelism at the genetic level than at the methylation level (Cayuela et al., 2021).

For the most part we did not detect a strong relationship between the overall methylation of a region and the expression of nearby genes, even though the methylation of most genes of interest and their expression levels separated the samples based on the morphs to an extent. As stated above, we evaluated methylation and expression at the whole embryonic level, which may have obscured tissue specific associations between methylation and transcription of nearby genes. However, four genes showed a negative (*MEGF9* and *MPP3*) or positive (*ARHGEF37-like* and *H3-like*) correlation between methylation and expression (Figure 7). The direction of correlation between methylation and expression has been shown to be dependent on the gene feature (exon, intron, promoter) where the methylation sites are located (Anastasiadi et al., 2018; Arechederra et al., 2018; Esteller, 2002). Our results seem to corroborate these trends; the methylation differences on *MEGF9* and *MPP3* are located in their respective promoter regions, while they are located on an exon for *ARHGEF37-like* and overlap both the promoter and exon of *H3-like*.

4.5 | Differential methylation in histone gene clusters between timepoints and morphs

Four genes in the highly covered histone clusters were differently methylated between timepoints, with two of them *H1* and *H2B*, being differentially methylated between morphs (Figure 8a). As histones are essential for chromatin structure, changes in the methylation and expression of histone genes between morphs may contribute to phenotypic differences. But many questions remain, for instance, are those patterns systemic across the embryo or driven by action in specific tissues? How early do they come on and do they persist in adults? Could the methylation differences be influenced by histone gene copy number variation between morphs?

Additionally, H1 showed high levels of methylation heterogeneity over its region, with a significant difference between the methylation state of its promoter and gene body, showcasing a good example of the subtle and precise establishment of methylation marks on different gene features (Figure 8b).

4.6 | Methylation and evolutionary divergence in Arctic charr

This study highlights many differences in the methylome of Arctic charr morphs in lake Thingvallavatn. Disentangling their importance for evolutionary divergence is not trivial and further studies are required to test whether variation in epigenetic regulation influences adaptive evolution. Methylation differences between morphs were observed mainly between benthic and limnetic individuals, reflecting the larger genetic differences between benthic and limnetic morphs (Han Xiao, Zophonias O. Jónsson, et al., unpublished data; Guðbrandsson et al., 2019). It is likely that the methylation differences were partially environmentally driven, as a result of the biotic or abiotic niche. However, it is not impossible that methylation differences helped influence niche colonization in the first place.

Correlation between methylation and expression was observed for genes in which methylation changes mainly over time rather than between morphs, which is not surprising given the strong methylation changes throughout development and the limited number of genes studied in detail. However, the DMR close to *MPP3* also shows significant differences in methylation between morphs. This DMR could thus have an impact on divergence by affecting the ontogenetic shift in expression of *MPP3*. Further studies on other regions, differentially methylated between morphs, would be needed to find genes specifically controlled by these methylation differences. However, while the most commonly proposed role of methylation is to regulate the expression of nearby genes, it can impact gene expression in other ways; a significant number of DMRs and differentially methylated CpGs were located in intergenic regions and were not investigated further (Figure S6). These regions may have biological significance; for example, they may affect gene expression by either encompassing enhancers or repressors (Aran et al., 2013) or by interacting with genomic regions located sterically close to them in terms of DNA quaternary structure (even if located far away in linear DNA sequence space) (Dekker, 2006). As procedures for chromatin conformation capture (Dekker, 2002) have not been reported for *S. alpinus*, the impact of methylation variation in these regions is harder to investigate, but it is very likely that methylation differences between morphs in these regions have an impact on the expression of genes playing a role in the setup of their phenotypic differences.

Finally, DMRs (both between morphs and between timepoints) were found to be enriched in genes encoding small RNAs, the regulation of which needs to be investigated further in some system, in regards to evolutionary divergence.

4.7 | Conclusion

This study highlights differences in methylation between recently diverged sympatric ecomorphs of Arctic charr at embryonic stages. While it is not yet possible to decipher in which proportions epigenomic changes arise following environmental triggers, to balance the effects of genetic mutations or to enhance them further, it is likely that both genetic and epigenetic layers interact to shape phenotypic diversity. Future research investigating epigenomic signals in evolutionarily interesting populations will be of great importance for a deeper understanding of adaptive divergence and speciation.

AUTHOR CONTRIBUTIONS

Zophonías O. Jónsson, Sigrídur Rut Franzdóttir, Arnar Pálsson and Sebastien Matlosz conceived and designed the experiments. Zophonías O. Jónsson, Sebastien Matlosz, Sigrídur Rut Franzdóttir and Arnar Pálsson were responsible for sampling and rearing. Sebastien Matlosz performed the experiments. Benjamín Sigurgeirsson, Zophonías O. Jónsson, and Arnar Pálsson contributed new analytical tools. Sebastien Matlosz and Benjamín Sigurgeirsson analysed the data. Sebastien Matlosz, Arnar Pálsson, Sigrídur Rut Franzdóttir, Benjamín Sigurgeirsson and Zophonías O. Jónsson were responsible for writing the manuscript.

ACKNOWLEDGEMENTS

This research was funded by the Icelandic Research Fund (grant number 163477-051 to Zophonías O. Jónsson) and a Teaching Assistant Scholarship from the University of Iceland. We acknowledge Sigurður Sveinn Snorrasson, Sigrún Reynisdóttir, Kalina Kapralova and Quentin Horta-Lacueva for assisting with the sampling of the parents, maintenance of and sampling of embryos. We thank Lea Jerman Plesec for developing the Sex-PCR method for charr, and her and Denis Warshan for their help with various programming tasks. We also thank the reviewers for their insightful comments that greatly improved this manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

All sequencing reads have been deposited on ENA (accession number: PRJEB45551). All scripts used for data processing, analyses and drawing of graphs have been made available at https://github.com/sebma-github/DNA-methylation-Charr.

ORCID

Sébastien Matlosz https://orcid.org/0000-0002-4995-6682
Zophonías O. Jónsson https://orcid.org/0000-0001-5798-9647

REFERENCES

Adams, C. E., Fraser, D., Huntingford, F. A., Greer, R. B., Askew, C. M., & Walker, A. F. (1998). Trophic polymorphism amongst Arctic charr from loch Rannoch, Scotland. *Journal of Fish Biology*, *52*(6), 1259–1271. https://doi.org/10.1111/j.1095-8649.1998.tb00970.x

Ahi, E. P., Guðbrandsson, J., Kapralova, K. H., Franzdóttir, S. R., Snorrason, S. S., Maier, V. H., & Jónsson, Z. O. (2013). Validation of reference genes for expression studies during craniofacial development in Arctic charr. *PLoS One*, *8*(6), e66389. https://doi.org/10.1371/journal.pone.0066389

Ahi, E. P., Kapralova, K. H., Pálsson, A., Maier, V. H., Gudbrandsson, J., Snorrason, S. S., Jónsson, Z. O., & Franzdóttir, S. R. (2014). Transcriptional dynamics of a conserved gene expression network associated with craniofacial divergence in Arctic charr. *EvoDevo*, *5*(1), 40. https://doi.org/10.1186/2041-9139-5-40

Akalin, A., Franke, V., Vlahoviček, K., Mason, C. E., & Schübeler, D. (2015). Genomation: A toolkit to summarize, annotate and visualize genomic intervals. *Bioinformatics*, *31*(7), 1127–1129. https://doi.org/10.1093/bioinformatics/btu775

Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F. E., Figueroa, M. E., Melnick, A., & Mason, C. E. (2012). MethylKit: A comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology*, *13*(10), R87. https://doi.org/10.1186/gb-2012-13-10-r87

Alexa A., & Rahnenfuhrer, J. (2022). topGO: Enrichment Analysis for Gene Ontology. R package version 2.48.0.

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

Andersen Ingrid, S., Reiner, A. H., Aanes, H., Alestrom, P., & Collas, P. (2012). Developmental features of DNA methylation during activation of the embryonic zebrafish genome. *Genome Biology*, *13*(7), R65. https://doi.org/10.1186/gb-2012-13-7-r65

Aran, D., Sabato, S., & Hellman, A. (2013). DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes. *Genome Biology*, *14*(3), R21. https://doi.org/10.1186/gb-2013-14-3-r21

Arbour, J., Hardie, D., & Hutchings, J. (2011). Morphometric and genetic analyses of two sympatric morphs of Arctic char (Salvelinus alpinus) in the Canadian High Arctic. *Canadian Journal of Zoology*, *89*, 19–30 https://doi.org/10.1139/Z10-100

Arechederra, M., Daian, F., Yim, A., Bazai, S. K., Richelme, S., Dono, R., Saurin, A. J., Habermann, B. H., & Maina, F. (2018). Hypermethylation of gene body CpG Islands predicts high dosage of functional oncogenes in liver cancer. *Nature Communications*, *9*(1), 3164. https://doi.org/10.1038/s41467-018-05550-5

Baerwald, M. R., Meek, M. H., Stephens, M. R., Nagarajan, R. P., Goodbla, A. M., Tomalty, K. M. H., Thorgaard, G. H., May, B., & Nichols, K. M. (2016). Migration-related phenotypic divergence is associated with epigenetic modifications in rainbow trout. *Molecular Ecology*, *25*(8), 1785–1800. https://doi.org/10.1111/mec.13231

Beck, S. V., Räsänen, K., Ahi, E. P., Kristjánsson, B. K., Skúlason, S., Jónsson, Z. O., & Leblanc, C. A. (2019). Gene expression in the phenotypically plastic Arctic charr (*Salvelinus alpinus*): A focus on growth and ossification at early stages of development. *Evolution & Development*, *21*(1), 16–30. https://doi.org/10.1111/ede.12275

Beck, S. V., Räsänen, K., Leblanc, C. A., Skúlason, S., Jónsson, Z. O., & Kristjánsson, B. K. (2020). Differences among families in craniofacial shape at early life-stages of Arctic charr (*Salvelinus alpinus*). *BMC Developmental Biology*, *20*(1), 21. https://doi.org/10.1186/s12861-020-00226-0

Bird, A. P., & Wolffe, A. P. (1999). Methylation-induced repression--belts, braces, and chromatin. *Cell*, *99*(5), 451–454.

Brandt-Bohne, U., Keene, D. R., White, F. A., & Koch, M. (2007). MEGF9: A novel transmembrane protein with a strong and developmentally regulated expression in the nervous system. *Biochemical Journal*, *401*(Pt 2), 447–457. https://doi.org/10.1042/BJ20060691

Brawand, D., Wagner, C. E., Li, Y. I., Malinsky, M., Keller, I., Fan, S., Simakov, O., Ng, A. Y., Lim, Z. W., Bezault, E., Turner-Maier, J., Johnson, J., Alcazar, R., Noh, H. J., Russell, P., Aken, B., Alföldi, J., Amemiya, C., Azzouzi, N., … Di Palma, F. (2014). The genomic substrate for adaptive radiation in African cichlid fish. *Nature*, *513*(7518), 375–381 https://doi.org/10.1038/nature13726

Cayuela, H., Rougeux, C., Laporte, M., Mérot, C.,Normandeau, E., Leitwein, M., Dorant, Y., Præbel, K., Kenchington, E., Clément, M., Sirois, P., & Bernatchez, L. (2021). Genome-wide DNA methylation predicts environmentally-driven life historyvariation in a marine fish. *BioRxiv*. 2021.01.28.428603. https://doi.org/10.1101/2021.01.28.428603

Claus, R., Lucas, D. M., Stilgenbauer, S., Ruppert, A. S., Yu, L., Zucknick, M., Mertens, D., Bühler, A., Oakes, C. C., Larson, R. A., Kay, N. E., Jelinek, D. F., Kipps, T. J., Rassenti, L. Z., Gribben, J. G., Döhner, H., Heerema, N. A., Marcucci, G., Plass, C., & Byrd, J. C. (2012). Quantitative DNA methylation analysis identifies a single CpG

dinucleotide important for ZAP-70 expression and predictive of prognosis in chronic lymphocytic leukemia. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *30*(20), 2483–2491. https://doi.org/10.1200/JCO.2011.39.3090

Dekker, J. (2002). Capturing chromosome conformation. *Science*, *295*(5558), 1306–1311. https://doi.org/10.1126/science.1067799

Dekker, J. (2006). The three "C" s of chromosome conformation capture: Controls, controls, controls. *Nature Methods*, *3*(1), 17–21. https://doi.org/10.1038/nmeth823

Doenecke, D., Albig, W., Bode, C., Drabent, B., Franke, K., Gavenis, K., & Witt, O. (1997). Histones: Genetic diversity and tissue-specific gene expression. *Histochemistry and Cell Biology*, *107*(1), 1–10. https://doi.org/10.1007/s004180050083

Esteller, M. (2002). CpG Island hypermethylation and tumor suppressor genes: A booming present, a brighter future. *Oncogene*, *21*(35), 5427–5440. https://doi.org/10.1038/sj.onc.1205600

Fraga, M. F., & Esteller, M. (2002). DNA methylation: A profile of methods and applications. *BioTechniques*, *33*(3), 632–634, 636–649. https://doi.org/10.2144/02333rv01

Frommer, M., McDonald, L. E., Millar, D. S., Collis, C. M., Watt, F., Grigg, G. W., Molloy, P. L., & Paul, C. L. (1992). A genomic sequencing protocol that yields a positive display of 5-Methylcytosine residues in individual DNA strands. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(5), 1827–1831.

Fuzinatto, C. F., Flohr, L., Melegari, S. P., & Matias, W. G. (2015). Oxidative stress and hypermethylation induced by exposure of *Oreochromis niloticus* to complex environmental mixtures of river water from Cubatão do Sul, Brazil. *Ecotoxicology and Environmental Safety*, *114*, 190–197. https://doi.org/10.1016/j.ecoenv.2015.01.025

Gavery, M. R., Nichols, K. M., Goetz, G. W., Middleton, M. A., & Swanson, P. (2018). Characterization of genetic and epigenetic variation in sperm and red blood cells from adult hatchery and natural-origin steelhead, *Oncorhynchus mykiss*. *G3 (Bethesda, Maryland)*, *8*(11), 3723–3736. https://doi.org/10.1534/g3.118.200458

Geng, Y., Gao, L., & Yang, J.. (2013). Epigenetic flexibility underlying phenotypic plasticity. In U. Lüttge, W. Beyschlag, D. Francis, & J. Cushman (Eds.), *Progress in botany* (Vol. 74, pp. 153–163). Springer. https://doi.org/10.1007/978-3-642-30967-0_5.

Gorodilov, Y. N. (1996). Description of the early ontogeny of the Atlantic Salmon, *Salmo salar*, with a novel system of interval (state) identification. *Environmental Biology of Fishes*, *47*, 109–127. https://doi.org/10.1007/BF00005034

Grewal, S. S. (2015). Why should cancer biologists care about tRNAs? tRNA synthesis, mRNA translation and the control of growth. *Biochimica et Biophysica Acta*, *1849*(7), 898–907. https://doi.org/10.1016/j.bbagrm.2014.12.005

Gudbrandsson, J., Ahi, E. P., Franzdottir, S. R., Kapralova, K. H., Kristjansson, B. K., Steinhaeuser, S. S., Maier, V. H., Johannesson, I. M., Snorrason, S. S., Jonsson, Z. O., & Palsson, A. (2016). The developmental transcriptome of contrasting Arctic charr (Salvelinus alpinus) morphs. F1000Research, 4, 136. https://doi.org/10.12688/f1000research.6402.3

Guðbrandsson, J., Franzdóttir, S. R., Kristjánsson, B. K., Ahi, E. P., Maier, V. H.,Kapralova, K. H., Snorrason, S. S., Jónsson, Z. O., & Pálsson, A. (2018).Differential gene expression during early development in recently evolved andsympatric Arctic charr morphs. *PeerJ*, *6*, e4345.https://doi.org/10.7717/peerj.4345

Guðbrandsson, J., Kapralova, K. H., Franzdóttir, S. R., Bergsveinsdóttir, Þ. M., Hafstað, V., Jónsson, Z. O., Snorrason, S. S., & Pálsson, A. (2019). Extensive genetic differentiation between recently evolved sympatric Arctic charr morphs. *Ecology and Evolution*, *9*(19), 10964–10983. https://doi.org/10.1002/ece3.5516

Han, F., Lamichhaney, S., Grant, B. R., Grant, P. R., Andersson, L., & Webster, M. T. (2017). Gene flow, ancient polymorphism, and ecological adaptation shape the genomic landscape of divergence among Darwin's finches. *Genome Research*, *27*(6), 1004–1015. https://doi.org/10.1101/gr.212522.116

Hanson, G., & Coller, J. (2018). Codon optimality, bias and usage in translation and mRNA decay. *Nature Reviews Molecular Cell Biology*, *19*(1), 20–30. https://doi.org/10.1038/nrm.2017.91

Hartig, F. (2022). *DHARMa: Residual diagnostics for hierarchical (multi-level/mixed) regression models*. R package version 0.4.5. http://florianhartig.github.io/DHARMa/

Heckwolf, M. J., Meyer, B. S., Häsler, R., Höppner, M. P., Eizaguirre, C., & Reusch, T. B. H. (2020). Two different epigenetic information channels in wild three-spined sticklebacks are involved in salinity adaptation. *Science Advances*, *6*(12), eaaz1138. https://doi.org/10.1126/sciadv.aaz1138

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F., & Vandesompele, J. (2007). QBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biology*, *8*(2), R19. https://doi.org/10.1186/gb-2007-8-2-r19

Hindar, K., & Jonsson, B. (1982). Habitat and food segregation of dwarf and normal Arctic charr (Salvelinus alpinus) from Vangsvatnet Lake, Western Norway. *Canadian Journal of Fisheries and Aquatic Sciences*., *39*(7), 1030–1045 https://doi.org/10.1139/f82-138

Hoegg, S., Brinkmann, H., Taylor, J. S., & Meyer, A. (2004). Phylogenetic timing of the fish-specific genome duplication correlates with the diversification of teleost fish. *Journal of Molecular Evolution*, *59*(2), 190–203. https://doi.org/10.1007/s00239-004-2613-z

Hu, J., Wuitchik, S. J. S., Barry, T. N., Jamniczky, H. A., Rogers, S. M., & Barrett, R. D. H. (2021). Heritability of DNA methylation in threespine stickleback (*Gasterosteus aculeatus*). *Genetics*, *217*(1), iyab001. https://doi.org/10.1093/genetics/iyab001

Hubley, R., Finn, R. D., Clements, J., Eddy, S. R., Jones, T. A., Bao, W., Arian, F., Smit, A., & Wheeler, T. J. (2016). The Dfam database of repetitive DNA families. *Nucleic Acids Research*, *44*(D1), D81–D89. https://doi.org/10.1093/nar/gkv1272

Irizarry, R. A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., Cui, H., Gabo, K., Rongione, M., Webster, M., Ji, H., Potash, J., Sabunciyan, S., & Feinberg, A. P. (2009). Genome-wide methylation analysis of human colon cancer reveals similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nature Genetics*, *41*(2), 178–186. https://doi.org/10.1038/ng.298

Jacobs, A., Carruthers, M., Yurchenko, A., Gordeeva, N. V., Alekseyev, S. S., Hooker, O., Leong, J. S., Minkley, D. R., Rondeau, E. B., Koop, B. F., Adams, C. E., & Elmer, K. R. (2020). Parallelism in eco-morphology and gene expression despite variable evolutionary and genomic backgrounds in a Holarctic fish. *PLoS Genetics*, *16*(4), e1008658. https://doi.org/10.1371/journal.pgen.1008658

Jiang, L., Zhang, J., Wang, J.-J., Wang, L., Zhang, L., Li, G., Yang, X., Ma, X., Sun, X., Cai, J., Zhang, J., Huang, X., Yu, M., Wang, X., Liu, F., Wu, C.-I., He, C., Zhang, B., Ci, W., & Liu, J. (2013). Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. *Cell*, *153*(4), 773–784. https://doi.org/10.1016/j.cell.2013.04.041

Jiang, L., Zhou, X., Xu, K., Hu, P., Bao, J., Li, J., Zhu, J., & Lidong, W. (2021). MiR-7/EGFR/MEGF9 axis regulates cartilage degradation in osteoarthritis via PI3K/AKT/MTOR signaling pathway. *Bioengineered*, *12*(1), 8622–8634. https://doi.org/10.1080/21655979.2021.1988362

Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M. C., White, S., Birney, E., Searle, S., Schmutz, J., Grimwood, J., Dickson, M. C., Myers, R. M., Miller, C. T., Summers, B. R., Knecht, A. K., … Kingsley, D. M. (2012). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, *484*(7392), 55–61. https://doi.org/10.1038/nature10944

Jones, P. A. (1999). The DNA methylation paradox. *Trends in Genetics*, *15*(1), 34–37. https://doi.org/10.1016/s0168-9525(98)01636-9

Jonsson, B., Skúlason, S., Snorrason, S. S., Sandlund, O. T., Malmquist, H. J., Jónasson, P. M., Cydemo, R., & Lindem, T. (1988). Life history variation of polymorphic Arctic charr (*Salvelinus alpinus*) in Thingvallavatn, Iceland. *Canadian Journal of Fisheries and Aquatic Sciences*, *45*(9), 1537–1547. https://doi.org/10.1139/f88-182

Joshi, N., Hamdan, A. M., & Fakhouri, W. D. (2014). Skeletal malocclusion: A developmental disorder with a life-long morbidity. *Journal of Clinical Medicine Research*, *6*(6), 399–408. https://doi.org/10.14740/jocmr1905w

Kapralova, K. H., Morrissey, M. B., Kristjánsson, B. K., Ólafsdóttir, G. Á., Snorrason, S. S., & Ferguson, M. M. (2011). Evolution of adaptive diversity and genetic connectivity in Arctic charr (*Salvelinus alpinus*) in Iceland. *Heredity*, *106*(3), 472–487. https://doi.org/10.1038/hdy.2010.161

Kapralova, K. H., Franzdóttir, S. R., Jónsson, H., Snorrason, S. S., & Jónsson, Z. O. (2014). Patterns of miRNA expression in Arctic charr development. *PLoS One*, *9*(8), e106084. https://doi.org/10.1371/journal.pone.0106084

Karimi, M. M., Goyal, P., Maksakova, I. A., Bilenky, M., Leung, D., Tang, J. X., Shinkai, Y., Mager, D. L., Jones, S., Hirst, M., & Lorincz, M. C. (2011). DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements and chimaeric transcripts in mouse ES cells. *Cell Stem Cell*, *8*(6), 676–687. https://doi.org/10.1016/j.stem.2011.04.004

Kitazawa, R., & Kitazawa, S. (2007). Methylation status of a single CpG locus 3 bases upstream of TATA-box of receptor activator of nuclear factor-KappaB ligand (RANKL) gene promoter modulates cell- and tissue-specific RANKL expression and Osteoclastogenesis. *Molecular Endocrinology (Baltimore, Maryland)*, *21*(1), 148–158. https://doi.org/10.1210/me.2006-0205

Klironomos, F. D., Berg, J., & Collins, S. (2013). How epigenetic mutations can affect genetic evolution: Model and mechanism. *BioEssays*, *35*(6), 571–578. https://doi.org/10.1002/bies.201200169

Kornberg, R. D. (1974). Chromatin structure: A repeating unit of histones and DNA. *Science (New York, N.Y.)*, *184*(4139), 868–871. https://doi.org/10.1126/science.184.4139.868

Krueger, F., & Andrews, S. R. (2011). Bismark: A flexible aligner and methylation caller for bisulfite-seq applications. *Bioinformatics (Oxford, England)*, *27*(11), 1571–1572. https://doi.org/10.1093/bioinformatics/btr167

Laporte, M., Le Luyer, J., Rougeux, C., Dion-Côté, A.-M., Krick, M., & Bernatchez, L. (2019). DNA methylation reprogramming, TE derepression, and postzygotic isolation of nascent animal species. *Science Advances*, *5*(10), eaaw1644. https://doi.org/10.1126/sciadv.aaw1644

Le Luyer, J., Laporte, M., Beacham, T. D., Kaukinen, K. H., Withler, R. E., Leong, J. S., Rondeau, E. B., Koop, B. F., & Bernatchez, L. (2017). Parallel epigenetic modifications induced by hatchery rearing in a Pacific Salmon. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(49), 12964–12969. https://doi.org/10.1073/pnas.1711229114

Lind, M. I., & Spagopoulou, F. (2018). Evolutionary consequences of epigenetic inheritance. *Heredity*, *121*(3), 205–209. https://doi.org/10.1038/s41437-018-0113-y

Liu, H., Dong, H., Wong, S., Xu, N., Zhao, H., & Feng, H. (2013). Rs929387 of GLI3 is involved in tooth agenesis in Chinese Han population. *PLoS One*, *8*(11), e80860. https://doi.org/10.1371/journal.pone.0080860

Liu, Y., Zhang, Y., Tao, S., Guan, Y., Zhang, T., & Wang, Z. (2016). Global DNA methylation in gonads of adult zebrafish danio rerio under bisphenol a exposure. *Ecotoxicology and Environmental Safety*, *130*, 124–132. https://doi.org/10.1016/j.ecoenv.2016.04.012

Locati, M. D., Pagano, J. F. B., Girard, G., Ensink, W. A., van Olst, M., van Leeuwen, S., Nehrdich, U., Spaink, H. P., Rauwerda, H., Jonker, M. J., Dekker, R. J., & Breit, T. M. (2017). Expression of distinct maternal and somatic 5.8S, 18S, and 28S rRNA types during zebrafish development. *RNA*, *23*(8), 1188–1199. https://doi.org/10.1261/rna.061515.117

Macqueen, D. J., Kristjánsson, B. K., Paxton, C. G. M., Vieira, V. L. A., & Johnston, I. A. (2011). The parallel evolution of dwarfism in Arctic charr is accompanied by adaptive divergence in MTOR-pathway gene expression. *Molecular Ecology*, *20*(15), 3167–3184. https://doi.org/10.1111/j.1365-294X.2011.05172.x

Macqueen, D., & Johnston, I. (2014). A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. *Proceedings. Biological Sciences/The Royal Society*, *281*, 20132881. https://doi.org/10.1098/rspb.2013.2881

Malmquist, H. J., Snorrason, S. S., Skúlason, S., Jonsson, B., Sandlund, O. T., & Jonasson, P. M. (1992). Diet differentiation in polymorphic Arctic charr in Thingvallavatn, Iceland. *Journal of Animal Ecology*, *61*(1), 21–35. https://doi.org/10.2307/5505

Marañón-Vásquez, G. A., Dantas, B., Kirschneck, C., Arid, J., Cunha, A., de Ramos, A. G. C., Omori, M. A., Rodrigues, A. S., Teixeira, E. C., Levy, S. C., Schroeder, A., Matsumoto, M. A. N., Proff, P., Antunes, L. A. A., Vieira, A. R., Antunes, L. S., & Küchler, E. C. (2019). Tooth agenesis-related GLI2 and GLI3 genes may contribute to craniofacial skeletal morphology in humans. *Archives of Oral Biology*, *103*, 12–18. https://doi.org/10.1016/j.archoralbio.2019.05.008

Markevich, G., Esin, E., & Anisimova, L. (2018). Basic description and some notes on the evolution of seven sympatric morphs of Dolly *Varden salvelinus* Malma from the Lake Kronotskoe Basin. *Ecology and Evolution*, *8*(5), 2554–2567. https://doi.org/10.1002/ece3.3806

Martinez, F., Marín-Reina, P., Sanchis-Calvo, A., Perez-Aytés, A., Oltra, S., Roselló, M., Mayo, S., Monfort, S., Pantoja, J., & Orellana, C. (2015). Novel mutations of NFIX gene causing Marshall-smith syndrome or Sotos-like syndrome: One gene, two phenotypes. *Pediatric Research*, *78*(5), 533–539. https://doi.org/10.1038/pr.2015.135

Meissner, A., Gnirke, A., Bell, G. W., Ramsahoye, B., Lander, E. S., & Jaenisch, R. (2005). Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Research*, *33*(18), 5868–5877. https://doi.org/10.1093/nar/gki901

Merlo, A., Herman, J. G., Mao, L., Lee, D. J., Gabrielson, E., Burger, P. C., Baylin, S. B., & Sidransky, D. (1995). 5′ CpG Island methylation is associated with transcriptional silencing of the tumour suppressor P16/CDKN2/MTS1 in human cancers. *Nature Medicine*, *1*(7), 686–692. https://doi.org/10.1038/nm0795-686

Meyer, A., & Van de Peer, Y. (2005). From 2R to 3R: Evidence for a fish-specific genome duplication (FSGD). *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *27*(9), 937–945. https://doi.org/10.1002/bies.20293

Murthy, A., Gonzalez-Agosti, C., Cordero, E., Pinney, D., Candia, C., Solomon, F., Gusella, J., & Ramesh, V. (1998). NHE-RF, a regulatory cofactor for Na(+)-H+ exchange, is a common interactor for Merlin and ERM (MERM) proteins. *The Journal of Biological Chemistry*, *273*(3), 1273–1276. https://doi.org/10.1074/jbc.273.3.1273

Navarro-Martín, L., Viñas, J., Ribas, L., Díaz, N., Gutiérrez, A., Di Croce, L., & 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 Genetics*, *7*(12), e1002447. https://doi.org/10.1371/journal.pgen.1002447

Nichols, J. T., Blanco-Sánchez, B., Brooks, E. P., Parthasarathy, R., Dowd, J., Subramanian, A., Nachtrab, G., Poss, K. D., Schilling, T. F., & Kimmel, C. B. (2016). Ligament versus bone cell identity in the zebrafish hyoid skeleton is regulated by Mef2ca. *Development (Cambridge, England)*, *143*(23), 4430–4440. https://doi.org/10.1242/dev.141036

Oh, Y.-S., Jo, N. W., Choi, J. W., Kim, H. S., Seo, S.-W., Kang, K.-O., Hwang, J.-I., Heo, K., Kim, S.-H., Kim, Y.-H., Kim, I.-H., Kim, J. H., Banno, Y., Ryu, S. H., & Suh, P.-G. (2004). NHERF2 specifically interacts with LPA2 receptor and defines the specificity and efficiency of

receptor-mediated phospholipase C-beta3 activation. *Molecular and Cellular Biology*, *24*(11), 5069–5079. https://doi.org/10.1128/MCB.24.11.5069-5079.2004

Palmer, D. H., & Kronforst, M. R. (2015). Divergence and gene flow among Darwin's finches: A genome-wide view of adaptive radiation driven by interspecies allele sharing. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *37*(9), 968–974. https://doi.org/10.1002/bies.201500047

Peng, H., Shi, J., Zhang, Y., Zhang, H., Liao, S., Li, W., Lei, L., Han, C., Ning, L., Cao, Y., Zhou, Q., Chen, Q., & Duan, E. (2012). A novel class of tRNA-derived small RNAs extremely enriched in mature mouse sperm. *Cell Research*, *22*(11), 1609–1612. https://doi.org/10.1038/cr.2012.141

Perez, M. F., & Lehner, B. (2019). Intergenerational and transgenerational epigenetic inheritance in animals. *Nature Cell Biology*, *21*(2), 143–151. https://doi.org/10.1038/s41556-018-0242-9

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Research*, *29*(9), e45.

Putnam, H. M., Davidson, J. M., & Gates, R. D. (2016). Ocean acidification influences host DNA methylation and phenotypic plasticity in environmentally susceptible corals. *Evolutionary Applications*, *9*(9), 1165–1178. https://doi.org/10.1111/eva.12408

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for ComparingGenomic features. *Bioinformatics (Oxford, England)*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Richardson, M. K. (1999). Vertebrate evolution: The developmental origins of adult variation. *BioEssays*, *21*(7), 604–613. https://doi.org/10.1002/(SICI)1521-1878(199907)21:7<604::AID-BIES9>3.0.CO;2-U

Roberts, S. B., & Gavery, M. R. (2012). Is there a relationship between DNA methylation and phenotypic plasticity in invertebrates? *Frontiers in Physiology*, *2*, 116. https://doi.org/10.3389/fphys.2011.00116

Rooney, A. P., Piontkivska, H., & Nei, M. (2002). Molecular evolution of the nontandemly repeated genes of the histone 3 multigene family. *Molecular Biology and Evolution*, *19*(1), 68–75. https://doi.org/10.1093/oxfordjournals.molbev.a003983

Sandlund, O. T., Jonsson, B., Malmquist, H. J., Gydemo, R., Lindem, T., Skúlason, S., Snorrason, S. S., & Jónasson, P. M. (1987). Habitat use of Arctic charr *Salvelinus alpinus* in Thingvallavatn, Iceland. *Environmental Biology of Fishes*, *20*(4), 263–274. https://doi.org/10.1007/BF00005297

Sandlund, O. T., Gunnarsson, K., Jónasson, P. M., Jonsson, B., Lindem, T., Magnússon, K. P., Malmquist, H. J., Sigurjónsdóttir, H., Skúlason, S., & Snorrason, S. S. (1992). The Arctic charr *Salvelinus alpinus* in Thingvallavatn. *Oikos*, *64*(1/2), 305–351. https://doi.org/10.2307/3545056

Schmitz, R. J., Schultz, M. D., Lewsey, M. G., O'Malley, R. C., Urich, M. A., Libiger, O., Schork, N. J., & Ecker, J. R. (2011). Transgenerational epigenetic instability is a source of novel methylation variants. *Science (New York, N.Y.)*, *334*(6054), 369–373. https://doi.org/10.1126/science.1212959

Schröder, C., Leitão, E., Wallner, S., Schmitz, G., Klein-Hitpass, L., Sinha, A., Jöckel, K.-H., Heilmann-Heimbach, S., Hoffmann, P., Nöthen, M. M., Steffens, M., Ebert, P., Rahmann, S., & Horsthemke, B. (2017). Regions of common inter-individual DNA methylation differences in human monocytes: Genetic basis and potential function. *Epigenetics & Chromatin*, *10*(1), 37. https://doi.org/10.1186/s13072-017-0144-2

Skinner, M. K., Gurerrero-Bosagna, C., Haque, M. M., Nilsson, E. E., Koop, J. A. H., Knutie, S. A., & Clayton, D. H. (2014). Epigenetics and the evolution of Darwin's finches. *Genome Biology and Evolution*, *6*(8), 1972–1989. https://doi.org/10.1093/gbe/evu158

Skúlason, S., Snorrason, S. S., Noakes, D. L. G., Ferguson, M. M., & Malmquist, H. J. (1989). Segregation in spawning and early life history among polymorphic Arctic charr, *Salvelinus alpinus*, in Thingvallavatn, Iceland. *Journal of Fish Biology*, *35*, 225–232. https://doi.org/10.1111/j.1095-8649.1989.tb03065.x

Skúlason, S., Noakes, D. L. G., & Snorrason, S. S. (1989). Ontogeny of trophic morphology in four sympatric morphs of Arctic charr *Salvelinus alpinus* in Thingvallavatn, Iceland. *Biological Journal of the Linnean Society*, *38*(3), 281–301. https://doi.org/10.1111/j.1095-8312.1989.tb01579.x

Smit, A. F. A., Hubley, R., & Green, P., & Arian, F.A. (2013). *RepeatMasker Open-4.0*. http://www.repeatmasker.org

SourceForge. *BBMap*. https://sourceforge.net/projects/bbmap/

Tamayo, A.-M. P., Devineau, O., Præbel, K., Kahilainen, K. K., & Østbye, K. (2020). A brain and a head for a different habitat: Size variation in four morphs of Arctic charr (*Salvelinus alpinus* [L.]) in a deep oligotrophic Lake. *Ecology and Evolution*, *10*(20), 11335–11351. https://doi.org/10.1002/ece3.6771

Varriale, A., & Bernardi, G. (2006). DNA methylation and body temperature in fishes. *Gene*, *385*, 111–121. https://doi.org/10.1016/j.gene.2006.05.031

Venney, C., Love, O., Drown, E., & Heath, D. (2020). *DNA methylation profiles suggest intergenerational transfer of maternal effects*. https://doi.org/10.1093/molbev/msz244

Venney, C. J., Wellband, K. W., Normandeau, E., Houle, C., Garant, D., Audet, C., & Bernatchez, L. (2022). Thermal regime during parental sexual maturation, butnot during offspring rearing, modulates DNA methylation in brook charr(Salvelinus fontinalis). Proceedings. *Biological Sciences*, *289*(1974), 20220670. https://doi.org/10.1098/rspb.2022.0670

Verhoeven Koen, J. F., von Holdt, B. M., & Sork, V. L. (2016). Epigenetics in ecology and evolution: What we know and what we need to know. *Molecular Ecology*, *25*(8), 1631–1638. https://doi.org/10.1111/mec.13617

Vernaz, G., Malinsky, M., Svardal, H., Du, M., Tyers, A. M., Santos, M. E., Durbin, R., Genner, M. J., Turner, G. F., & Miska, E. A. (2021). Mapping epigenetic divergence in the massive radiation of Lake Malawi cichlid fishes. *Nature Communications*, *12*(1), 5870. https://doi.org/10.1038/s41467-021-26166-2

Viplav, A., Saha, T., Huertas, J., Selenschik, P., Ebrahimkutty, M. P., Grill, D., Lehrich, J., Hentschel, A., Biasizzo, M., Mengoni, S., Ahrends, R., Gerke, V., Cojocaru, V., Klingauf, J., & Galic, M. (2019). ArhGEF37 assists dynamin 2 during clathrin-mediated endocytosis. *Journal of Cell Science*, *132*(9), jcs226530. https://doi.org/10.1242/jcs.226530

Wilkins, A. S. (2001). *The evolution of developmental pathways*. Oxford University Press http://global.oup.com/ushe/product/the-evolution-of-developmental-pathways-9780878939169

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Matlosz, S., Sigurgeirsson, B., Franzdóttir, S. R., Pálsson, A., & Jónsson, Z. O. (2022). DNA methylation differences during development distinguish sympatric morphs of Arctic charr (*Salvelinus alpinus*). *Molecular Ecology*, *00*, 1–23. https://doi.org/10.1111/mec.16620

Supplementary Data

Figures:

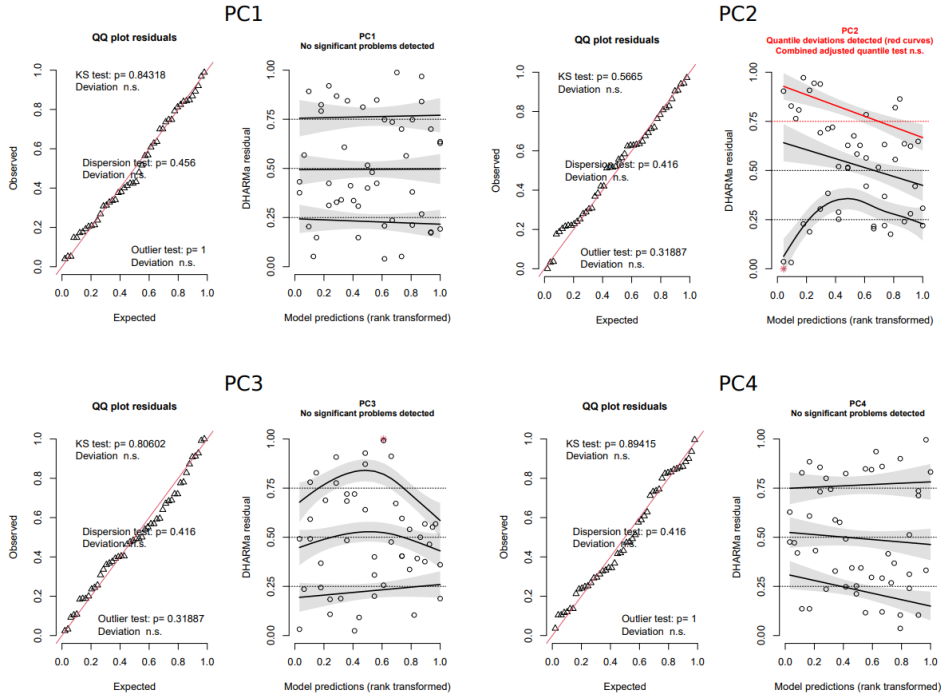


Figure S1: Results of validation analyses of linear models for the first 4 PCs (see labels), using the DHARMa package. For each of the PC's two plots are depicted, on the left a QQ plot with the results of three tests for deviation from normality (with assessment of deviation with associated significance) and on the right plot of residuals against the predicted values from the model

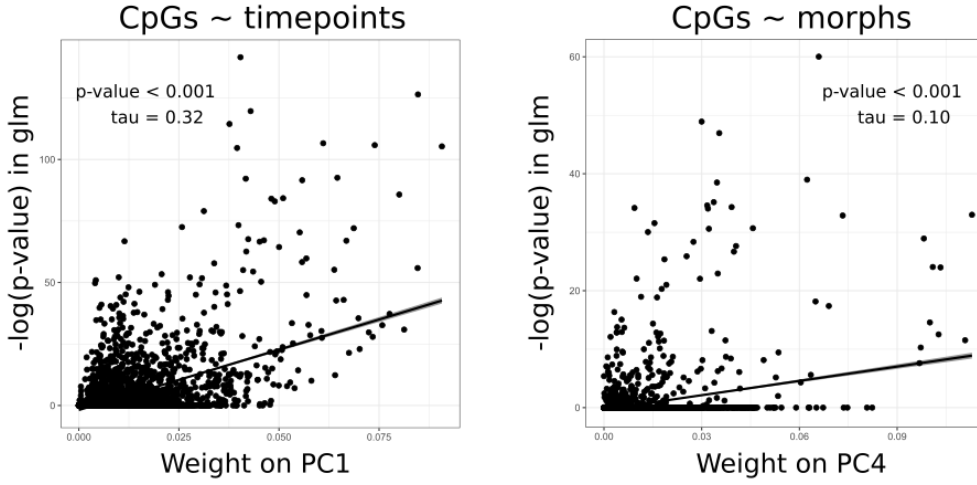
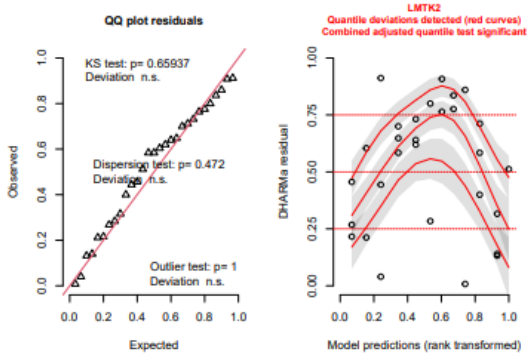
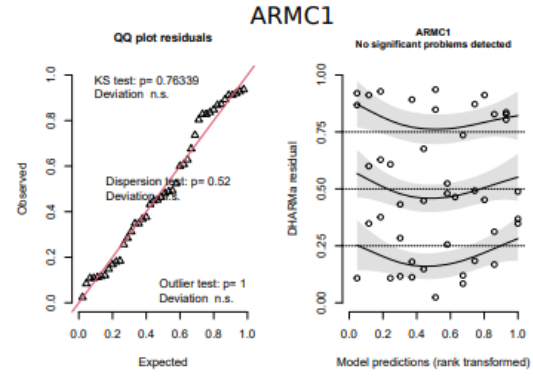


Figure S2: Comparisons between glm and PCA analyses. Contribution of each of the 10340 CpGs to timepoints and morphs. For each CpG, the absolute weight to specific PCs is plotted against the –log(p-value) in the glm. The regression lines were obtained with kendall correlation. Kendall test results (p-values and tau) are displayed on the graphs.

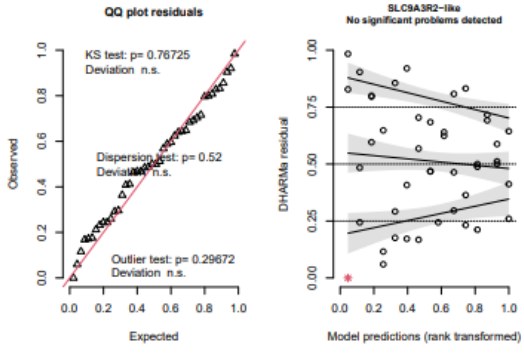
LMTK2



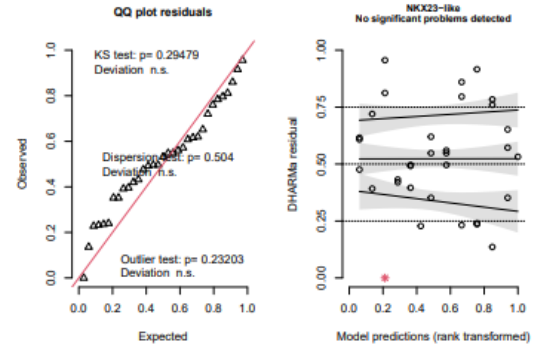
ARMC1



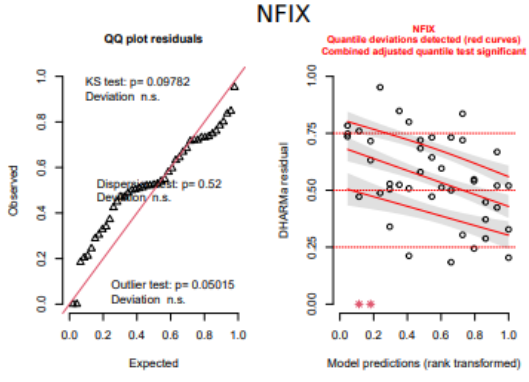
SLC9A3R2-like



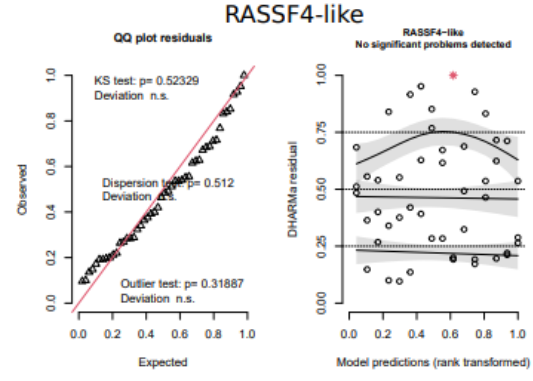
NKX23-like



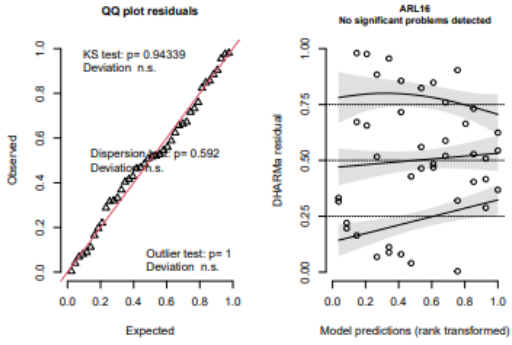
NFIX



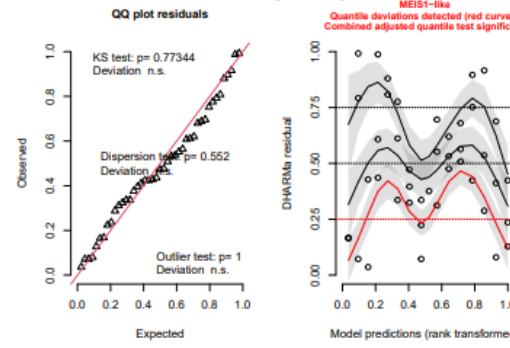
RASSF4-like



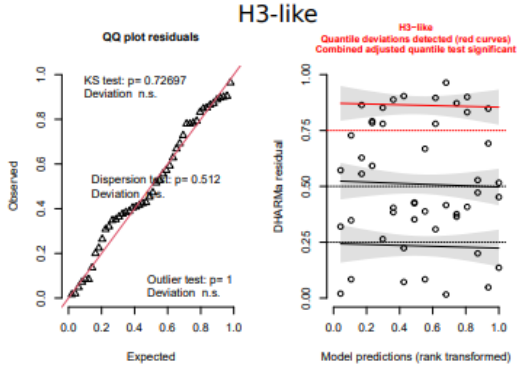
ARL16



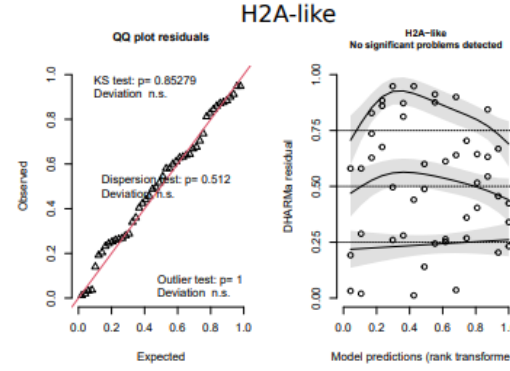
MEIS1-like



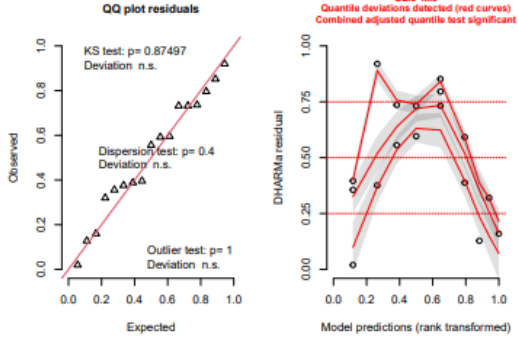
H3-like



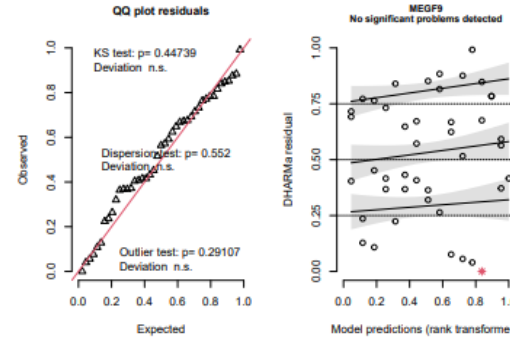
H2A-like



GLI3-like



MEGF9



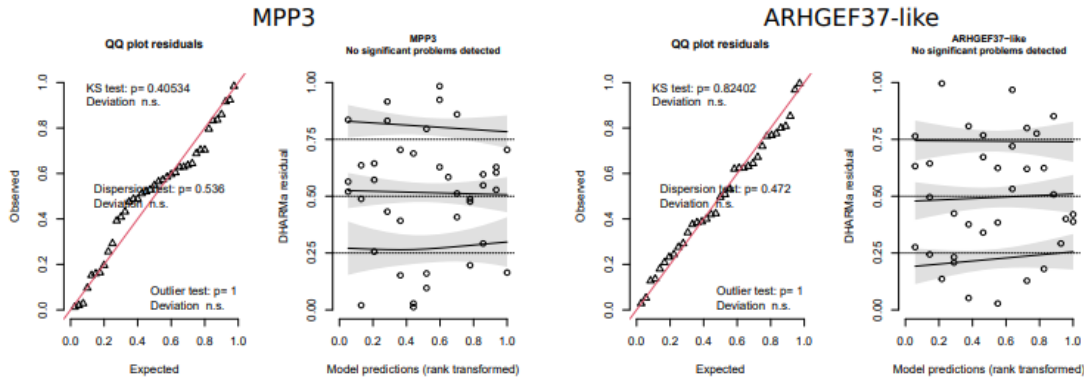
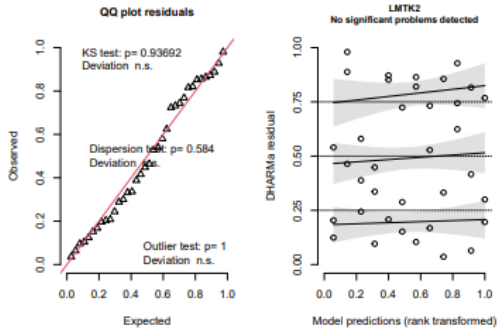
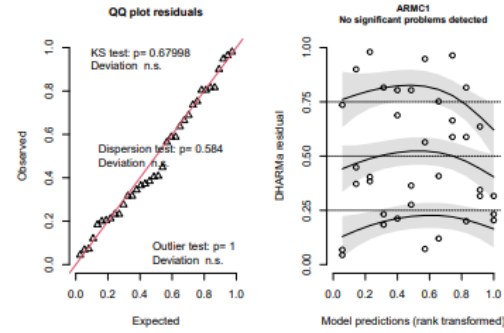


Figure S3: Results of validation analyses of linear models for methylation data on the 14 genes of interest (see labels), using the DHARMa package. For each of the genes two plots are depicted, on the left a QQ plot with the results of three tests for deviation from normality (with assessment of deviation with associated significance) and on the right plot of residuals against the predicted values from the model.

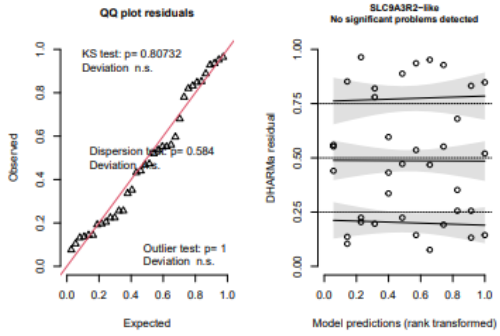
LMTK2



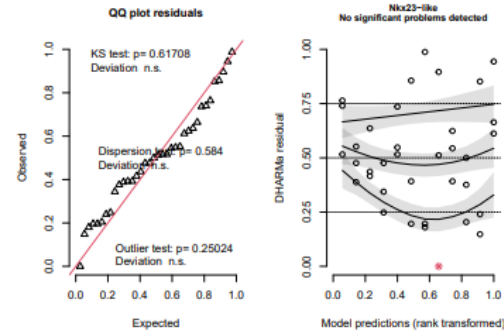
ARMC1



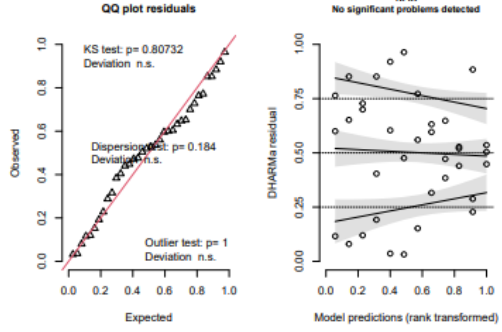
SLC9A3R2-like



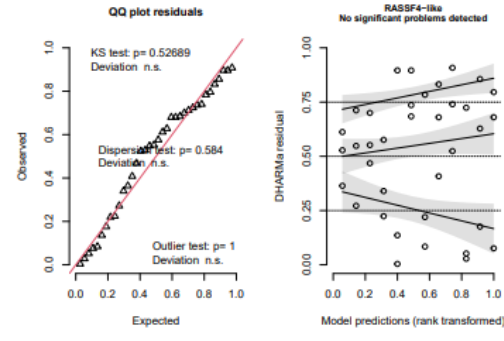
NKX23-like



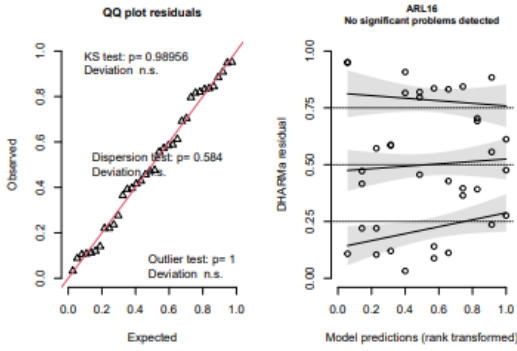
NFIX



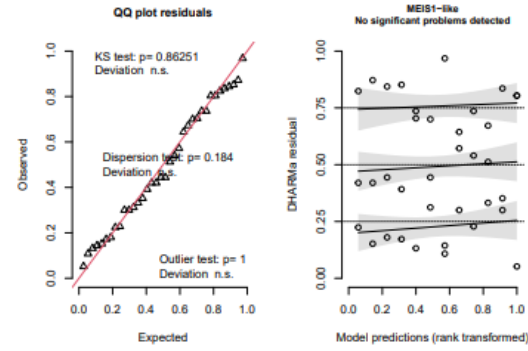
RASSF4-like



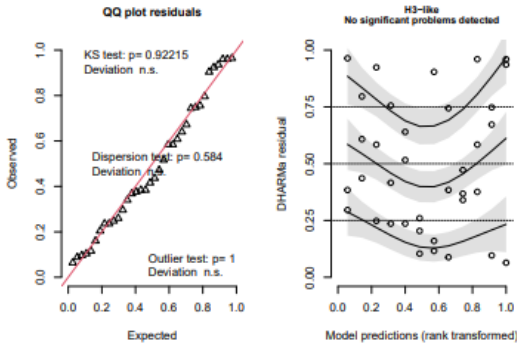
ARL16



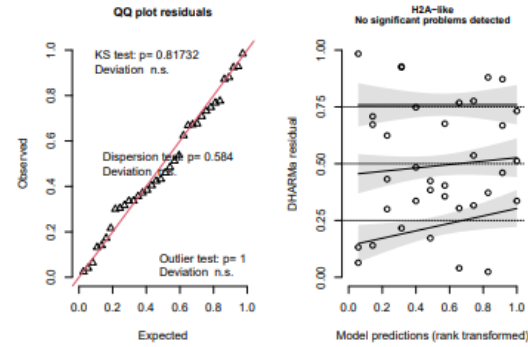
MEIS1-like



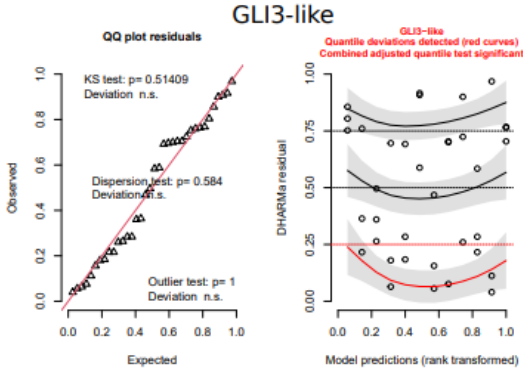
H3-like



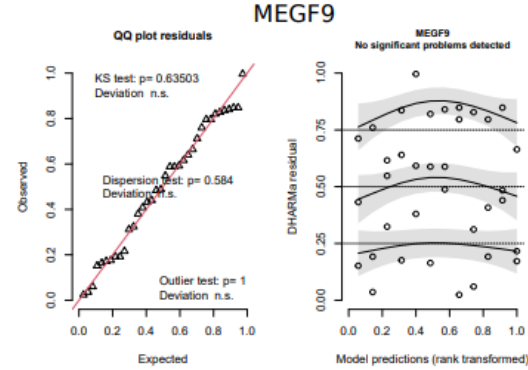
H2A-like



GLI3-like



MEGF9



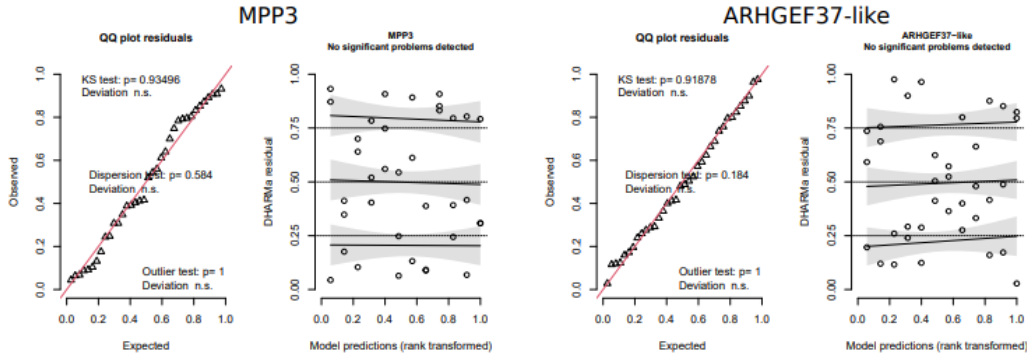


Figure S4: Results of validation analyses of linear models for expression data on the 14 genes of interest (see labels), using the DHARMa package. For each of the genes two plots are depicted, on the left a QQ plot with the results of three tests for deviation from normality (with assessment of deviation with associated significance) and on the right plot of residuals against the predicted values from the model.

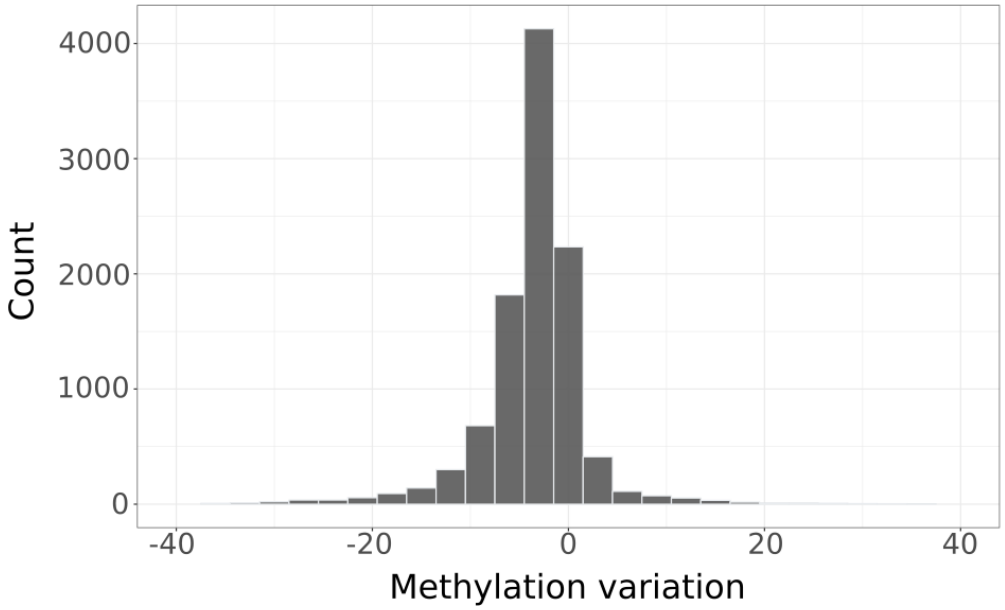


Figure S5: Changes in methylation of 10340 CpGs between 50 and 200 ts. Most of the cytosines show a decrease in methylation over this period, with the mean methylation change being -3.6%.

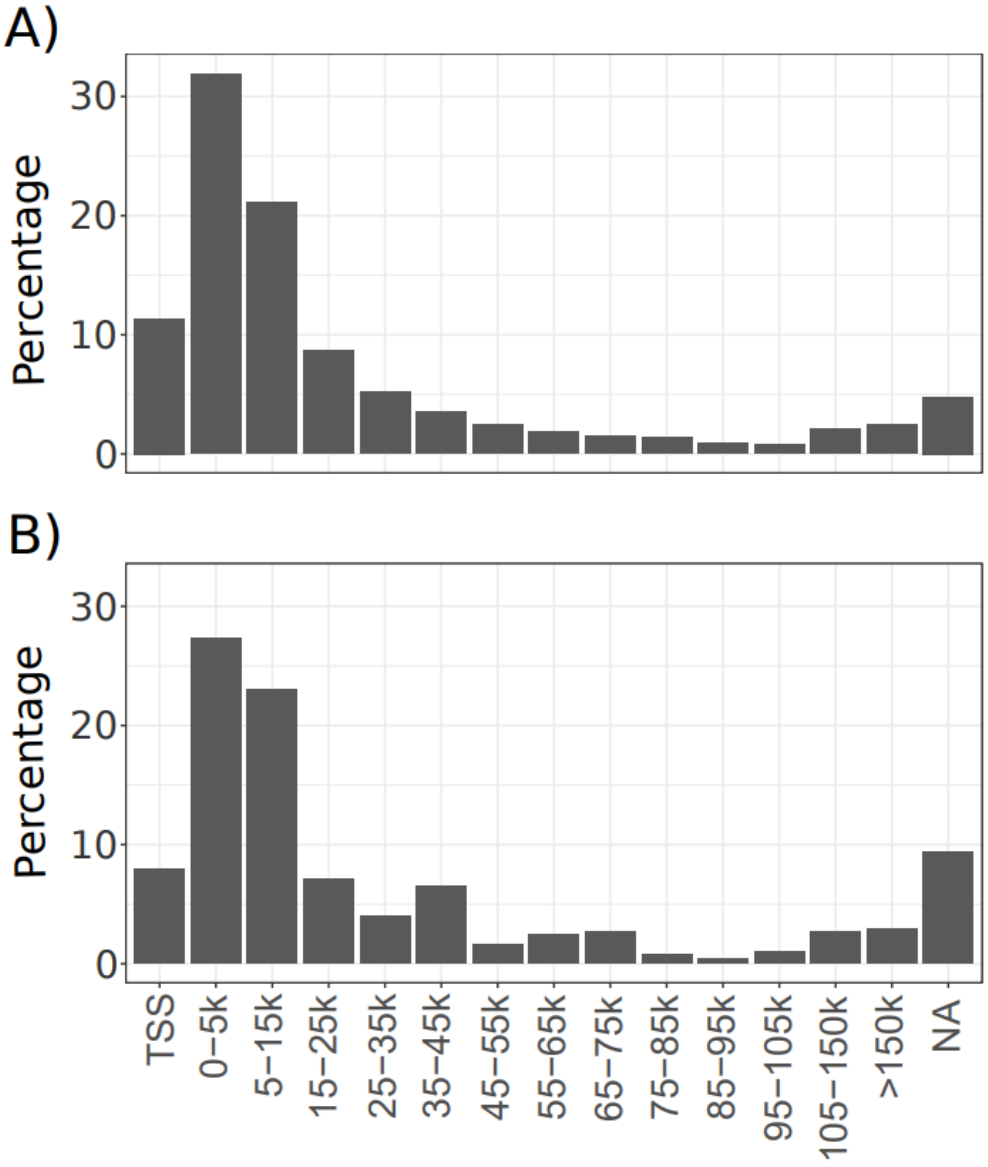


Figure S6: Similar distribution of DMRs and non-DMRs in relation to transcription start sites. Distribution of methylated and differentially methylated regions in the genome, represented as percentage of regions in each category, binned by their distance to the closest transcription start site (in bp). A) For the 18469 non-DMRs. B) For the 479 DMRs. It appears that differences in methylation arise in CpG islands regardless of their distance to gene bodies.

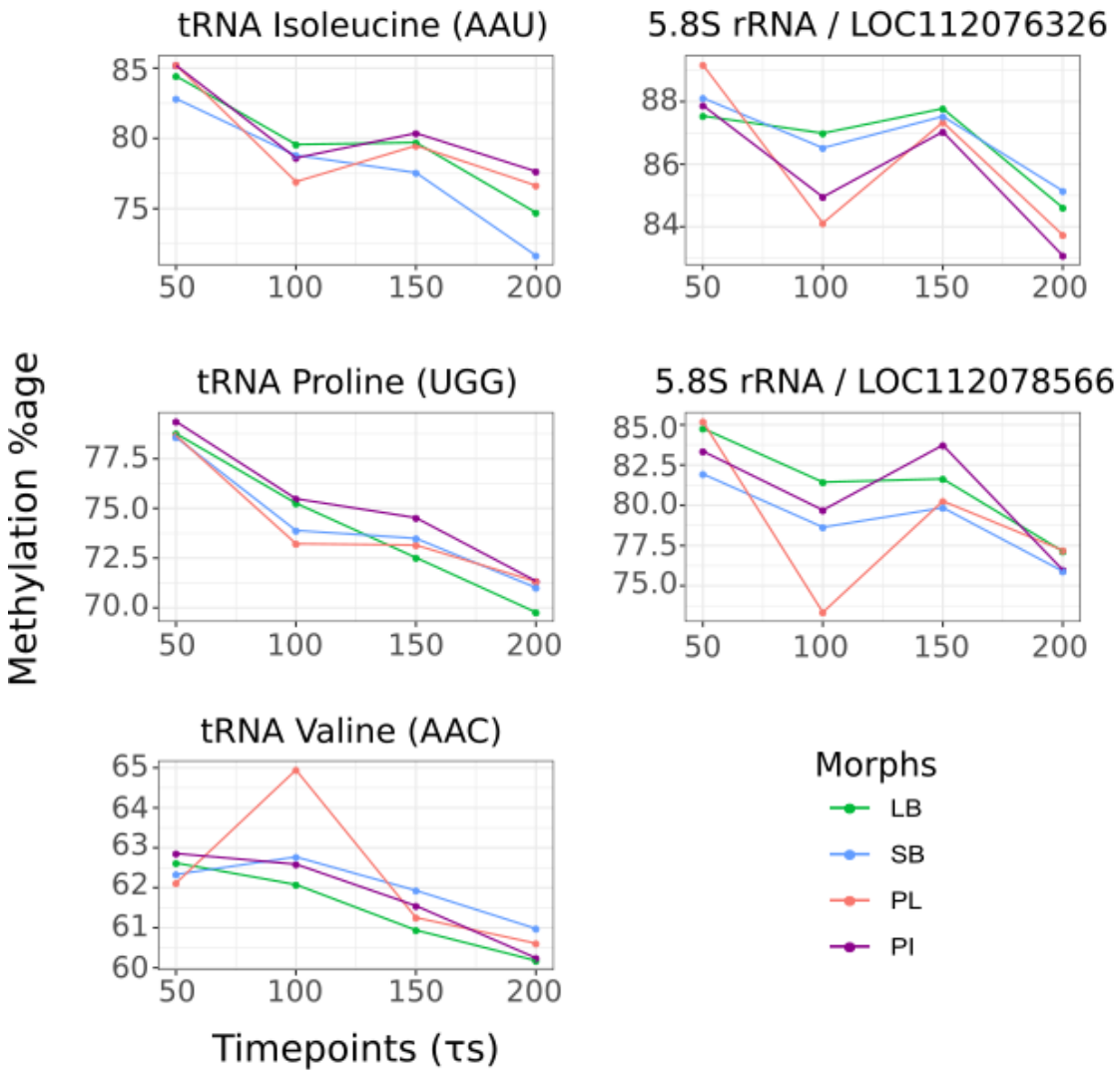
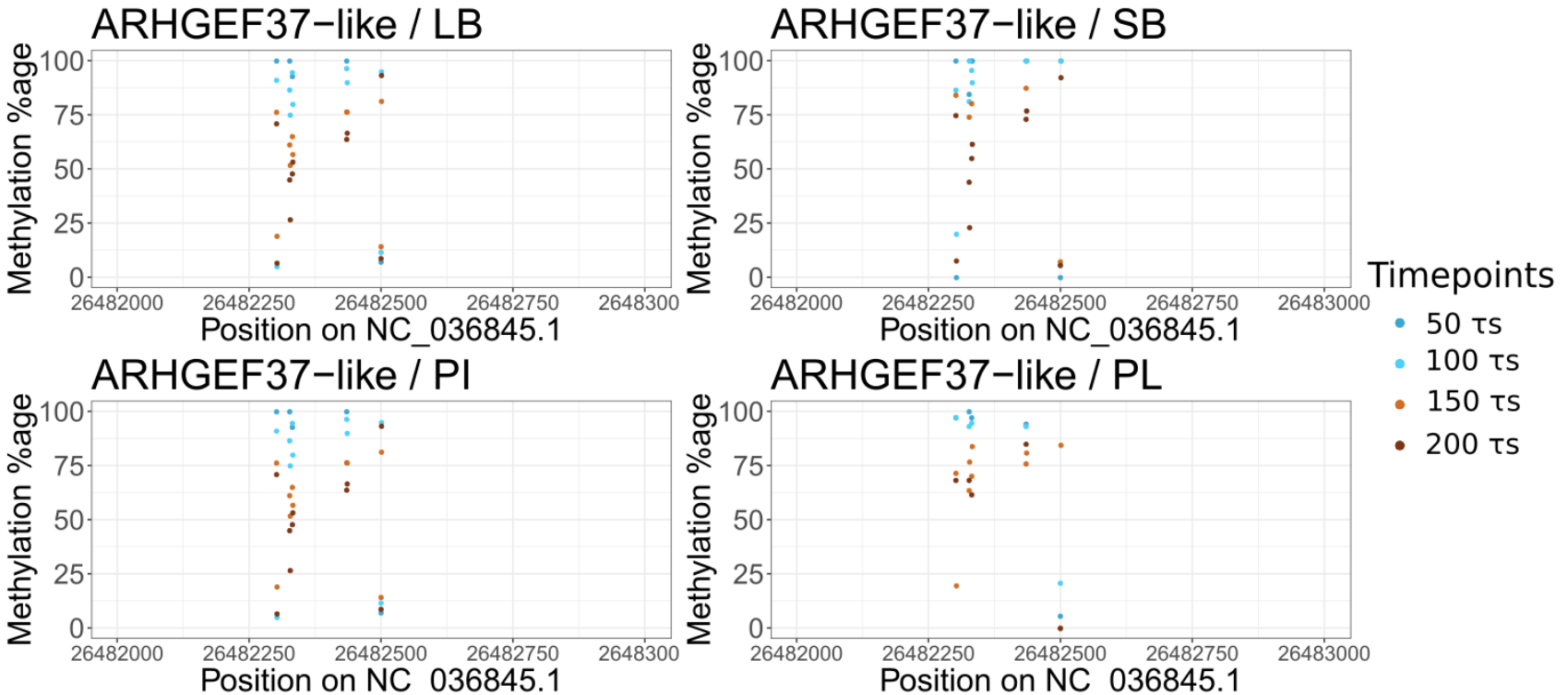
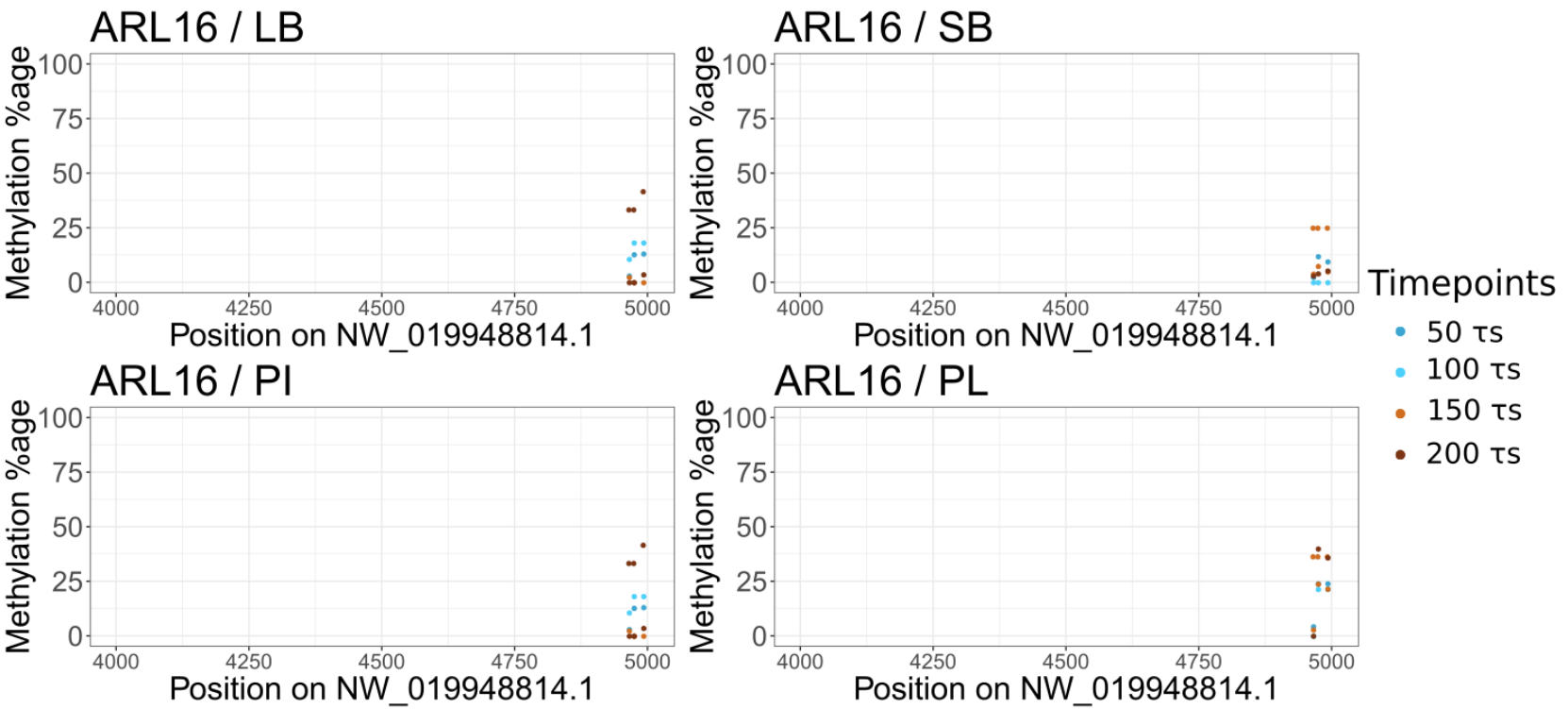
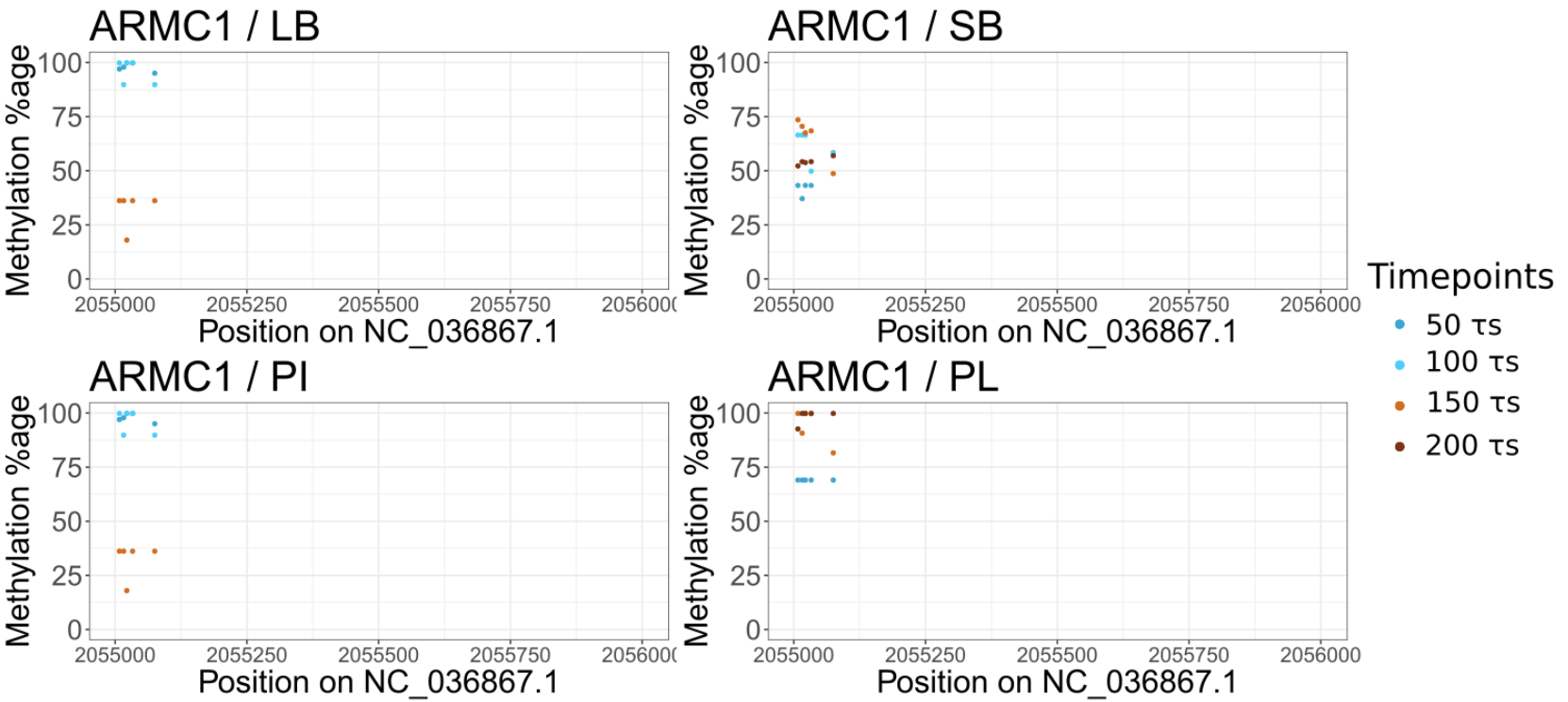
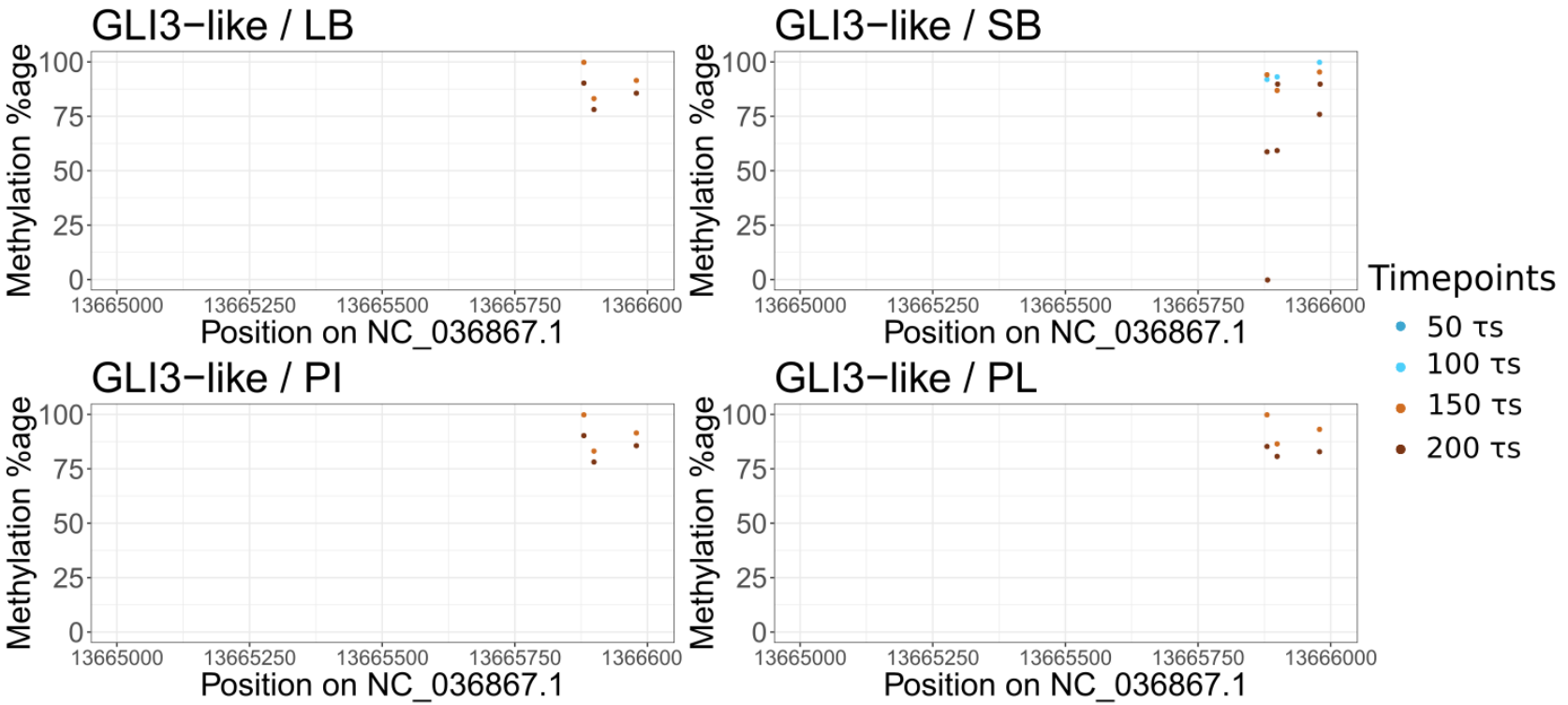


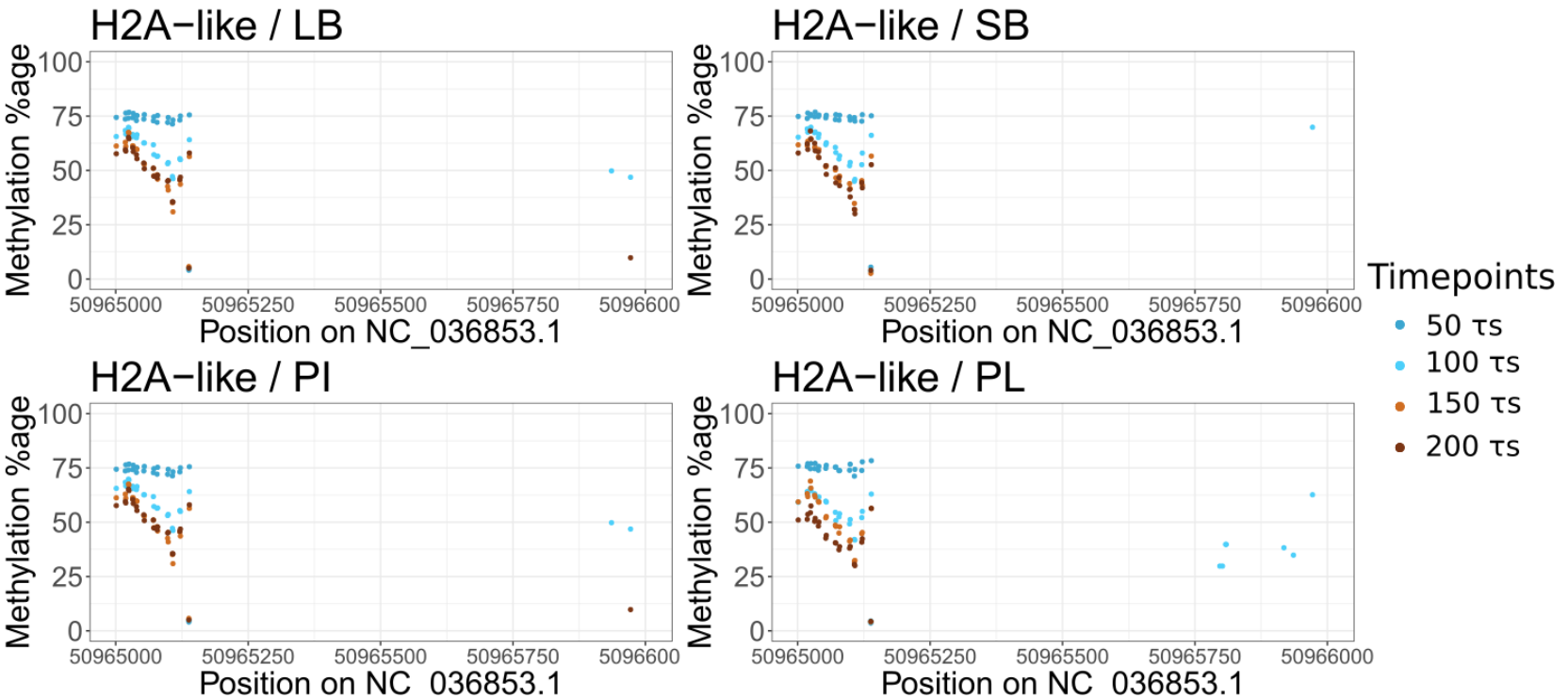
Figure S7: Average methylation of five DMRs located close to tRNA and rRNA sequences for each morph and timepoint.

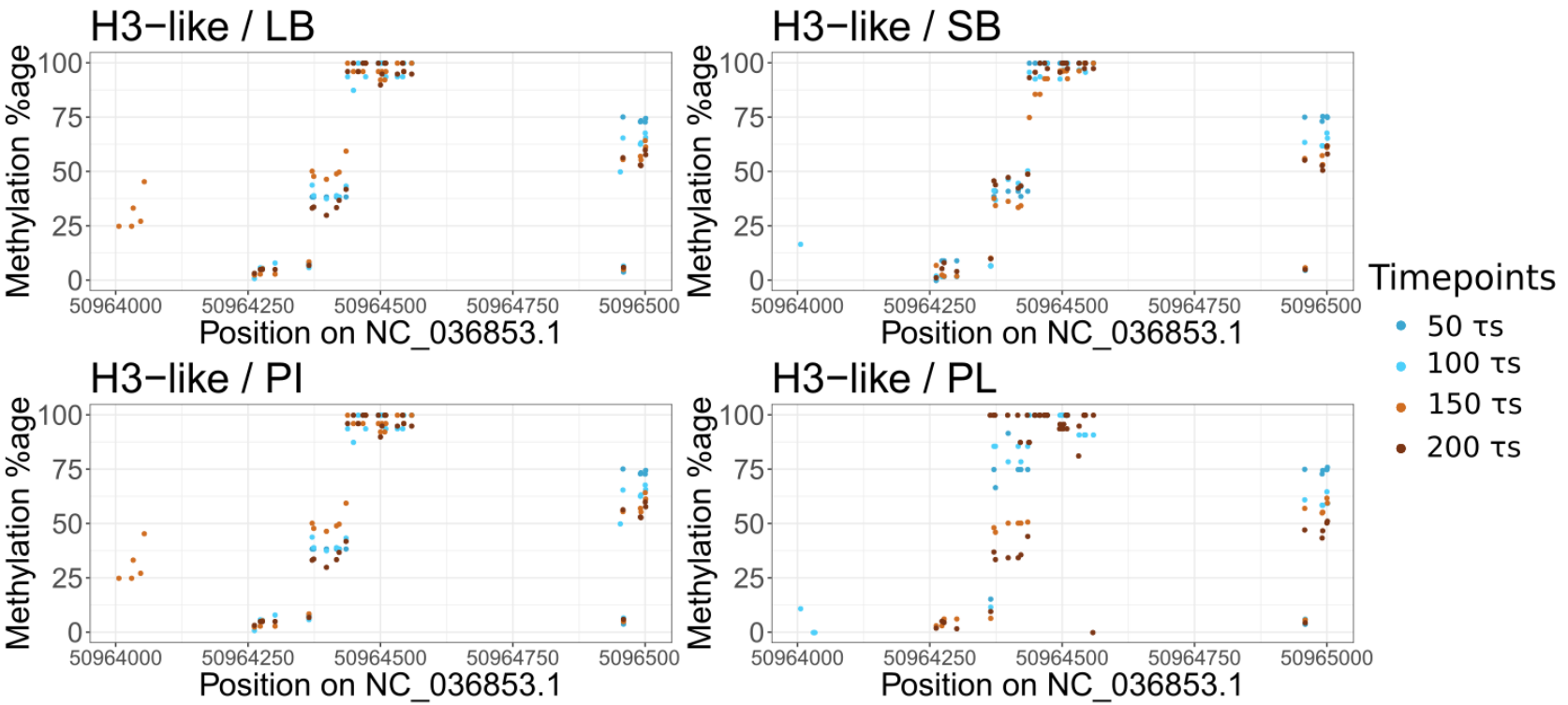

ARHGEF37−like / LB
ARHGEF37−like / SB
ARHGEF37−like / PI
ARHGEF37−like / PL
Methylation %age
Position on NC_036845.1
26482000
26482250
26482500
26482750
2648300
26483000
0
25
50
75
100
Timepoints
50 τs
100 τs
150 τs
200 τs

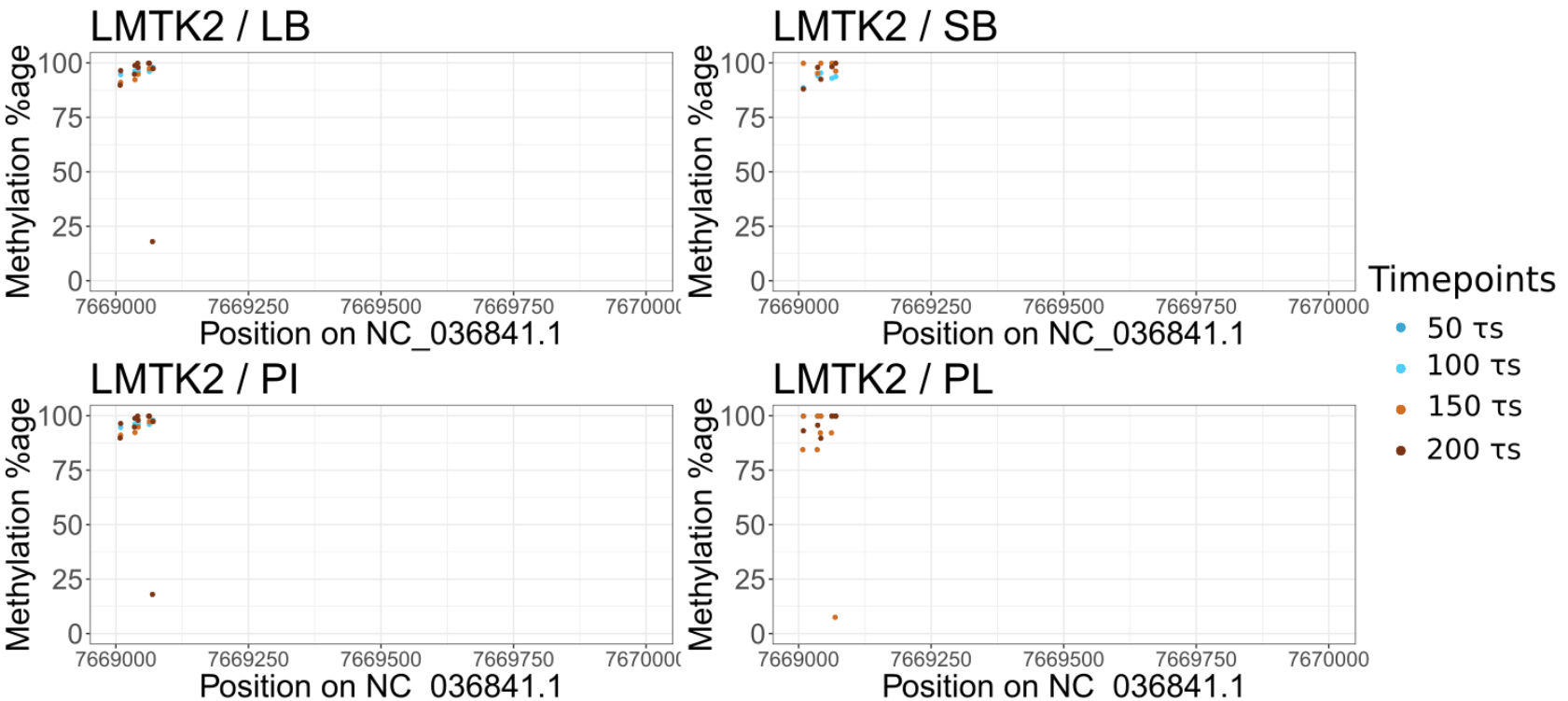

ARL16 / LB
ARL16 / SB
ARL16 / PI
ARL16 / PL
Methylation %age
100
75
50
25
0
4000
4250
4500
4750
5000
Position on NW_019948814.1
Timepoints
50 τs
100 τs
150 τs
200 τs

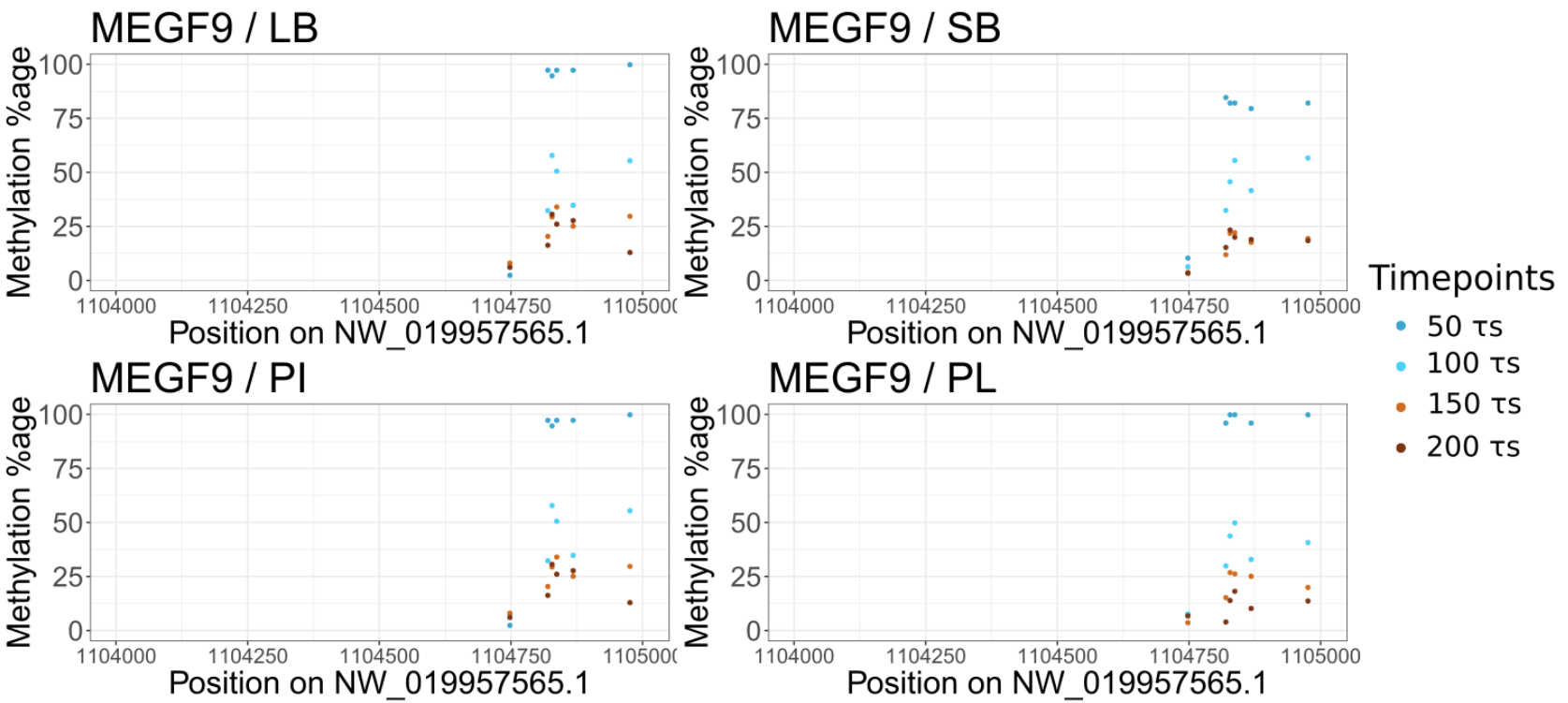

ARMC1 / LB
ARMC1 / SB
ARMC1 / PI
ARMC1 / PL
Methylation %age
Position on NC_036867.1
2055000
2055250
2055500
2055750
2056000
0
25
50
75
100
Timepoints
50 τs
100 τs
150 τs
200 τs

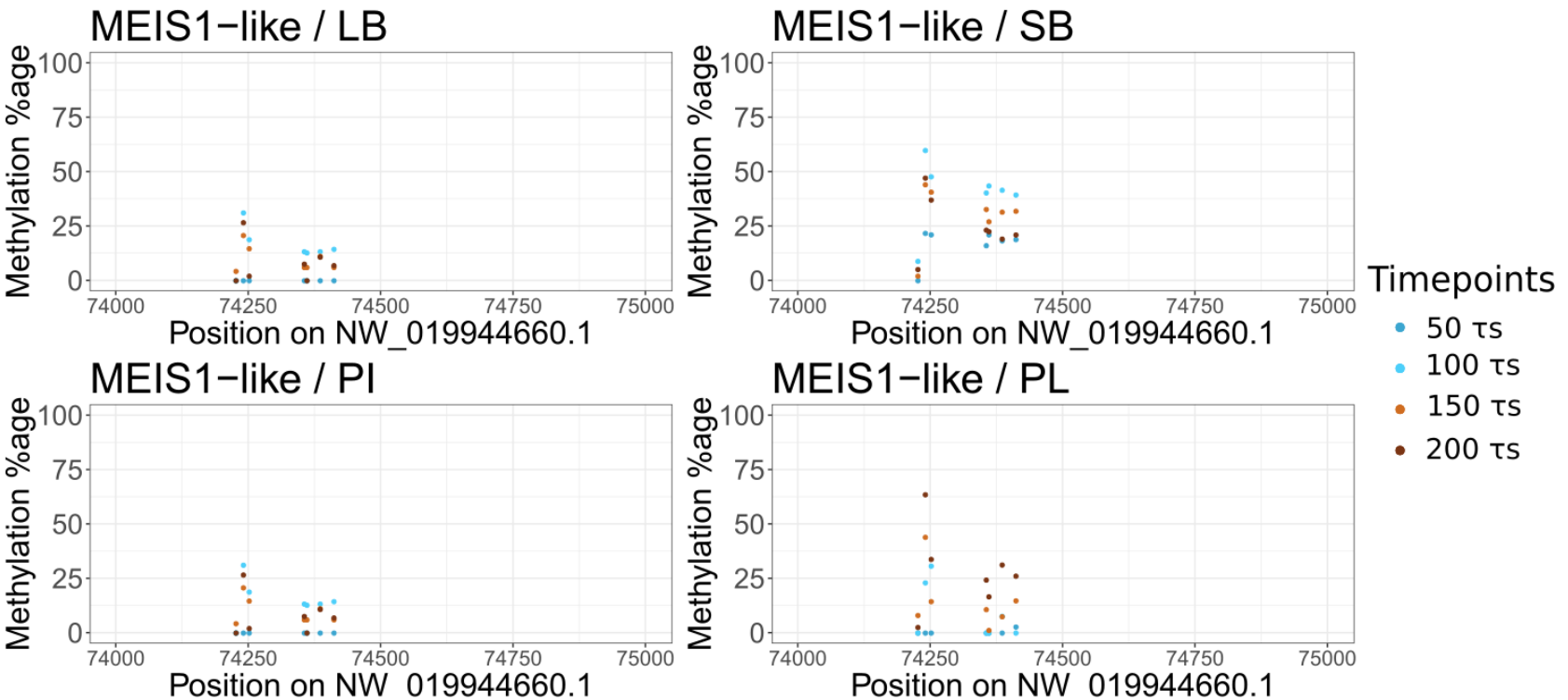

GLI3−like / LB
GLI3−like / SB
GLI3−like / PI
GLI3−like / PL
Methylation %age
Position on NC_036867.1
13665000
13665250
13665500
13665750
13666000
0
25
50
75
100
Timepoints
50 τs
100 τs
150 τs
200 τs

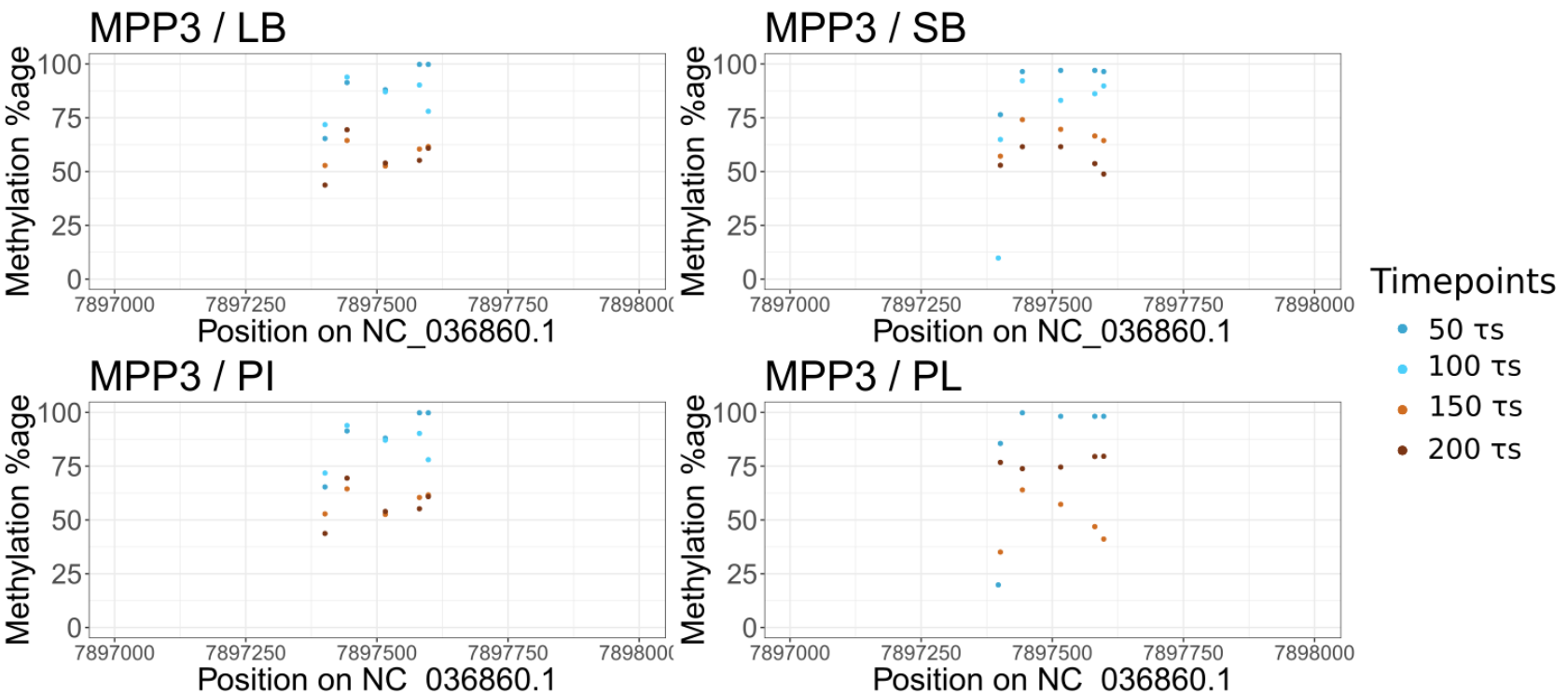

H2A−like / LB
H2A−like / SB
H2A−like / PI
H2A−like / PL
Methylation %age
100
75
50
25
0
50965000
50965250
50965500
50965750
5096600
50966000
Position on NC_036853.1
Timepoints
50 τs
100 τs
150 τs
200 τs

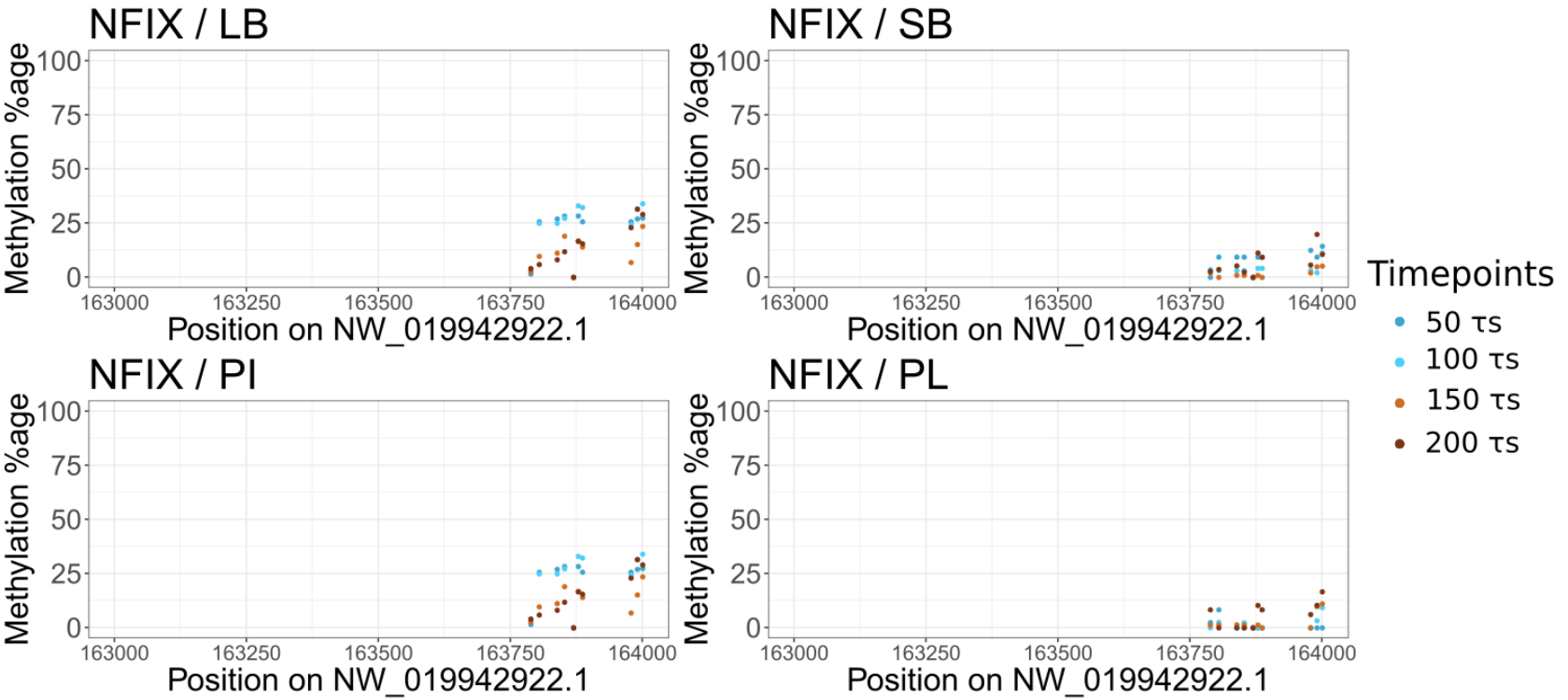

H3−like / LB
H3−like / SB
H3−like / PI
H3−like / PL
Methylation %age
Position on NC_036853.1
0
25
50
75
100
50964000
50964250
50964500
50964750
5096500
50965000
Timepoints
50 τs
100 τs
150 τs
200 τs

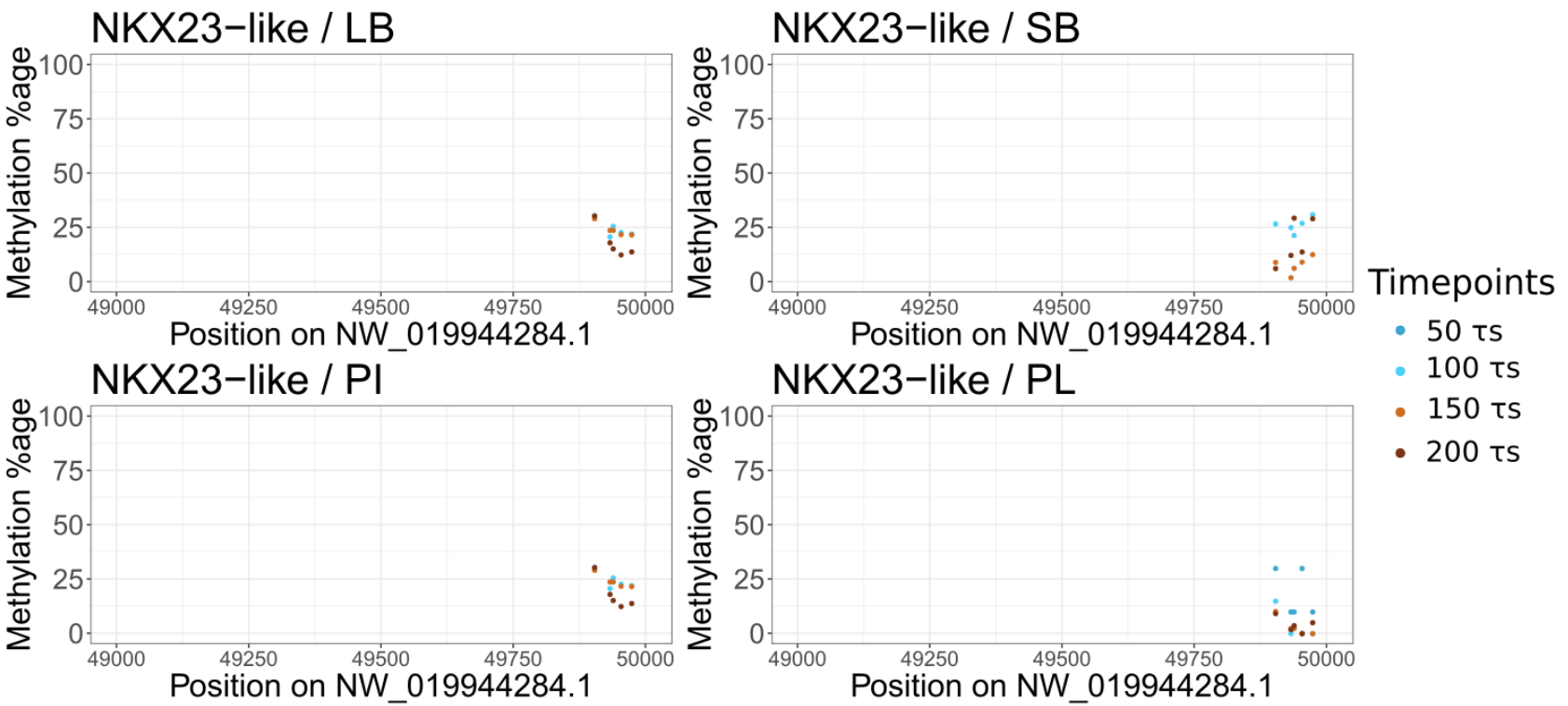

LMTK2 / LB
LMTK2 / SB
LMTK2 / PI
LMTK2 / PL
Methylation %age
100
75
50
25
0
7669000
7669250
7669500
7669750
7670000
Position on NC_036841.1
Timepoints
50 τs
100 τs
150 τs
200 τs

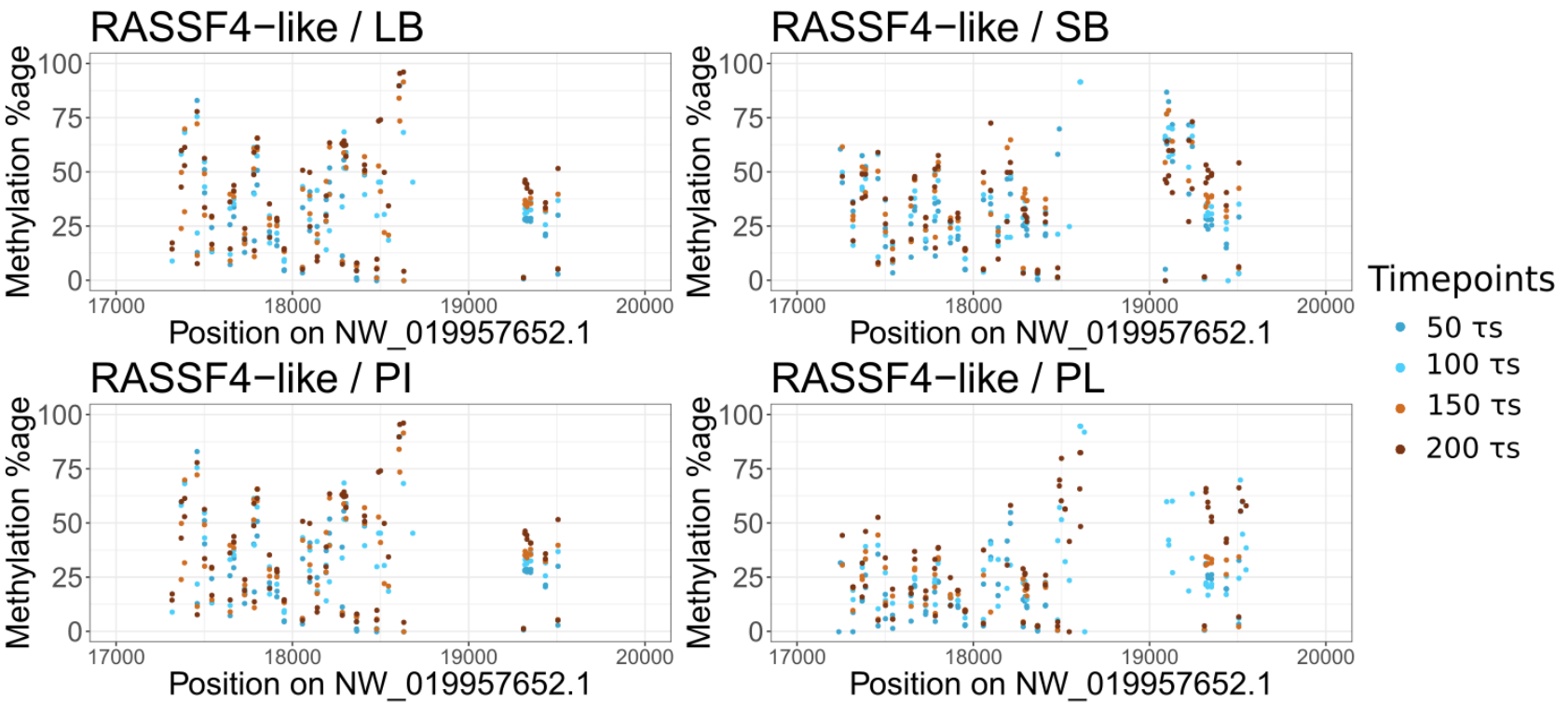

MEGF9 / LB
MEGF9 / SB
MEGF9 / PI
MEGF9 / PL
Methylation %age
Position on NW_019957565.1
1104000
1104250
1104500
1104750
1105000
0
25
50
75
100
Timepoints
50 τs
100 τs
150 τs
200 τs

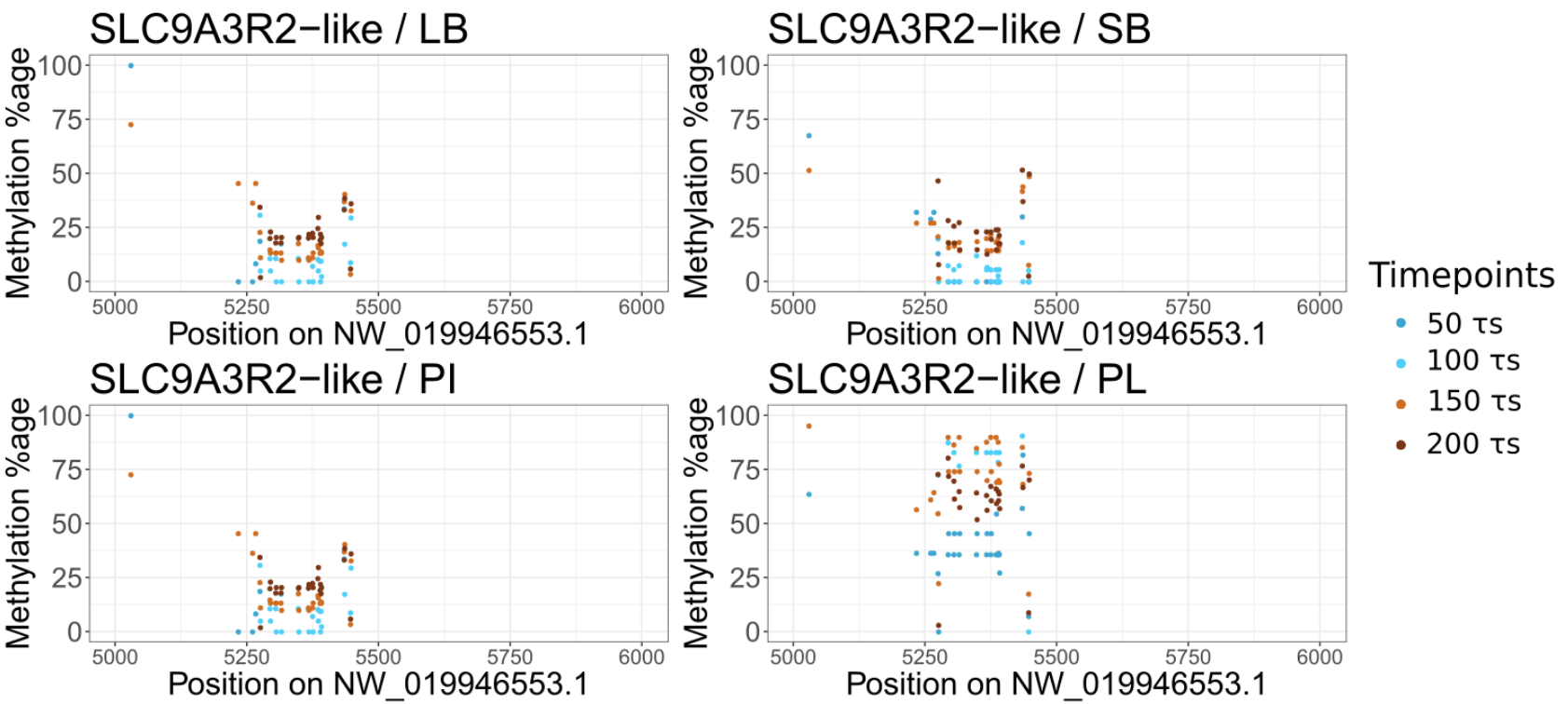

MEIS1−like / LB
MEIS1−like / SB
MEIS1−like / PI
MEIS1−like / PL
Methylation %age
100
75
50
25
0
74000
74250
74500
74750
75000
Position on NW_019944660.1
Timepoints
50 τs
100 τs
150 τs
200 τs


MPP3 / LB
MPP3 / SB
MPP3 / PI
MPP3 / PL
Methylation %age
100
75
50
25
0
7897000
7897250
7897500
7897750
7898000
Position on NC_036860.1
Timepoints
50 τs
100 τs
150 τs
200 τs


NFIX / LB
NFIX / SB
NFIX / PI
NFIX / PL
Methylation %age
100
75
50
25
0
163000
163250
163500
163750
164000
Position on NW_019942922.1
Timepoints
50 τs
100 τs
150 τs
200 τs


NKX23−like / LB
NKX23−like / SB
NKX23−like / PI
NKX23−like / PL
Methylation %age
100
75
50
25
0
49000
49250
49500
49750
50000
Position on NW_019944284.1
Timepoints
50 τs
100 τs
150 τs
200 τs


RASSF4−like / LB
RASSF4−like / SB
RASSF4−like / PI
RASSF4−like / PL
Methylation %age
Position on NW_019957652.1
17000
18000
19000
20000
0
25
50
75
100
Timepoints
50 τs
100 τs
150 τs
200 τs


SLC9A3R2−like / LB
SLC9A3R2−like / SB
SLC9A3R2−like / PI
SLC9A3R2−like / PL
Methylation %age
Position on NW_019946553.1
0
25
50
75
100
5000
5250
5500
5750
6000
Timepoints
50 τs
100 τs
150 τs
200 τs

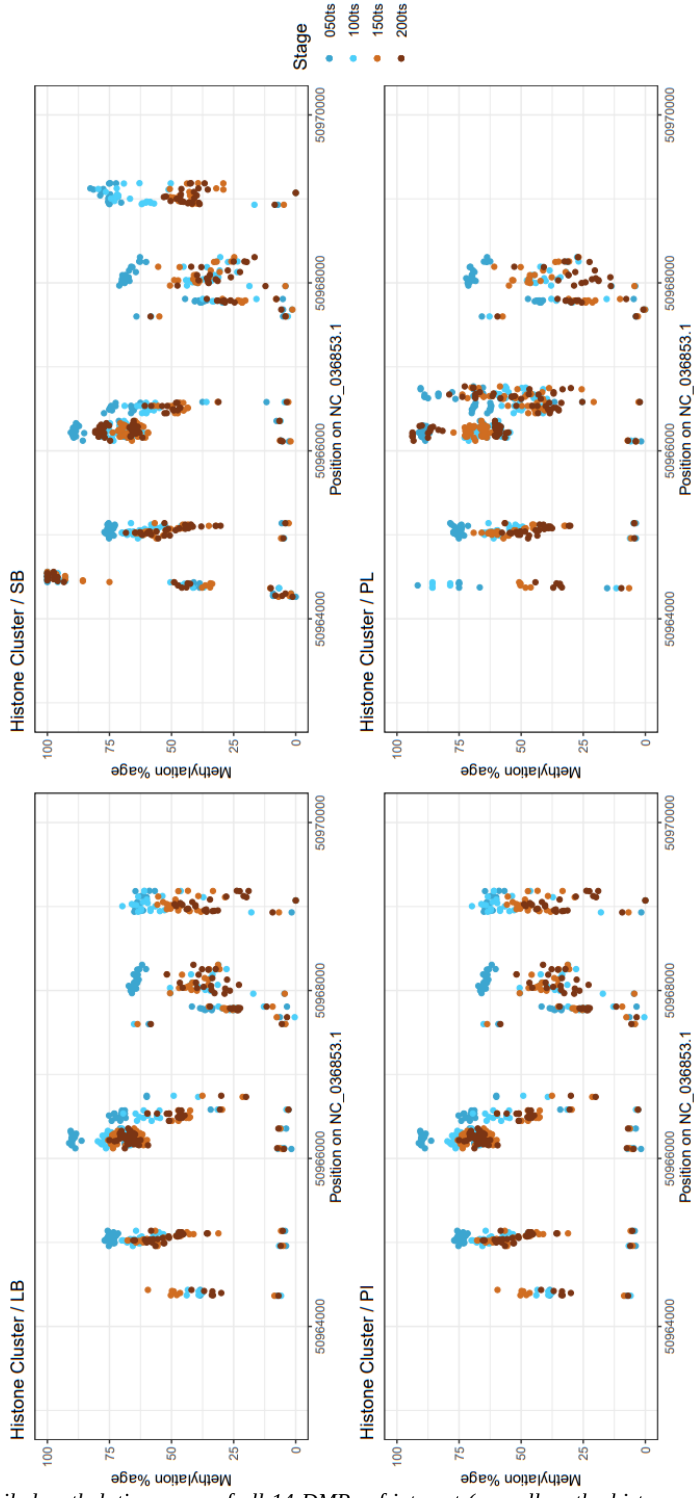


Figure S8: Detailed methylation maps of all 14 DMRs of interest (as well as the histone gene clusters), for each morph and timepoint. Residue-level methylation profiles of regions close to ARHGEF37-like, ARL16, ARMC1, GLI3-like, H2A-like, H3-like, LMTK2, MEGF9, MEIS1-like, MPP3, NKX23-like, NFIX, RASSF4-like, SLC9A3R2-like and the histone gene clusters.

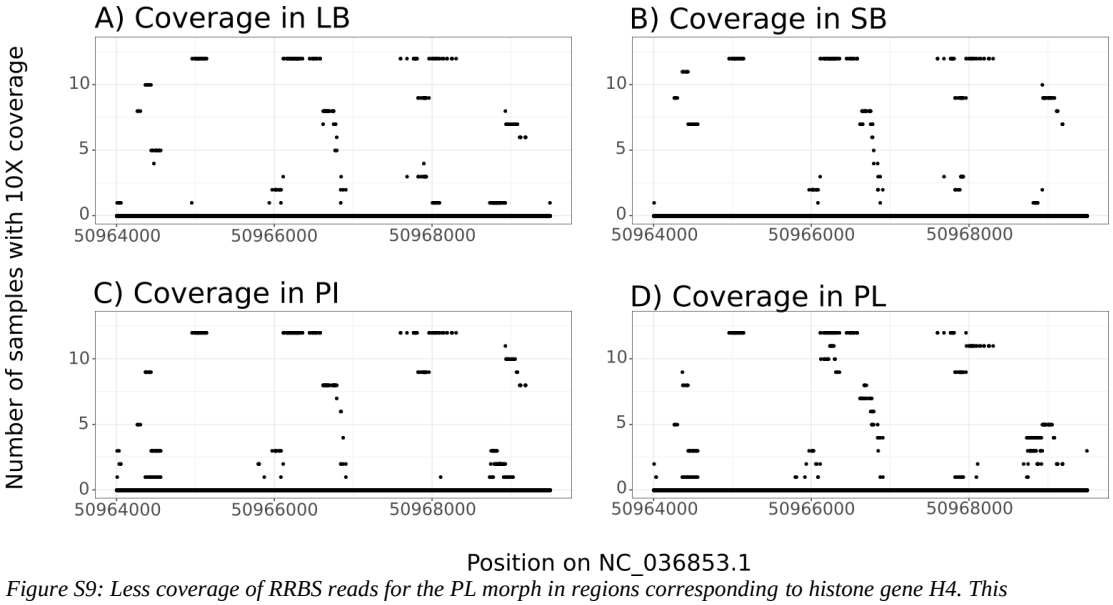


Figure S9: Less coverage of RRBS reads for the PL morph in regions corresponding to histone gene H4. This graph represents the number of samples that have more than 10X coverage for each cytosine in the histone gene cluster. The maximum number of samples is 12 for each morph. H4 is located to the far right of the graph (for reference see Figure 8), and the coverage for the PL morph in this region is less than for other morphs.

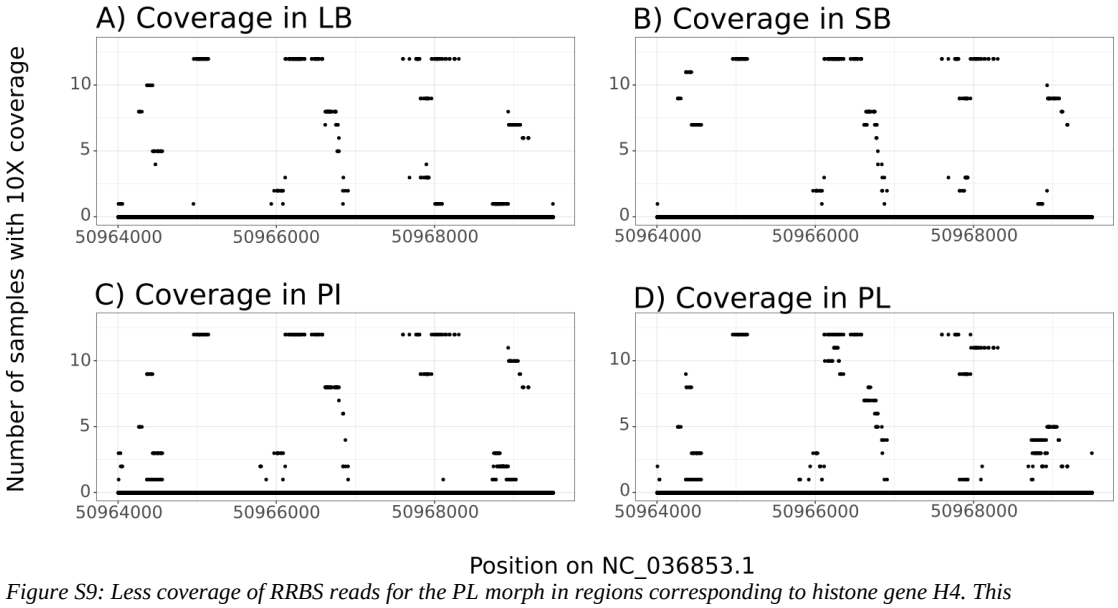


Figure S9: Less coverage of RRBS reads for the PL morph in regions corresponding to histone gene H4. This graph represents the number of samples that have more than 10X coverage for each cytosine in the histone gene cluster. The maximum number of samples is 12 for each morph. H4 is located to the far right of the graph (for reference see Figure 8), and the coverage for the PL morph in this region is less than for other morphs.

Tables:

Table S1: Samples of the embryos used for DNA (RRBS) and RNA extraction (qPCR), grouped by morph, time, family and sex. Library 5 is the one creating technical effects driving separation in PC2.*

Use	Morph	Timepoint	Family	Sample_ID	Sequencing Library	Sex
RRBS	PL	200	17PL30	S1	1	F
RRBS	PL	200	17PL53	S2	1	M
RRBS	PL	200	17PL54	S3	1	M
RRBS	LB	200	17LB1	S4	1	M
RRBS	LB	200	17LB2	S5	1	F
RRBS	LB	200	17LB3	S6	1	M
RRBS	PI	200	17PI1	S7	2	M
RRBS	PI	200	17PI3	S8	2	M
RRBS	PI	200	17PI5	S9	2	M
RRBS	SB	200	17SB1	S10	2	M
RRBS	SB	200	17SB3	S11	2	F
RRBS	SB	200	17SB35	S12	2	F
RRBS	PL	150	17PL30	S13	3	F
RRBS	PL	150	17PL53	S14	3	NA
RRBS	PI	150	17PI5	S15	3	F
RRBS	LB	150	17LB1	S16	3	M
RRBS	LB	150	17LB2	S17	3	M
RRBS	SB	150	17SB35	S18	3	F
RRBS	PI	150	17PI1	S19	4	M
RRBS	PI	150	17PI3	S20	4	F
RRBS	PL	150	17PL54	S21	4	M
RRBS	SB	150	17SB1	S22	4	F
RRBS	SB	150	17SB3	S23	4	M
RRBS	LB	150	17LB3	S24	4	F
RRBS	PL	100	17PL30	S25	5*	F
RRBS	PL	100	17PL34	S26	5*	F
RRBS	PI	100	17PI5	S27	5*	M
RRBS	LB	100	17LB1	S28	5*	F
RRBS	LB	100	17LB2	S29	5*	M
RRBS	SB	100	17SB35	S30	5*	M
RRBS	PI	100	17PI1	S31	6	M
RRBS	PI	100	17PI3	S32	6	F
RRBS	PL	100	17PL54	S33	6	F
RRBS	SB	100	17SB3	S34	6	M
RRBS	SB	100	17SB20	S35	6	NA
RRBS	LB	100	17LB3	S36	6	F
RRBS	PL	50	17PL23	S37	7	F
RRBS	PL	50	17PL25	S38	7	F
RRBS	PI	50	17PI5	S39	7	NA
RRBS	LB	50	17LB1	S40	7	NA
RRBS	LB	50	17LB2	S41	7	NA

RRBS	SB	50	17SB35	S42	7	NA
RRBS	PI	50	17PI1	S43	8	NA
RRBS	PI	50	17PI3	S44	8	NA
RRBS	PL	50	17PL11	S45	8	F
RRBS	SB	50	17SB1	S46	8	NA
RRBS	SB	50	17SB3	S47	8	NA
RRBS	LB	50	17LB3	S48	8	NA
qPCR	PL	200	17PL30	S49	-	-
qPCR	PL	200	17PL23	S50	-	-
qPCR	PL	200	17PL32	S51	-	-
qPCR	LB	200	17LB1	S52	-	-
qPCR	LB	200	17LB2	S53	-	-
qPCR	LB	200	17LB3	S54	-	-
qPCR	PI	200	17PI1	S55	-	-
qPCR	PI	200	17PI3	S56	-	-
qPCR	PI	200	17PI5	S57	-	-
qPCR	SB	200	17SB1	S58	-	-
qPCR	SB	200	17SB3	S59	-	-
qPCR	SB	200	17SB4	S60	-	-
qPCR	PL	150	17PL30	S61	-	-
qPCR	PL	150	17PL53	S62	-	-
qPCR	PI	150	17PI5	S63	-	-
qPCR	LB	150	17LB1	S64	-	-
qPCR	LB	150	17LB2	S65	-	-
qPCR	SB	150	17SB4	S66	-	-
qPCR	PI	150	17PI1	S67	-	-
qPCR	PI	150	17PI3	S68	-	-
qPCR	PL	150	17PL54	S69	-	-
qPCR	SB	150	17SB1	S70	-	-
qPCR	SB	150	17SB3	S71	-	-
qPCR	LB	150	17LB3	S72	-	-
qPCR	PL	100	17PL32	S73	-	-
qPCR	PL	100	17PL34	S74	-	-
qPCR	PI	100	17PI5	S75	-	-
qPCR	LB	100	17LB1	S76	-	-
qPCR	LB	100	17LB2	S77	-	-
qPCR	SB	100	17SB35	S78	-	-
qPCR	PI	100	17PI1	S79	-	-
qPCR	PI	100	17PI3	S80	-	-
qPCR	PL	100	17PL26	S81	-	-
qPCR	SB	100	17SB3	S82	-	-
qPCR	SB	100	17SB20	S83	-	-
qPCR	LB	100	17LB3	S84	-	-

Table S2: List of pairwise comparisons for DMR analysis.

Pairwise comparisons DMR analysis	**Number of samples**
Benthic VS Limnetic	24 VS 24
Benthic VS Limnetic at 50ts	6 VS 6
Benthic VS Limnetic at 100ts	6 VS 6
Benthic VS Limnetic at 150ts	6 VS 6
Benthic VS Limnetic at 200ts	6 VS 6
SB VS PL	12 VS 12
SB VS PL at 50ts	3 VS 3
SB VS PL at 100ts	3 VS 3
SB VS PL at 150ts	3 VS 3
SB VS PL at 200ts	3 VS 3
SB VS PI	12 VS 12
SB VS PI at 50ts	3 VS 3
SB VS PI at 100ts	3 VS 3
SB VS PI at 150ts	3 VS 3
SB VS PI at 200ts	3 VS 3
SB VS LB	12 VS 12
SB VS LB at 50ts	3 VS 3
SB VS LB at 100ts	3 VS 3
SB VS LB at 150ts	3 VS 3
SB VS LB at 200ts	3 VS 3
LB VS PI	12 VS 12
LB VS PI at 50ts	3 VS 3
LB VS PI at 100ts	3 VS 3
LB VS PI at 150ts	3 VS 3
LB VS PI at 200ts	3 VS 3
LB VS PL	12 VS 12
LB VS PL at 50ts	3 VS 3
LB VS PL at 100ts	3 VS 3
LB VS PL at 150ts	3 VS 3
LB VS PL at 200ts	3 VS 3
PL VS PI	12 VS 12
PL VS PI at 50ts	3 VS 3
PL VS PI at 100ts	3 VS 3
PL VS PI at 150ts	3 VS 3
PL VS PI at 200ts	3 VS 3

Table S3: Primers used for qPCR and sexing.

Name	Sequence	Product size (bp)	Use	PCR Efficiency
actb-Fw	5-GAA GAT CAA GAT CAT CGC CC-3	122	qPCR (control)	1.95
actb-Rv	5-CAG ACT CGT CGT ACT CCT GCT-3			
Ub2l3-Fw	5-CGA GAA GGG ACA GGT GTG TC-3	96	qPCR (control)	1.93
Ub2l3-Rv	5-ACC AAC GCA ATC AGG GAC T-3			
HiH2a-F	5-GCA AAG ACA CGT TCA TCC CG-3	187	qPCR	1.99
HiH2a-R	5-TCT TCT TGT TGT CAC GGG CA-3			
Lmtk2-F	5-TTT TCG ACC AAG AGA CCC CC-3	110	qPCR	1.9
Lmtk2-R	5-CCT ATT CAG GTA GCT GGC CG-3			
Nkx23-F2	5-CCG GAG GAT CAG GAA AAC AA-3	156	qPCR	1.93
Nkx23-R2	5-CCG CTC TAA CTC GAA CAC CT-3			
SLC9A3R2-F2	5-GCA CAC ATG TCA AAG AGG TC-3	138	qPCR	1.95
SLC9A3R2-R2	5-CGT TTG TGA TTG GTG AGC CC-3			
ARL16-F1	5-CTA CCT ACG GTG GGC ACC AA-3	145	qPCR	2.49
ARL16-R1	5-CAT GGT ACA AGG GAG GTC CGA-3			
Meis1l-F2	5-CTG AAG AGG GCT GGC TTG AA-3	197	qPCR	2.2
Meis1l-R2	5-GTG GGG CAA GTC TTC GTA CC-3			
HiH3l-F2	5-CAG GAG GCT AGC GAG GCT TA-3	159	qPCR	2.17
HiH3l-R2	5-GGG GGA TTT TGG AGA TCA GGT C-3			
ARMC1-F1	5-TCT CTG CTT CTC ACA ACT GCT T-3	115	qPCR	1.92
ARMC1-R1	5-ATC AAA CCA GCC TCT TCA CAC T-3			
Gli3-F1	5-GCG GTT TTG GAG TTG GTG TC-3	156	qPCR	2.06
Gli3-R1	5-ACG TCA TGG AGA CAG TAC ACG-3			
RASSF4-F2	5-CGA CGA CAA CGA GAG ACT CC-3	143	qPCR	2.21
RASSF4-R2	5-CGT CGT CCT CCT CTA GAC TAC T-3			
NFIX-F1	5-TTT CCC TGC TCC ACA ATG GC-3	197	qPCR	2.17
NFIX-R1	5-GTG TAG GAG AAG GCA CGG AC-3			
MEGF9-F1	5-CCA CCA CTG ACA CGC AAC AAA C-3	191	qPCR	1.95
MEGF9-R1	5-GTC GCA GCT ACA GGG GAT G-3			
MAGUK-F2	5-ACC TAC GAG GAG GTG ACT CG-3	168	qPCR	2.02
MAGUK-R2	5-CTT CTT GGG CCT GGT GGT ATG-3			
Rhoguanin-F1	5-TGA GGC ACA GCT CTG TTA GTG-3	172	qPCR	1.92
Rhoguanin-R1	5-CTT TCC ATG CTG AGG CTG TC-3			
SdY-Fw	5 -CCC AGC ACT GTT TTC TTG TCT CA-3	177	Sexing	-
SdY-Rv2	5 -CTT AAA ACC ACT CCA CCC TCC AT-3			

Table S4: Statistical differences in methylation distribution between morphs, for each timepoint. p-values for both Kruskal-Wallis (KW) and pairwise Kolmogorov-Smirnoff (KS) tests are indicated. KW p-values for distribution differences between timepoints were always < 2.2E-16.

50 ts	**KW = 3.901E-16**		
KS pairwise	LB	SB	PL
LB	-		
SB	7.72E-06	-	
PL	3.49E-10	3.71E-03	-
PI	< 2.2E-16	5.06E-10	4.12E-07
100ts	**KW < 2.2E-16**		
KS pairwise 100ts	LB	SB	PL
LB	-		
SB	4.81E-09	-	
PL	6.82E-09	< 2.2E-16	-
PI	2.67E-06	< 2.2E-16	0.055
150ts	**KW = 3.571E-08**		
KS pairwise 150ts	LB	SB	PL
LB	-		
SB	7.32E-05	-	
PL	3.18E-10	9.58E-03	-
PI	1.91E-08	0.240	0.059
200ts	**KW = 0.0034**		
KS pairwise 200ts	LB	SB	PL
LB	-		
SB	0.093	-	
PL	7.80E-05	4.42E-04	-
PI	0.036	0.943	1.73E-04

Table S5: Variance explained by each principal component.

Principal component	Variance explained (%)	Cumulative variance (%)
PC1	8	8
PC2	3.52	11.52
PC3	3.19	14.71
PC4	2.92	17.63
PC5	2.6	20.23
PC6	2.51	22.74
PC7	2.49	25.23
PC8	2.45	27.68
PC9	2.39	30.07
PC10	2.37	32.44
PC11	2.31	34.75
PC12	2.26	37.01
PC13	2.23	39.24
PC14	2.17	41.41
PC15	2.15	43.56
PC16	2.12	45.68
PC17	2.05	47.73
PC18	2.04	49.77
PC19	2.01	51.78
PC20	2	53.78
PC21	1.96	55.74
PC22	1.94	57.68
PC23	1.91	59.59
PC24	1.89	61.48
PC25	1.86	63.34
PC26	1.86	65.2
PC27	1.84	67.04
PC28	1.81	68.85
PC29	1.8	70.65
PC30	1.77	72.42
PC31	1.75	74.17
PC32	1.75	75.92
PC33	1.72	77.64
PC34	1.71	79.35
PC35	1.69	81.04
PC36	1.67	82.71
PC37	1.66	84.37
PC38	1.63	86
PC39	1.62	87.62
PC40	1.6	89.22
PC41	1.58	90.8
PC42	1.58	92.38
PC43	1.57	93.95
PC44	1.56	95.51
PC45	1.54	97.05
PC46	1.51	98.56
PC47	1.46	100.02

Table S6: ANOVA results for multiple regressions on PCs, using 3 explanatory variables (Morph, Time and Sex). P-values have been corrected for multiple testing with Bonferroni.

Principal Component	Morph	Time	Sex	Morph X Time
PC1	1	1.06E-28	1	1
PC2	1	1.22E-07	1	1
PC3	0.052	0.003	1	0.369
PC4	9.80E-09	0.982	1	1
PC5	1	1	1	1
PC6	1	1	1	1
PC7	0.197	1	1	1
PC8	0.115	1	1	1
PC9	1	1	1	1
PC10	0.046	1	1	1
PC11	1	1	1	1
PC12	1	1	1	1
PC13	1	1	1	1
PC14	1	1	1	1
PC15	1	1	1	1
PC16	1	1	1	1
PC17	1	1	1	1
PC18	1	1	1	1
PC19	1	1	1	1
PC20	1	1	1	1
PC21	1	1	1	1
PC22	1	1	1	1
PC23	1	1	1	1
PC24	1	1	1	1
PC25	1	1	1	1
PC26	1	1	1	1
PC27	1	1	1	1
PC28	1	1	1	1
PC29	1	1	1	1
PC30	1	1	1	1
PC31	1	1	1	1
PC32	1	1	1	1
PC33	1	1	1	1
PC34	1	1	1	0.369
PC35	1	1	1	1
PC36	1	1	1	1
PC37	1	1	1	1
PC38	1	1	1	1
PC39	1	1	1	1
PC40	1	1	1	1
PC41	1	1	1	1
PC42	1	1	1	1
PC43	1	1	1	1
PC44	1	1	1	1
PC45	1	1	1	1
PC46	1	1	1	1
PC47	1	1	1	1

Table S7: Full list of GO enrichment analysis for genes near residues that contributed to timepoint and morph differences (with PCA and glm). GO categories displayed have a p-value < 0.01 with Fisher's test.

Genes near CpGs separating timepoints (PC1)					
GO.ID	**Term**	**Annotated**	**Significant**	**Expected**	**weightedFisher**
GO:0006334	nucleosome assembly	152	5	0.46	0.00011
GO:0006189	'de novo' IMP biosynthetic process	10	2	0.03	0.00041
GO:0016559	peroxisome fission	26	2	0.08	0.00287
GO:0045638	negative regulation of myeloid cell differentiation	325	5	0.99	0.00327
GO:0007567	parturition	98	3	0.3	0.00342
GO:0009113	purine nucleobase biosynthetic process	29	2	0.09	0.00356
GO:0035584	calcium-mediated signaling using intracellular calcium source	110	3	0.34	0.00472
GO:0090197	positive regulation of chemokine secretion	34	2	0.1	0.00487
GO:0050847	progesterone receptor signaling pathway	35	2	0.11	0.00516
GO:0030511	positive regulation of transforming growth factor beta receptor signaling pathway	133	3	0.41	0.00797
GO:1900273	positive regulation of long-term synaptic potentiation	134	3	0.41	0.00814
GO:0006525	arginine metabolic process	46	2	0.14	0.00878
Genes near CpGs separating timepoints (glm)					
GO.ID	**Term**	**Annotated**	**Significant**	**Expected**	**weightedFisher**
GO:0006334	nucleosome assembly	152	5	0.47	0.00011
GO:0042742	defense response to bacterium	640	9	1.98	0.00017
GO:0006189	'de novo' IMP biosynthetic process	10	2	0.03	0.00042
GO:1900275	negative regulation of phospholipase C activity	14	2	0.04	0.00084
GO:0046148	pigment biosynthetic process	171	4	0.53	0.00199
GO:0015824	proline transport	23	2	0.07	0.00229
GO:0051599	response to hydrostatic pressure	28	2	0.09	0.00339
GO:0032328	alanine transport	29	2	0.09	0.00363
GO:0051262	protein tetramerization	491	6	1.52	0.00424
GO:0033209	tumor necrosis factor-mediated signaling pathway	356	5	1.1	0.00502
GO:1904468	negative regulation of tumor necrosis factor secretion	35	2	0.11	0.00526
GO:0015816	glycine transport	35	2	0.11	0.00526
GO:0007076	mitotic chromosome condensation	42	2	0.13	0.00751

Genes near CpGs separating morphs (PC4)					
GO.ID	**Term**	**Annotated**	**Significant**	**Expected**	**weightedFisher**
GO:0006334	nucleosome assembly	152	5	0.53	0.0002
GO:0016340	calcium-dependent cell-matrix adhesion	9	2	0.03	0.00043
GO:0046548	retinal rod cell development	51	3	0.18	0.00077
GO:0021963	spinothalamic tract morphogenesis	12	2	0.04	0.00078
GO:0032780	negative regulation of ATPase activity	53	3	0.19	0.00086
GO:0106028	neuron projection retraction	14	2	0.05	0.00107
GO:0099557	trans-synaptic signaling by trans-synaptic complex, modulating synaptic transmission	14	2	0.05	0.00107
GO:0055069	zinc ion homeostasis	59	3	0.21	0.00118
GO:1904782	negative regulation of NMDA glutamate receptor activity	16	2	0.06	0.00141
GO:0007216	G protein-coupled glutamate receptor signaling pathway	63	3	0.22	0.00142
GO:1904783	positive regulation of NMDA glutamate receptor activity	18	2	0.06	0.00179
GO:0031938	regulation of chromatin silencing at telomere	21	2	0.07	0.00244
GO:0010961	cellular magnesium ion homeostasis	22	2	0.08	0.00268
GO:0051389	inactivation of MAPKK activity	23	2	0.08	0.00293
GO:1902723	negative regulation of skeletal muscle satellite cell proliferation	24	2	0.08	0.00318
GO:1902725	negative regulation of satellite cell differentiation	24	2	0.08	0.00318
GO:0036109	alpha-linolenic acid metabolic process	24	2	0.08	0.00318
GO:0022038	corpus callosum development	85	3	0.3	0.00334
GO:0032490	detection of molecule of bacterial origin	25	2	0.09	0.00345
GO:0021554	optic nerve development	86	3	0.3	0.00346
GO:0071679	commissural neuron axon guidance	95	3	0.33	0.00457
GO:0045075	regulation of interleukin-12 biosynthetic process	29	2	0.1	0.00463
GO:0045671	negative regulation of osteoclast differentiation	102	3	0.36	0.00557
GO:0050847	progesterone receptor signaling pathway	35	2	0.12	0.0067
GO:0014719	skeletal muscle satellite cell activation	35	2	0.12	0.0067
GO:0043052	thermotaxis	37	2	0.13	0.00746
GO:0051262	protein tetramerization	491	6	1.72	0.00771
GO:0031648	protein destabilization	120	3	0.42	0.0087

Genes near CpGs separating morphs (glm)					
GO.ID	**Term**	**Annotated**	**Significant**	**Expected**	**weightedFisher**
GO:0006334	nucleosome assembly	152	4	0.11	4.20E-06
GO:0042742	defense response to bacterium	640	5	0.46	8.20E-05
GO:0051295	establishment of meiotic spindle localization	30	2	0.02	0.00022
GO:1901223	negative regulation of NIK/NF-kappaB signaling	69	2	0.05	0.00115
GO:0030071	regulation of mitotic metaphase/anaphase transition	118	2	0.09	0.0033
GO:0008205	ecdysone metabolic process	5	1	0	0.00362
GO:0051673	membrane disruption in other organism	5	1	0	0.00362
GO:0000964	mitochondrial RNA 5'-end processing	5	1	0	0.00362
GO:0051663	oocyte nucleus localization involved in oocyte dorsal/ventral axis specification	6	1	0	0.00434
GO:1902956	regulation of mitochondrial electron transport, NADH to ubiquinone	6	1	0	0.00434
GO:0051758	homologous chromosome movement towards spindle pole in meiosis I anaphase	6	1	0	0.00434
GO:1901208	negative regulation of heart looping	7	1	0.01	0.00506
GO:2000321	positive regulation of T-helper 17 cell differentiation	7	1	0.01	0.00506
GO:1901856	negative regulation of cellular respiration namespace	7	1	0.01	0.00506
GO:0000965	mitochondrial RNA 3'-end processing	7	1	0.01	0.00506
GO:1901211	negative regulation of cardiac chamber formation	7	1	0.01	0.00506
GO:0072068	late distal convoluted tubule development	7	1	0.01	0.00506
GO:0045478	fusome organization	8	1	0.01	0.00579
GO:0031952	regulation of protein autophosphorylation	162	2	0.12	0.00611
GO:1905832	positive regulation of spindle assembly	9	1	0.01	0.00651
GO:0003167	atrioventricular bundle cell differentiation	9	1	0.01	0.00651
GO:0046604	positive regulation of mitotic centrosome separation	9	1	0.01	0.00651
GO:1904801	positive regulation of neuron remodeling	9	1	0.01	0.00651
GO:0075259	spore-bearing structure development	10	1	0.01	0.00723
GO:0070900	mitochondrial tRNA modification	10	1	0.01	0.00723
GO:0071169	establishment of protein localization to chromatin	11	1	0.01	0.00795
GO:0000349	generation of catalytic spliceosome for first transesterification step	11	1	0.01	0.00795
GO:003262	interleukin-18 production	11	1	0.01	0.00795

1					
GO:0042780	tRNA 3'-end processing	12	1	0.01	0.00867
GO:0090235	regulation of metaphase plate congression	12	1	0.01	0.00867
GO:0045653	negative regulation of megakaryocyte differentiation	12	1	0.01	0.00867
GO:0072019	proximal convoluted tubule development	13	1	0.01	0.00939
GO:0072385	minus-end-directed organelle transport along microtubule	13	1	0.01	0.00939
GO:0070649	formin-nucleated actin cable assembly	13	1	0.01	0.00939

Table S8: List of all 478 DMRs between timepoints, morphs, or both. The first column is the scaffold name. The second column is the position at which the 1000bp window starts (in base pairs from the beginning of the scaffold). The third column is the distance from the DMR to the closest annotated feature (in base pairs). The fourth column is the Locus-ID of this feature, corresponding to the Salvelinus sp. IW2-2015 (assembly ASM291031v2). The fifth column is a more accessible name for each feature/gene.

Consult **online version** for more information.

Table S9: Distribution of DMRs and non-DMRs on linkage groups from the Salvelinus sp. IW2-2015 (assembly ASM291031v2).

	DMRs between timepoints	**DMRs between morphs**	**DMRs between timepoints and morphs**	**non-DMRs**
total	338	91	49	18469
LG1	11	5	1	577
LG2	10	2	0	394
LG3	3	0	2	300
LG4p	0	2	0	236
LG4q.1:29	14	5	0	904
LG4q.2	6	2	0	260
LG5	5	1	0	355
LG6.1	4	1	0	268
LG6.2	3	1	1	227
LG7	6	0	0	343
LG8	4	5	1	519
LG9	7	2	1	300
LG10	4	0	1	225
LG11	8	1	3	496
LG12	1	0	0	127
LG13	8	1	4	475
LG14	6	1	1	449
LG15	6	1	0	616
LG16	3	0	1	346
LG17	13	1	1	400
LG18	7	3	4	678
LG19	5	1	0	368
LG20	14	1	1	748
LG21	2	0	0	53
LG22	7	1	0	348
LG23	7	2	0	461
LG24	1	1	0	76
LG25	4	1	1	225
LG26	4	0	1	421
LG27	2	2	1	382
LG28	5	4	0	283
LG30	3	1	0	209
LG31	7	1	1	295
LG32	5	0	1	340
LG33	11	1	0	349
LG34	1	0	0	101
LG35	3	0	0	192
LG36	7	1	1	413
LG37	3	0	2	179
Unplaced	118	40	19	4531

Table S10: Results of ANOVA´s and Tukey's HSD tests on methylation data. The first column states the name of the gene of interest. Columns 2, and 7 show ANOVA p-values for each variable (Morph, and Time). Columns 3-6 as well as 8-10 show HSD test results in a compact letter display format: groups with the same letter are not significantly different

Name	**Morph *p*-value**	**LB**	**SB**	**PI**	**PL**	**Time *p*-value**	**50 ts**	**100 ts**	**150 ts**	**200 ts**
LMTK2	0.677	a	a	a	a	0.425	a	a	a	a
H2A-like	0.089	a	a	a	a	3.63E-21	a	b	c	d
SLC9A3R2-like	7.42E-05	b	b	ab	a	0.405	a	a	a	a
NKX23-like	0.073	a	a	a	a	0.673	a	a	a	a
NFIX	7.01E-06	a	bc	b	c	0.648	a	a	a	a
RASSF4-like	2.64E-06	a	ab	bc	c	5.39E-06	c	b	ab	a
ARL16	0.002	ab	b	b	a	0.689	a	a	a	a
MEIS1-like	0.001	b	a	b	ab	0.106	a	a	a	a
H3-like	0.093	a	a	a	a	2.60E-16	a	b	c	d
ARMC1	0.301	a	a	a	a	0.830	a	a	a	a
GLI3-like	0.662	a	a	a	a	0.001	NA	a	a	b
MEGF9	0.056	a	a	a	a	5.53E-16	a	b	c	c
MPP3	0.031	ab	a	b	ab	3.86E-07	a	a	b	b
ARHGEF37-like	0.146	a	a	a	a	5.01E-10	a	a	b	c

Table S11: Results of ANOVA´s and Tukey's HSD tests on gene expression data. The first column states the name of the gene of interest. Columns 2, 7 and 11 show ANOVA p-values for each variable (Morph, Time and Morph x Time interaction). Columns 3,4,5,6 as well as 8,9,10 show HSD test results in a compact letter display format: groups with the same letter are not significantly different.

Name	**Morph *p*-value**	**LB**	**SB**	**PI**	**PL**	**Time *p*-value**	**100 ts**	**150 ts**	**200 ts**	**Interaction *p*-value**
LMTK2	0.220	a	a	a	a	0.020	b	a	b	0.140
H2A-like	7.32E-06	c	a	b	ab	3.68E-06	b	a	c	0.650
SLC9A3R2-like	2.07E-03	b	a	b	b	1.10E-12	c	b	a	0.970
NKX23-like	0.270	a	a	a	a	9.02E-08	a	a	b	0.620
NFIX	0.030	ab	ab	a	b	1.99E-28	c	b	a	4.47E-03
RASSF4-like	0.300	a	a	a	a	2.06E-05	b	a	b	0.170
ARL16	3.46E-08	c	b	b	a	1.93E-05	b	a	c	0.590
MEIS1-like	0.110	a	a	a	a	6.17E-07	c	b	a	4.48E-03
H3-like	2.69E-05	b	a	b	b	4.43E-08	a	b	c	0.420
ARMC1	1.16E-05	c	a	b	ab	0.210	a	a	a	0.350
GLI3-like	0.020	ab	b	a	ab	0.190	a	a	a	0.050
MEGF9	2.62E-06	b	a	a	a	1.60E-07	b	a	a	0.570
MPP3	0.390	a	a	a	a	2.68E-14	b	a	a	0.060
ARHGEF37-like	4.91E-07	b	a	a	a	5.60E-09	a	b	c	3.51E-03

Table S12: Tests for correlations between average methylation and gene expression. For GLI3-like and NFIX, the data was not parametric so the Kendall coefficient and test were used. For the other 12 genes, Pearson's correlation and t-tests were implemented.

Gene Name	**R**	**p**
H2A	0.1415	0.661
LMTK2	0.0505	0.8827
GLI3-like	0.0556	0.9195
ARHGEF37-like	0.7471	0.0052
MPP3	-0.7846	0.0042
RASSF4-like	0.0212	0.948
H3-like	0.6287	0.0285
ARMC1	-0.1112	0.7309
MEGF9	-0.643	0.0241
NFIX	0.0909	0.7373
NKX23-like	0.0387	0.905
SLC9A3R2-like	-0.0417	0.8977
ARL16	0.3265	0.3002
MEIS1-like	-0.2364	0.4595

Paper II

Patterns of genetic, methylomic, and expression divergence among three sympatric resource morphs of Arctic charr

S. Matlosz, L. Jerman-Plesec, H. Xiao, A. Guðjónsson, S.R. Franzdóttir, A. Pálsson, Z.O. Jónsson

Manuscript.

Author contribution: S. Matlosz performed comparative analyses between the previously generated datasets and wrote the first version of the manuscript.

Patterns of genetic, methylomic, and expression divergence among three sympatric resource morphs of Arctic charr

Authors **S. Matlosz (1), L. Jerman-Plesec (1), H. Xiao (1), A. Guðjónsson(1,2), S.R. Franzdóttir (1), A. Pálsson (1), Z.O. Jónsson (1)**

Email: slt2@hi.is

Affiliations:
(1) Institute of Life and Environmental Sciences, University of Iceland, Reykjavík, Iceland
(2) Clinical Microbiomics, Copenhagen, Denmark

Abstract:

Adaptive evolution favors changes in genes and developmental processes that allow organisms to utilize specific resources, and the availability of different niches frequently leads to discrete resource polymorphisms within species. Such resource polymorphism is often found among salmonids that colonized new habitats after the last ice-age, with repeatable morphological divergence that relates to specialization to benthic vs pelagic niches. These systems are ideal to study genetic and developmental changes associated with rapid evolution. Arctic charr is one of the more polymorphic salmonids and in lake Thingvallavatn four morphs coexist in sympatry. Population genomics have indicated monophyly of these morphs and revealed genetic separation among three of them. In this study we ask if whole genome sequencing of two individuals per morph (PL-, LB and SB-charr) can highlight loci and genomic regions that differentiate morphs, by comparison to available SNPs from ddRADseq data. Furthermore, we ask if genes differentially expressed between these morphs (during early development) are over-represented in these genomic regions of differentiation, and if similar patterns are seen for residues with differential methylation between morphs. The results reveal overlap of genetic signals and in some morphs differential expression. The bulk of the differentially expressed genes however do not associate with regions with high genetic divergence, illustrating that most are affected by changes in *trans*-factors, upstream elements or more complex modulations of the developmental system. Additionally, we ask whether differentially methylated CpGs lie in these regions of divergence. Finally, the data provide numerous candidate loci that may associate with morph separation and the specific adaptations that characterize these morphs.

I. Introduction

Adaptive evolution favors changes in genes and developmental processes that allow organisms to utilize specific resources or survive under specific environmental conditions (Olson-Manning et al., 2012). Identifying the key traits, developmental and physiological processes that selection favors, can be achieved by comparison of anatomy, physiology and genetics between groups. While higher taxonomic units are differentiated by major adaptations like photosynthesis (Oliver et al., 2023), neurons (Moroz and Kohn, 2016), cranium (Morris and Abzhanov, 2021) and gynoeciums (Reyes-Olalde, Aida and de Folter, 2023), identifying the key changes responsible for emergence of these traits is problematic due to the cumulative impact of genetic drift and selection on multiple traits over millions of generations. Identifying alleles and specific genes favored during adaptive evolution by comparing recently diverged species or populations is a more manageable endeavor, particularly by focusing on rapidly evolving ecologically important traits, such as those related to resource use and feeding (Huber et al., 2007).

Adaptive evolution is usually due to changes in multiple genes and pathways, as even the simplest traits are shaped by complex molecular mechanisms (Martin and Courtier Orgogozo, 2017)). On the other hand, several recent examples show that major alleles can segregate within species, and enable rapid evolutionary responses to ecological shifts (Thompson et al., 2020, Almen et al., 2016). In the last 15 years, genetic analyses with high-throughput sequencing have enabled studies of the genetic differences between groups and species, and of functional differences including responses to environmental factors through changes in mRNA levels and methylation patterns (Major et al., 2020, Pierron et al., 2021). Nonetheless, while genetics, environment and chance can influence mRNA levels and methylation patterns, identifying the causative relationships is far from easy. Testing for overlap of regions of genetic differentiation, genes differently expressed, and differently methylated regions, can implicate the influence of *cis* and *trans*-factors on these intermediate phenotypes. Are genes showing expression divergence affected by genetic divergence? And if so, in a *cis*- or *trans*- manner? We argue that recently diverged populations undergoing adaptive divergence are useful models for addressing such questions.

In many species of salmonids, anadromous ancestors have given rise to resident spring, stream and/or lake populations that rely entirely on freshwater niches and in many cases have diverged to produce resource polymorphism (Foote, Wood and Withler, 1989). These, often quite specialized, morphs have adapted to particular niches and food types resulting in size, shape and life-history polymorphism. These morphs are often found in isolated locales reflecting allopatric divergence, but sympatric polymorphism is also quite common, particularly in *S. malma* and *S. alpinus*, for which anadromous populations became landlocked following the end of the last ice age 10-20,000 years ago (Kapralova et al., 2011, Weinstein et al., 2024). Most commonly, two morphs that tend to specialize in pelagic and limnetic habitats are found per lake, but some watersheds are inhabited by more than two morphs. The most dramatic and best studied are the seven sympatric morphs of *S. malma* in Lake Kronotskoe (Markevich et al., 2018) and the four sympatric morphs of *S. alpinus* in Lake Thingvallavatn Iceland. In Thingvallavatn, these morphs, large benthic (LB), small benthic (SB), planktivorous (PL) and piscivorous (PI) differ by their habitat use (Sandlund et al., 1987), life-history (Skúlason, Noakes, & Snorrason, 1989; Skúlason, Snorrason, et al., 1989) and size. These morphs also differ in feeding specializations along the benthic-limnetic axis (Malmquist et al., 1992), with benthic morphs eating gastropods and arthropods while limnetic morphs feed on plankton or fish. Relating to this difference in diet, the four morphs display contrasting trophic apparatus

structures: SB and LB display a subterminal mouth and blunt snout while PI and PL display a terminal mouth and pointed snout (Sandlund et al., 1992). The internal skeletal variations driving these external characteristics have recently been established by geometric morphometric analyses to assess variation in shape and size of six upper and lower jaw bones, and the dentary, articular-angular, maxilla and premaxilla were found to display shape divergence between morphs along a benthic-limnetic axis (Jónsdóttir et al., 2023). Hybridization studies between the SB and PL morphs have been done in common garden (Horta-Lacueva et al., 2021), demonstrating that hybrid offspring display transgressive, intermediate and parental-like phenotypes (De la Camara et al., 2023). However, in the wild the spatial and temporal reproductive barriers stemming from differences in life-history between the morphs, as well as post-zygotic barriers such as fertilization failure or early-development mortality (Horta-Lacueva, 2022), limit gene flow between these morphs.

Prior population genetic studies suggested genetic differentiation of the morphs, although some doubt was cast about the genetic status of the PI-charr (Gíslason, 1998, reviewed by Guðbrandsson 2019). With the advance of genomic methods, and their application to studies in non-model species population, higher coverage datasets have been generated in S. alpinus.

SNPs identified in RNA-seq data showed that the PL, SB and LB morphs were clearly genetically separated (Guðbrandsson et al., 2019, Brachmann et al., 2022), and KASP genotyping of outlier SNPs identified in that data suggests PI-charr has similarity to PL-charr but may be either recently originated and/or admixed with LB-charr. More recent analysis of variations from ddRAD-seq confirmed the separation of the three morphs, and that PI-charr is more genetically heterogeneous (Xiao et al., unpublished data). Molecular studies have also been conducted, including mRNA, miRNA and methylation sequencing, as well as qPCR and in situ analyses on candidate genes, in order to get a handle on the genes and systems that may relate to morph differences. Importantly mRNA sequencing during early development revealed many genes differentially expressed between PL, SB and LB morphs, with the PL being most diverged in expression from the other two (Guðbrandsson et al., 2018). Candidate gene studies identified differential expression between benthic and limnetic morphs (with some of the genes associated with bone development): for instance *Mmp2* and *Sparc*, two genes associated with matrix remodeling were found to be differentially expressed between benthic and limnetic morphs (Ahi et al., 2013), as was the transcription factor *Ets2*, which might regulate these two genes (Ahi et al., 2014), and *Eif4ebp1*, which encodes a repressor of translation initiation (Guðbrandsson et al., 2017). In a master's thesis (Eskafi 2022), investigation of ~65 differently expressed genes from Ahi et al., (2014, 2015) revealed some overlaps between differently expressed genes and regions of genetic differentiation, although no enrichment was found (Eskafi, 2022). More recently, DNA methylation differences were also detected between the morphs during early embryonic development (from gastrulation to late organogenesis), adding an epigenetic dimension to the list of observed molecular differences (Matlosz et al., 2022).

However, apart from the study by Eskafi (2022), for most of these differentially expressed genes it is unknown whether they show genetic divergence that matches their expression divergence, or if genetic differences in upstream factors are responsible. In other words, the genetic variation, methylation and transcriptomic differences between morphs, have so far only been analysed in isolation, and genomic overlaps between differentially expressed genes, differently methylated CpGs and regions of genomic divergence have never been investigated. While no genome sequence is currently available for *S. alpinus*, a *Salvelinus*

sp. genome assembly from closely related *S. malma* or *S. malma*/*S. alpinus* hybrid (ASM291031v2, (https://www.ncbi.nlm.nih.gov/assembly/GCF_002910315.2/)) enables comparison of genomic patterns of differentiation for this non-model species.

Here, we take advantage of a diverse combination of datasets available for the sympatric morphs of Arctic charr from lake Thingvallavatn to compare genomic, epigenomic and transcriptomic patterns and investigate their correlations. These datasets include previously unpublished high coverage whole genome sequences from 2 individuals (female and male) per morph and markers from 98 individuals from ddRAD sequencing (Xiao et al., unpublished), both used to detect genetic variation, as well as already published functional genomics datasets from early embryonic development, including mRNA sequencing (Guðbrandsson et al., 2016) and reduced representation bisulfite sequencing (RRBS) (Matlosz et al., 2022) (**Figure 1**). As the fourth morph, PI charr, is genetically heterogeneous, presumably of recent PL and LB ancestry with continuing admixture (Guðbrandsson 2019, Xiao et al., unpublished data), it was not included in the analyses.

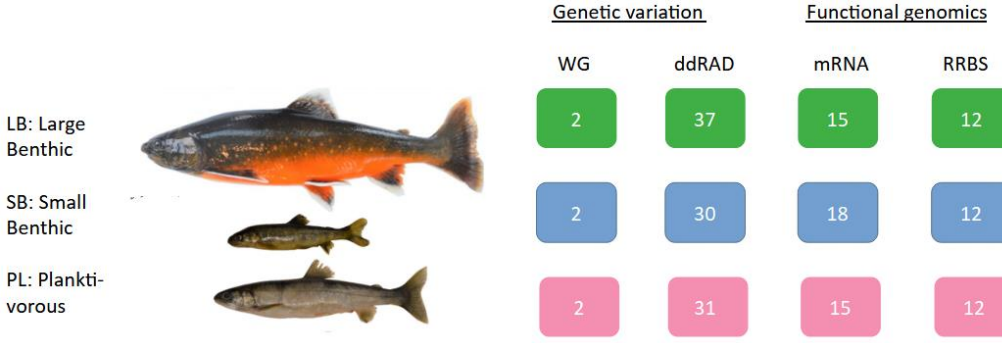


Figure 1: Phenotypic differences between three charr morphs and the genomic datasets used in this study. Represented are three of the four sympatric morphs in Lake Thingvallavatn that differ in size, body and head shape. Numbers for each dataset indicate numbers of individuals (samples). Whole genome sequences were generated for this study and SNPs from ddRAD from a population sample are from Xiao et al., (unpublished). mRNA abundance from RNA sequencing (Guðbrandsson et al., 2018) and the methylation status of CpGs obtained from RRBS (Matlosz et al., 2022) came from common garden experiments, generated from crosses made within each morph with samples taken over different timepoints during early development. Pictures taken by Arnar Pálsson and Sarah E. Steele. The coloring scheme indicating morphs (green: LB-, blue: SB-, pink: PL-charr) will be retained throughout the manuscript.

The aim of this study is to investigate patterns of genetic differentiation between three of the four recently evolved sympatric morphs (Small Benthic - SB, Large Benthic - LB and Planktivorous - PL) and characterize the potential functionality of these regions of differentiation.

The first research question is methodological, how well does whole genome sequencing of few individuals capture divergence in specific loci or regions? To answer this, we compared SNPs from WGS data with available ddRAD-seq data.

The remaining research questions, on the other hand, focus on a) whether regions of genetic differentiation overlap with genes with differential expression; b) whether regions of genetic differentiation overlap with differently methylated CpGs and c) whether genes (differently expressed or not) located in these regions possess biological function related to the observed phenotypic differences between morphs. These are addressed by performing

permutation tests to measure whether overlaps between datasets happen significantly more (or less) than expected by chance. Several hypotheses were put forth.

i) We hypothesize that genes with differential expression between morphs will be enriched in regions of genetic divergence.

ii) We predict congruence of signals, e.g. that genes with differential expression in PL-charr vs the other morphs will be enriched in genomic regions with PL-specific alleles or haplotypes.

iii) We hypothesize a similar degree of congruence between differential expression and regions of divergence for the three genetically distinct morphs in the lake.

iv) We hypothesize that differently methylated CpGs between morphs will be enriched in regions of genetic divergence.

The analysis of the expression data largely agreed with these hypotheses, but morph specific patterns were revealed indicating varying contribution of *cis*- and *trans*- effects on expression by morphs. Morph specific congruences could not be estimated for methylation and genomic data, due to the dynamic nature of the former dataset. Finally, we set out to further explore the methylation patterns and expression of loci within peaks of genetic differentiation, to identify candidate functional systems that may differ between morphs. These results provide further evidence of the complex genetic, expression and methylation divergence and the potential interplay of these sources of variation in closely related sympatric morphs.

II. Material and Methods

Whole genome sequencing data:

DNA was extracted from approximately 10 mg pieces of ethanol-preserved fin tissue from two individuals of each of the four "morphs", a male and a female. The fin clips were air dried briefly and then soaked in 125 µl digestion buffer (100 mM EDTA, 25 mM Tris pH 7.5, 1% SDS). The clips were then transferred to 125 µl of the same buffer freshly supplemented with 25mM DTT and 1.75 µg of Proteinase K. After incubation with gentle shaking at 55C overnight, samples were spun at 20.000 g in a benchtop centrifuge and 115µl of the supernatant transferred to another tube with 115 µl of freshly prepared binding buffer (2.5 M NaCl, 20% PEG 8000, 25mM DTT) and 40µl of NucleoMag beads (Machery-Nagel) added. After 5 minutes of incubation the pelleted magnetic beads were washed three times with 1 ml 80% Ethanol, dried briefly and eluted in 120µl of 10 mM Tris pH 7.5. High quality DNA was prepared for sequencing using the Illumina TruSeq DNA PCR-free sample preparation kit and sequenced on Illumina HiSeq X ten at deCODE genetics, Iceland. Reads were filtered for quality with *cutadapt* (Martin, 2011), trimmed with *TrimGalore* (Krueger et al., 2023) and mapped to the reference genome using *BWA* (Li and Durbin, 2009).

As a *S. alpinus* genome was not available when this work was conducted, reads were mapped to the *Salvelinus sp*. genome (ASM291031v2, (https://www.ncbi.nlm.nih.gov/assembly/GCF_002910315.2/)), which comes from S. malma or S. malma/S. alpinus hybrid individuals. The same applied to the other datasets analyzed.

The --*weir-fst-pop* command from the *vcftools* software (Danecek et al., 2011) was used to identify F_{ST} indexes for the SB, PL and LB morphs by comparing one morph to the other two (for instance SB vs LB and PL). Contrarily to ddRAD data where these calculations were done on a per-site basis, F_{ST} calculations for WGS data were done on a windowed basis, with an --*fst-window-size* of 100000 and an --*fst-window-step* of 10000.

A genomic window was considered a region of genomic divergence when the highest F_{ST} index value of all comparisons for this window lay outside 2σ of the overall F_{ST} distribution. A higher number of significant windows had a PL-specific signal (**Supplementary Figure 1**), but regions of genomic differentiation from all three morphs were still found to be significant (**Supplementary Table 1**).

ddRAD data:

A second set of genetic variant calls used for this study was derived from ddRAD seq data (Xiao et al., unpublished). Briefly the procedure was as follows. DNA was extracted from fin clips of 122 individuals from lake Thingvallavatn, from each of the four morphs (see **Supplementary Table 2**). Two restriction enzymes *Ape*KI and *Bam*HI HF were used to process genomic DNA from each individual, and then ligated with a unique combination of 4bp *Ape*KI barcodes and 6 bp *Bam*HI HF barcodes (as described in (Lagunas et al., 2023)). This allowed for multiplexing of 96 individuals in a sequencing library. A mixture of fragments from different individuals was size-selected by Pippin prep (Sage Science) collecting fragments in the range of 360 to 440 bp, then amplified using PCR. The concentration of the library was checked with The Qubit™ 3.0 Fluorometer (Thermo Fisher Scientific), and agarose electrophoresis (2% of agarose gel) was used to verify fragment size range. Sequencing was performed on Illumina Hi-seq X-Ten by BGI, China.

Demultiplexing and filtering was done with the Stacks software (Catchen et al., 2011) with the parameters *-p 3 -r 0.66 --min-mac 3 --max-obs-het 0.6*, meaning that a SNP at a genomic position was kept if present in at least 3 morphs out of the 4 originally analyzed, in at least 66% of individuals per morph, with a minimum minor allele count of 3, and a maximum observed heterozygosity of 0.6. Data from samples "ThSB_60" and "ThPL_34" (which were suspected to be hybrids) as well as all PI samples (a morph which is genetically ambiguous) was then set aside and the remaining samples were used to calculate, for each genomic position, F_{ST} indexes for the SB, PL and LB morphs by comparing one morph to the other two (for instance SB vs LB and PL). Similar to the analysis of WGS data, this was done with the --*weir-fst-pop* command from the vcftools software (Danecek et al., 2011).

A genomic position was considered a SNP outlier when the highest F_{ST} index value of all comparisons for that position lay outside 2σ of the overall F_{ST} distribution. A higher number of SNPs came from PL-specific signals, for which the F_{ST} distribution was higher overall (**Supplementary Figure 2**), but SNPs from all three morphs could still be found (**Supplementary Table 1**).

Methylation data:

DNA extraction- and RRBS protocols are detailed in (Matlosz et al., 2022). Methylation data came from 27 samples, more specifically from three biological replicates of the three clearly separated morphs (SB, LB and PL) at three developmental timepoints (50ts, 150ts, 200ts). As with WGS and ddRAD data, samples from the PI morph were not included in the current methylation analysis, as the PI is genetically ambiguous and as the

transcriptomic dataset only included the three other morphs. Moreover, samples at the 100ts developmental timepoint were also removed from the methylation analyses as reads from these samples had the lowest coverage in the RRBS sequencing, due to a technical sequencing bias (Matlosz et al., 2022). Our filtering criterium for keeping CpGs for further methylation analyses was that each CpG needed to have a 10X coverage in each of the samples studied. Thus, removing the PI morph and individuals at the 100ts timepoint increased the number of CpGs passing filtering tests, expanding the methylation dataset from 10,340 to 14,427 CpGs with more than 10X coverage in every sample (see **Supplementary Table 3**). A generalized mixed model (in R, using glm) was used to identify CpGs with methylation levels influenced by the different biological variables (this was the same as in our previous study, Matlosz et al., 2022):

CpG=Morph+Timepoint+Sex+Morph×Timepoint+Morph×Sex+Timepoint×Sex + error

This model was run for each of the 14,427 positions, assuming Gaussian distributions. Corrections for multiple testing were done with the Bonferroni method and CpGs with a corrected p-value <0.05 were considered significant. While sex and timepoints were part of the model, the focus here was CpGs with significant methylation differences between morphs, that is CpGs with statistically significant terms "Morph", "Morph x Time" and "Morph x Sex". The numbers of CpGs with a significantly different methylation observed for each of these terms, and whether some CpGs were found in multiple of these terms at once, is shown in **Supplementary Figure 3**. The general patterns were congruent with results from Matlosz et al. (2022).

Analyses of methylation percentage distribution (**Figure 8**) were done by calculating, per residue, the mean methylation percentage for all individuals of the same category (i.e. LB-morph, 150ts). Significance was evaluated with KS tests (**Supplementary Table 9**).

Transcriptome data:

For consistency, raw transcriptomic reads from Guðbrandsson et al., (2018) (see paper for sequencing methods) were re-aligned to the *Salvelinus sp.* genome IW2-2015, assembly ASM291031v2, (https://www.ncbi.nlm.nih.gov/assembly/GCF_002910315.2/) used for all other datasets. The raw numbers of reads for each gene were counted with the HTSeq R package (Anders et al., 2015) and parameters *-i gene_name -t exon -f bam -s no -r pos*. These raw counts were then used to perform a differential expression analysis with a likelihood ratio test comparing the model "time + morph" against the reduced model "time". This was done with the *DESeqDataSetFromHTSeqCount()* function from the DEseq2 package in R (Love et al., 2014). Samples from the 160ts developmental timepoint that suffered from 3' bias (see Guðbrandsson et al., 2018) were not included in these analyses.

Tests of overlap of genomic datasets:

Whole datasets (all CpGs analysed with RRBS, all potential SNPs/regions covered by ddRADseq and WGS, all transcripts sequenced with transcriptome analyses) as well as subsets deemed significant (differentially methylated CpGs, genomic SNPs/regions of differentiation, differently expressed genes) with Salvelinus sp. genome coordinates were converted to GRanges objects with the GenomicRanges R package (Lawrence et al., 2013) for further analyses. Because of the ambiguity of the *Salvelinus sp.* genome, and potential genomic rearrangements and changes to synteny, we chose to focus on intersects in narrow genomic regions. For the window sizes in the WGS data, all 100kb regions, whether

genetically different or not, were converted to GRanges objects. For ddRAD data, each SNP (regardless of significance of genetic differentiation), was considered the middle of a 20kb region (including 10kb upstream and 10kb downstream of the position), and these regions were converted to GRanges. For transcriptomic data, windows for each gene were considered to start 10kb upstream and end 10kb downstream of the annotated gene start and end, in order to include proximal promoters/enhancers. Lastly, the CpGs in the RRBS data, whether differently methylated or not, were represented by a 2bp region which was converted to GRanges. Only data aligning to placed scaffolds (excluding the mitochondrial genome) were analysed. The names and lengths of these placed scaffolds are provided in **Supplementary Table 4**. GRanges objects were then used in overlap analyses with the RegioneR package in R (Gel et al., 2016). The function *overlapPermTest()* was used with the parameters *evaluate.function=numOverlaps* and *randomize.function=resampleRegions* to generate, out of the whole datasets, 1000 random subsets of data with the same parameters (number of SNPs, size of regions, number of genes, etc) as the significant subsets. This function then compared genomic overlaps between two significant subsets (I.e. significant CpGs and differently expressed genes) to the various overlaps obtained between resampled subsets. The resampling is important as all the datasets except for WGS cover only a small part of the genome.

III. Results

WGS and ddRAD data display similar trends, and congruence of results

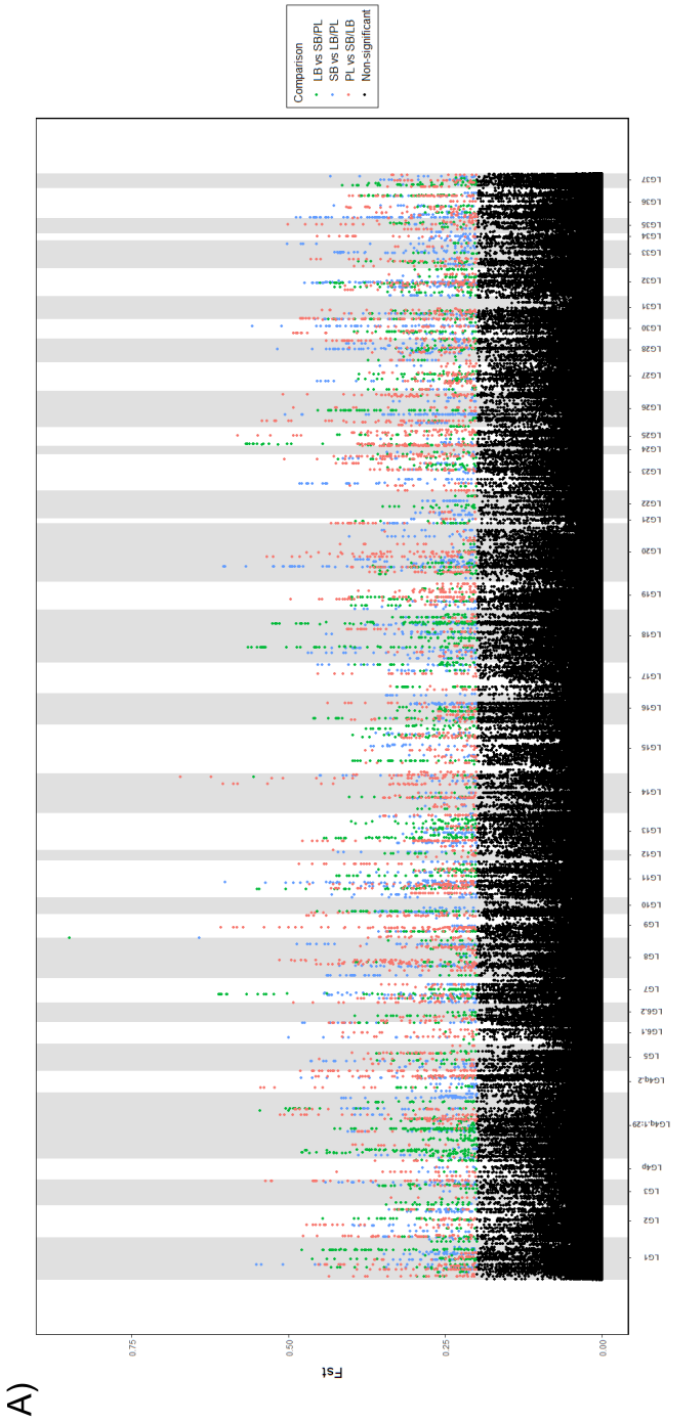
Investigating overlaps between genetic variation and functional genomics requires robust genetic data. To screen for and identify the genomic regions and loci which relate with morph divergence, we analyzed whole genome sequencing data from two individuals from each of the three sympatric morphs and ddRAD sequencing data from 30 to 37 individuals also from each (Ntotal = 98, see **Supplementary Table 2**, Xiao et al., unpublished). We used F_{ST} values from comparisons of one morph vs two others (e.g. PL vs SB and LB) to identify genomic regions of separation (pairwise comparisons gave very similar results, data not shown). As only two individuals per morph were whole genome sequenced, the differentiation was analyzed by calculating average F_{ST} values in 100kb sliding windows (shifted 10kb, here after called "windows", see methods). For the ddRAD data, which contained more individuals per morph, F_{ST} values were calculated for each of the 8472 polymorphic SNPs, and we assumed each of them tagged 20kb regions for the overlap analyses. For both datasets we identified the most significantly diverged regions by two methods, i) arbitrary threshold value (if window/SNP $F_{ST} > 0.2$), and ii) screening for outliers in each morph based on variance (if window/SNP $F_{ST} > 2\sigma$).

Both whole genome and ddRAD sequencing data identified multiple regions of genomic differentiation between morphs. The first approach ($F_{ST} > 0.2$) revealed more windows of divergence in the PL-morph for the WGS data, consistent with prior studies (Guðbrandsson et al., 2019, Xiao et al., unpublished) (**Supplementary Figure 4A**). The ddRAD results were in agreement, 1182 SNPs surpassed the $F_{ST} > 0.2$ threshold in the PL, but only 723 in SB and 466 in LB-charr (**Supplementary Figure 4B**). For both datasets, SNPs/regions with F_{ST} above 0.5 were found in all three morphs (**Supplementary Figure 4**). Thus, the WGS and ddRAD data identify similar trends.

The second approach ($F_{ST} > 0.199$ (2σ of the distribution)) using the WGS data revealed 8283 regions (out of 151540 on placed scaffolds), while from the ddRADseq data 530 SNPs (out of 8472 locations on placed scaffolds) were found to have an F_{ST} above 0.415 (**Supplementary Table 1**, **Supplementary Figure 1** and **2**). The same enrichment in PL-

charr was seen (374, vs 97 SB- and 59 LB-charr SNPs). Going forward, we will use this second approach to consider significant genomic differentiation at WGS windows or ddRAD positions.

The Salvelinus sp. genome assembly (assembly ASM291031v2) has 39 linkage groups. Morph specific markers or regions were found on all of them (**Figure 2**), though for the ddRAD SNPs some of the shorter linkage groups lacked markers associating with a particular morph. We attribute this to the lower genome coverage of ddRAD seq, as highly divergent WGS windows were found on all linkage groups (for each morph) confirming earlier conclusions (Gudbrandsson 2019, Brachmann 2022) that multiple genomic regions and loci differentiate these three morphs.


A)
Fst
0.00
0.25
0.50
0.75
Linkage group
Comparison
LB vs SB/PL
SB vs LB/PL
PL vs SB/LB
Non-significant

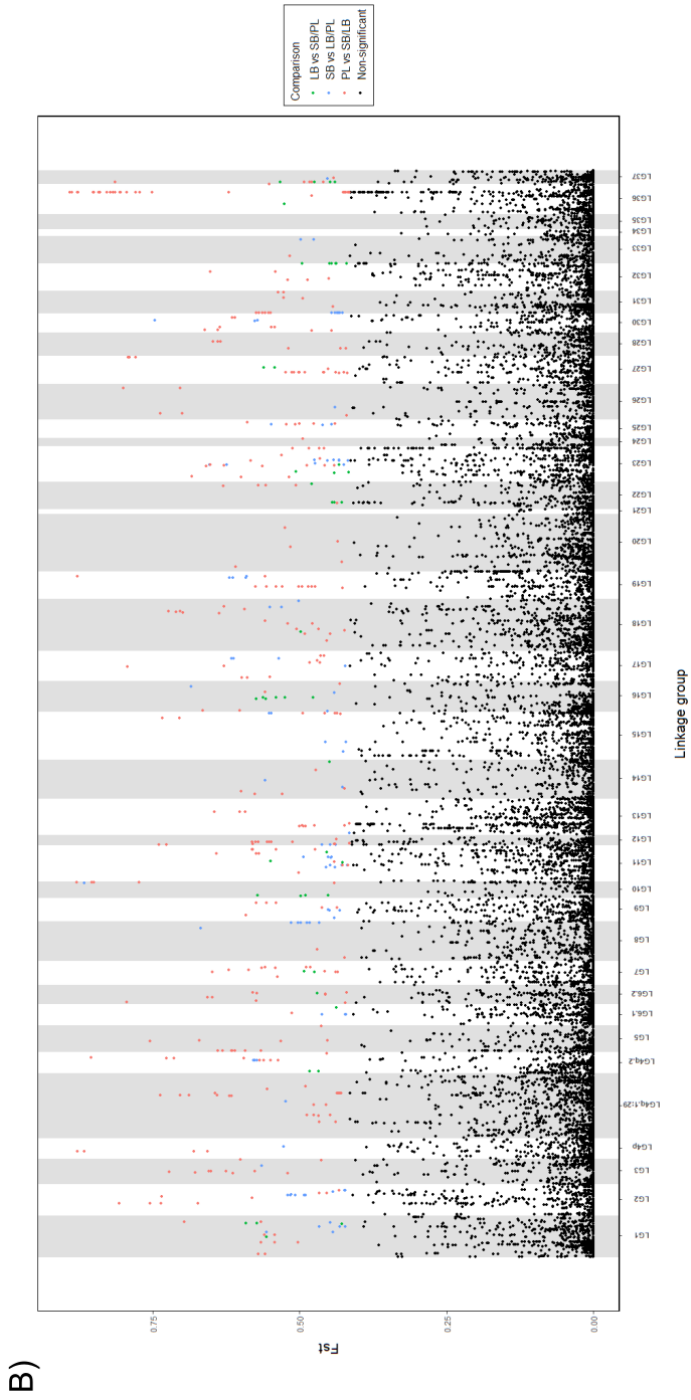


Figure 2: Genomic distribution of genetic differentiation (F_{ST}) between morphs, for A) WGS data and B) ddRADseq data. A) Each point represents a 100kb window, with color (green, blue or pink) applied to windows for which the F_{ST} value exceeded 2σ of the distribution (0.199) for one of the morphs. If multiple morphs had significant F_{ST} value for the same region, color represented the morph for which the F_{ST} value was the highest. B) Each point represents a genomic position, now of a SNP (same color scheme). The significant outliers had FST above 2σ of the distribution (0.415). Black points represent "non-significant" A) windows or B) SNPs.

The stepwise approach used in analyzing WGS data implies that some regions of differentiation are likely to span multiple neighboring windows. Thus, we asked how many independent genomic outlier regions were revealed by the WGS windows, and investigated their size. The lowest number of independent regions was in LB- (809) whereas PL-charr had the highest (953), and the majority of them was in the 100kb category (nearly 57% in all morphs, **Figure 3A** and **Supplemental Table 10**), indicating that most regions were contained inside a single window. Less than 10% of these regions were larger than 200kb, and the largest just over 500kb. There was no indication of differences in the WGS window size distributions by morph ($X^2 = 11.9$, df = 10, p = 0.29, **Supplementary Figure 10**). Several of the larger regions had high F_{ST} values. From the list of most extreme F_{ST} values we picked three examples that represent curious patterns. Region on linkage group NC_036849.1 had an extended peak with high FST (max = 0.61), that includes a dozen genes (**Figure 3B**). The region on linkage group NC_036860.1 represents a more complex case, as two peaks exceeded the threshold in SB-charr, but the F_{ST} values fluctuated around the 2σ threshold (**Figure 3C**). Finally on the tip of linkage group NC_036848.1 we saw a sharp peak with a signal that indicates a three-way separation of the morphs, as for each of them a window above 2σ was revealed. These are clear examples of regions of differentiation spanning multiple overlapping windows. However, as these are the minority, and because in some cases it is hard to define what is or isn't a wider peak of differentiation, we used all 100kb windows for the overlap analyses.

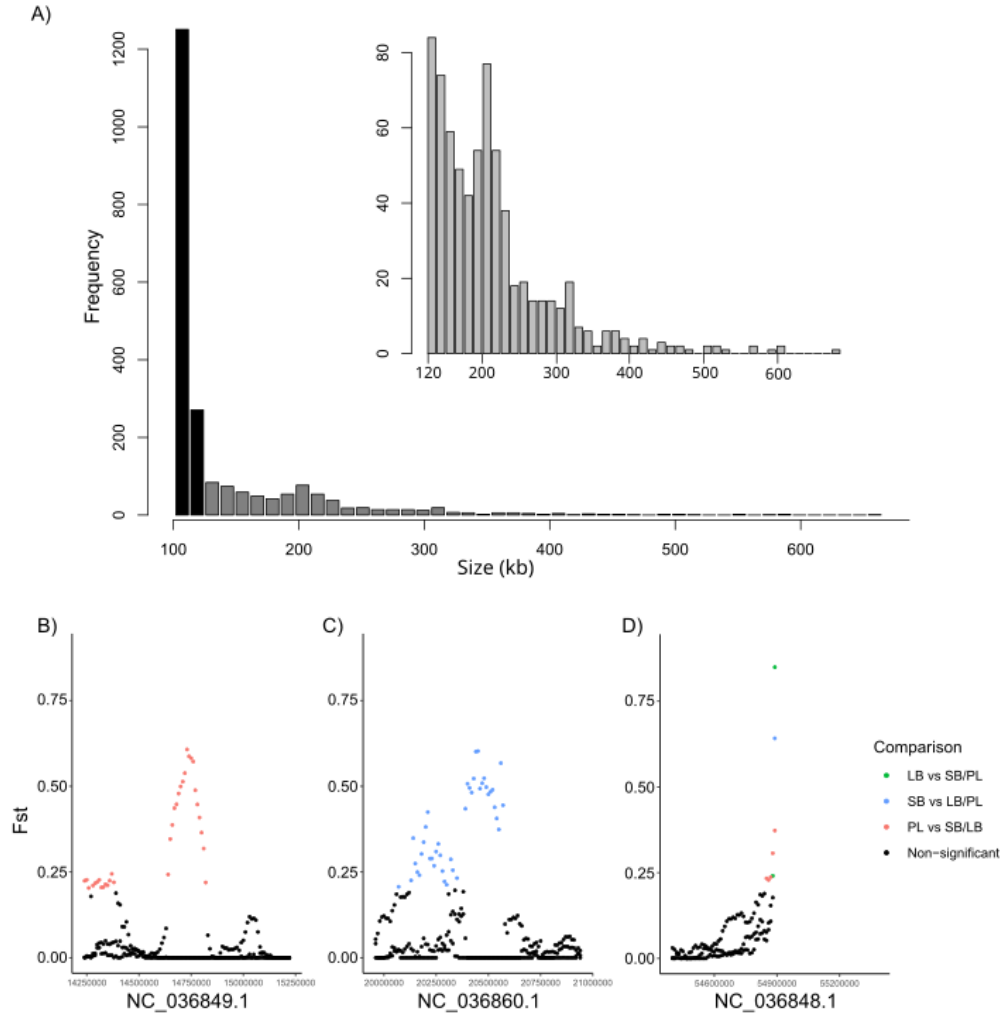


Figure 3: A) Distribution of sizes of regions showing genetic differentiation in the WGS data. Windows were 100kb, sliding 10kb, outlier regions had a single 100kb window. Inset are the rarer windows, 120 kb up to 600kb+. These distributions did not differ by morph, see Supplementary Figure 10. B-D Three regions of interest show wider peaks of differentiation spanning multiple overlapping windows. Colour code for B-D) follows Figure 2.

The SNPs from the ddRAD sequencing have high statistical power to reveal morph differences (because of the 30+ individuals sampled per morph), while the WGS windows add genomic coverage. There was very strong overlap of regions of genetic separation in WGS- and ddRAD-seq data ($p < 0.0001$, shown with permutation analyses, **Figure 4A**), which means that the two methods pick up many of the same genomic regions of differentiation between the three morphs.

Next, we tested if these regions of differentiation in the ddRAD and WGS datasets were congruent, for instance, if a PL SNP found in ddRAD data mapped to a PL-region in the WGS data. The analyses showed significant congruence within all morphs (**Figure 4B-D**). It is worth noting that for some regions or SNPs there was differentiation in more than one morph, sometimes in all three. For those cases, a SNP was categorized as derived in the

morph that had the highest F_{ST} value. A list of these markers can be found in **Supplementary Table 5**.

The highest Z-score, suggesting highest power between diverging ddRAD SNPs and diverging WGS regions, was for the PL morph, followed by SB and then LB-charr. This was predictable, as F_{ST} values were calculated for one morph against the other two morphs and as the PL morph is the only limnetic morph out of the three.

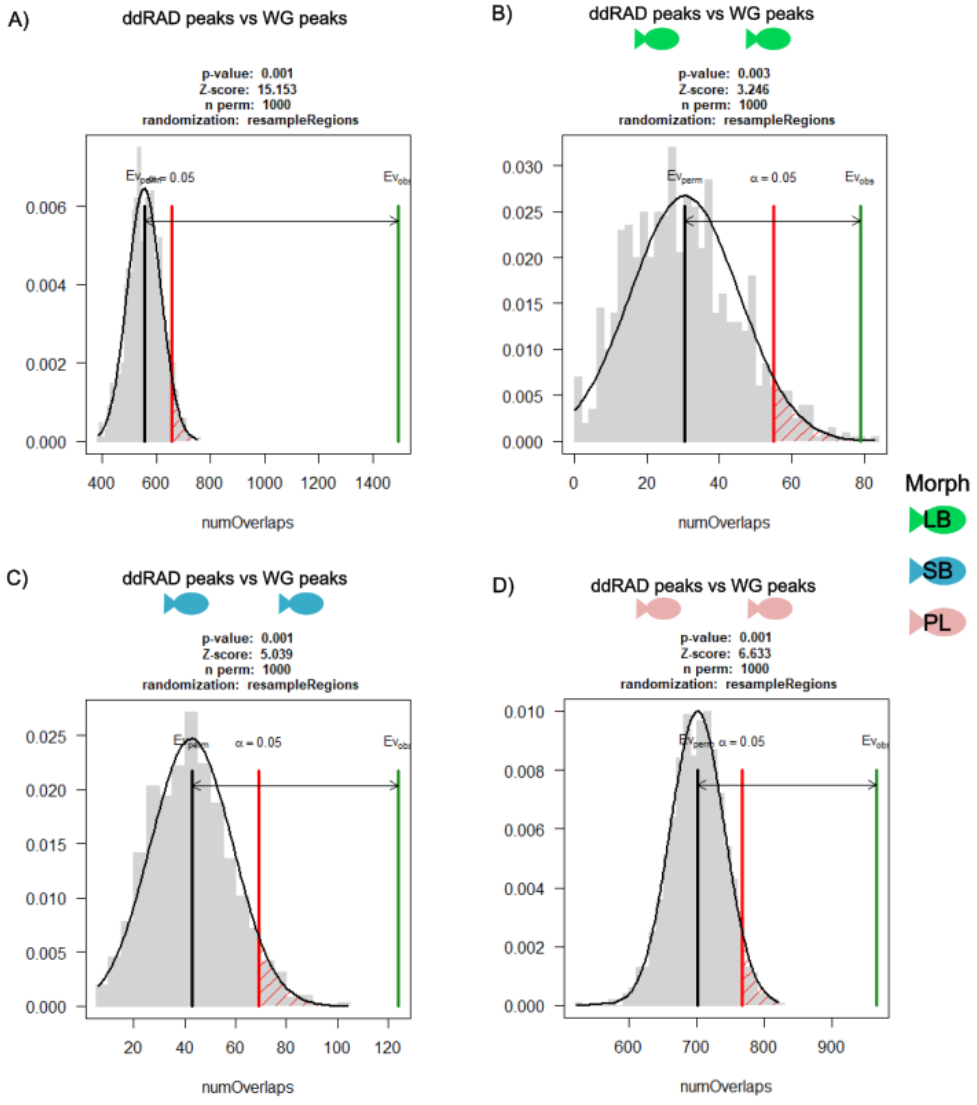


Figure 4: A) Regions of genetic differentiation between morphs identified through ddRAD-seq and WGS overlap more than expected by chance. B-D) Moreover, congruence was observed between regions of genetic differentiation identified by WGS and SNPs from ddRADseq for all three sympatric morphs (i.e. PL specific ddRAD SNPs and PL specific WGS regions of differentiation). For each of the morph-derived markers there was a significant excess of overlaps between WGS and ddRAD seq signals, compared to lists of markers and regions from different morphs (i.e. PL specific ddRAD SNPs and LB specific WGS regions). B) LB specific markers, C) SB specific markers and D) PL specific markers. Shown are the results of the permutation tests between ddRAD-seq SNPs and regions of differentiation identified through WGS. The expected and observed numbers of overlaps are shown with black (Ev perm) and green (Ev obs) lines, respectively. The red line (α = 0.05) represents the limit above which the observed number of overlaps is considered significant.

The overlap of the two datasets was, however, not complete. Of the 530 ddRAD outliers, 341 (~65%) were not located in WGS outlier regions. More obviously, 93% of the WGS outlier regions were not tagged by ddRAD outliers (**Supplementary Table 11**). This percentage, the strong correspondence between ddRAD and WGS data, and the genome-wide span of the latter, argues that F_{ST} windows analyses of the WGS data can be used to test for overlaps with functional data, despite the low number of samples sequenced. Thus, we set out to identify whether differentially expressed genes or differentially methylated regions were enriched with regions of genomic differentiation between morphs in the WGS data.

Overall enrichment of differentially expressed genes in diverging regions in some morphs but not all

First, we compared genes with differential expression between morphs during early development (from 100ts to 200ts, spanning the formation of key craniofacial structures, see Guðbrandsson et al., 2018) with regions of differentiation from the WGS data. For the entire dataset (not by specific morph) there was significant enrichment ($p < 0.001$, permutation test) of DE genes in these regions (**Supplementary Figure 5A**). Of the 1141 DE genes, 184 were found in these regions of divergence. This suggest just over 15% of the DE genes may be influenced by *cis*-variants, and the rest most likely *trans*- or longer range *cis*-elements. The same type of analyses with SNPs from the ddRAD-data did not yield significant enrichment (while the direction of effects was the same, towards enrichment (**Supplementary Figure 5B**)). The numbers of overlaps were quite low in this case, because of the relatively small genome coverage of each dataset.

We expected morph congruences between these datasets (e.g. genomic region with PL divergence to contain genes with DE in PL vs the other two). Curiously, such congruence was only observed for one morph, SB-charr, but not the other two. SB-specific regions of differentiation in the WGS data and SB-specific DE genes showed a significant overlap (**Figure 5**), that involved 22 unique genes (**Table 1**). Among them, genes coding for “Transcobolamin-2”, “ATPase family AAA domain-containing protein 2-like" and “collagen alpha-1(XIV) chain-like” were located in the 3 regions with the highest F_{ST} values out of this list ($F_{ST} \sim 0.42$). The LB-charr comparison yielded no significant overlap, but the PL-charr comparison was suggestive ($p = 0.09$), with the observed numbers leaning towards enrichment (**Figure 5**). As DE genes in those regions can be regarded as candidates for loci with functional divergence, we provide lists of both LB-DE genes in LB-regions and PL-DE genes in PL-regions in **Supplementary Tables 6** and **7**, respectively.

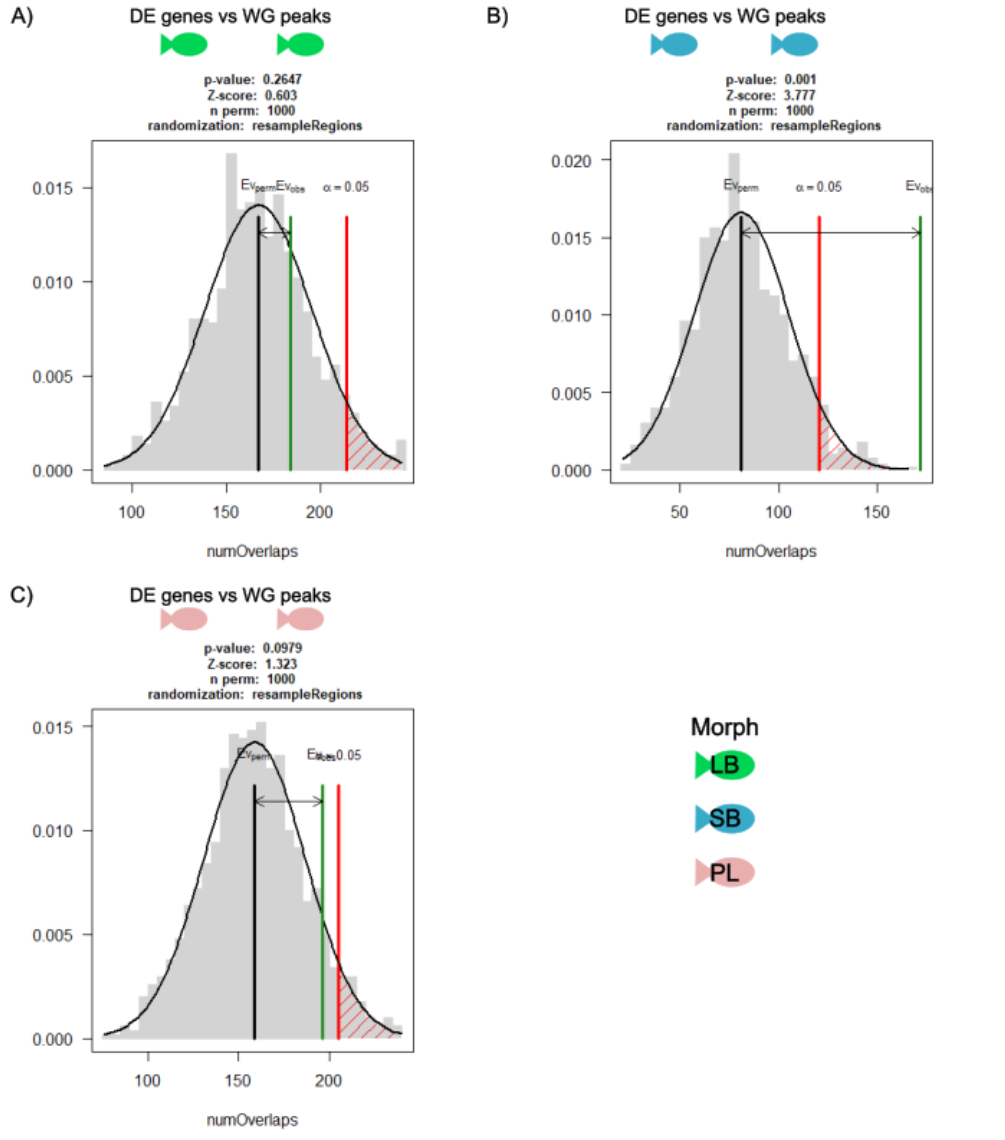


Figure 5: Tests of congruence between DE genes and regions of genetic differentiation (WGS), for the three morphs LB (A), SB (B) and PL (C). Shown are the results of the permutation tests for specific regions of differentiation and morph specific DE genes. Only for the SB-charr was this overlap more than expected by chance. The expected and observed numbers of overlaps are shown with black (Ev perm) and green (Ev obs) lines, respectively. The red line (α = 0.05) represents the limit above which the observed number of overlaps is considered significant.

Table 1: List of genes differentially expressed in SB-charr, located in regions of genetic differentiation that separate SB from LB- and PL-charr. These 22 genes reside on 11 distinct linkage groups.

LOC_ID	Gene name	Linkage group	Start	End	Gene ID
hdac8	histone deacetylase 8	NC_036840.1	19205169	19261361	111951965
LOC111955048	protein ALP1-like	NC_036840.1	32182266	32220902	111955048
ddx4	DEAD-box helicase 4	NC_036842.1	37005430	37036707	111961941
LOC111961502	NF-kappa-B inhibitor-like protein 1	NC_036842.1	13512918	13534965	111961502
LOC111962416	proline-serine-threonine phospha-tase-interacting pro-tein 1-like	NC_036842.1	68419535	68459614	111962416
LOC111962658	dickkopf-related pro-tein 3	NC_036842.1	83308871	83341355	111962658
rbm41	RNA binding motif protein 41	NC_036842.1	45161489	45201848	111962033
LOC111967243	stimulator of inter-feron genes protein	NC_036848.1	3319672	3343534	111967243
LOC111975342	volume-regulated anion channel subunit LRRC8C	NC_036856.1	28387655	28448518	111975342
LOC111980302	C3a anaphylatoxin chemotactic receptor	NC_036860.1	61752414	61773953	111980302
LOC111949715	transmembrane pro-tease serine 4	NC_036862.1	7356480	7389595	111949715
LOC111951657	brorin	NC_036865.1	5473629	5516981	111951657
LOC111952428	septin-7-like	NC_036866.1	17218669	17279110	111952428
LOC111952429	C-C motif chemokine 4	NC_036866.1	17259106	17280619	111952429
LOC111952507	large neutral amino acids transporter small subunit 1	NC_036866.1	16697635	16744298	111952507
LOC111953064	anoctamin-10-like	NC_036866.1	16922362	16949931	111953064
LOC111954731	ATPase family AAA domain-containing protein 2-like	NC_036869.1	17351485	17382035	111954731
LOC111955050	collagen alpha-1(XIV) chain-like	NC_036869.1	17201185	17318928	111955050
LOC111955204	zinc fingers and ho-meoboxes protein 1-like	NC_036869.1	17345996	17370352	111955204
LOC111956548	transmembrane pro-tein 56-B-like	NC_036871.1	13015356	13049196	111956548
LOC111957038	uncharacterized LOC111957038	NC_036871.1	28610236	28631735	111957038
LOC111957651	Transcobalamin-2	NC_036872.1	20524417	20556516	111957651

CpGs with differential methylation by morph may be enriched within regions of genetic separation

To test for co-occurence of genetic and methylation differentiation in charr morphs, we used developmental RRBS methylation data from the three clearly separated morphs SB-, LB- and PL-charr at three developmental timepoints: 50ts, 150ts, 200ts, spanning from gastrulation to late organogenesis (see methods, (Matlosz et al., 2022)). Using methods similar to the DE analyses above, we tested for overlaps of differentially methylated CpGs (between morphs) with regions of genetic differentiation. The results showed a trend that was not statistically significant, towards enrichment of differentially methylated CpGs in regions of genetic differentiation (from WGS data, $p = 0.0519$, **Supplementary Figure 6**).

We could not test for the morph congruence between methylation and genetic data, as the direction of methylation changes by morph varied in many cases through development (that is the morph separating from the others varied depending on the timepoint, Matlosz et al., 2022). We had previously identifed CpGs with different methylation levels by timepoints in development. This list we regarded as a control, and it did not show enrichment in regions of genetic separation (**Supplementary Figure 7**). While we wanted to test for enrichment of overlaps between differently expressed genes and differentially methylated CpGs, overlaps between both datasets were not numerous enough to draw conclusions (overlap observed = 1, average overlap from permutations = 4.5)(**Supplementary Figure 9**), and this question awaits further investigation with higher coverage datasets.

However, on a larger scale, 35 genes were found to reside in regions of differential methylation and allele frequency by all morphs (**Supplementary Table 8**). Among them, genes coding for "26S proteasome regulatory subunit 10B", "protein LBH" and "lysocardiolipin acyltransferase 1" were located in the region with the highest F_{ST} value out of this list (F_{ST} ~ 0.41, PL-specific signal).

We set out to confirm these patterns by also testing for enrichment of morph methylation CpGs near divergent loci identified by the ddRAD-sequencing. Compared to the average CpG in the methylation dataset, differentially methylated CpGs between the morphs overlapped more than expected by chance with SNPs (**Figure 6A**). However, the biological significance of this pattern is unclear as differentially methylated CpGs were found to overlap with all genomic positions sequenced by ddRAD, regardless of whether they were outlier SNPs or not (**Figure 6B**).

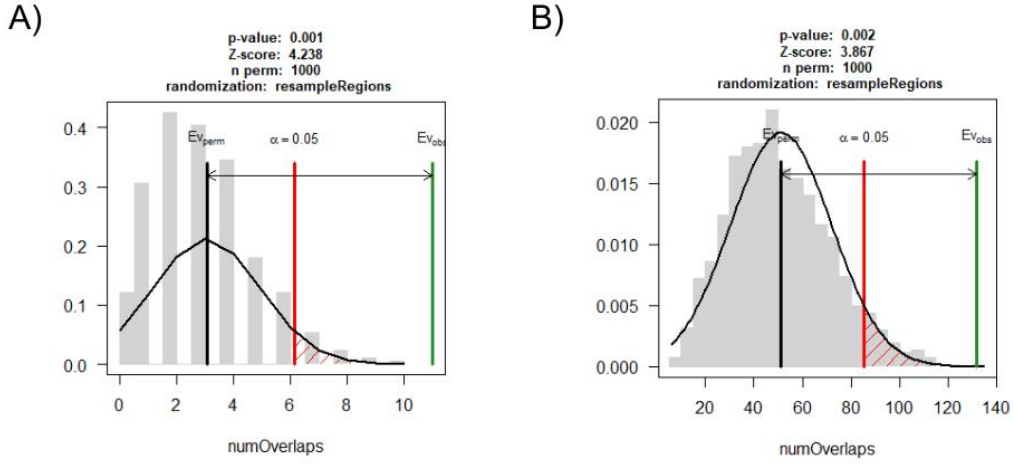


Figure 6: Differently methylated CpGs are enriched in all regions sequenced by RADseq, regardless of F_{ST} value. A) Differentially methylated CpGs between morphs are enriched close to outlier SNPs from the ddRAD. B) Differentially methylated CpGs between morphs are also enriched close to all genomic positions sequenced by ddRAD, regardless of F_{ST} value. The expected and observed numbers of overlaps are shown with black (Ev perm) and green (Ev obs) lines, respectively. The red line (α = 0.05) represents the limit above which the observed number of overlaps is considered significant.

The explanation for this somewhat counter-intuitive observation may be the protocol for ddRAD which was designed to bias the sequencing away from transposon derived sequences. Specifically, one of the enzymes commonly used for ddRAD sequencing, *Ape*KI, is blocked by CpG methylation and, as transposons and other repeat elements are

repressed by methylation, this leads to an increase in sequencing yield and coverage in non-repeat regions (Elshire et al., 2011; Schilling et al., 2014).

On the other hand, the RRBS dataset is not biased in the same way, as it does include CpGs that are fully methylated, no matter the cell type, such as those located in transposable elements that need to be repressed for genome safekeeping purposes. However, the subset of CpGs differentially methylated between the morphs are more likely located in regions that are unmethylated in at least some embryonic tissues (although a small proportion of repeat elements was shown to be differentially methylated between morphs (Matlosz et al., 2022)). This particular model suggests that differentially methylated CpGs are located less frequently in repeat element sequences and may be overrepresented in regions sequenced by the ddRAD approach (**Figure 7**).

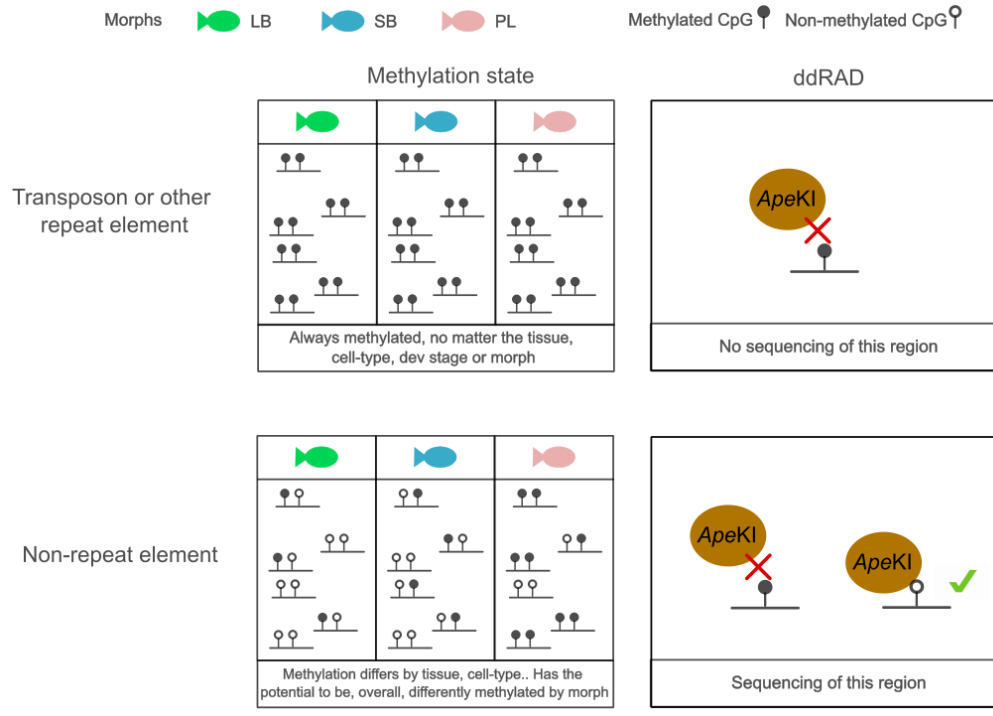


Figure 7: By design, ddRAD sequencing using the methylation sensitive enzyme ApeKI avoids the sequencing of repeat elements, therefore increasing the yield and coverage of non-repeat regions, which are of particular interest in most analysis of genomic variation (Elshire et al., 2011;Schilling et al., 2014). However, RRBS sequences CpGs regardless of genomic location (also sampling repeat elements), and as differential methylation is unlikely in repeat regions, then we expect CpGs with differential methylation to be more likely sequenced by ddRAD compared to the average non-differentially methylated CpG.

This model leads to a simple prediction about the distribution of average methylation per residue for two subsets of the RRBS data. Indeed, the differentially methylated CpGs between morphs possessed on average a lower methylation state than the whole dataset (**Figure 8**, **Supplementary Table 9**). Importantly, this pattern held for all three developmental timepoints.

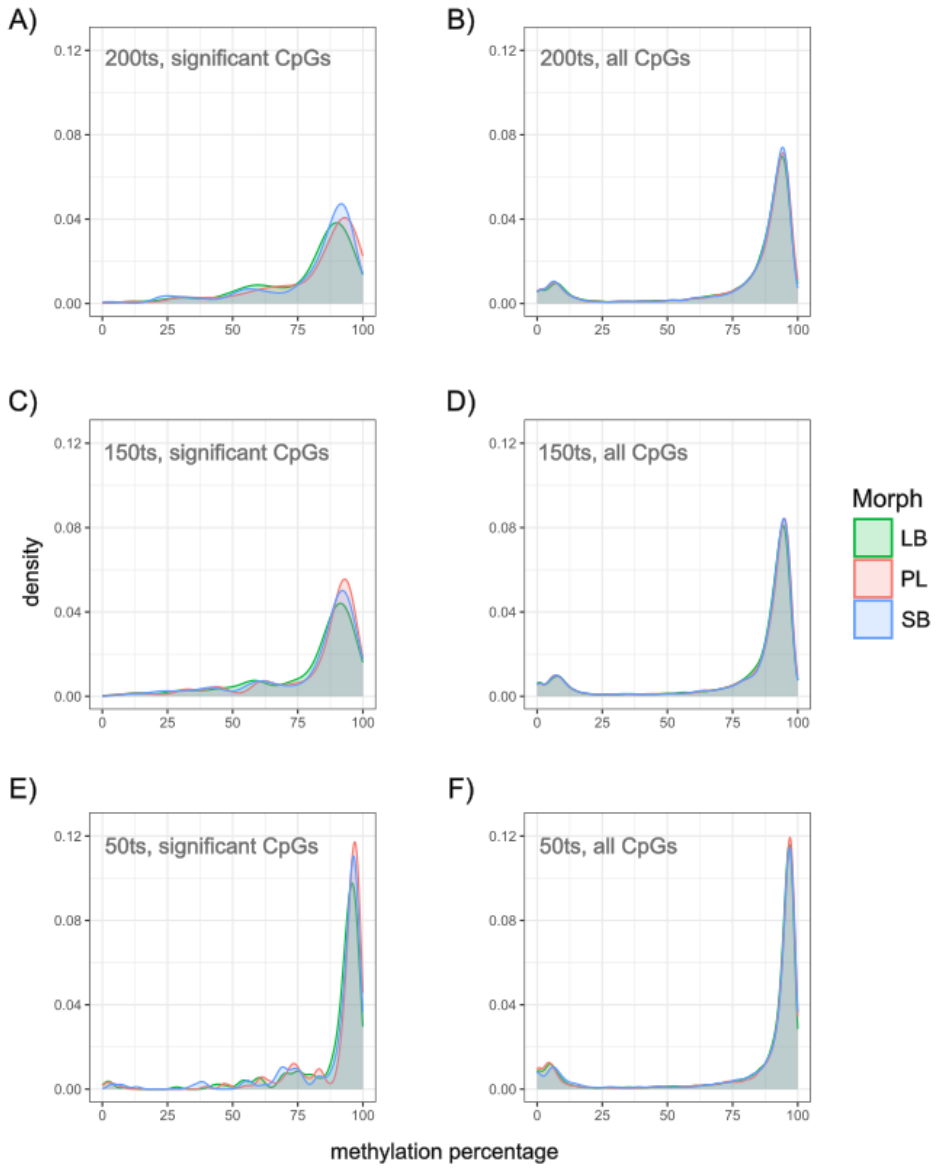


Figure 8: The distribution of average methylation percentage per residue varied between the significant CpGs (217 CpGs differently methylated by morph, left-side panels) and the entire RRBS dataset (9719 CpGs, right-side panels). This pattern held for three developmental timepoints: A) B) 200ts, C) D) 150ts, E) F) 50ts. Differently methylated CpGs had on average lower methylation than the entire RRBS dataset. KS tests of all CpGs and the differentially methylated revealed significant differences, Supplementary Table 9.

This bias set aside, we investigated whether ddRAD SNPs overlapped more than the whole ddRAD dataset with differently methylated CpGs between morphs. This addresses the question "Do differently methylated CpGs and SNPs overlap?" but asked in a way to avoid the previously described technical bias. The result showed no enrichment of overlaps between ddRAD SNPs and differentially methylated CpGs, compared with permutations on the whole ddRAD dataset (**Supplementary Figure 8**).

In sum, differently methylated CpGs were not enriched close to SNPs identified with ddRAD sequencing, and while they seemed to be enriched with regions of differentiation

identified through WGS, the signal was weak. An independent assessment of the methylation status over the whole genome is needed to confirm or refute such enrichment.

Two regions show genetic divergence as well as expression and methylation differences:

As both differentially methylated CpGs and DE genes showed, on their own, overlap with regions of genetic differentiation (WGS), we looked for regions containing both differentially methylated CpGs and DE genes. Interestingly, and despite the sparseness of the methylation and expression data, three genes were identified in overlaps from all three datasets (WGS window of genetic differentiation, DE genes and differentially methylated CpGs): *ALP1-like*, *PTCHD3-like* and *YME1L* (**Figure 9**). Importantly, all three genes show morph congruence between transcription and genetic variation, with *ALP1-like* located in a SB- diverged region and displaying differential expression in the SB morph (compared to PL and LB), while *PTCHD3-like* and *YME1L* are located in the same PL- specific region of differentiation and are differently expressed in the PL morph (compared to LB and SB). However, the methylation data in these regions was not congruent with the genetic and expression patterns. For instance, in the *ALP1-like* region the SB-charr diverged genetically from the other two morphs, but it was the LB-charr that had lower methylation levels at 200ts (**Figure 9**).

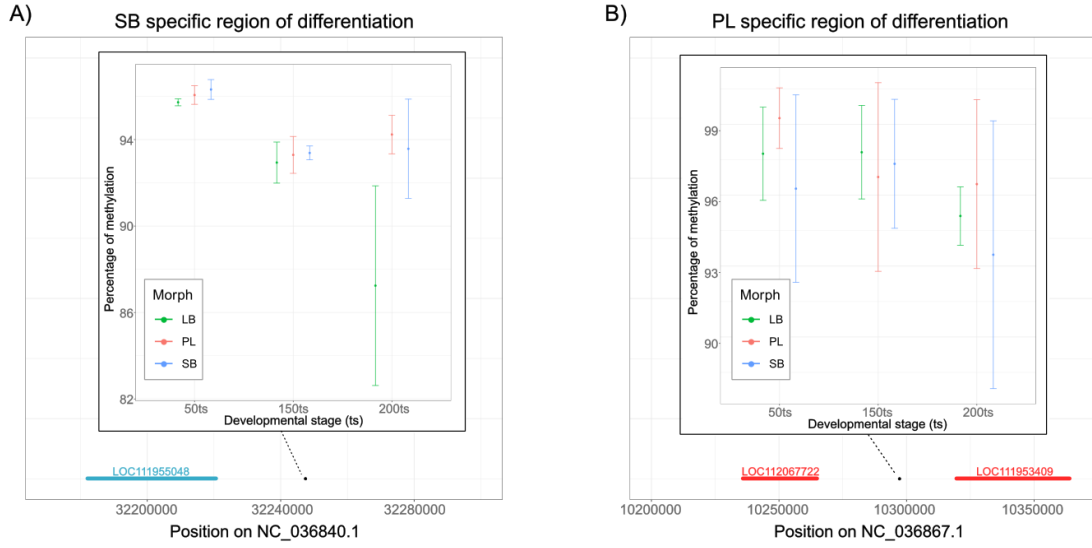


Figure 9: Regions of genetic differentiation containing differently methylated CpGs and DE genes. Methylation information is displayed for all three morphs and developmental stages. A) LOC111955048 (coding for protein ALP1-like) is differentially expressed in SB (compared to PL and LB). B) LOC112067722 (coding for patched domain-containing protein 3-like) and LOC111953409 (coding for ATP-dependent zinc metalloprotease YME1L1) are differentially expressed in PL (compared to SB and LB).

IV. Discussion

Studies of recently diverged populations, such as sympatric morphs of a given species, allows for the observation of evolutionary processes on a shorter time-scale, enabling identification of molecular differences (alleles, epi-alleles, etc..) that may be lost over longer periods of time. The ecology and morphology of the sympatric Arctic charr in Thingvallavatn has been studied for decades (Sandlund et al., 1992), but only with the advent of high throughput molecular methods has it been possible to systematically tackle

the genetic and developmental correlates of morph divergence (Macqueen et al., 2011; Kapralova et al., 2014; Salisbury et al., 2020). The molecular processes driving the phenotypic difference observed between the morphs have been addressed by genomic, transcriptomic and epigenomic studies (Guðbrandsson et al., 2018; Brachmann et al.,2022; Guðbrandsson et al., 2019; Matlosz et al., 2022), whereas the interactions between these layers have not been evaluated. Here, we take advantage of these datasets to investigate their overlaps.

Whole genome sequencing of two individuals per morph captures patterns of genetic divergence

Analyses of genetic variation (mostly SNPs) ascertained by reduced representation sequencing (eg ddRADseq) has revolutionized population genetic studies and enabled analyses of many species that lack proper reference genomes (Peterson et al., 2012). Whole genome sequencing is a stronger tool, as it captures more markers and the increased resolution greatly enhances the power to reveal other variations (inversions, CNVs and other complex chromosomal changes). The downside of WGS sequencing is cost, which limits its practicality for large scale population studies and species with large genomes, although bulk segregant or low coverage sequencing can leverage some of the power at lower cost (Pelletier et al., 2023; Frézal et al., 2018).

Two genome scale studies have addressed the divergence among the sympatric morphs in Lake Thingvallavatn.

First, Gudbrandsson et al (2019) estimated SNPs from a developmental transcriptome from the three morphs in focus here (PL-, LB and SB-charr). In total, 19,252 transcriptome SNPs were identified in 7,968 genes and just over 2300 were found to exceed a > 0.2 cutoff (note, the estimator used is a modified-F_{ST} and thus not comparable to regular F_{ST}, as each developmental transcriptome sample came from 3 pooled individuals and allele frequency estimates from mRNA reads are imprecise). KASP genotyping of 21 markers identified in the transcriptome confirmed these patterns, but also revealed that the fourth morph (PI-charr) is closely related to PL-charr, with some indications of gene flow with LB-charr. More recent analysis of ddRAD data confirms that picture (Xiao et al, unpublished).

Second, Brachmann et al (2022) utilized a 97k SNP chip, with about 15k markers being polymorphic across the four Icelandic lakes studied. The data confirmed the monophyly of the morphs within the lake previously revealed with microsatellite analysis (Kapralova et al 2011) exemplified by the fact that just under 4500 SNPs segregated within Lake Thingvallavatn. The data revealed the same patterns of genetic separation among the three morphs revealed by Guðbrandsson (2019), and demographic modeling was most consistent with an initial phase of allopatric separation and then secondary contact (Brachmann et al., 2022).

In this study we analysed SNPs derived from whole genome sequencing of 2 individuals of each morph (excluding the PI-charr), and found the data to confirm the hypothesis that WG SNPs picked up regions of morph differentiation revealed by ddRADseq derived SNP. We chose to analyze genetic differentiation in 100kb windows, that captured just under 40% of the peaks of differentiation that the ddRAD method highlighted. Several factors may be at play. A) Sampling effects on allele frequencies are expected to be very substantial, with only two individuals and thus four chromosomes per morph studied. Thus, regions with many linked markers with large allele frequency differences between morphs were most

likely to yield a signal. B) We settled for 100kb windows reflecting the assumption that regions with large allele frequency differences will be extended due to the impact of positive selection (leading to longer haplotypes, LD above average for other non-positively selected regions in the genome). This could be explored more systematically. C) The genome used is not from the study species *S. alpinus* (most likely a close relative, *S. malma*). Since the fourth whole genome duplication occurred in salmonids (Macqueen and Johnston, 2014) and because salmonid genomes have undergone considerable chromosomal evolution, it is possible that differences and/or errors in the genome used here could affect windows in the WGS analyses and thus blur some signals of divergence. Despite that, the WGS approach has clear merit as 93% of the outlier genomic regions it detects were not identified by ddRAD. Moreover, regions of differentiation identified by WGS and ddRADseq overlapped significantly, and did so in morph specific manner (PL-specific regions identified with ddRADseq overlapped with PL-specific regions identified with WGS). Reassuringly, both WGS and ddRADseq data confirmed the clear genetic separation of three morphs (Guðbrandsson 2019, Brachmann 2022), and that the PL-charr is the most diverged of the three (the two benthic morphs, LB- and SB-charr being more similar).

All this justified the application of the WGS data for comparisons with other datasets, such as differentially expressed genes and differentially methylated CpGs.

Morph specific congruence of expression and genetic divergence

While genetic and expression divergence can correlate (Kirst et al., 2005), this is not the rule (Wagner, 2000). We wondered whether expression and genetic separation would colocate generally in this dataset. Indeed they did, as differentially expressed genes are enriched in regions of genetic differentiation, with 184 out of the 1141 DE genes located in such regions. Despite the enrichment, this also means that the remaining DE genes are likely influenced by a) genetic variations in more distant *cis*--elements, b) direct transacting factors (e.g. sequence specific TFs or miRNAs), c) factors acting at earlier points (further upstream) in the developmental hierarchy or d) epigenetic signals acting in *cis*- or *trans*-.

These analyses did not consider the direction of effects, so we next asked about congruence. For example, do genes higher (or lower) expressed in SB-charr tend to be overrepresented in genomic regions that separate SB-charr from the other two? Much to our surprise congruence was only seen in one morph, the SB-charr, showing significant enrichment of DE genes in regions with genetic divergence. This suggests a higher proportion of expression divergence in this morph is driven by *cis*- genetic variation, relative to the other two morphs. Differences of this type have been reported in other systems (Hill, Zande and Wittkopp, 2021). For instance, functional differences of developmental genes in fruitflies influence the relative importance of *cis* vs *trans* effect on expression (Cartwright and Lott, 2020). Further analysis with better genomic coverage of genetic variants, finer scale analyses of expression (tissue and timepoint specific) are needed to confirm this (for example, see Garfield et al., 2013).

One of our main aims was to identify genes that might be involved in morph divergence. While many alleles can mediate fitness effect in the absence of expression divergence (Good, Hayden and Wheeler, 2006), we draw particular attention to differently expressed genes in regions with concordant pattern of morph divergence. Some of those may be biologically important regarding the craniofacial differences observed between the morphs, in head shape, lower jaw length and maxilla (Sandlund et al., 1992; Jónsdóttir et al., 2023).

For instance, *brorin* codes for a protein shown to inhibit the activity of the bone morphogenetic protein 2 (BMP2) (Koike et al., 2007), dickkopf-related protein 3 is part of a family of proteins shown to impact craniofacial bone formation during early embryonic development in mammals (Nie, 2005), and type X collagen plays a role in endochondral ossification (Shen, 2005), suggesting that *collagen alpha-1(XIV) chain-like* is also an interesting candidate gene for further studies. Investigating the specific locations or impact of SNPs and other genetic variants in these regions of differentiation, with respect to potential effects on expression, could provide valuable information.

Overlaps between regions of methylation and genetic divergence

Similar to genetic and expression correlations, interplays between genetic and methylation differentiation have been described. For instance, two thirds of differentially methylated sites between stickleback populations adapted to different salinities were found to be independent from this environmental variable and were instead suggested to be influenced by genetic variation *cis*- or *trans*-factors (Heckwolf et al., 2020). Similarly, methylome comparisons between the sympatric Gobio occitaniae and Phoxinus phoxinus suggest that genetic variation is responsible for most epigenetic differences observed between the two species (Fargeot et al., 2021). However, as methylation is dynamic, measuring such correlation is no easy task and studies investigating whether this type of interplay happens in *cis*- or *trans*- are rare. Moreover, methylation changes can either stay a short-term acclimatory response, appearing and disappearing within the lifetime of an individual, or become a more long-term signal in populations (Venney et al., 2023). Can genetic variation lead to both types of changes? Is the fixation of an epi-allele in species or groups the result of CNVs or other types of structural variants? Complicating matters further, 5mC, through its spontaneous deamination to thymine has been shown to be a source of genetic variation (Nabel et al., 2012), prompting the question of causality between both molecular layers. A recent study by Venney et al., (2024) using WGS and WGBS (Whole Genome Bisulfite Sequencing) found that outlier SNPs are enriched close to differentially methylated CpGs in the sympatric species Coregonus clupeaformis and Coregonus lavaretus, suggesting a more important *cis*- regulation between these molecular layers. However, more studies are needed to establish whether this trend is observed in more species and lineages. Only by understanding in which context genetic variants and methylation differences colocalize will we be able to answer the aforementioned questions.

For the Thingvallavatn Arctic charr system, our hypothesis was that genetic variation in regions of biological importance may have led to methylation differences due to evolutionary divergence (either due to positive selection or drift because of relaxation of purifying selection in certain lineages). The overlap analyses did not show significant overlap of SNPs and methylation signals, although differentially methylated CpGs tended towards enrichment in regions of genetic differentiation between the morphs. While this result needs to be confirmed, a certain number of regions showcasing methylation and genomic differences were unveiled and need further study, as does the expression of the 35 genes located in their vicinity. Perhaps even more interestingly, two differentially methylated CpGs were located in regions showing congruent genomic and expression divergence (regions close to *ALP1-like*, *PTCHD3-like* and *YME1L*), although the morph differences in methylation signals were not congruent with morph differences in genetic and transcriptomic signals. However, this is not so surprising as we previously showed that, out of 14 genes located in differentially methylated regions, only 4 showed correlation between methylation and expression (Matlosz et al., 2022). Furthermore, the differentially methylated CpGs closest to the three aforementioned genes are not so close (>20kb), and although they could still have an impact on enhancers, this is not necessarily the case.

It is again important to note that these datasets were mapped to the *Salvelinus sp*. genome, which differs from that of *Salvelinus alpinus*. Moreover, the assembly used is not complete, with thousands of unplaced scaffolds limiting analyses. Thus, reanalysis of the data using a more complete and correct assembly would help with addressing these research questions further. Importantly, we predict that most regions of overlap/enrichment identified here will be confirmed when a better *Salvelinus alpinus* genome assembly is completed, although their specific placement on chromosomes is likely to differ.

V. Author contributions

Zophonías O. Jónsson and Sigríður Franzdóttir prepared samples for genome sequencing. Han Xiao performed ddRAD-sequencing. Sébastien Matlosz performed reduced representation bisulfite sequencing. Alexander Guðjónsson remapped the RNA-sequencing data to the Salvelinus sp. genome and performed differential expression analysis. Lea Jerman-Plesec and Han Xiao analysed ddRAD sequencing data. Sébastien Matlosz reanalyzed methylation data to encompass only three morphs at three developmental timepoints. Arnar Pálsson did statistical analyses. Sébastien Matlosz performed overlap analyses and consolidated other analyses. Sébastien Matlosz, Arnar Pálsson, Sigríður Franzdóttir and Zophonías O. Jónsson wrote the manuscript.

References

Ahi, Ehsan Pashay, Jóhannes Guðbrandsson, Kalina H. Kapralova, Sigríður R. Franzdóttir, Sigurður S. Snorrason, Valerie H. Maier, and Zophonías O. Jónsson. 2013. 'Validation of Reference Genes for Expression Studies during Craniofacial Development in Arctic Charr'. PloS One 8 (6): e66389. https://doi.org/10.1371/journal.pone.0066389.

Ahi, Ehsan Pashay, Kalina Hristova Kapralova, Arnar Pálsson, Valerie Helene Maier, Jóhannes Gudbrandsson, Sigurdur S. Snorrason, Zophonías O. Jónsson, and Sigrídur Rut Franzdóttir. 2014. 'Transcriptional Dynamics of a Conserved Gene Expression Network Associated with Craniofacial Divergence in Arctic Charr'. EvoDevo 5 (1): 40. https://doi.org/10.1186/2041-9139-5-40.

Ahi, Ehsan Pashay, Sophie S. Steinhäuser, Arnar Pálsson, Sigrídur Rut Franzdóttir, Sigurdur S. Snorrason, Valerie H. Maier, and Zophonías O. Jónsson. 2015. 'Differential Expression of the Aryl Hydrocarbon Receptor Pathway Associates with Craniofacial Polymorphism in Sympatric Arctic Charr'. EvoDevo 6: 27. https://doi.org/10.1186/s13227-015-0022-6.

Almén, Markus Sällman, Sangeet Lamichhaney, Jonas Berglund, B. Rosemary Grant, Peter R. Grant, Matthew T. Webster, and Leif Andersson. 2016. 'Adaptive Radiation of Darwin's Finches Revisited Using Whole Genome Sequencing'. BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology 38 (1): 14–20. https://doi.org/10.1002/bies.201500079.

Anders, Simon, Paul Theodor Pyl, and Wolfgang Huber. 2015. 'HTSeq--a Python Framework to Work with High-Throughput Sequencing Data'. Bioinformatics (Oxford, England) 31 (2): 166–69. https://doi.org/10.1093/bioinformatics/btu638.

Brachmann, Matthew K., Kevin Parsons, Skúli Skúlason, Oscar Gaggiotti, and Moira Ferguson. 2022. 'Variation in the Genomic Basis of Parallel Phenotypic and Ecological Divergence in Benthic and Pelagic Morphs of Icelandic Arctic Charr (Salvelinus Alpinus)'. Molecular Ecology 31 (18): 4688–4706. https://doi.org/10.1111/mec.16625.

Cámara, Marina de la, Lieke Ponsioen, Quentin J. B. Horta-Lacueva, and Kalina H. Kapralova. 2023. 'The Dynamic Ontogenetic Shape Patterns of Adaptive Divergence and Sexual Dimorphism'. Evolutionary Biology 50 (2): 170–80. https://doi.org/10.1007/s11692-022-09592-y.

Cartwright, Emily L., and Susan E. Lott. 2020. 'Evolved Differences in Cis and Trans Regulation Between the Maternal and Zygotic mRNA Complements in the Drosophila Embryo'. Genetics 216 (3): 805–21. https://doi.org/10.1534/genetics.120.303626.

Catchen, Julian M., Angel Amores, Paul Hohenlohe, William Cresko, and John H. Postlethwait. 2011. 'Stacks: Building and Genotyping Loci de Novo from Short-Read Sequences'. G3 (Bethesda, Md.) 1 (3): 171–82. https://doi.org/10.1534/g3.111.000240.

Danecek, Petr, Adam Auton, Goncalo Abecasis, Cornelis A. Albers, Eric Banks, Mark A. DePristo, Robert E. Handsaker, et al. 2011. 'The Variant Call Format and VCFtools'. Bioinformatics (Oxford, England) 27 (15): 2156–58. https://doi.org/10.1093/bioinformatics/btr330.

Elshire, Robert J., Jeffrey C. Glaubitz, Qi Sun, Jesse A. Poland, Ken Kawamoto, Edward S. Buckler, and Sharon E. Mitchell. 2011. 'A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species'. PLOS ONE 6 (5): e19379. https://doi.org/10.1371/journal.pone.0019379.

Eskafi, Nahal. 2022. 'Study of Ecological and Developmental Divergence of Icelandic Arctic Charr (Salvelinus Alpinus) at Different Scales'.

Fargeot, Laura, Géraldine Loot, Jérôme G. Prunier, Olivier Rey, Charlotte Veyssière, and Simon Blanchet. 2021. 'Patterns of Epigenetic Diversity in Two Sympatric Fish Species: Genetic vs. Environmental Determinants'. Genes 12 (1): 107. https://doi.org/10.3390/genes12010107.

Foote, Chris J., Chris C. Wood, and Ruth E. Withler. 1989. 'Biochemical Genetic Comparison of Sockeye Salmon and Kokanee, the Anadromous and Nonanadromous Forms of Oncorhynchus Nerka'. Canadian Journal of Fisheries and Aquatic Sciences 46 (1): 149–58. https://doi.org/10.1139/f89-020.

Frézal, Lise, Emilie Demoinet, Christian Braendle, Eric Miska, and Marie-Anne Félix. 2018. 'Natural Genetic Variation in a Multigenerational Phenotype in C. Elegans'. Current Biology: CB 28 (16): 2588-2596.e8. https://doi.org/10.1016/j.cub.2018.05.091.

Garfield, David A., Daniel E. Runcie, Courtney C. Babbitt, Ralph Haygood, William J. Nielsen, and Gregory A. Wray. 2013. 'The Impact of Gene Expression Variation on the Robustness and Evolvability of a Developmental Gene Regulatory Network'. PLOS Biology 11 (10): e1001696. https://doi.org/10.1371/journal.pbio.1001696.

Gel, Bernat, Anna Díez-Villanueva, Eduard Serra, Marcus Buschbeck, Miguel A. Peinado, and Roberto Malinverni. 2016. 'regioneR: An R/Bioconductor Package for the Association

Analysis of Genomic Regions Based on Permutation Tests'. Bioinformatics 32 (2): 289–91. https://doi.org/10.1093/bioinformatics/btv562.

Gíslason, Daví? 1998. 'Genetic and Morphological Variation in Polymorphic Arctic Charr, Salvelinus Alpinus, from Icelandic Lakes'. University of Guelph. https://hdl.handle.net/10214/21675.

Good, Jeffrey M., Celine A. Hayden, and Travis J. Wheeler. 2006. 'Adaptive Protein Evolution and Regulatory Divergence in Drosophila'. Molecular Biology and Evolution 23 (6): 1101–3. https://doi.org/10.1093/molbev/msk002.

Gudbrandsson, Johannes, Ehsan P. Ahi, Sigridur R. Franzdottir, Kalina H. Kapralova, Bjarni K. Kristjansson, S. Sophie Steinhaeuser, Valerie H. Maier, et al. 2016. 'The Developmental Transcriptome of Contrasting Arctic Charr (Salvelinus Alpinus) Morphs'. F1000Research 4 (December): 136. https://doi.org/10.12688/f1000research.6402.3.

Guðbrandsson, Jóhannes, Sigríður Rut Franzdóttir, Bjarni Kristófer Kristjánsson, Ehsan Pashay Ahi, Valerie Helene Maier, Kalina Hristova Kapralova, Sigurður Sveinn Snorrason, Zophonías Oddur Jónsson, and Arnar Pálsson. 2018. 'Differential Gene Expression during Early Development in Recently Evolved and Sympatric Arctic Charr Morphs'. PeerJ 6: e4345. https://doi.org/10.7717/peerj.4345.

Guðbrandsson, Jóhannes, Sigríður Rut Franzdóttir, Bjarni Kristófer Kristjánsson, Ehsan Pashay Ahi, Valerie Helene Maier, Sigurður Sveinn Snorrason, Zophonías Oddur Jónsson, and Arnar Pálsson. 2017. 'Differential Gene Expression during Early Development in Recently Evolved and Sympatric Arctic Charr Morphs'. e3318v1. PeerJ Inc. https://doi.org/10.7287/peerj.preprints.3318v1.

Guðbrandsson, Jóhannes, Kalina H. Kapralova, Sigríður R. Franzdóttir, Þóra Margrét Bergsveinsdóttir, Völundur Hafstað, Zophonías O. Jónsson, Sigurður S. Snorrason, and Arnar Pálsson. 2019. 'Extensive Genetic Differentiation between Recently Evolved Sympatric Arctic Charr Morphs'. Ecology and Evolution 9 (19): 10964–83. https://doi.org/10.1002/ece3.5516.

Heckwolf, Melanie J., Britta S. Meyer, Robert Häsler, Marc P. Höppner, Christophe Eizaguirre, and Thorsten B. H. Reusch. 2020. 'Two Different Epigenetic Information Channels in Wild Three-Spined Sticklebacks Are Involved in Salinity Adaptation'. Science Advances 6 (12): eaaz1138. https://doi.org/10.1126/sciadv.aaz1138.

Hill, Mark S., Pétra Vande Zande, and Patricia J. Wittkopp. 2021. 'Molecular and Evolutionary Processes Generating Variation in Gene Expression'. Nature Reviews. Genetics 22 (4): 203–15. https://doi.org/10.1038/s41576-020-00304-w.

Horta-Lacueva, Quentin. 2022. 'Evolution of Reproductive Isolation in Sympatric Arctic Charr Morphs (Salvelinus Alpinus)'.

Horta-Lacueva, Quentin J.-B., Sigurður S. Snorrason, Michael B. Morrissey, Camille A.-L. Leblanc, and Kalina H. Kapralova. 2021. 'Multivariate Analysis of Morphology, Behaviour, Growth and Developmental Timing in Hybrids Brings New Insights into the Divergence of Sympatric Arctic Charr Morphs'. BMC Ecology and Evolution 21 (1): 170. https://doi.org/10.1186/s12862-021-01904-8.

Huber, Sarah K., Luis Fernando De León, Andrew P. Hendry, Eldredge Bermingham, and Jeffrey Podos. 2007. ‘Reproductive Isolation of Sympatric Morphs in a Population of Darwin’s Finches’. Proceedings. Biological Sciences 274 (1619): 1709–14. https://doi.org/10.1098/rspb.2007.0224.

Jónsdóttir, Guðbjörg Ósk, Laura-Marie von Elm, Finnur Ingimarsson, Samuel Tersigni, Sigurður Sveinn Snorrason, Arnar Pálsson, and Sarah Elisabeth Steele. 2023. ‘Diversity in the Internal Functional Feeding Elements of Sympatric Morphs of Arctic Charr (Salvelinus Alpinus)’. bioRxiv. https://doi.org/10.1101/2023.02.17.528955.

Kapralova, K. H., M. B. Morrissey, B. K. Kristjánsson, G. Á Ólafsdóttir, S. S. Snorrason, and M. M. Ferguson. 2011. ‘Evolution of Adaptive Diversity and Genetic Connectivity in Arctic Charr (Salvelinus Alpinus) in Iceland’. Heredity 106 (3): 472–87. https://doi.org/10.1038/hdy.2010.161.

Kapralova, Kalina H., Sigrídur Rut Franzdóttir, Hákon Jónsson, Sigurður S. Snorrason, and Zophonías O. Jónsson. 2014. ‘Patterns of miRNA Expression in Arctic Charr Development’. PloS One 9 (8): e106084. https://doi.org/10.1371/journal.pone.0106084.

Kirst, Matias, Christopher J. Basten, Alexander A. Myburg, Zhao-Bang Zeng, and Ronald R. Sederoff. 2005. ‘Genetic Architecture of Transcript-Level Variation in Differentiating Xylem of a Eucalyptus Hybrid’. Genetics 169 (4): 2295–2303. https://doi.org/10.1534/genetics.104.039198.

Koike, Naomi, Yoshiaki Kassai, Yuya Kouta, Hiroyuki Miwa, Morichika Konishi, and Nobuyuki Itoh. 2007. ‘Brorin, a Novel Secreted Bone Morphogenetic Protein Antagonist, Promotes Neurogenesis in Mouse Neural Precursor Cells*’. Journal of Biological Chemistry 282 (21): 15843–50. https://doi.org/10.1074/jbc.M701570200.

Krueger, Felix, Frankie James, Phil Ewels, Ebrahim Afyounian, Michael Weinstein, Benjamin Schuster-Boeckler, Gert Hulselmans, and sclamons. 2023. ‘FelixKrueger/TrimGalore’. Zenodo. https://doi.org/10.5281/zenodo.7598955.

Lagunas, Marcos, Arnar Pálsson, Benóný Jónsson, Magnús Jóhannsson, Zophonías O. Jónsson, and Sigurður S. Snorrason. 2023. ‘Genetic Structure and Relatedness of Brown Trout (Salmo Trutta) Populations in the Drainage Basin of the Ölfusá River, South-Western Iceland’. PeerJ 11 (September): e15985. https://doi.org/10.7717/peerj.15985.

Lawrence, Michael, Wolfgang Huber, Hervé Pagès, Patrick Aboyoun, Marc Carlson, Robert Gentleman, Martin T. Morgan, and Vincent J. Carey. 2013. ‘Software for Computing and Annotating Genomic Ranges’. PLOS Computational Biology 9 (8): e1003118. https://doi.org/10.1371/journal.pcbi.1003118.

Li, Heng, and Richard Durbin. 2009. ‘Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform’. Bioinformatics (Oxford, England) 25 (14): 1754–60. https://doi.org/10.1093/bioinformatics/btp324.

Love, Michael I., Wolfgang Huber, and Simon Anders. 2014. ‘Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2’. Genome Biology 15 (12): 550. https://doi.org/10.1186/s13059-014-0550-8.

Macqueen, Daniel J., and Ian A. Johnston. 2014. ‘A Well-Constrained Estimate for the Timing of the Salmonid Whole Genome Duplication Reveals Major Decoupling from Species Diversification’. Proceedings of the Royal Society B: Biological Sciences 281 (1778): 20132881. https://doi.org/10.1098/rspb.2013.2881.

Macqueen, Daniel J., Bjarni K. Kristjánsson, Charles G. M. Paxton, Vera L. A. Vieira, and Ian A. Johnston. 2011. ‘The Parallel Evolution of Dwarfism in Arctic Charr Is Accompanied by Adaptive Divergence in mTOR-Pathway Gene Expression’. Molecular Ecology 20 (15): 3167–84. https://doi.org/10.1111/j.1365-294X.2011.05172.x.

Major, Kaley M., Bethany M. DeCourten, Jie Li, Monica Britton, Matthew L. Settles, Alvine C. Mehinto, Richard E. Connon, and Susanne M. Brander. 2020. ‘Early Life Exposure to Environmentally Relevant Levels of Endocrine Disruptors Drive Multigenerational and Transgenerational Epigenetic Changes in a Fish Model’. Frontiers in Marine Science 7. https://www.frontiersin.org/articles/10.3389/fmars.2020.00471.

Malmquist, H. J., S. S. Snorrason, S. Skúlason, B. Jonsson, O. T. Sandlund, and P. M. Jonasson. 1992. ‘Diet Differentiation in Polymorphic Arctic Charr in Thingvallavatn, Iceland’. Journal of Animal Ecology 61 (1): 21–35. https://doi.org/10.2307/5505.

Markevich, Grigorii, Evgeny Esin, and Liudmila Anisimova. 2018. ‘Basic Description and Some Notes on the Evolution of Seven Sympatric Morphs of Dolly Varden Salvelinus Malma from the Lake Kronotskoe Basin’. Ecology and Evolution 8 (5): 2554–67. https://doi.org/10.1002/ece3.3806.

Martin, Arnaud, and Virginie Courtier-Orgogozo. 2017. ‘Morphological Evolution Repeatedly Caused by Mutations in Signaling Ligand Genes’. In Diversity and Evolution of Butterfly Wing Patterns: An Integrative Approach, edited by Toshio Sekimura and H. Frederik Nijhout, 59–87. Singapore: Springer. https://doi.org/10.1007/978-981-10-4956-9_4.

Martin, Marcel. 2011. ‘Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads’. EMBnet.Journal 17 (1): 10–12. https://doi.org/10.14806/ej.17.1.200.

Matlosz, Sébastien, Benjamín Sigurgeirsson, Sigrídur Rut Franzdóttir, Arnar Pálsson, and Zophonías O. Jónsson. 2022. ‘DNA Methylation Differences during Development Distinguish Sympatric Morphs of Arctic Charr (Salvelinus Alpinus)’. Molecular Ecology 31 (18): 4739–61. https://doi.org/10.1111/mec.16620.

Moroz, Leonid L., and Andrea B. Kohn. 2016. ‘Independent Origins of Neurons and Synapses: Insights from Ctenophores’. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 371 (1685): 20150041. https://doi.org/10.1098/rstb.2015.0041.

Morris, Zachary S., and Arhat Abzhanov. 2021. ‘Heading for Higher Ground: Developmental Origins and Evolutionary Diversification of the Amniote Face’. Current Topics in Developmental Biology 141: 241–77. https://doi.org/10.1016/bs.ctdb.2020.12.003.

Nabel, Christopher S., Sara A. Manning, and Rahul M. Kohli. 2012. ‘The Curious Chemical Biology of Cytosine: Deamination, Methylation and Oxidation as Modulators of

Genomic Potential'. ACS Chemical Biology 7 (1): 20–30. https://doi.org/10.1021/cb2002895.

Nie, Xuguang. 2005. 'Dkk1, -2, and -3 Expression in Mouse Craniofacial Development'. Journal of Molecular Histology 36 (5): 367–72. https://doi.org/10.1007/s10735-005-9008-3.

Oliver, Thomas, Tom D. Kim, Joko P. Trinugroho, Violeta Cordón-Preciado, Nitara Wijayatilake, Aaryan Bhatia, A. William Rutherford, and Tanai Cardona. 2023. 'The Evolution and Evolvability of Photosystem II'. Annual Review of Plant Biology 74 (May): 225–57. https://doi.org/10.1146/annurev-arplant-070522-062509.

Olson-Manning, Carrie F., Maggie R. Wagner, and Thomas Mitchell-Olds. 2012. 'Adaptive Evolution: Evaluating Empirical Support for Theoretical Predictions'. Nature Reviews Genetics 13 (12): 867–77. https://doi.org/10.1038/nrg3322.

Pelletier, Katie, William R. Pitchers, Anna Mammel, Emmalee Northrop-Albrecht, Eladio J. Márquez, Rosa A. Moscarella, David Houle, and Ian Dworkin. 2023. 'Complexities of Recapitulating Polygenic Effects in Natural Populations: Replication of Genetic Effects on Wing Shape in Artificially Selected and Wild-Caught Populations of Drosophila Melanogaster'. Genetics 224 (3): iyad050. https://doi.org/10.1093/genetics/iyad050.

Peterson, Brant K., Jesse N. Weber, Emily H. Kay, Heidi S. Fisher, and Hopi E. Hoekstra. 2012. 'Double Digest RADseq: An Inexpensive Method for de Novo SNP Discovery and Genotyping in Model and Non-Model Species'. PloS One 7 (5): e37135. https://doi.org/10.1371/journal.pone.0037135.

Pierron, Fabien, Sophie Lorioux, Débora Héroin, Guillemine Daffe, Bruno Etcheverria, Jérôme Cachot, Bénédicte Morin, Sylvie Dufour, and Patrice Gonzalez. 2021. 'Transgenerational Epigenetic Sex Determination: Environment Experienced by Female Fish Affects Offspring Sex Ratio'. Environmental Pollution (Barking, Essex: 1987) 277 (May): 116864. https://doi.org/10.1016/j.envpol.2021.116864.

Reyes-Olalde, J. Irepan, Mitsuhiro Aida, and Stefan de Folter. 2023. 'An Evo-Devo View of the Gynoecium'. Journal of Experimental Botany 74 (14): 3933–50. https://doi.org/10.1093/jxb/erad135.

Salisbury, Sarah J., Gregory R. McCracken, Robert Perry, Donald Keefe, Kara K.S. Layton, Tony Kess, Cameron M. Nugent, et al. 2020. 'Limited Genetic Parallelism Underlies Recent, Repeated Incipient Speciation in Geographically Proximate Populations of an Arctic Fish (Salvelinus Alpinus)'. Molecular Ecology 29 (22): 4280–94. https://doi.org/10.1111/mec.15634.

Sandlund, Odd Ter, Bror Jonsson, Hilmar J. Malmquist, Rolf Gydemo, Torfinn Lindem, Skúli Skúlason, Sigurdur S. Snorrason, and Pétur M. Jónasson. 1987. 'Habitat Use of Arctic charrSalvelinus Alpinus in Thingvallavatn, Iceland'. Environmental Biology of Fishes 20 (4): 263–74. https://doi.org/10.1007/BF00005297.

Sandlund, Odd Terje, Karl Gunnarsson, Pétur M. Jónasson, Bror Jonsson, Torfinn Lindem, Kristinn P. Magnússon, Hilmar J. Malmquist, Hrefna Sigurjónsdóttir, Skúli Skúlason, and Sigurdur S. Snorrason. 1992. 'The Arctic Charr Salvelinus Alpinus in Thingvallavatn'. Oikos 64 (1/2): 305–51. https://doi.org/10.2307/3545056.

Shen, G. 2005. ‘The Role of Type X Collagen in Facilitating and Regulating Endochondral Ossification of Articular Cartilage’. Orthodontics & Craniofacial Research 8 (1): 11–17. https://doi.org/10.1111/j.1601-6343.2004.00308.x.

Skúlason, S., S. S. Snorrason, D. L. G. Noakes, M. M. Ferguson, and H. J. Malmquist. 1989. ‘Segregation in Spawning and Early Life History among Polymorphic Arctic Charr, Salvelinus Alpinus, in Thingvallavatn, Iceland’. Journal of Fish Biology 35 (sA): 225–32. https://doi.org/10.1111/j.1095-8649.1989.tb03065.x.

SKÚLASON, SKÚLI, DAVID L. G. NOAKES, and SIGURDUR S. SNORRASON. 1989. ‘Ontogeny of Trophic Morphology in Four Sympatric Morphs of Arctic Charr Salvelinus Alpinus in Thingvallavatn, Iceland’. Biological Journal of the Linnean Society 38 (3): 281–301. https://doi.org/10.1111/j.1095-8312.1989.tb01579.x.

Thompson, Neil F., Eric C. Anderson, Anthony J. Clemento, Matthew A. Campbell, Devon E. Pearse, James W. Hearsey, Andrew P. Kinziger, and John Carlos Garza. 2020. ‘A Complex Phenotype in Salmon Controlled by a Simple Change in Migratory Timing’. Science 370 (6516): 609–13. https://doi.org/10.1126/science.aba9059.

Venney, Clare J, Dafni Anastasiadi, Maren Wellenreuther, and Louis Bernatchez. 2023. ‘The Evolutionary Complexities of DNA Methylation in Animals: From Plasticity to Genetic Evolution’. Genome Biology and Evolution 15 (12): evad216. https://doi.org/10.1093/gbe/evad216.

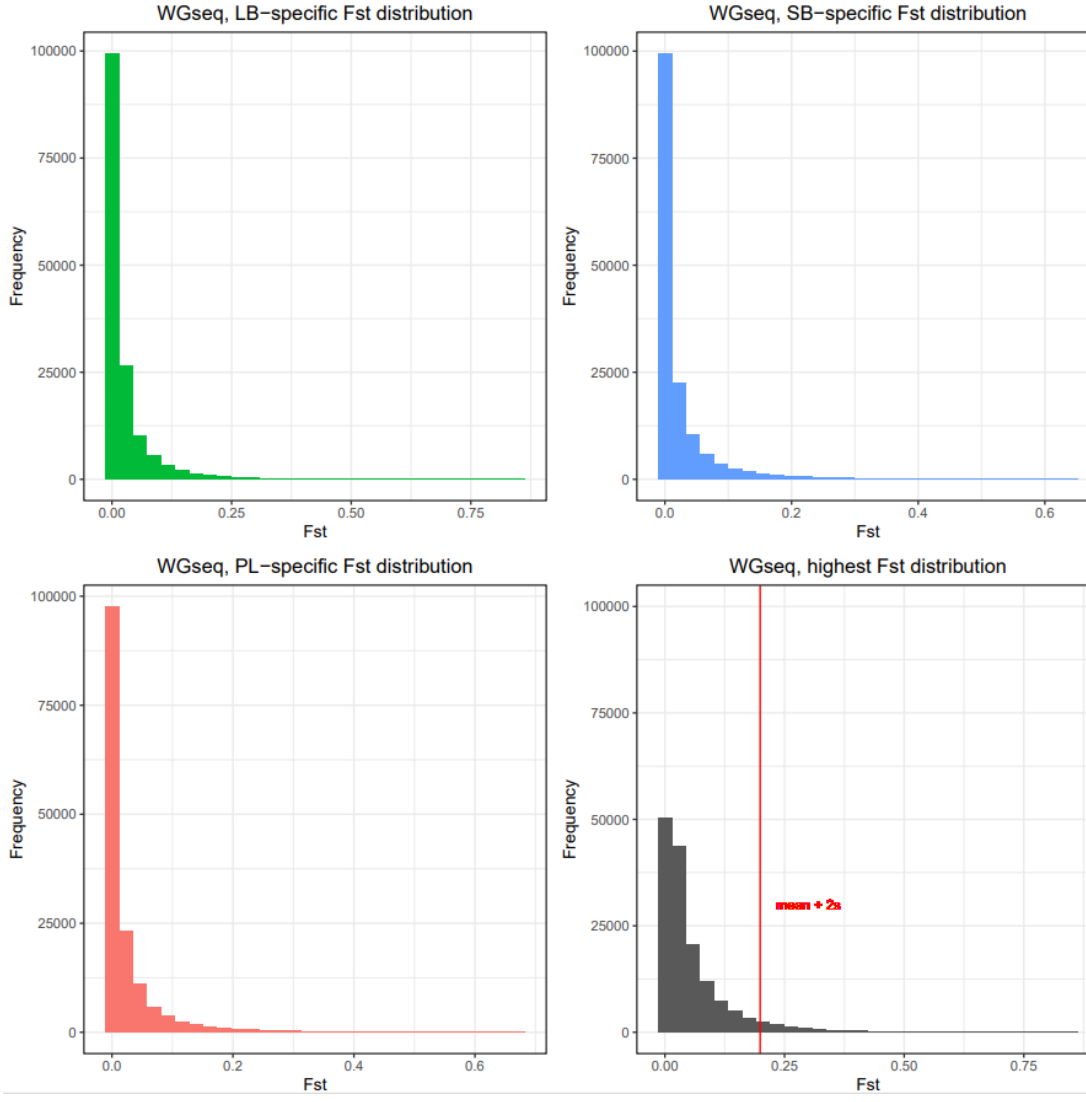
Venney, Clare J, Claire Mérot, Eric Normandeau, Clément Rougeux, Martin Laporte, and Louis Bernatchez. 2024. ‘Epigenetic and Genetic Differentiation Between Coregonus Species Pairs’. Genome Biology and Evolution 16 (2): evae013. https://doi.org/10.1093/gbe/evae013.

Wagner, A. 2000. ‘Decoupled Evolution of Coding Region and mRNA Expression Patterns after Gene Duplication: Implications for the Neutralist-Selectionist Debate’. Proceedings of the National Academy of Sciences of the United States of America 97 (12): 6579–84. https://doi.org/10.1073/pnas.110147097.

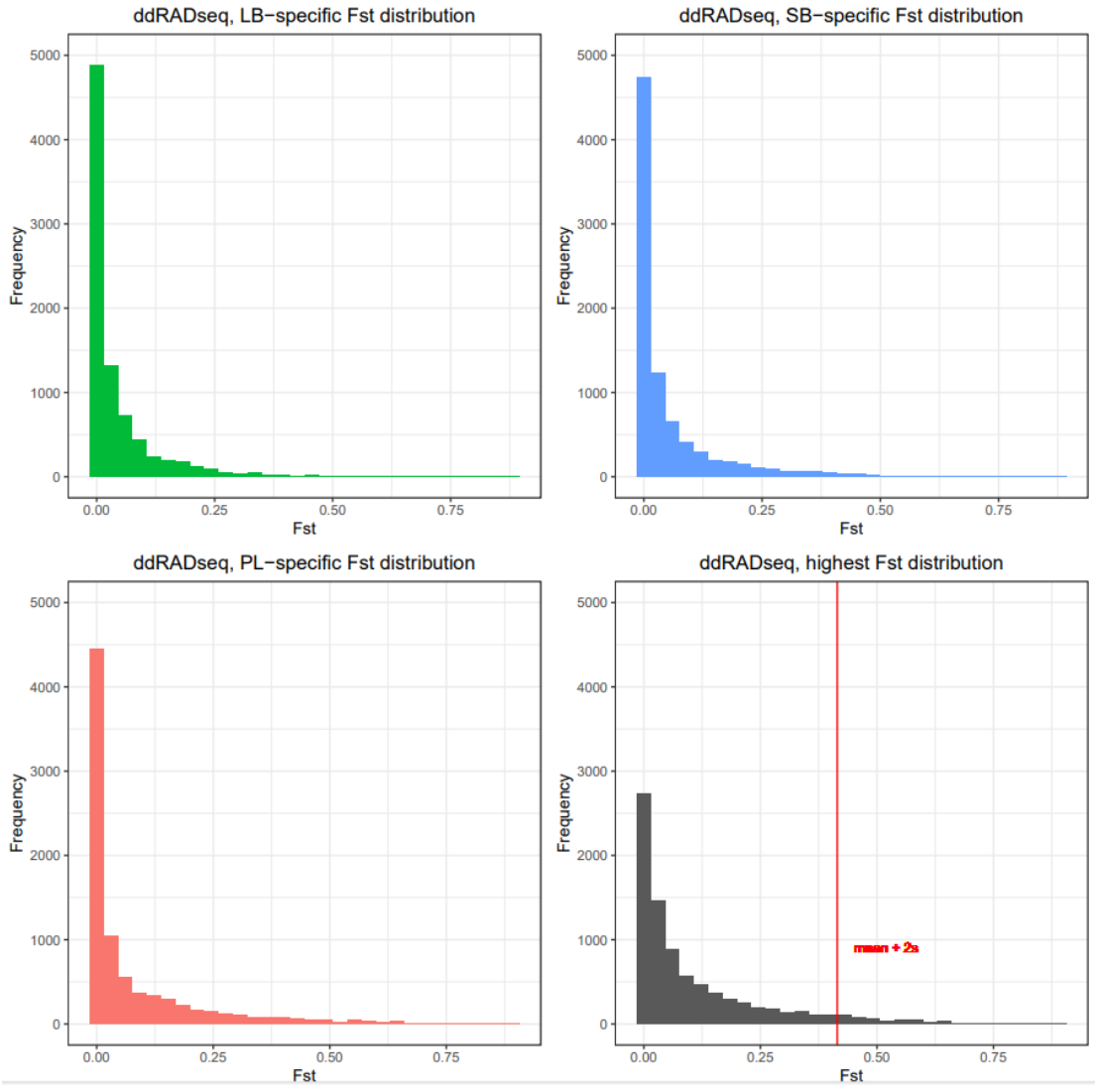
Weinstein, Spencer Y., Colin P. Gallagher, Matthew C. Hale, Tracey N. Loewen, Michael Power, James D. Reist, and Heidi K. Swanson. 2024. ‘An Updated Review of the Post-Glacial History, Ecology, and Diversity of Arctic Char (Salvelinus Alpinus) and Dolly Varden (S. Malma)’. Environmental Biology of Fishes 107 (1): 121–54. https://doi.org/10.1007/s10641-023-01492-0.

Supplementary Data

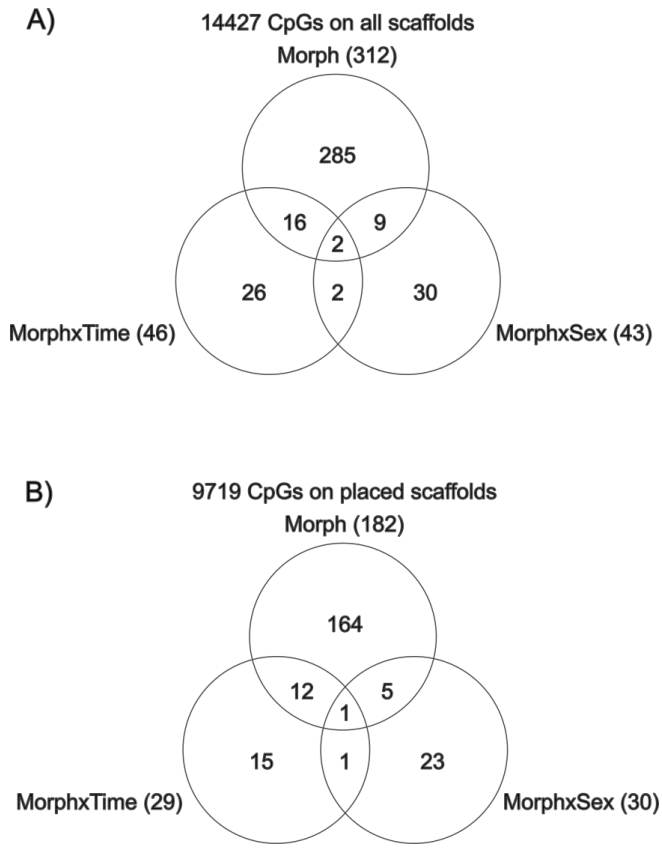
Figures:



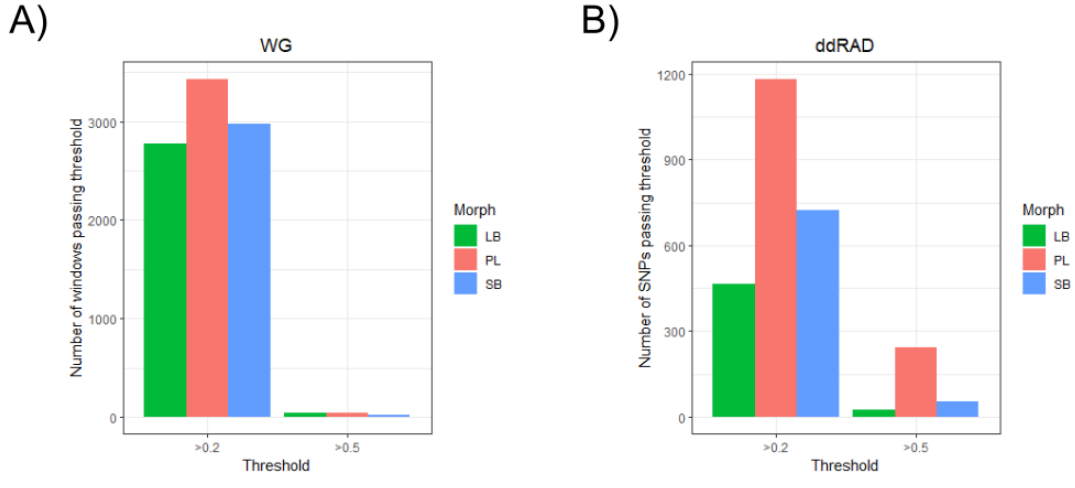
Supplementary Figure 1: Distributions of F_{ST} values across WGS 100kb windows. Shown are the distributions per morph comparison (for instance LB vs SB and PL in the top left panel). Bottom right panel represents the distribution of the highest F_{ST} value for each window (coming from one of the three morph comparisons).



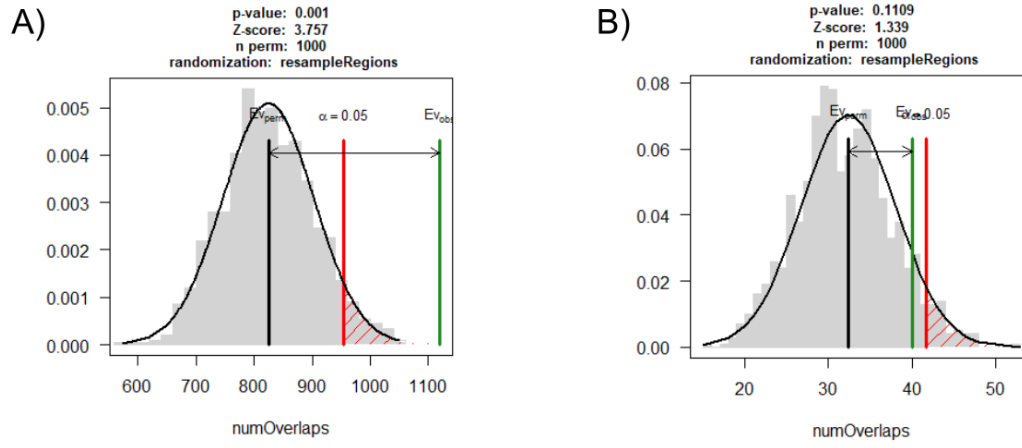
Supplementary Figure 2: Distributions of F_{ST} values across WGS 100kb windows. Shown are the distributions per morph comparison (for instance LB vs SB and PL in the top left panel). Bottom right panel represents the distribution of the highest F_{ST} value for each window (coming from one of the three morph comparisons).



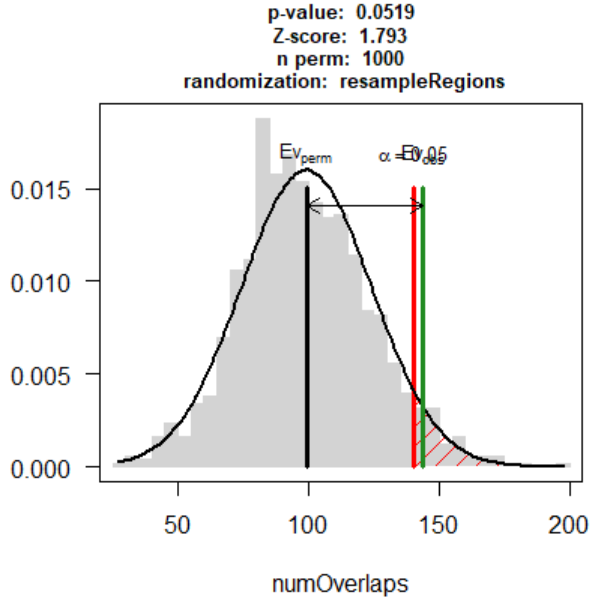
Supplementary Figure 3: Limited overlap of CpGs that differed by Morph, or were found to differ based on interactions of Morph and Time or Morph and Sex. The same pattern was seen for CpGs A) on all scaffolds or B) only on placed scaffolds. The significant methylation differences per CpG were identified with linear model (glm in R), and significance adjusted with Bonferroni for A) 14427 and B) 9719 tests (sites).



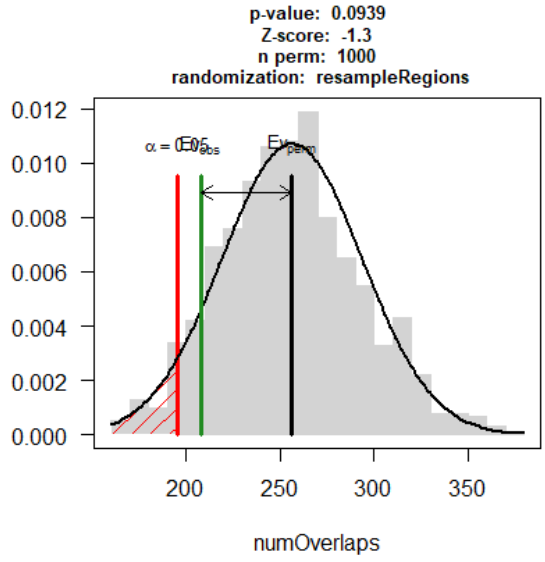
Supplementary Figure 4: PL-specific WG windows or ddRAD SNPs are more numerous than LB or SB-specific ones.



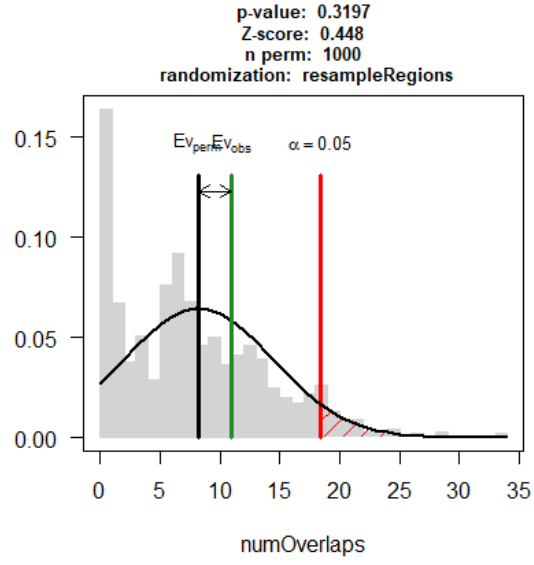
Supplementary Figure 5: A) Regions of genetic differentiation (WGS) and DE genes overlap more than expected by chance. (resampling on all WG regions) B) Regions of genetic differentiation (ddRAD) and DE genes tend towards overlap enrichment. The expected and observed numbers of overlaps are shown with black (Ev perm) and green (Ev obs) lines, respectively. The red line (α = 0.05) represents the limit above which the observed number of overlaps is considered significant.



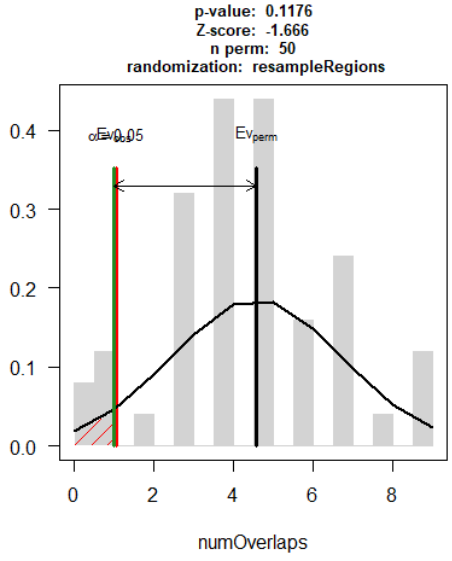
Supplementary Figure 6: Tests for overlap of CpGs differently methylated by morphs (any of the three morphs, tested with linear model in R, see Matlosz 2022) and regions of genetic differentiation (from WGS data). The p-value was not formally significant (sits on the cusp), but the direction of effects suggests enrichment of overlaps. The expected and observed numbers of overlaps are shown with black (Ev perm) and green (Ev obs) lines, respectively. The red line (α = 0.05) represents the limit above which the observed number of overlaps is considered significant.



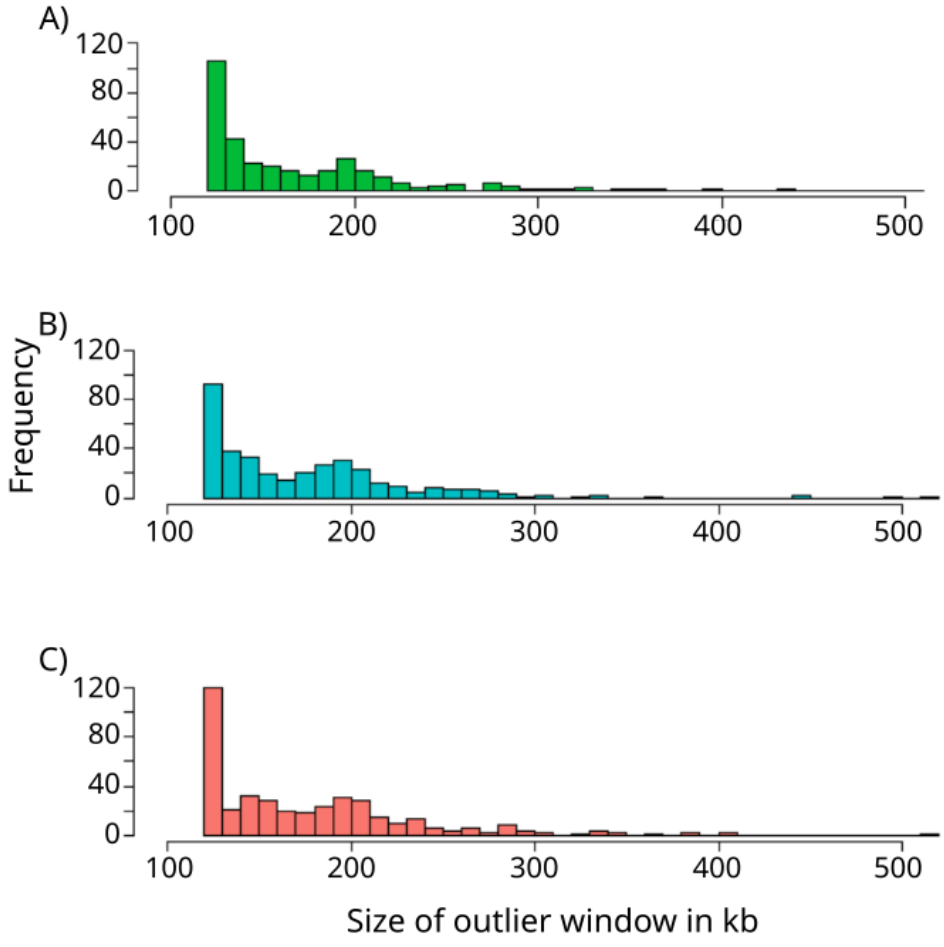
Supplementary Figure 7: CpGs differently methylated by timepoint (Timepoint, TimepointxMorph and TimepointxSex) are not enriched in regions of genetic differentiation (WGS). The expected and observed numbers of overlaps are shown with black (Ev perm) and green (Ev obs) lines, respectively. The red line (α = 0.05) represents the limit above which the observed number of overlaps is considered significant.



Supplementary Figure 8: ddRAD SNPs are not enriched in regions close to differentially methylated CpGs, compared with the rest of the ddRAD dataset. The expected and observed numbers of overlaps are shown with black (Ev perm) and green (Ev obs) lines, respectively. The red line (α = 0.05) represents the limit above which the observed number of overlaps is considered significant.



Supplementary Figure 9: Differently expressed genes and differently methylated CpGs do not overlap more or less than expected. The expected and observed numbers of overlaps are shown with black (Ev perm) and green (Ev obs) lines, respectively. The red line (α = 0.05) represents the limit below which the observed number of overlaps is considered significant.



Supplementary Figure 10: Similar distributions of window sizes for regions of genetic differentiation per morph in the WGS data. For visualization we excluded the smallest windows (100 kb, that did not overlap with others), but the counts were 466, 479, 543 for LB- (A), SB-(B) and PL-charr (C) respectively (See Supplementary Table 10).

Tables:

Supplementary Table 1: Proportion of PL-, LB- and SB- specific SNPs (lying outside 2σ of the overall distribution). Note, the raw number of WGS regions is overestimated by the step-wise windowed approach, but the proportions still hold.

ddRADsea data				
Total	Outlier SNPs	PL-SNPs	SB-SNPs	LB-SNPs
8472	530	374	97	59
WGS data				
Total	Outlier windows	PL-windows	SB-windows	LB-windows
151540	8283	3138	2648	2497

Supplementary Table 2: ddRAD samples. All samples were used to generate genomic positions. ThSB_60 and ThPL_34, suspected to be hybrids, and all PI-charr samples were removed from F_{ST} calculations.

Name	NCBI_ID	Length
LG1	NC_036838.1	58017395
LG2	NC_036839.1	43538721
LG3	NC_036840.1	36001405
LG4p	NC_036841.1	28292882
LG4q.1:29	NC_036842.1	90519428
LG4q.2	NC_036843.1	29595583
LG5	NC_036844.1	37080635
LG6.1	NC_036845.1	30249148
LG6.2	NC_036846.1	26025374
LG7	NC_036847.1	34303021
LG8	NC_036848.1	54842065
LG9	NC_036849.1	32654316
LG10	NC_036850.1	22457292
LG11	NC_036851.1	51124027
LG12	NC_036852.1	13980584
LG13	NC_036853.1	50975424
LG14	NC_036854.1	54096485
LG15	NC_036855.1	67329100
LG16	NC_036856.1	42871064
LG17	NC_036857.1	41841263
LG18	NC_036858.1	72741121
LG19	NC_036859.1	38228754
LG20	NC_036860.1	79996362
LG21	NC_036861.1	6905391

LG22	NC_036862.1	37604395
LG23	NC_036863.1	49632736
LG24	NC_036864.1	11432800
LG25	NC_036865.1	26198113
LG26	NC_036866.1	49931436
LG27	NC_036867.1	38733064
LG28	NC_036868.1	32734159
LG30	NC_036869.1	26193892
LG31	NC_036870.1	32006513
LG32	NC_036871.1	38480802
LG33	NC_036872.1	38084510
LG34	NC_036873.1	8958605
LG35	NC_036874.1	21595701
LG36	NC_036875.1	41232801
LG37	NC_036876.1	19546989

Supplementary Table 3: Number of CpGs passing filtering (10X coverage in each sample) for further methylation analyses, based on the selected samples.

Number of samples	**48 samples (Matlosz et al., 2022)**	**27 samples (noPI, no 100ts)**
Number of CpGs passing filtering	10340	14427

Supplementary Table 4: Names and lengths of placed scaffolds used in the analyses.

ID	**Morph**	**Sex**	**Library**
ThSB_36	Thingvallavatn_SB	M	Han_8
ThSB_37	Thingvallavatn_SB	M	Han_8
ThSB_38	Thingvallavatn_SB	F	Han_8
ThSB_39	Thingvallavatn_SB	M	Han_8
ThSB_40	Thingvallavatn_SB	M	Han_8
ThSB_41	Thingvallavatn_SB	M	Han_13
ThSB_42	Thingvallavatn_SB	M	Han_13
ThSB_43	Thingvallavatn_SB	F	Han_13
ThSB_44	Thingvallavatn_SB	F	Han_13
ThSB_45	Thingvallavatn_SB	F	Han_13

ThSB_46	Thingvallavatn_SB	F	Han_13
ThSB_47	Thingvallavatn_SB	M	Han_13
ThSB_48	Thingvallavatn_SB	M	Han_13
ThSB_49	Thingvallavatn_SB	M	Han_13
ThSB_50	Thingvallavatn_SB	M	Han_13
ThSB_51	Thingvallavatn_SB	M	Han_13
ThSB_52	Thingvallavatn_SB	M	Han_13
ThSB_53	Thingvallavatn_SB	M	Han_13
ThSB_54	Thingvallavatn_SB	F	Han_13
ThSB_55	Thingvallavatn_SB	M	Han_13
ThSB_56	Thingvallavatn_SB	F	Han_13
ThSB_57	Thingvallavatn_SB	F	Han_13
ThSB_58	Thingvallavatn_SB	F	Han_13
ThSB_59	Thingvallavatn_SB	F	Han_13
ThSB_60	Thingvallavatn_SB	M	Han_13
ThSB_61	Thingvallavatn_SB	M	Han_13
ThSB_62	Thingvallavatn_SB	M	Han_13
ThSB_63	Thingvallavatn_SB	M	Han_13
ThSB_P01	Thingvallavatn_SB	M	Plast_2
ThSB_P02	Thingvallavatn_SB	M	Plast_2
ThSB_P03	Thingvallavatn_SB	M	Plast_2
ThLB_17	Thingvallavatn_LB	F	Han_10
ThLB_18	Thingvallavatn_LB	F	Han_10
ThLB_19	Thingvallavatn_LB	F	Han_10
ThLB_20	Thingvallavatn_LB	M	Han_10
ThLB_21	Thingvallavatn_LB	M	Han_10
ThLB_22	Thingvallavatn_LB	M	Han_10
ThLB_23	Thingvallavatn_LB	M	Han_10
ThLB_24	Thingvallavatn_LB	F	Han_10
ThLB_25	Thingvallavatn_LB	F	Han_10
ThLB_26	Thingvallavatn_LB	F	Han_10
ThLB_27	Thingvallavatn_LB	F	Han_10
ThLB_28	Thingvallavatn_LB	F	Han_10
ThLB_29	Thingvallavatn_LB	M	Han_13
ThLB_30	Thingvallavatn_LB	M	Han_13
ThLB_31	Thingvallavatn_LB	M	Han_13
ThLB_32	Thingvallavatn_LB	F	Han_13
ThLB_33	Thingvallavatn_LB	F	Han_13
ThLB_34	Thingvallavatn_LB	F	Han_13
ThLB_35	Thingvallavatn_LB	F	Han_13
ThLB_36	Thingvallavatn_LB	F	Han_13
ThLB_37	Thingvallavatn_LB	M	Han_13
ThLB_38	Thingvallavatn_LB	M	Han_13
ThLB_39	Thingvallavatn_LB	M	Han_13
ThLB_40	Thingvallavatn_LB	M	Han_13
ThLB_41	Thingvallavatn_LB	F	Han_13
ThLB_42	Thingvallavatn_LB	F	Han_13

ThLB_43	Thingvallavatn_LB	F	Han_13
ThLB_P02	Thingvallavatn_LB	M	Plast_1
ThLB_P05	Thingvallavatn_LB	M	Plast_1
ThLB_P06	Thingvallavatn_LB	F	Plast_2
ThLB_P07	Thingvallavatn_LB	F	Plast_2
ThLB_P08	Thingvallavatn_LB	F	Plast_2
ThLB_P09	Thingvallavatn_LB	F	Plast_2
ThLB_P10	Thingvallavatn_LB	F	Plast_2
ThLB_P13	Thingvallavatn_LB	M	Plast_2
ThLB_P14	Thingvallavatn_LB	M	Plast_2
ThLB_P17	Thingvallavatn_LB	M	Plast_2
ThPI_16	Thingvallavatn_PI	F	Han_10
ThPI_17	Thingvallavatn_PI	M	Han_10
ThPI_18	Thingvallavatn_PI	F	Han_10
ThPI_19	Thingvallavatn_PI	M	Han_10
ThPI_20	Thingvallavatn_PI	F	Han_10
ThPI_21	Thingvallavatn_PI	F	Han_10
ThPI_22	Thingvallavatn_PI	M	Han_10
ThPI_23	Thingvallavatn_PI	M	Han_10
ThPI_24	Thingvallavatn_PI	F	Han_10
ThPI_25	Thingvallavatn_PI	M	Han_10
ThPI_26	Thingvallavatn_PI	F	Han_10
ThPI_27	Thingvallavatn_PI	F	Han_10
ThPI_28	Thingvallavatn_PI	M	Han_14
ThPI_29	Thingvallavatn_PI	F	Han_14
ThPI_30	Thingvallavatn_PI	F	Han_14
ThPI_31	Thingvallavatn_PI	M	Han_14
ThPI_32	Thingvallavatn_PI	M	Han_14
ThPI_33	Thingvallavatn_PI	M	Han_14
ThPI_34	Thingvallavatn_PI	F	Han_14
ThPI_35	Thingvallavatn_PI	M	Han_15
ThPI_36	Thingvallavatn_PI	M	Han_15
ThPI_37	Thingvallavatn_PI	F	Han_15
ThPL_31	Thingvallavatn_PL	F	Han_8
ThPL_32	Thingvallavatn_PL	F	Han_8
ThPL_33	Thingvallavatn_PL	F	Han_8
ThPL_34	Thingvallavatn_PL	F	Han_8
ThPL_35	Thingvallavatn_PL	F	Han_8
ThPL_36	Thingvallavatn_PL	M	Han_8
ThPL_37	Thingvallavatn_PL	M	Han_8
ThPL_38	Thingvallavatn_PL	M	Han_8
ThPL_39	Thingvallavatn_PL	F	Han_8
ThPL_40	Thingvallavatn_PL	F	Han_13
ThPL_41	Thingvallavatn_PL	M	Han_13
ThPL_42	Thingvallavatn_PL	M	Han_13
ThPL_43	Thingvallavatn_PL	F	Han_13
ThPL_44	Thingvallavatn_PL	F	Han_13

ThPL_45	Thingvallavatn_PL	F	Han_13
ThPL_46	Thingvallavatn_PL	F	Han_13
ThPL_47	Thingvallavatn_PL	M	Han_13
ThPL_48	Thingvallavatn_PL	M	Han_13
ThPL_49	Thingvallavatn_PL	M	Han_13
ThPL_50	Thingvallavatn_PL	M	Han_13
ThPL_51	Thingvallavatn_PL	M	Han_13
ThPL_P01	Thingvallavatn_PL	F	Plast_1
ThPL_P02	Thingvallavatn_PL	F	Plast_1
ThPL_P03	Thingvallavatn_PL	M	Plast_1
ThPL_P04	Thingvallavatn_PL	M	Plast_1
ThPL_P05	Thingvallavatn_PL	NA	Plast_1
ThPL_P06	Thingvallavatn_PL	F	Plast_2
ThPL_P07	Thingvallavatn_PL	M	Plast_2
ThPL_P08	Thingvallavatn_PL	NA	Plast_2
ThPL_P10	Thingvallavatn_PL	F	Plast_2
ThPL_P11	Thingvallavatn_PL	F	Plast_2
ThPL_P12	Thingvallavatn_PL	M	Plast_2

Supplementary Table 5: List of WG windows or ddRAD SNPs for which multiple morphs have an F_{ST} value $>2\sigma$.
Consult **online version** at https://github.com/sebma-github/PhD_Thesis_Supplements

Supplementary Table 6: List of LB-specific DE genes located in LB-specific regions of genetic differentiation.

LOC ID	Gene name	Linkage group	Start	End	Gene ID
LOC111964131	Saxiphilin-like	NC_036838.1	28137063	28164665	111964131
LOC111964178	fibroblast growth factor receptor substrate 2-like	NC_036838.1	28176448	28210857	111964178
LOC111967700	rho guanine nucleotide exchange factor 10-like protein	NC_036838.1	40942815	41077314	111967700
lgi2	leucine rich repeat LGI family member 2	NC_036840.1	17431891	17483901	111951731
LOC111961360	B-cell CLL/lymphoma 7 protein family member A	NC_036842.1	8138916	8167357	111961360
LOC111961461	unconventional myosin-Ih	NC_036842.1	12533925	12562955	111961461
LOC111961725	interleukin-31 receptor subunit alpha-like	NC_036842.1	25768723	25809519	111961725
LOC111961768	DTW domain-containing protein 2	NC_036842.1	27586667	27617078	111961768
LOC111961829	KN motif and ankyrin repeat domain-containing protein 1-like	NC_036842.1	32636184	32683448	111961829
LOC111964087	serine/threonine-protein phosphatase 2A 55 kDa regulatory subunit B alpha isoform	NC_036844.1	24116522	24152595	111964087
LOC111964088	BCL2/adenovirus E1B 19 kDa protein-interacting protein 3-like	NC_036844.1	24104713	24136205	111964088
LOC111966404	protein piccolo-like	NC_036847.1	12281029	12366337	111966404
LOC111967092	transforming growth factor beta-2-like	NC_036848.1	3694006	3726027	111967092
LOC111970336	acidic mammalian chitinase-like	NC_036851.1	38110939	38135192	111970336
LOC111971218	desmoplakin	NC_036853.1	30523914	30593492	111971218
LOC111971429	tripartite motif-containing protein 16-like	NC_036853.1	38907616	38937176	111971429
LOC111972213	acyl-CoA-binding domain-containing protein 6	NC_036853.1	17075845	17131175	111972213
LOC111972257	tetratricopeptide repeat protein 39A-like	NC_036853.1	16551086	16596349	111972257
LOC111972551	pro-opiomelanocortin B	NC_036854.1	21760117	21787868	111972551
LOC111975601	calcium/calmodulin-dependent protein kinase type IV-like	NC_036856.1	7412065	7461604	111975601
LOC11197606	PWWP domain-	NC_036856.	9210111	9238469	11197606

6	containing protein MUM1L1	1			6
LOC111980145	exocyst complex component 7-like	NC_036860.1	79933994	79965402	111980145
LOC111951072	relaxin receptor 1	NC_036863.1	31595300	31645982	111951072
LOC111952124	multidrug and toxin extrusion protein 1-like	NC_036866.1	8785186	8834073	111952124
LOC111952435	multiple C2 and transmembrane domain-containing protein 2-like	NC_036866.1	22455518	22824017	111952435
LOC111952454	homer protein homolog 3	NC_036866.1	9255629	9313992	111952454
LOC111954277	sodium-dependent multivitamin transporter	NC_036868.1	15976401	16030278	111954277
LOC111955012	retinoic acid receptor beta-like	NC_036869.1	9269319	9419438	111955012
LOC111955076	NF-kappa-B inhibitor-interacting Ras-like protein 1	NC_036869.1	9703061	9728153	111955076
LOC111955658	acyl-coenzyme A thioesterase 1-like	NC_036870.1	541399	571040	111955658
rsrc2	arginine and serine rich coiled-coil 2	NC_036872.1	19817192	19856438	111957700
LOC111959474	insulin-like growth factor-binding protein 5	NC_036875.1	16170613	16209669	111959474

Supplementary Table 7: List of PL-specific DE genes located in PL-specific regions of genetic differentiation.

LOC ID	Gene name	Linkage group	Start	End	Gene ID
LOC111965023	acyl-coenzyme A thioesterase 4	NC_036838.1	31159241	31196621	111965023
LOC111965130	diacylglycerol kinase alpha	NC_036838.1	31301211	31355505	111965130
LOC111955272	echinoderm microtubule-associated protein-like 3	NC_036840.1	33883625	33943683	111955272
LOC111959085	elastin	NC_036841.1	3815610	3917498	111959085
tmem25	transmembrane protein 25	NC_036841.1	9657590	9684693	111960402
LOC111963187	pre-mRNA-splicing factor syf2	NC_036843.1	6848222	6881717	111963187
LOC111963478	tyrosine-protein kinase receptor TYRO3	NC_036843.1	22048900	22100839	111963478
maco1	Macoilin 1	NC_036843.1	6892954	6938166	111963191
LOC111963852	prostaglandin E2 receptor EP4 subtype	NC_036844.1	7277181	7302847	111963852
LOC111964089	leucine-rich repeat transmembrane neuronal protein 4-like	NC_036844.1	24602065	24768516	111964089
LOC111968115	Myomesin-2	NC_036848.1	52361677	52419146	111968115
LOC111969697	copine-5	NC_036851.1	5846896	6077466	111969697
LOC111970313	isotocin receptor-like	NC_036851.1	36275265	36332482	111970313
LOC111970355	retinoblastoma-binding protein 5-like	NC_036851.1	39405721	39448357	111970355
LOC111970717	uncharacterized LOC111970717	NC_036851.1	39381565	39416163	111970717
LOC111970978	contactin-associated protein-like 5	NC_036852.1	7549446	7586902	111970978
etfbkmt	electron transfer flavoprotein beta subunit lysine methyltransferase	NC_036853.1	5078932	5100572	111972396
LOC111973502	piezo-type mechanosensitive ion channel component 2-like	NC_036854.1	27559046	27667504	111973502
LOC111977727	eukaryotic translation initiation factor 3 subunit L	NC_036858.1	22732557	22787009	111977727
LOC111978671	putative protein MSS51 homolog	NC_036858.1	13988223	14013009	111978671
LOC111978888	dipeptidyl peptidase 9	NC_036859.	31323267	31367803	111978888

7		1			7
LOC111978989	uncharacterized LOC111978989	NC_036859.1	17237690	17262440	111978989
LOC111950536	glutamate receptor 3-like	NC_036863.1	28397924	28516794	111950536
LOC111950589	ribonucleoside-diphosphate reductase large subunit	NC_036863.1	27970487	27997002	111950589
LOC111950748	neuroligin-3-like	NC_036863.1	28325307	28387695	111950748
LOC111951159	sodium/calcium exchanger 1	NC_036863.1	37310716	37350475	111951159
LOC111951612	fibroblast growth factor-binding protein 2	NC_036865.1	20140749	20162674	111951612
LOC111951680	kynurenine 3-monooxygenase	NC_036865.1	5970158	6011324	111951680
LOC111952641	uncharacterized LOC111952641	NC_036866.1	26665312	26687400	111952641
LOC111953409	ATP-dependent zinc metalloprotease YME1L1	NC_036867.1	10319512	10364287	111953409
LOC112067722	patched domain-containing protein 3-like	NC_036867.1	10235653	10265181	112067722
LOC111956152	tumor necrosis factor alpha-induced protein 8-like protein 2	NC_036870.1	11165473	11202196	111956152
LOC111956715	THAP domain-containing protein 1-like	NC_036871.1	21416263	21438691	111956715
LOC111958361	fibrinogen gamma chain	NC_036873.1	5281441	5309297	111958361
LOC111959258	leukocyte cell-derived chemotaxin 1	NC_036875.1	14645630	14669973	111959258
LOC111960039	butyrophilin subfamily 1 member A1	NC_036876.1	1578763	1617206	111960039

Supplementary Table 8: List of genes located in regions of genetic and methylomic differentiation between morphs.

Gene name	Linkage group	Start of gene	End of gene	Start of window	End of window	Highest F_{ST}	Morph with highest F_{ST}
Caveolin-2	NC_036838.1	35229859	35234254	35200001	35300000	0,263503	SB
probable G-protein coupled receptor	NC_036838.1	35241816	35249857	35200001	35300000	0,263503	SB
poly [ADP-ribose] polymerase 3-like	NC_036838.1	35253843	35275932	35200001	35300000	0,263503	SB
metabotropic glutamate receptor 2	NC_036838.1	35301677	35369590	35210001	35310000	0,275684	SB
testis-expressed protein 264	NC_036838.1	35374927	35515188	35280001	35380000	0,24642	SB
zinc finger and BTB domain-containing protein 38	NC_036840.1	32171383	32186820	32170001	32270000	0,25821	SB
protein ALP1-like	NC_036840.1	32192266	32210902	32170001	32270000	0,25821	SB
UDP-glucuronosyltransferase 2A1	NC_036840.1	32194001	32197013	32170001	32270000	0,25821	SB
equilibrative nucleoside transporter 2	NC_036840.1	32248100	32263713	32170001	32270000	0,25821	SB
reticulon-4 receptor-like	NC_036842.1	12124601	12163785	12030001	12130000	0,294607	LB
low-density lipoprotein receptor class A domain-containing protein 3	NC_036842.1	77660919	77829224	77670001	77770000	0,205779	LB
Protocadherin-19	NC_036848.1	24247938	24324397	24150001	24250000	0,302474	PL
uncharacterized protein C14orf132-like	NC_036849.1	11700087	11704485	11660001	11760000	0,201643	PL
B2 bradykinin receptor-like	NC_036849.1	11765727	11766764	11670001	11770000	0,230225	PL
psychosine receptor-like	NC_036849.1	11776827	11777798	11680001	11780000	0,261069	PL
potassium channel subfamily K member 10	NC_036849.1	11809475	11854576	11710001	11810000	0,276798	PL

26S proteasome regulatory subunit 10B	NC_036849.1	13219868	13232899	13210001	13310000	0,410696	PL
protein LBH	NC_036849.1	13238776	13248723	13210001	13310000	0,410696	PL
lysocardiolipin acyltransferase 1	NC_036849.1	13257341	13278095	13210001	13310000	0,410696	PL
solute carrier family 7 member 9	NC_036849.1	13299251	13302618	13210001	13310000	0,410696	PL
RNA-binding Raly-like protein	NC_036859.1	35463991	35497954	35480001	35580000	0,214044	PL
gastrula zinc finger protein xFG20-1	NC_036859.1	35591956	35597647	35500001	35600000	0,22862	PL
uncharacterized oxidoreductase ZK1290.5	NC_036859.1	35598430	35608768	35500001	35600000	0,22862	PL
exostosin like glycosyltransferase 2	NC_036859.1	35608873	35614555	35510001	35610000	0,26297	PL
peroxidasin	NC_036859.1	35619457	35625883	35520001	35620000	0,253205	PL
uncharacterized LOC111953959	NC_036867.1	10137244	10200884	10200001	10300000	0,240053	PL
retinol dehydrogenase 12	NC_036867.1	10212694	10226317	10200001	10300000	0,240053	PL
probable polyketide synthase 1	NC_036867.1	10229732	10239052	10200001	10300000	0,240053	PL
patched domain-containing protein 3-like	NC_036867.1	10240792	10244504	10200001	10300000	0,240053	PL
patched domain-containing protein 3-like	NC_036867.1	10245653	10255181	10200001	10300000	0,240053	PL
probable polyketide synthase 1	NC_036867.1	10285222	10311234	10200001	10300000	0,240053	PL
patched domain-containing protein 3-like	NC_036867.1	10312940	10319547	10220001	10320000	0,245811	PL
mycocerosic acid synthase-like	NC_036867.1	10320715	10328652	10230001	10330000	0,249808	PL
ATP-dependent zinc metalloprotease YME1L1	NC_036867.1	10329512	10354287	10230001	10330000	0,249808	PL

microtubule associated se-rine/threonine kinase like	NC_036867.1	10356994	10364152	10260001	10360000	0,208757	PL

Supplementary Table 9: Significance of results from asymptotic two-samples KS test between methylation distributions of all CpGs and differentially methylated CpGs at 200ts, 150ts and 50ts.

KS tests	**200ts**	**150ts**	**50ts**
D	0.12168	0.13313	0.10398
p-value	1.13E-08	2.68E-10	1.90E-06

Supplementary Table 10: Numbers of unique genomic outlier regions per morph and region size. In this summary some windows were counted twice (or three times), if they passed the threshold in two (or three) morphs.

	Size of regions of differentiation						
Morph	100 kb	110-200kb	210-300kb	310-400kb	410 - 500kb	> 510kb	Total
LB	466	264	58	16	4	1	809
SB	479	275	81	6	3	1	845
PL	543	293	99	14	2	1	952

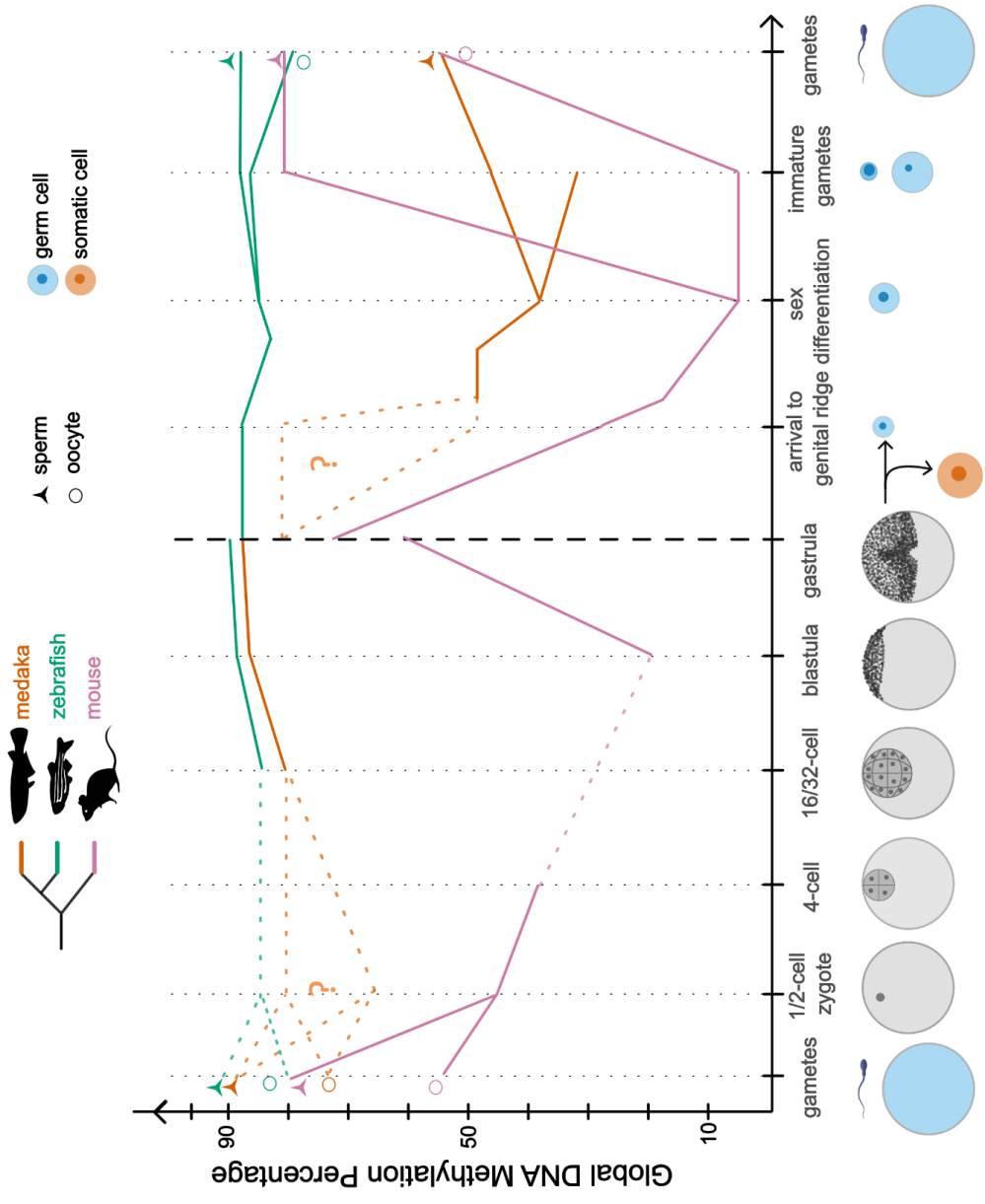
Paper III

DNA methylation reprogramming in teleosts

Matlosz, Sébastien, Sigríður R. Franzdóttir, Arnar Pálsson, and Zophonías O. Jónsson.

Evolution & Development 2024, e12486.
https://doi.org/10.1111/ede.12486.

Author contribution: S. Matlosz wrote the manuscript.



<u>Graphical abstract</u>: DNA methylation reprogramming dynamics in the early embryonic stages (left) and primordial germ cells (right) of zebrafish, medaka and mouse highlight the diversity within and between teleost groups.

Received: 23 September 2023 | Accepted: 14 May 2024

DOI: 10.1111/ede.12486

REVIEW ARTICLE

WILEY

DNA methylation reprogramming in teleosts

Sébastien Matlosz | **Sigríður R. Franzdóttir** | **Arnar Pálsson** | **Zophonías O. Jónsson**

Institute of Life and Environmental Sciences, University of Iceland, Reykjavík, Iceland

Correspondence
Sébastien Matlosz, Institute of Life and Environmental Sciences, University of Iceland, Reykjavík, Iceland.
Email: slt2@hi.is

Abstract

Early embryonic development is crucially important but also remarkably diverse among animal taxa. Axis formation and cell lineage specification occur due to both spatial and temporal control of gene expression. This complex system involves various signaling pathways and developmental genes such as transcription factors as well as other molecular interactants that maintain cellular states, including several types of epigenetic marks. 5mC DNA methylation, the chemical modification of cytosines in eukaryotes, represents one such mark. By influencing the compaction of chromatin (a high-order DNA structure), DNA methylation can either repress or induce transcriptional activity. Mammals exhibit a reprogramming of DNA methylation from the parental genomes in the zygote following fertilization, and later in primordial germ cells (PGCs). Whether these periods of methylation reprogramming are evolutionarily conserved, or an innovation in mammals, is an emerging question. Looking into these processes in other vertebrate lineages is thus important, and teleost fish, with their extensive species richness, phenotypic diversity, and multiple rounds of whole genome duplication, provide the perfect research playground for answering such a question. This review aims to present a concise state of the art of DNA methylation reprogramming in early development in fish by summarizing findings from different research groups investigating methylation reprogramming patterns in teleosts, while keeping in mind the ramifications of the methodology used, then comparing those patterns to reprogramming patterns in mammals.

KEYWORDS

DNA methylation, epigenetics, evolution, reprogramming, teleosts

1 | INTRODUCTION

Early embryonic development is complex, driven by a multitude of molecular pathways orchestrating the timing and dynamics of cell differentiation and migration, bringing about the makings of a functioning organism from a totipotent zygote. The processes at play are quite diverse in vertebrates, particularly from first cleavage up to organ specification, and this likely reflects multiple different systems involved in regulating the expression of genes at key developmental time points. Out of these different systems, epigenetic regulation, in particular DNA and histone modification, has received increasing attention over the past years. Diversity in these epigenetic mechanisms may explain in part the diversity of vertebrates, and highly diverse vertebrates,

such as fishes, can be used as models to study adaptive divergence in these systems.

Histone modifications play a role in multiple aspects of gene- and genome regulation. Histone genes are highly conserved in eukaryotes: four core histones (H2A, H2B, H3, and H4) form octameric structures around which genomic DNA is wrapped, creating the nucleosomes, which associate to form chromatin, a higher order structure of DNA allowing for efficient compaction of genetic material (Kornberg, 1974). The N-terminal tails of the core histone proteins protrude from the nucleosome complex and are subject to a range of post-translational chemical modifications that affect inter-nucleosomal interactions, stabilize or destabilize the chromatin structure, and induce changes in gene expression (Bannister & Kouzarides, 2011). Because of how highly conserved the core histone proteins are, especially within vertebrates, a common nomenclature is used to describe these modification marks; comprised of the name of the histone protein, the amino acid abbreviation, the position of the amino acid on the protein, the type of chemical modification and the number of modifications on that particular residue. For example, H3K4me3 refers to a tri-methylation of Lysine 4 on the histone H3 protein.

While histone proteins can be modified by a range of different chemical groups, direct chemical modification of the eukaryotic DNA is predominantly by methylation. DNA methylation has a reciprocal relationship with histone modifications and chromatin structure, and happens almost exclusively on the fifth carbon of cytosines (5mC), first and foremost on cytosines located next to guanines (CpG dinucleotides) (Cooper, 1983). These methylation marks can be replicated during cell division and, through their impact on chromatin structure, have been shown to influence many mechanisms, including gene expression (Ng & Bird, 1999), silencing of transposable elements (Karimi et al., 2011), gene imprinting and X chromosome inactivation in mammals (Lee, 2003). The enzymes responsible for the maintenance of methylation marks through rounds of DNA replication and their de novo establishment in most eukaryotes are Dnmt1 (Bestor et al., 1988) and Dnmt3A/Dnmt3B (Okano et al., 1998), respectively. While these marks can be established de novo and maintained throughout cell divisions, they can also be removed through either passive or active demethylation processes. Passive demethylation happens when methylation marks are not maintained by Dnmt1 through DNA replication, leading to a dilution of the methylation signal with each cell division. Active demethylation, on the other hand, involves the conversion of 5mC residues by TET enzymes (reviewed in Wu & Zhang, 2017) into derived bases: hydroxymethylcytosine (5hmC), formylcytosine (5fC) and carboxycytosine (5CaC) which can then be removed by thymine DNA glycosylase (TDG) followed by base excision repair of the resulting abasic site (Maiti et al., 2013).

In mammals, most of the parental methylome undergoes strict reprogramming during early embryonic development, as two instances of genome-wide DNA methylation reprogramming happen during this period (Morgan et al., 2005). The first wave happens immediately following fertilization (Santos et al., 2002), when the highly differentiated parental germ cells give rise to a totipotent zygote. More specifically, global demethylation happens in the zygote following fertilization and during the cleavage stage, where methylation levels drop from ~80% (of all CpGs) for the paternal genome and ~55% for the maternal genome to around 20% at the early inner cell mass stage (Wang et al., 2014). During this time, the paternal genome is actively demethylated through the activity of TET enzymes (Oswald et al., 2000; Seisenberger et al., 2013), while the maternal genome is believed to be mainly passively demethylated by dilution (Zeng & Chen, 2019), even though partial active demethylation of the maternal DNA has also been observed (Wang et al., 2014). This demethylation phase is followed by a phase of de novo methylation starting at pre-implantation, where global methylation eventually rises back to ~70% (Wang et al., 2014). Importantly, some genomic regions bypass methylation reprogramming in mammals. These include parentally imprinted genes (Morgan et al., 2005), as well as repeat elements such as Intracisternal A-particles (IAP) (Lane et al., 2003).

This first instance of methylation reprogramming often overlaps with the maternal-to-zygotic transition (MZT), a fundamental developmental process during which newly formed zygotes switch on their own transcriptional activity (zygotic genome activation or ZGA) and stop relying on maternal mRNA transcripts deposited in the egg, which will eventually be degraded (Baroux et al., 2008). Heterochrony in the timing of ZGA is observed between vertebrate groups (Jukam et al., 2017), and thus, the first wave of methylation reprogramming can either precede or accompany ZGA, but might play a role in the orchestration of this fundamental process in either case.

After the mammalian zygote has switched on its own transcriptional activity, a second wave of methylation reprogramming takes place in the primordial germ cells (PGCs), a cell lineage that will eventually give rise to the gametes (Hajkova et al., 2002). During this process, genomic imprints need to be reset, and the methylation of imprinted genes, which escape the first wave of demethylation and are expressed in the zygote in a

mono-allelic manner based on their parental origin, has been shown to be removed in PGCs by active demethylation driven by TET enzymes (Yamaguchi et al., 2013) before being re-established later during gamete formation in a sex-specific manner.

These instances of demethylation-remethylation in mammals are referred to as "waves" because of the global patterns observed and their amplitude. However, while the term "wave" might make sense in the mammalian context, this is not necessarily the case in other vertebrate groups where the reprogramming may be more subtle or become heterogeneous between cell types or genomic regions.

Understanding mammalian DNA methylation is of importance for biomedical research, but a question arises: how well conserved are these reprogramming processes evolutionarily? Thus, exploring them in other vertebrate groups is of interest. With around 30,000 described species, teleost fishes are the largest and most diverse group of vertebrates (Nelson et al., 2016). Their phenotypic and behavioral diversity, as well as multiple whole-genome duplication (WGD) events in their history (Dehal & Boore, 2005; Macqueen & Johnston, 2014; Meyer & Van de Peer, 2005) which may have altered their methylation machinery (Best et al., 2018), make them particularly valuable for addressing this. Methylation differences between teleosts have been described in both broad and narrow taxonomic contexts. Indeed, methylation differences have been identified between species (Varriale & Bernardi, 2006), and we found that developmental methylation patterns even differ between sympatric morphs of arctic charr (Matlosz et al., 2022).

Many environmental variables have been shown to impact methylation patterns in fish, for instance, pollutants (Major et al., 2020), hydrogen sulfide-rich springs (Kelley et al., 2021), and hypoxia (Lai et al., 2019). Additionally, environmentally induced marks highlighted in these studies were found to persist in subsequent generations, even when the external stimuli were removed from the environment, which begs the question whether methylome reprogramming in teleosts differs from mammals. Thus, the extent to which methylation is reprogrammed during early development in teleosts has gained attention. This review aims to summarize, compare, and discuss patterns of methylation reprogramming in different species of teleosts. Importantly, it also discusses the pitfalls of the different methods currently in use for measuring global levels of methylation, as methodological factors may give the impression of, or mask, methylation differences between species. After summarizing these methodological differences, the review is structured around the mammalian wave concepts, starting with the first reprogramming events in the teleost zygote, before describing the patterns of DNA methylation related to ZGA (which happens later in fishes than in mammals), and finally investigating the second occurrence of reprogramming in teleost PGCs. While most studies of teleost DNA methylation reprogramming involve zebrafish, a highly practical model species because of its short generation time, rapid development, high fecundity, and ease of genetic manipulation, studies on DNA methylation reprogramming in other teleosts have recently been published. However, there are important knowledge gaps in this field and additional species and vertebrate groups need to be investigated in this regard.

2 | GLOBAL METHYLATION ANALYSIS METHODS

As mentioned above, different methods exist to measure global DNA methylation, each with its own advantages and pitfalls. One of the main differences is the capability of methods to distinguish between methylation contexts: nonspecific methods such as 5mC fluorometric ELISA assays or LUMA (Karimi et al., 2006) cannot distinguish between genomic and organelle DNA methylation, while gDNA-specific sequencing methods (WGBS and RRBS) allow for the separation of genomic and organelle DNA methylation levels but cannot distinguish between 5mC and 5hmC.

It is important to take these differences into consideration when looking at gametes and early embryonic development. Indeed, as previously pointed out by Ortega-Recalde and Hore (2019), Ross et al. (2023), and others, the inability of nonspecific methods to distinguish between genomic and mitochondrial DNA introduces a strong bias in the oocyte as the amount of mitochondrial DNA greatly surpasses that of genomic DNA. For instance, zebrafish oocytes contain on average 19.0 million mtDNA molecules while, during embryogenesis, non-PGCs only possess ~50 (Otten et al., 2016). Similarly, mtDNA reads account for ~63% of all reads in the medaka oocyte, with mtDNA methylation measured at ~0% (Ross et al., 2023). Importantly, this bias lowers the measured methylation levels not only in the oocyte but in the zygote as well, as mtDNA lingers in great quantities in the embryo (mtDNA reads still account for ~43% of all reads at the medaka 32-cell stage; Ross et al., 2023). Note that the gradual decrease in mtDNA/gDNA ratio as the embryo develops and as maternal contributions are degraded can create the illusion of global remethylation. Thus, the use of nonspecific methods to measure global methylation in early fish embryos (and other complex organism with similar variation in

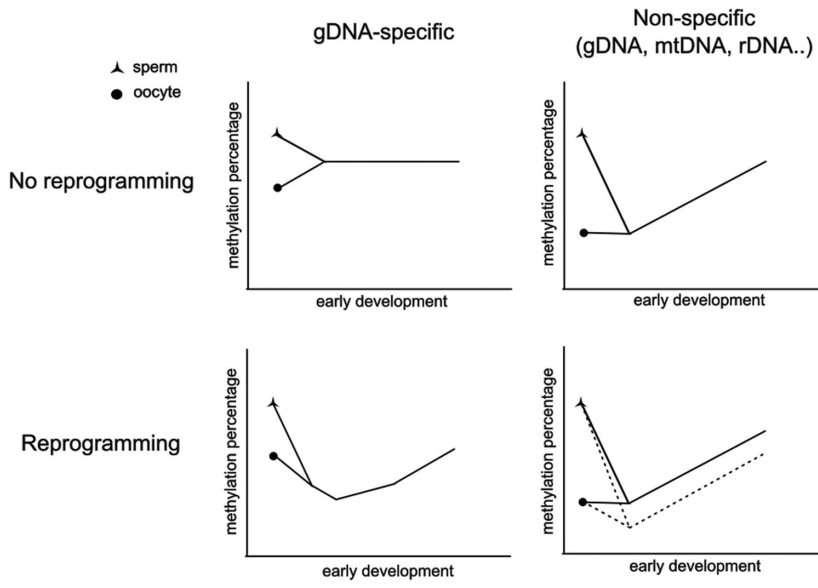


FIGURE 1 Hypothetical scenarios for changes in global methylation during early development, and differences between methods. Shown are the expected measured global methylation trends in cases of reprogramming (below), or lack thereof (above), with gDNA-specific (left) and nonspecific methods (right). We have displayed the sperm as more highly methylated than the oocyte as this seems to be a recurring trend in mammals and teleosts.

mtDNA/gDNA ratio) will lead to low measured methylation levels in the oocyte and zygote, followed by a gradual increase in measured methylation due to the decrease of mtDNA/gDNA ratio, regardless of whether reprogramming happens or not (Figure 1). In case of a major gDNA demethylation event following fertilization, we do not rule out the possibility that nonspecific methods have the capability to observe zygotic methylation levels dropping below those of the oocyte. However, in most cases, this signal is likely too subtle to be observed, or considered nonsignificant.

In short, while nonspecific methods can be a good indication of methylation levels in differentiated cells or tissues, these methods are not ideally suited for the study of methylation in early embryos.

3 | METHYLATION REPROGRAMMING IN THE EARLY FISH EMBRYO

Methylation dynamics during early embryonic development in fish were first investigated in zebrafish by Macleod et al. (1999), who observed the methylation status of three randomly chosen single-copy genes using bisulfite sequencing and found them to stay highly methylated from the gametes to organogenesis. Through this very targeted approach, they hypothesized that zebrafish embryos might not undergo methylation reprogramming as mammals do. Needless to say, three genes are not necessarily representative of genome-wide patterns, however, later studies using more extensive bisulfite-based methods confirmed these trends on a larger scale: Jiang et al. (2013) found the sperm DNA to be more highly methylated than oocyte DNA, and methylation levels in the 16-cell stage embryo matched the average between the global methylation levels of both gametes (Figure 2). While a demethylation/remethylation event could still happen before the 16-cell stage, Potok et al. (2013) pooled embryos from the 2-cell to 16-cell stages and found their methylation level to be at 83%, the average between the sperm and oocyte (with global methylation at 95% and 75%, respectively). Consistently, no expression of TET enzymes was detected from the two-cell stage until the onset of organogenesis in zebrafish (Almeida et al., 2012). Thus, there appears to be no "wave" of gDNA demethylation following fertilization in zebrafish (Jiang et al., 2013), in contrast to that seen in mammals. To confirm this, more detailed measurements of global gDNA methylation levels in the one-cell zygote and first-cell divisions are needed.

During the zebrafish cleavage stage, the maternal genome is then progressively reprogrammed to match the paternal methylation pattern and by the 1000-cell stage (around mid-blastula), the zygote's methylation

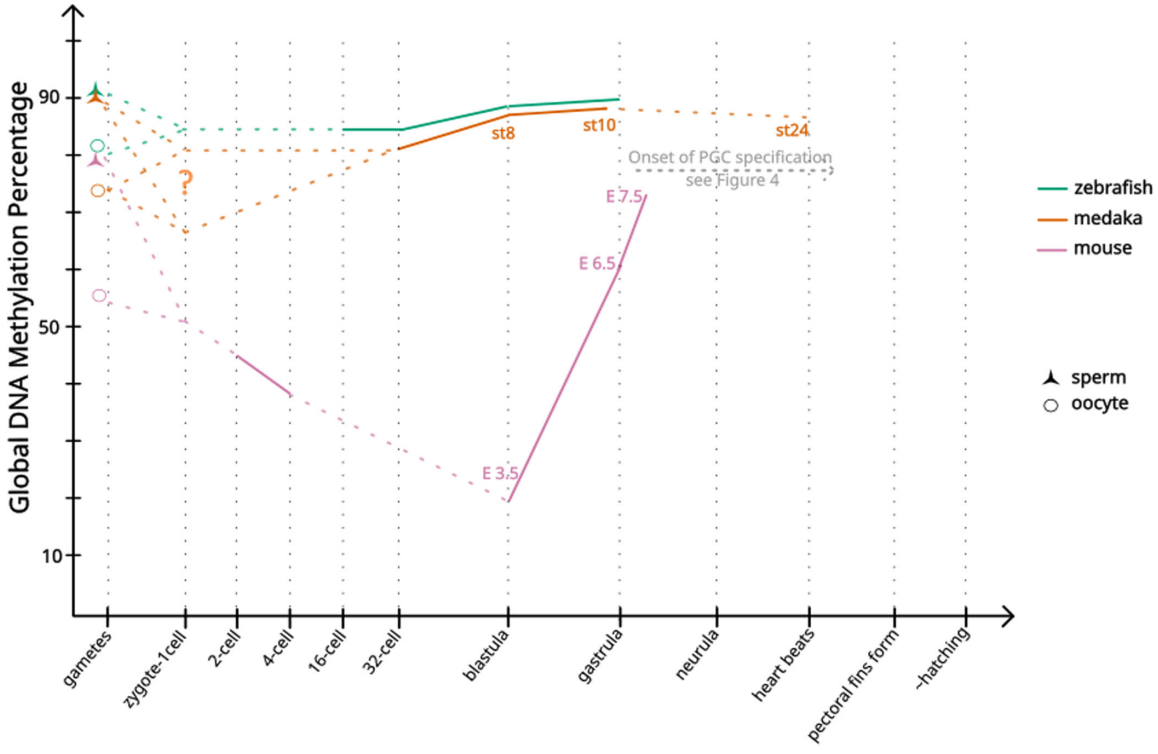


FIGURE 2 DNA methylation reprogramming dynamics in the early embryonic stages of zebrafish, medaka, and mouse. Adapted from Jiang et al. (2013) (zebrafish), Ross et al. (2023) (medaka), and Wang et al. (2014) (mouse). The *x*-axis showing the different developmental stages is not to scale. Dotted lines represent lack of data. Zebrafish, medaka, and mouse methylation levels in the zygote (one-cell) are here arbitrary placed to depict the possibilities of reprogramming or lack thereof (which results in an average between egg and sperm genomic methylation percentage upon fertilization), and are not based on any measurement. As the extent of reprogramming in the medaka zygote is debated (Ross et al., 2023; Wang & Bhandari, 2019), two hypotheses are displayed on the figure, one with a drop in methylation after fertilization and the other with zygotic methylation levels matching the average between methylation levels in the gametes. [Color figure can be viewed at wileyonlinelibrary.com]

level is comparable to that of the sperm (Jiang et al., 2013) (Figure 2).

While global methylation levels are higher in the sperm, at a finer scale, the oocyte contains both hypo- and hypermethylated regions (Potok et al., 2013) and genes that are differentially methylated between egg and sperm belong to distinct functional classes. For instance, genes required in germline development and function, as well as early developmental genes, are methylated in the zebrafish oocyte (the methylome of which resembles more that of terminally differentiated muscle cells), while hypermethylated genes in sperm tend to be related to differentiation. In other words, the oocyte methylation pattern resembles the methylation state of later development, while the sperm methylation pattern is that of early development. Thus, when the maternal methylome gets reprogrammed to match that of the paternal methylome before ZGA, early development/germline genes become demethylated and ready (or poised) for transcription, and genes involved in late development/differentiation are methylated and silenced (Potok et al., 2013).

Why the paternal methylome is already primed for ZGA while the maternal methylome needs to undergo reprogramming following fertilization in zebrafish needs to be further investigated. Zhang et al. (2018) suggest that the programming to the zygotic state might occur earlier in sperm due to the differences in lengths of transcriptionally quiescent states following gametogenesis (a few hours in the oocyte following late-stage oogenesis vs. more than 2 weeks in the sperm following spermatogenesis). This difference could be a consequence of the requirement for the oocyte to provide the zygotic cells with important factors before the ZGA.

To investigate whether reprogramming of the maternal methylome post-fertilization required the presence of

the paternal genome, Potok et al. (2013) created maternal haploids using parthenogenesis, by UV-inactivating the sperm before fertilization. Interestingly, the maternal methylome in these haploids underwent reprogramming in the same way as it does in normal diploids, suggesting that the paternal genome is not required for this process.

Potential mechanisms underlying sex-related differences in gamete methylation and reprogramming in the zebrafish zygote were studied by Murphy et al. (2018), who highlighted the existence of "Placeholder" nucleosomes which, along with associated histone marks, have the property to establish and maintain hypomethylated DNA regions. Indeed, a specific histone variant called H2AFV (zebrafish ortholog of the mammalian H2A.Z), deposited by the chromatin remodeling complex SRCAP and removed by the histone chaperone Anp32e, was shown to be colocated with hypomethylated DNA regions in the zebrafish sperm and zygote. These nucleosomes were often associated with the permissive histone modification marks H3K4me1 and H3K14ac (Murphy et al., 2018), as well as H3K27ac, especially at the 256-cell stage (Zhang et al., 2018).

Interestingly, H2AFV along with its corresponding H3K4me1 mark was located at the same hypomethylated loci in both sperm and pre-ZGA embryos, matching similarities in observed methylation patterns between the two. Zebrafish *anp32e*$^{-/-}$ mutants lacking expression of Anp32e exhibit a clear expansion of H2AFV in intergenic regions and most regions that acquired H2AFV were found to lose methylation, supporting a role for the placeholder nucleosomes in maintaining hypomethylated regions. Similar to Potok et al. (2013), Murphy et al. (2018) generated haploid parthenotes and showed that the paternal genome was not required for the proper establishment of placeholder nucleosomes in the maternal genome to match the patterns observed in the sperm. Moreover, placeholder-occupied sites show enrichment in specific transcription factor (TF) binding motifs, suggesting that these TFs likely interact with SRCAP to maintain placeholders in the paternal genome (or install them during spermatogenesis) and insert new placeholders in the maternal genome after fertilization in zebrafish.

To summarize, in zebrafish, the histone variant H2AFV is inserted by SRCAP at regions that need to undergo demethylation and removed by Anp32e at regions that will stay/become methylated. In addition, it plays a role in bringing the methylome from both parents to the same methylation state (Murphy et al., 2018).

More recently, DNA methylation reprogramming in the early zygote has also been studied in other teleost species. Medaka is the second most studied in this regard, with Wang and Bhandari (2019) and Ross et al. (2023) investigating reprogramming patterns in this species. Interestingly, these two studies display contrasting results: while the former team finds the medaka zygote to undergo reprogramming in a fashion that resembles mammals, the latter team observes a pattern that is not unlike what is seen in zebrafish. These drastic differences do not come from individual or rearing-based differences but reflect the previously mentioned methodological differences: these two studies used different methods to measure CpG methylation levels. Ross et al. (2023) used bisulfite sequencing (which is gDNA-specific), while Wang and Bhandari (2019) used an ELISA assay (nonspecific) (Table 1).

The most important difference in measurement can be seen in the oocyte, where one estimates global methylation levels at ~25% (Wang & Bhandari, 2019) while the other at 75% (Ross et al., 2023). Ross et al. (2023) estimated mtDNA reads in the oocyte account for 60%–65% of the total reads, and measured the global methylation levels of mtDNA at ~0%. This likely caused the low estimates of methylation levels by Wang and Bhandari (2019) in the egg and zygote following fertilization, as the nonspecific method used could not distinguish genomic from mitochondrial DNA.

TABLE 1 References of studies measuring global methylation levels in the early development of teleosts (and mouse), their species of interest and methods used.

Study	Species	Method	Type
Macleod et al. (1999)	Zebrafish	Bisulfite seq (targeted)	gDNA-specific
Jiang et al. (2013)	Zebrafish	WGBS	gDNA-specific
Potok et al. (2013)	Zebrafish	WGBS	gDNA-specific
Wang et al. (2014)	Mouse	WGBS	gDNA-specific
Ross et al. (2023)	Medaka	WGBS	gDNA-specific
Wang and Bhandari (2019)	Medaka	ELISA	Nonspecific
Fellous et al. (2018)	Mangrove rivulus	LUMA	Nonspecific
Fellous et al. (2022)	Stickleback	ELISA	Nonspecific

Interestingly, the methylation patterns in medaka sperm were observed to match embryonic methylome patterns at the blastula through gastrula stages, which is reminiscent of what was observed in zebrafish, suggesting that the maternal methylome gets reprogrammed to match the paternal methylome (Ross et al., 2023) in this group of fishes. In a broader taxonomic context, large-scale maternal-to-paternal epigenome remodeling was recently observed with gDNA-specific methods in lamprey, a jawless fish, which could imply that this remodeling might not be specific to teleosts, but more basal for vertebrates in general (Angeloni et al., 2024).

As Ross et al. (2023) did not investigate DNA methylation levels in the window from fertilization to the 32-cell stage, it is currently unknown whether reprogramming occurs in these early stages in medaka. In contrast to zebrafish, expression of TET genes has been observed in medaka which suggests that some reprogramming occurs post-fertilization in this species (Wang & Bhandari, 2019). TET2 and TET3 mRNAs seem to be largely provided by the sperm and could therefore provide a source for initiating demethylation upon fertilization, together with TET1 which, although expressed at a lower level in oocyte and sperm is the most highly expressed of the three throughout early development until gastrulation (Wang & Bhandari, 2019). Subsequent studies assessing genomic methylation levels in the zygote and cleavage stage are required to assess the extent of this demethylation.

Because of the differences in methods, we represented in Figure 2, the gDNA 5mC levels observed by Ross et al. (2023), alongside zebrafish (Jiang et al., 2013) and mouse (Wang et al., 2014) methylation levels that were also obtained through similar gDNA-specific methods (Table 1). The expression of TET enzymes in medaka gametes and post-fertilization suggests some level of demethylation early in development. Thus, we present two possible scenarios for the state of methylation reprogramming in this species in Figure 2. In addition to zebrafish and medaka, global methylation levels in early embryonic development have been measured for stickleback (Fellous et al., 2022) and mangrove rivulus (Fellous et al., 2018), but only with nonspecific methods, and results from these two studies have been displayed on Figure 3, alongside methylation levels measured in medaka by Wang and Bhandari (2019) through a similar nonspecific method.

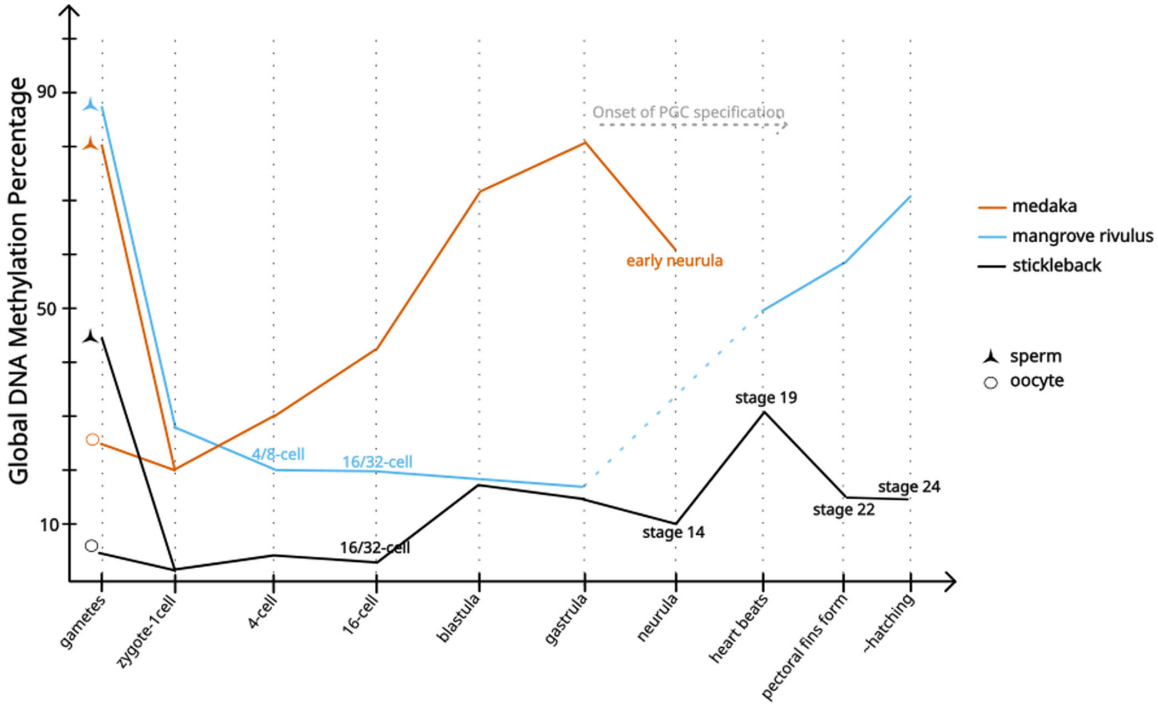


FIGURE 3 Measures of global DNA methylation reprogramming dynamics in the early embryonic stages of medaka, mangrove rivulus, and stickleback, with nonspecific methods. Adapted from Wang and Bhandari (2019) (medaka), Fellous et al. (2018) (mangrove rivulus), and Fellous et al. (2022) (stickleback). The *x*-axis showing the different developmental stages is not to scale. [Color figure can be viewed at wileyonlinelibrary.com]

Measured with an ELISA assay, stickleback sperm appears to have higher global methylation than the oocyte, the zygote exhibits methylation levels similar to those of the oocyte and then seems to be gradually remethylated until gastrulation (Fellous et al., 2022). This again is reminiscent of a bias created by varying mtDNA levels. Similar to medaka, while the possibility that demethylation occurs in the stickleback genome post-fertilization cannot be rejected, this cannot be inferred from this study alone and further studies with gDNA-specific methods are required to confirm or dispute this. Post gastrula, more dynamic patterns of demethylation/remethylation were observed, potentially reflecting tissue-specific methylation (Figure 3). Interestingly, the dynamics of methylation reprogramming from the morula stage onwards varied when sticklebacks were reared at different temperatures, suggesting environmental factors could impact methylation in the early embryo (Fellous et al., 2022).

Finally, Fellous et al. (2018) used LUMA to investigate the state of methylation reprogramming in the self-fertilizing mangrove rivulus fish (*Kryptolebias marmoratus*). Hermaphrodite individuals of this species most often self-fertilize by internal fecundation, even though outcrossing with males can also occur. Male testes exhibit higher levels of CpG DNA methylation (~87%) than hermaphrodite ovotestes (~80%), which is reminiscent of the hypermethylation of sperm compared to oocytes observed in the other described species. While it is not possible to measure the methylation levels of the female and male parts of the ovotestes separately, Fellous et al. (2018) hypothesized that this ~80% methylation level is characteristic of the ovarian tissue as the ovotestes contain less spermatogenic than ovogenic tissue. Following fertilization, the one-cell zygotes exhibit low methylation levels (~28%), highlighting a drastic drop in methylation compared to that of the ovotestes, suggesting that both male and female methylomes undergo demethylation (Figure 1). However, the analysis did not detect strong expression of TET proteins during this period (Fellous et al., 2018), and the alternate explanation could be that the mangrove rivulus oocytes are less methylated than what was hypothesized. Once again, further studies with gDNA-specific methods are needed to unravel what truly happens postfertilization in this species. With regard to patterns and timing, the estimated methylation levels in the mangrove rivulus differ from that of the other species investigated with similar nonspecific methods, as no significant methylation changes are observed between fertilization and the gastrula stage (even though there is a small trend toward methylation decrease between these stages). An observed "remethylation" occurs post-gastrulation, with global methylation levels rising to ~70% until the liver formation stage, after which methylation stays relatively stable until at least the 20th day post-hatching (Fellous et al., 2018). These differences could stem from heterochrony in mtDNA dilution, in proportion to gDNA, in this species.

In sum, while several labs have now started to investigate methylation reprogramming post-fertilization in different teleost species, the studies described above highlight the important technical issues caused by variation in the mtDNA/gDNA ratio in early development, which leaves many questions unanswered. In addition to these methodological issues, differences between studies in both the developmental stages analyzed, and potential divergence in these stages between taxa make comparisons between distant species extremely difficult. More detailed studies, with gDNA-specific methods, both among closely related species and across diverse fish taxa, with a focus on the zygote and first cell divisions, are needed to confirm the pattern and diversity of methylation reprogramming, or lack thereof, in fish, and test hypotheses about its evolution in teleosts and derived tetrapod groups.

If differences in reprogramming patterns were to be confirmed between teleost species, they could stem from various causes such as, for example, heterochrony of developmental events, the state/function of the (de) methylation machinery in each species (differences in *TET* or *DNMT* copy numbers, sub- or neo-functionalization of the genes, etc.), and differences in DNA compaction in the sperm of these fish species. While the DNA in zebrafish sperm is compacted with histones (Wu et al., 2011), this is not the case for all teleosts, and the sperm DNA of several other fish species has been shown instead to be condensed by protamines (a family of arginine-rich proteins) or protamine-like proteins (Saperas et al., 1994), similar to mammalian sperm (Steven Ward & Coffey, 1991). This is for example the case for stickleback (Lemke, 1985) and medaka (Tamura et al., 1994), and it would be interesting to see whether this property has an impact on the molecular pathways orchestrating methylation reprogramming patterns postfertilization. A study by Saperas et al. (1994) concluded that the exchange of histones for protamines in the sperm was not a single evolutionary event but occurred multiple times in teleost history. Further studies of methylation reprogramming in other teleost species are needed to unravel possible links between protamine- or histone-packed sperm DNA and methylation reprogramming patterns.

4 | ZYGOTIC GENOME ACTIVATION

The ZGA is a process that is both gradual and highly species-specific (Jukam et al., 2017). In zebrafish, ZGA is sometimes thought to happen in concert with the

mid-blastula transition (MBT), when the cell cycle lengthens and cell divisions stop being synchronous. However, while a large-scale increase in embryonic transcription is indeed observed in zebrafish at the MBT (Jukam et al., 2017), some zygotic miRNAs were shown to be expressed as early as the 64-cell stage, and protein-coding genes were expressed from the 128-cell stage onwards (Heyn et al., 2014). As a very complicated and rather poorly understood phenomenon, ZGA relies on a range of different molecular mechanisms such as nucleocytoplasmic (N-C) ratio (the size of the nucleus compared to the size of a cell's cytoplasm), transcriptional activators, and chromatin dynamics (Jukam et al., 2017), the last of which is most relevant for this review.

In zebrafish, Potok et al. (2013) only found 40 regions differentially methylated between the 256-cell stage (before the large-scale increase in zygotic transcription) and the sphere stage (post-ZGA), suggesting that methylation barely changes during this process. However, the changes in methylation before the start of ZGA (i.e., the matching of the maternal and paternal methylomes) could still prime the genome for specific gene activation during ZGA.

Liu et al. (2018) found an increase in accessible chromatin regions at promoters during ZGA in zebrafish and highlighted an association between the promoters that become accessible during this period and regions of DNA hypomethylation (high number of unmethylated CpG dinucleotides). Cfp1, a CXXC domain-containing factor, has been shown to bind to unmethylated CpG regions and recruit H3K4 methyltransferase in mammals (Thomson et al., 2010) which sets up the permissive H3K4me3 mark. By knocking down Cxxc1b, the zebrafish homolog of Cfp1, Liu et al. (2018) observed a reduction in accessible chromatin at promoters during ZGA, suggesting that unmethylated CpG regions at promoters prime the accessibility of chromatin through recruitment of Cxxc1b, and, consequently, H3K4 methyltransferases.

Interestingly, while chromatin accessibility correlated with regions of low methylation at promoters in this study, it correlated with regions of high methylation at sites distal from transcription start sites (Liu et al., 2018). By looking into TF DNA-binding motifs in these distal regions, the researchers found motifs associated with a high methylation state, leading them to identify a second molecular pathway involving the TFs Pou5f3 and Nanog, which are expressed early in ZGA where they play an important role (Lee et al., 2013). These pioneer TFs were found to be required for the accessibility of chromatin at distal sites with high methylation, regions deemed likely to be enhancers, during ZGA (Liu et al., 2018).

However, the importance of Pou5f3, Nanog as well as Sox TFs in the priming of accessible chromatin regions during ZGA in zebrafish was recently verified by others (Pálfy et al., 2020; Veil et al., 2019). Veil et al. (2019) did not find the chromatin binding of Pou5f3 or Nanog to depend on DNA methylation levels, and Pálfy et al. (2020) showed that Pou5f3, Sox19b, and Nanog can also regulate chromatin accessibility at hypomethylated promoters.

These results suggest that hypomethylated regions recruit Cfp1 which subsequently recruits H3K4 methyltransferase to drive chromatin accessibility at regions of low methylation. On the other hand, regulation of chromatin accessibility by the pioneer factors Pou5f3, and Nanog may only require methylation marks in a context-dependent manner and needs to be further investigated.

Both studies from Liu et al. (2018) and Murphy et al. (2018) show correlations between hypomethylated regions and chromatin features in zebrafish. The former links hypomethylated regions with accessible chromatin, while the latter links the previously mentioned H2AFV placeholder nucleosomes with hypomethylated regions. It is tempting to infer causal relationship from these correlations. Do H2AFV placeholder nucleosomes identified through one study and accessible chromatin defined through the other overlap? However, Liu et al. (2018) defined accessible chromatin regions, observed through ATAC-seq, as nucleosome-free. Thus, it is unclear whether the regions showing a high number of unmethylated CpGs in the two studies correlate. If this was the case, it could be hypothesized that the H2AFV placeholder nucleosomes, which maintain/establish hypomethylated regions before ZGA, could subsequently impact chromatin accessibility during ZGA through the recruitment of Cfp1 at these hypomethylated regions.

Whether or not this is the case, H2AFV placeholder nucleosomes have been shown to play a role during ZGA. Developmental genes associated with placeholder nucleosomes, and the permissive chromatin environment they constitute, can become either transcriptionally active during ZGA or poised for later activation. Hickey et al. (2022) showed that TFs recruit the polycomb complex Rnf2/PRC1 to sites bearing placeholder nucleosomes. This complex modifies the nucleosome through H2A monoubiquitination (H2Aub1), subsequently recruiting a second complex Aebp2/PRC2, which catalyzes repressive methylation of Histone H3 (H3K27me3) (Hickey et al., 2022). Ablation of the H2Aub1 ubiquitination mark (via Rnf2 inhibition) can then later induce derepression and allow the expression of these placeholder-marked poised genes (Hickey et al., 2022).

To summarize, the extent to which patterns of DNA methylation/hypomethylation are involved in priming the accessibility of chromatin during ZGA in zebrafish has not been fully resolved and further studies are

needed to unravel the order of events and molecular pathways surrounding chromatin accessibility, including the interplay with pre-established DNA methylation patterns. Furthermore, it would be curious to assess how diverse the ZGA is in other fish species, and whether the influence of DNA methylation reprogramming on its orchestration varies in importance and timing.

5 | METHYLATION REPROGRAMMING IN PGCs

PGCs, the precursors of the gametes (Nikolic et al., 2016), separate from their somatic counterparts very early in development. In mammals, PGCs undergo extensive reprogramming involving DNA demethylation, targeted remethylation, and chromatin modification (Zeng & Chen, 2019) which resets genomic imprints. In fish, cases of genomic imprinting have not been reported, although patterns of mono-allelic expression were recently uncovered in channel catfish (*Ictalurus punctatus*), where the gene *hydin-1* is differentially methylated and expressed between X and Y chromosomes and appears to be necessary for driving sex determination (Wang et al., 2022). Note that the studies mentioned in this section all use gDNA-specific methods (Table 2).

The mechanisms behind germline specification are different in mammals and fish. In mice, PGCs derive from epiblast cells (Ohinata et al., 2009) and undergo a distinct molecular program during which expression of several key transcriptional regulators such as *Stella* and *Blimp1* is turned on, which subsequently inhibits the expression of *Hox* genes, allowing these cells to escape a somatic cell fate (Kurimoto et al., 2008). In zebrafish, PGCs are not conditionally specified via cellular interactions but autonomously specified via asymmetric distribution of maternally provided cytoplasmic determinants (Raz, 2003), which relocalize subcellularly during PGC migration, under the influence of the germplasm determinant Tdrd7, and prevent somatic cell fate (D'Orazio et al., 2021).

However, although the mechanism of how cells acquire a PGC fate differs, there are multiple similarities in the cellular events leading to gonad establishment in these different groups of vertebrates, including the PGC's reliance on the chemokine Cxcl12 to reach the position of gonad development (Paksa & Raz, 2015). Indeed, both in mammals and teleosts, PGCs migrate. In mammals, it is during this proliferation and migration to the genital ridge (where the gonads will form) that PGCs undergo drastic demethylation, after which new DNA methylation patterns are established differently for male and female germ cells (Smallwood & Kelsey, 2012) (Figure 4).

The timing of reprogramming of the PGCs relative to migration, and the extent of such reprogramming, seems to be different in zebrafish from mice. Skvortsova et al. (2019) and Wang et al. (2021) observed that the methylome of PGCs does not change during the migration period between 4 and 24 hpf, and is thus largely kept in the zygotic configuration established before ZGA (Figure 2). However, starting at 36 hpf, after the PGC's arrival in the genital ridge, Wang et al. (2021) observed a ~6% decrease in methylation until 9 dpf. Ortega-Recalde et al. (2019) also observed a trend toward slightly decreasing methylation during this period, although their study did not find it statistically significant. Interestingly, this decrease in methylation observed in PGCs around 36 hpf seems to mirror methylation dynamics in somatic cells at the same period, suggesting that this event might not be directly linked to the reprogramming of sex-specific imprints (Ortega-Recalde et al., 2019).

However, the establishment of such imprints through methylation changes have been shown to occur at different stages of gametogenesis (Chaillet et al., 1991). Thus, it is important to follow germ cells through their maturation. It is important to stress again that although mtDNA is amplified during oogenesis and could, therefore, bias global methylation studies, all studies mentioned here employ gDNA-specific methods. Wang

TABLE 2 Studies measuring global methylation levels in the PGCs of teleosts (and mouse), their species of interest and methods used.

Study	Species	Method	Type
Skvortsova et al. (2019)	Zebrafish	WGBS	gDNA-specific
Ortega-Recalde et al. (2019)	Zebrafish	WGBS	gDNA-specific
D'Orazio et al. (2021)	Zebrafish	RRBS	gDNA-specific
Wang et al. (2021)	Zebrafish	Bisulfite-seq	gDNA-specific
Wang and Bhandari (2020)	Medaka	WGBS	gDNA-specific
von Meyenn et al. (2016)	Mouse	WGBS	gDNA-specific
Wang et al. (2014)	Mouse	WGBS	gDNA-specific

Abbreviations: PGC, primordial germ cell.

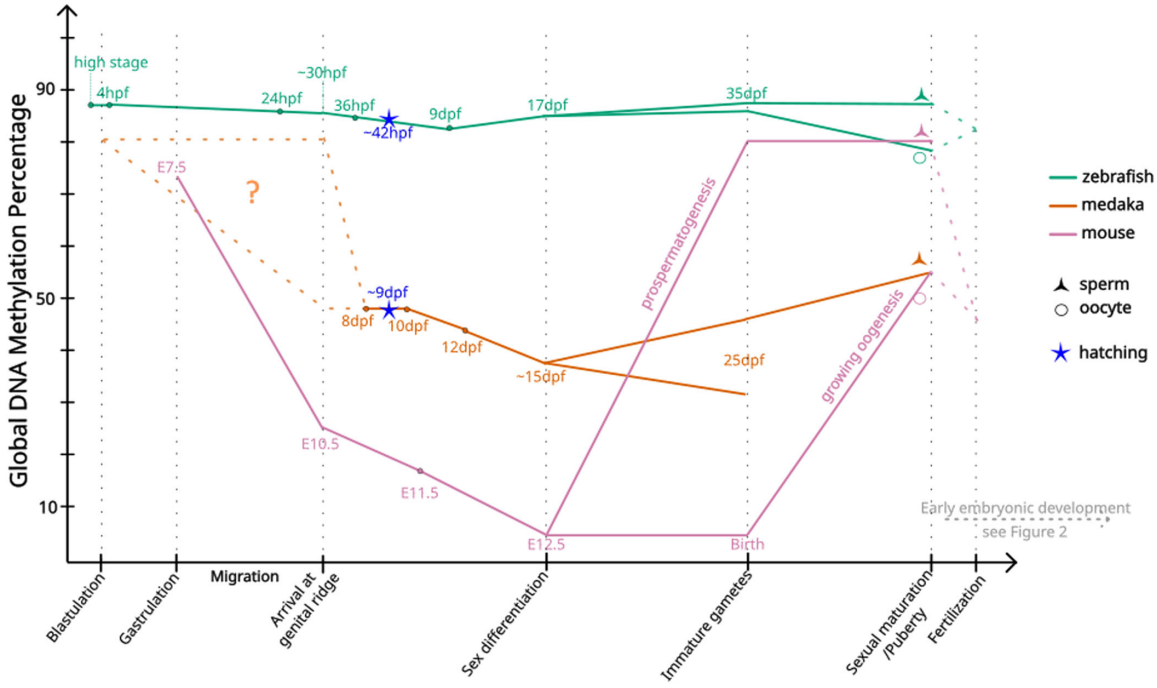


FIGURE 4 DNA methylation reprogramming dynamics in the PGCs of zebrafish, medaka, and mouse. Adapted from D'Orazio et al. (2021), Skvortsova et al. (2019), Wang et al. (2021) (zebrafish), Wang and Bhandari (2020) (medaka), and Smallwood and Kelsey (2012), von Meyenn et al. (2016), and Wang et al. (2014) (mouse). The x-axis showing the different developmental stages is not to scale. Developmental stages shown on the graph represent the stages as defined in the papers cited. [Color figure can be viewed at wileyonlinelibrary.com]

et al. (2021) studied methylation patterns through zebrafish gametogenesis and observed that the average methylation level of 35 dpf male germ cells is comparable to that of sperm, while the average methylation level of 35 dpf female germ cells differed from that of the oocyte. The fact that the final sperm methylome is already established at this early stage of germ cell development while the female germ cell genome needs to undergo further methylation reprogramming to reach the levels of the oocyte likely reflects the aforementioned fact that the sperm methylome is primed for ZGA before fertilization while the oocyte undergoes reprogramming following fertilization.

Therefore, while PGCs possibly undergo reprogramming following their arrival in the genital ridge in zebrafish, this seems to be much less pronounced (~6% decrease in global methylation levels) than the demethylation event happening during PGC migration in mammals (where global methylation levels drop from ~75% to ~25%).

In medaka, PGCs also get demethylated, although this time in a more drastic two-step demethylation event, with an initial loss of 30% of global methylation during progression from the blastula to 8 dpf, and a second drop from 10 to 12 dpf where an additional 10% of methylation is lost (Figure 2). At the 25 dpf stage, which coincides with gonadal sex determination and gonadogenesis, de novo methylation occurs in male PGCs, while female PGCs stay hypomethylated (Wang & Bhandari, 2020). Unfortunately, this study only investigated methylation signals in early PGCs at the blastula and 8 dpf stages, the latter being already past the migration of PGCs to the gonadal region, as in medaka PGCs migrate to the lateral plate mesoderm by ~2 dpf (Kurokawa et al., 2006). It is thus unclear whether the first step of the demethylation event seen in medaka PGCs happens during migration to the genital ridge (as it does in mice) or following the arrival of the PGCs in the genital ridge (as it does in zebrafish). Additionally, sperm methylation levels were measured at around 55% by Wang and Bhandari (2020), while they were measured at around 90% by Ross et al. (2023). As both teams used gDNA-specific methods, this discrepancy calls for further measurements of gametic methylation in this species.

Thus, while strict reprogramming of the DNA methylome is seen in mammals, much weaker

demethylation is seen in PGCs of the only two fish species where this has been studied to date, and the extent of methylation reprogramming seems to vary between these species. Higher resolution (spatial and temporal) developmental studies of more species are needed to discover the general trends and diversity of methylation reprogramming patterns in teleost PGCs and reveal the ancestral state and patterns of divergence for this mechanism.

6 | CONCLUSION

Identifying and understanding the mechanisms behind DNA methylation reprogramming in zygotes or PGCs is challenging. It is clear from the above that most of the previously described differences in patterns of reprogramming in the zygote stem from differences in methods used to measure global methylation levels and future studies should aim for the use of nuclear genome specific methods to avoid this issue.

The current understanding is that zebrafish does not exhibit a large-scale demethylation event following fertilization, but although no expression of TET enzymes was detected during early embryonic development in this species, global methylation levels between the 1-cell and 16-cell stages have never been measured at high temporal resolution. Further studies at these particular stages are needed to confirm the absence of demethylation in this developmental window. Similarly, the recent analysis of global methylation levels in early development in medaka suggests that this species may also escape large-scale demethylation following fertilization, but the zygotic global methylation levels in this study have not been assessed, so this remains a hypothesis for now.

While other studies tackled and characterized methylation reprogramming patterns post-fertilization in stickleback and the self-fertilizing fish mangrove rivulus, the methods used were not specific to gDNA and the results are therefore likely to have been impacted by the presence of large amounts of mtDNA in the oocyte and during early cleavage. Thus, it is not possible to know whether these species undergo reprogramming like mammals or escape reprogramming like zebrafish (and possibly medaka).

Alongside stickleback and mangrove rivulus, it would be informative to examine the patterns and mechanisms responsible for methylation reprogramming in other species of teleosts, including Salmonids which, on account of their extra round of WGD (Ss4R) (Macqueen & Johnston, 2014) that may have affected their methylation machinery, may undergo methylation reprogramming in ways that differ from zebrafish or medaka. Additionally, a detailed comparison of these DNA methylation reprogramming events within groups of closely related fishes that differ in adult phenotypes or life history would be interesting. Studying more basal teleosts, such as the spotted gar, which has not experienced the teleost-specific WGD (Ts3R) event (Amores et al., 2011), would also be of great interest to shed light on ancestral and derived states for these important developmental events.

If future studies reveal that lack of reprogramming in teleosts (i.e., in species other than zebrafish) is common, this could explain why a growing number of studies have shown environmentally-induced methylation marks to be transmitted from the parents to the next generation in fish (Kelley et al., 2021; Lai et al., 2019; Major et al., 2020), and the elucidation of the molecular mechanisms underlying which loci are targeted for or escape reprogramming may help reveal the potential role of DNA methylation in the notable phenotypic plasticity within many teleost species and the extraordinary diversity among them.

With respect to the reprogramming of PGCs in fishes, data is lacking to draw strong conclusions: zebrafish PGCs do seem to undergo a weak reprogramming event (compared to mice), while medaka PGCs exhibit a more drastic drop in methylation, but whether that occurs during or after the migration period to the site of the gonads still needs to be established. Investigating methylation patterns post-gastrula in different somatic cell lineages is also of interest, as tissue-specific patterns become more relevant than global methylation patterns with the growing complexity of the embryo. This would for example allow for the comparison of the methylation reprogramming patterns in PGCs with that of their somatic counterparts.

Additionally, how methylation patterns relate to or help orchestrate ZGA is another important aspect of the field. Better characterizing this event (spatially and temporally) in more teleost species, as well as identifying the epigenetic interactants mediating it, is needed to understand the possible interplays between methylation reprogramming and genome activation.

Interestingly, evolution may have shaped DNA methylation patterns differently for these three developmental events (first divisions post-fertilization, ZGA, and PGC specification). Purifying selection has maintained the methylation machinery, but the genes at play and their function may have been shaped by positive selection to build specific molecular adaptations in particular groups of fishes or tetrapods. To take an example, the variation in methylation levels and patterns in the PGCs of two fish species and mouse could reflect positive selection on these systems in particular lineages.

This could be studied by finer grained taxonomic sampling and more detailed analyses of methylation, also down to the specific genes that are reprogrammed (or not) in particular species or lineages.

Finally, while this review focuses on 5mC, the most common type of DNA methylation, this modification is not the only type found in eukaryotes. The potential biological relevance of 5mC derivatives should be further investigated, as 5hmC has been shown to be more than a transient intermediate in mammals (Bachman et al., 2014), and as correlations between hydroxymethylation and myeloma disease severity have been found in humans (Alberge et al., 2020). Regarding teleosts, genes involved in somatotropic growth were found to be differentially hydroxymethylated between tissues in Nile tilapia (Konstantinidis et al., 2021). Non-CpG (mCH) methylation reprogramming is another interesting avenue of research, as a type of sequence-specific mCH found at mosaic satellite repeats in zebrafish seems to undergo demethylation and reestablishment during early embryogenesis in this species (Ross et al., 2020). While we do recommend the use of bisulfite sequencing techniques and methods focusing specifically on gDNA, we emphasize that these methods currently cannot distinguish between 5mC and 5hmC. In the future, methods that can distinguish these marks, such as oxidative bisulfite sequencing and nanopore sequencing, might become more common in this field. Although 5hmC is present in small amount compared to 5mC, it is still possible that future studies using more precise methods will observe slightly different reprogramming dynamics from the ones we currently observe, due to the current methods' inability to distinguish between modified bases, and biological interpretations will have to take this possible bias into account.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Sébastien Matlosz http://orcid.org/0000-0002-4995-6682

REFERENCES

Alberge, J. B., Magrangeas, F., Wagner, M., Denié, S., Guérin-Charbonnel, C., Campion, L., Attal, M., Avet-Loiseau, H., Carell, T., Moreau, P., Minvielle, S., & Sérandour, A. A. (2020). DNA hydroxymethylation is associated with disease severity and persists at enhancers of oncogenic regions in multiple myeloma. *Clinical Epigenetics*, *12*(1), 163. https://doi.org/10.1186/s13148-020-00953-y

Almeida, R. D., Loose, M., Sottile, V., Matsa, E., Denning, C., Young, L., Johnson, A. D., Gering, M., & Ruzov, A. (2012). 5-Hydroxymethyl-Cytosine enrichment of non-committed cells is not a universal feature of vertebrate development. *Epigenetics*, *7*(4), 383–389. https://doi.org/10.4161/epi.19375

Amores, A., Catchen, J., Ferrara, A., Fontenot, Q., & Postlethwait, J. H. (2011). Genome evolution and meiotic maps by massively parallel DNA sequencing: Spotted gar, an outgroup for the teleost genome duplication. *Genetics*, *188*(4), 799–808. https://doi.org/10.1534/genetics.111.127324

Angeloni, A., Fissette, S., Kaya, D., Hammond, J. M., Gamaarachchi, H., Deveson, I. W., Klose, R. J., Li, W., Zhang, X., & Bogdanovic, O. (2024). Extensive DNA methylome rearrangement during early lamprey embryogenesis. *Nature Communications*, *15*(1), 1977. https://doi.org/10.1038/s41467-024-46085-2

Bachman, M., Uribe-Lewis, S., Yang, X., Williams, M., Murrell, A., & Balasubramanian, S. (2014). 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nature Chemistry*, *6*(12), 1049–1055. https://doi.org/10.1038/nchem.2064

Bannister, A. J., & Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Research*, *21*(3), 381–395. https://doi.org/10.1038/cr.2011.22

Baroux, C., Autran, D., Gillmor, C. S., Grimanelli, D., & Grossniklaus, U. (2008). The maternal to zygotic transition in animals and plants. *Cold Spring Harbor Symposia on Quantitative Biology*, *73*, 89–100. https://doi.org/10.1101/sqb.2008.73.053

Best, C., Ikert, H., Kostyniuk, D. J., Craig, P. M., Navarro-Martin, L., Marandel, L., & Mennigen, J. A. (2018). Epigenetics in teleost fish: From molecular mechanisms to physiological phenotypes. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, *224*, 210–244. https://doi.org/10.1016/j.cbpb.2018.01.006

Bestor, T., Laudano, A., Mattaliano, R., & Ingram, V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. *Journal of Molecular Biology*, *203*(4), 971–983. https://doi.org/10.1016/0022-2836(88)90122-2

Chaillet, J. R., Vogt, T. F., Beier, D. R., & Leder, P. (1991). Parental-specific methylation of an imprinted transgene is established during gametogenesis and progressively changes during embryogenesis. *Cell*, *66*(1), 77–83. https://doi.org/10.1016/0092-8674(91)90140-t

Cooper, D. N. (1983). Eukaryotic DNA methylation. *Human Genetics*, *64*(4), 315–333. https://doi.org/10.1007/BF00292363

Dehal, P., & Boore, J. L. (2005). Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biology*, *3*(10), e314. https://doi.org/10.1371/journal.pbio.0030314

D'Orazio, F. M., Balwierz, P. J., González, A. J., Guo, Y., Hernández-Rodríguez, B., Wheatley, L., Jasiulewicz, A., Hadzhiev, Y., Vaquerizas, J. M., Cairns, B., Lenhard, B., & Müller, F. (2021). Germ cell differentiation requires Tdrd7-dependent chromatin and transcriptome reprogramming marked by germ plasm relocalization. *Developmental Cell*, *56*(5), 641–656.e5. https://doi.org/10.1016/j.devcel.2021.02.007

Fellous, A., Labed-Veydert, T., Locrel, M., Voisin, A. S., Earley, R. L., & Silvestre, F. (2018). DNA methylation in adults and during development of the self-fertilizing mangrove

rivulus, *Kryptolebias marmoratus*. *Ecology and Evolution*, *8*(12), 6016–6033. https://doi.org/10.1002/ece3.4141

Fellous, A., Wegner, K. M., John, U., Mark, F. C., & Shama, L. N. S. (2022). Windows of opportunity: Ocean warming shapes temperature-sensitive epigenetic reprogramming and gene expression across gametogenesis and embryogenesis in marine stickleback. *Global Change Biology*, *28*(1), 54–71. https://doi.org/10.1111/gcb.15942

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J., & Surani, M. A. (2002). Epigenetic reprogramming in mouse primordial germ cells. *Mechanisms of Development*, *117*(1–2), 15–23. https://doi.org/10.1016/s0925-4773(02)00181-8

Heyn, P., Kircher, M., Dahl, A., Kelso, J., Tomancak, P., Kalinka, A. T., & Neugebauer, K. M. (2014). The earliest transcribed zygotic genes are short, newly evolved, and different across species. *Cell Reports*, *6*(2), 285–292. https://doi.org/10.1016/j.celrep.2013.12.030

Hickey, G. J., Wike, C. L., Nie, X., Guo, Y., Tan, M., Murphy, P. J., & Cairns, B. R. (2022). Establishment of developmental gene silencing by ordered polycomb complex recruitment in early zebrafish embryos. *eLife*, *11*, e67738. https://doi.org/10.7554/eLife.67738

Jiang, L., Zhang, J., Wang, J. J., Wang, L., Zhang, L., Li, G., Yang, X., Ma, X., Sun, X., Cai, J., Zhang, J., Huang, X., Yu, M., Wang, X., Liu, F., Wu, C. I., He, C., Zhang, B., Ci, W., & Liu, J. (2013). Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. *Cell*, *153*(4), 773–784. https://doi.org/10.1016/j.cell.2013.04.041

Jukam, D., Shariati, S. A. M., & Skotheim, J. M. (2017). Zygotic genome activation in vertebrates. *Developmental Cell*, *42*(4), 316–332. https://doi.org/10.1016/j.devcel.2017.07.026

Karimi, M., Johansson, S., Stach, D., Corcoran, M., Grandér, D., Schalling, M., Bakalkin, G., Lyko, F., Larsson, C., & Ekström, T. J. (2006). LUMA (LUminometric Methylation Assay)—A high throughput method to the analysis of genomic DNA methylation. *Experimental Cell Research*, *312*(11), 1989–1995. https://doi.org/10.1016/j.yexcr.2006.03.006

Karimi, M. M., Goyal, P., Maksakova, I. A., Bilenky, M., Leung, D., Tang, J. X., Shinkai, Y., Mager, D., Jones, S., Hirst, M., & Lorincz, M. (2011). DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell*, *8*(6), 676–687. https://doi.org/10.1016/j.stem.2011.04.004

Kelley, J. L., Tobler, M., Beck, D., Sadler-Riggleman, I., Quackenbush, C. R., Rodriguez, L. A., & Skinner, M. K. (2021). Epigenetic inheritance of DNA methylation changes in fish living in hydrogen sulfide–rich springs. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(26), e2014929118. https://doi.org/10.1073/pnas.2014929118

Konstantinidis, I., Anastasiadi, D., Sætrom, P., Nedoluzhko, A. V., Mjelle, R., Podgorniak, T., Piferrer, F., & Fernandes, J. M. O. (2021). Epigenetic mapping of the somatotropic axis in Nile tilapia reveals differential DNA hydroxymethylation marks associated with growth. *Genomics*, *113*(5), 2953–2964. https://doi.org/10.1016/j.ygeno.2021.06.037

Kornberg, R. D. (1974). Chromatin structure: A repeating unit of histones and DNA. *Science*, *184*(4139), 868–871. https://doi.org/10.1126/science.184.4139.868

Kurimoto, K., Yabuta, Y., Ohinata, Y., Shigeta, M., Yamanaka, K., & Saitou, M. (2008). Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice. *Genes & Development*, *22*(12), 1617–1635. https://doi.org/10.1101/gad.1649908

Kurokawa, H., Aoki, Y., Nakamura, S., Ebe, Y., Kobayashi, D., & Tanaka, M. (2006). Time-lapse analysis reveals different modes of primordial germ cell migration in the medaka *Oryzias Latipes*. *Development, Growth & Differentiation*, *48*(3), 209–221. https://doi.org/10.1111/j.1440-169X.2006.00858.x

Lai, K. P., Wang, S. Y., Li, J. W., Tong, Y., Chan, T. F., Jin, N., Tse, A., Zhang, J. W., Wan, M. T., Tam, N., Au, D., Lee, B. Y., Lee, J. S., Wong, A., Kong, R., & Wu, R. (2019). Hypoxia causes transgenerational impairment of ovarian development and hatching success in fish. *Environmental Science & Technology*, *53*(7), 3917–3928. https://doi.org/10.1021/acs.est.8b07250

Lane, N., Dean, W., Erhardt, S., Hajkova, P., Surani, A., Walter, J., & Reik, W. (2003). Resistance of IAPs to methylation reprogramming May provide a mechanism for epigenetic inheritance in the mouse, *Genesis (New York, N.Y.: 2000)*, *35*(2), 88–93. https://doi.org/10.1002/gene.10168

Lee, J. T. (2003). Molecular links between X-Inactivation and autosomal imprinting: X-Inactivation as a driving force for the evolution of imprinting? *Current Biology*, *13*(6), R242–R254. https://doi.org/10.1016/s0960-9822(03)00162-3

Lee, M. T., Bonneau, A. R., Takacs, C. M., Bazzini, A. A., DiVito, K. R., Fleming, E. S., & Giraldez, A. J. (2013). Nanog, Pou5f1 and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition. *Nature*, *503*(7476), 360–364. https://doi.org/10.1038/nature12632

Lemke, M. J. (1985). *Variability of testis-specific proteins in Gasterosteus aculeatus L. and related species*. University of British Columbia. https://doi.org/10.14288/1.0096157

Liu, G., Wang, W., Hu, S., Wang, X., & Zhang, Y. (2018). Inherited DNA methylation primes the establishment of accessible chromatin during genome activation. *Genome Research*, *28*(7), 998–1007. https://doi.org/10.1101/gr.228833.117

Macleod, D., Clark, V. H., & Bird, A. (1999). Absence of genome-wide changes in DNA methylation during development of the zebrafish. *Nature Genetics*, *23*(2), 139–140. https://doi.org/10.1038/13767

Macqueen, D., & Johnston, I. (2014). A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20132881. https://doi.org/10.1098/rspb.2013.2881

Maiti, A., Michelson, A. Z., Armwood, C. J., Lee, J. K., & Drohat, A. C. (2013). Divergent mechanisms for enzymatic excision of 5-Formylcytosine and 5-Carboxylcytosine from DNA. *Journal of the American Chemical Society*, *135*(42), 15813–15822. https://doi.org/10.1021/ja406444x

Major, K. M., DeCourten, B. M., Li, J., Britton, M., Settles, M. L., Mehinto, A. C., Connon, R. E., & Brander, S. M. (2020). Early life exposure to environmentally relevant levels of endocrine disruptors drive multigenerational and transgenerational epigenetic changes in a fish model. *Frontiers in Marine Science*, *7*. https://doi.org/10.3389/fmars.2020.00471

Matlosz, S., Sigurgeirsson, B., Franzdóttir, S. R., Pálsson, A., & Jónsson, Z. O. (2022). DNA methylation differences during development distinguish sympatric morphs of Arctic charr (*Salvelinus Alpinus*). *Molecular Ecology*, *31*(18), 4739–4761. https://doi.org/10.1111/mec.16620

Meyer, A., & Van de Peer, Y. (2005). From 2R to 3R: Evidence for a fish-specific genome duplication (FSGD). *BioEssays*, *27*(9), 937–945. https://doi.org/10.1002/bies.20293

Morgan, H. D., Santos, F., Green, K., Dean, W., & Reik, W. (2005). Epigenetic reprogramming in mammals. *Human Molecular Genetics*, *14*, R47–R58. https://doi.org/10.1093/hmg/ddi114

Murphy, P. J., Wu, S. F., James, C. R., Wike, C. L., & Cairns, B. R. (2018). Placeholder nucleosomes underlie germline-to-embryo DNA methylation reprogramming. *Cell*, *172*(5), 993–1006.e13. https://doi.org/10.1016/j.cell.2018.01.022

Nelson, J., Grande, T., & Wilson, M. (2016). *Fishes of the world* (5th ed.). John Wiley & Sons, Inc. https://doi.org/10.1002/9781119174844

Ng, H. H., & Bird, A. (1999). DNA methylation and chromatin modification. *Current Opinion in Genetics & Development*, *9*(2), 158–163. https://doi.org/10.1016/s0959-437x(99)80024-0

Nikolic, A., Volarevic, V., Armstrong, L., Lako, M., & Stojkovic, M. (2016). Primordial germ cells: Current knowledge and perspectives. *Stem Cells International*, *2016*, 1–8. https://doi.org/10.1155/2016/1741072

Ohinata, Y., Ohta, H., Shigeta, M., Yamanaka, K., Wakayama, T., & Saitou, M. (2009). A signaling principle for the specification of the germ cell lineage in mice. *Cell*, *137*(3), 571–584. https://doi.org/10.1016/j.cell.2009.03.014

Okano, M., Xie, S., & Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (Cytosine-5) methyltransferases. *Nature Genetics*, *19*(3), 219–220. https://doi.org/10.1038/890

Ortega-Recalde, O., Day, R. C., Gemmell, N. J., & Hore, T. A. (2019). Zebrafish preserve global germline DNA methylation while sex-linked rDNA is amplified and demethylated during feminisation. *Nature Communications*, *10*(1), 3053. https://doi.org/10.1038/s41467-019-10894-7

Ortega-Recalde, O., & Hore, T. A. (2019). DNA methylation in the vertebrate germline: Balancing memory and erasure. *Essays in Biochemistry*, *63*(6), 649–661. https://doi.org/10.1042/EBC20190038

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., & Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. *Current Biology*, *10*(8), 475–478. https://doi.org/10.1016/S0960-9822(00)00448-6

Otten, A. B. C., Theunissen, T. E. J., Derhaag, J. G., Lambrichs, E. H., Boesten, I. B. W., Winandy, M., van Montfoort, A. P. A., Tarbashevich, K., Raz, E., Gerards, M., Vanoevelen, J., van den Bosch, B., Muller, M., & Smeets, H. (2016). Differences in strength and timing of the mtDNA bottleneck between zebrafish germline and non-germline cells. *Cell Reports*, *16*(3), 622–630. https://doi.org/10.1016/j.celrep.2016.06.023

Paksa, A., & Raz, E. (2015). Zebrafish germ cells: Motility and guided migration. *Current Opinion in Cell Biology*, *36*, 80–85. https://doi.org/10.1016/j.ceb.2015.07.007

Pálfy, M., Schulze, G., Valen, E., & Vastenhouw, N. L. (2020). Chromatin accessibility established by Pou5f3, Sox19b and nanog primes genes for activity during zebrafish genome activation. *PLoS Genetics*, *16*(1), e1008546. https://doi.org/10.1371/journal.pgen.1008546

Potok, M. E., Nix, D. A., Parnell, T. J., & Cairns, B. R. (2013). Reprogramming the maternal zebrafish genome after fertilization to match the paternal methylation pattern. *Cell*, *153*(4), 759–772. https://doi.org/10.1016/j.cell.2013.04.030

Raz, E. (2003). Primordial germ-cell development: The zebrafish perspective. *Nature Reviews Genetics*, *4*(9), 690–700. https://doi.org/10.1038/nrg1154

Ross, S. E., Angeloni, A., Geng, F. S., de Mendoza, A., & Bogdanovic, O. (2020). Developmental remodelling of non-CG methylation at satellite DNA repeats. *Nucleic Acids Research*, *48*(22), 12675–12688. https://doi.org/10.1093/nar/gkaa1135

Ross, S. E., Vázquez-Marín, J., Gert, K. R. B., González-Rajal, Á., Dinger, M. E., Pauli, A., Martínez-Morales, J. R., & Bogdanovic, O. (2023). Evolutionary conservation of embryonic DNA methylome remodelling in distantly related teleost species. *Nucleic Acids Research*, *51*(18), 9658–9671. https://doi.org/10.1093/nar/gkad695

Santos, F., Hendrich, B., Reik, W., & Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. *Developmental Biology*, *241*(1), 172–182. https://doi.org/10.1006/dbio.2001.0501

Saperas, N., Ausio, J., Lloris, D., & Chiva, M. (1994). On the evolution of protamines in bony fish: Alternatives to the "retroviral horizontal transmission" hypothesis. *Journal of Molecular Evolution*, *39*(3), 282–295. https://doi.org/10.1007/BF00160152

Seisenberger, S., Peat, J. R., Hore, T. A., Santos, F., Dean, W., & Reik, W. (2013). Reprogramming DNA methylation in the mammalian life cycle: Building and breaking epigenetic barriers. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *368*(1609), 20110330. https://doi.org/10.1098/rstb.2011.0330

Skvortsova, K., Tarbashevich, K., Stehling, M., Lister, R., Irimia, M., Raz, E., & Bogdanovic, O. (2019). Retention of paternal DNA methylome in the developing zebrafish germline. *Nature Communications*, *10*(1), 3054. https://doi.org/10.1038/s41467-019-10895-6

Smallwood, S. A., & Kelsey, G. (2012). De novo DNA methylation: A germ cell perspective. *Trends in Genetics*, *28*(1), 33–42. https://doi.org/10.1016/j.tig.2011.09.004

Steven Ward, W., & Coffey, D. S. (1991). DNA packaging and organization in mammalian spermatozoa: Comparison with somatic cell. *Biology of Reproduction*, *44*(4), 569–574. https://doi.org/10.1095/biolreprod44.4.569

Tamura, M., Yamamoto, H., & Onitake, K. (1994). Cloning of protamine cDNA of the medaka (*Oryzias Latipes*) and its expression during spermatogenesis: (Protamine/gene expression/spermatogenesis/in situ hybridization/medaka). *Development, Growth & Differentiation*, *36*(4), 419–425. https://doi.org/10.1111/j.1440-169X.1994.00419.x

Thomson, J. P., Skene, P. J., Selfridge, J., Clouaire, T., Guy, J., Webb, S., Kerr, A. R. W., Deaton, A., Andrews, R., James, K. D., Turner, D. J., Illingworth, R., & Bird, A. (2010). CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature*, *464*(7291), 1082–1086. https://doi.org/10.1038/nature08924

Varriale, A., & Bernardi, G. (2006). DNA methylation and body temperature in fishes. *Gene*, *385*, 111–121. https://doi.org/10.1016/j.gene.2006.05.031

Veil, M., Yampolsky, L. Y., Grüning, B., & Onichtchouk, D. (2019). Pou5f3, SoxB1, and nanog remodel chromatin on high nucleosome affinity regions at zygotic genome activation. *Genome Research*, *29*(3), 383–395. https://doi.org/10.1101/gr.240572.118

von Meyenn, F., Berrens, R. V., Andrews, S., Santos, F., Collier, A. J., Krueger, F., Osorno, R., Dean, W., Rugg-Gunn, P. J., & Reik, W. (2016). Comparative principles of DNA methylation reprogramming during human and mouse in vitro primordial germ cell specification. *Developmental Cell*, *39*(1), 104–115. https://doi.org/10.1016/j.devcel.2016.09.015

Wang, L., Zhang, J., Duan, J., Gao, X., Zhu, W., Lu, X., Yang, L., Zhang, J., Li, G., Ci, W., Li, W., Zhou, Q., Aluru, N., Tang, F., He, C., Huang, X., & Liu, J. (2014). Programming and inheritance of parental DNA methylomes in mammals. *Cell*, *157*(4), 979–991. https://doi.org/10.1016/j.cell.2014.04.017

Wang, W., Yang, Y., Tan, S., Zhou, T., Liu, Y., Tian, C., Bao, L., Xing, D., Su, B., Wang, J., Zhang, Y., Liu, S., Shi, H., Gao, D., Dunham, R., & Liu, Z. (2022). Genomic imprinting-like monoallelic paternal expression determines sex of channel catfish. *Science Advances*, *8*(51), eadc8786. https://doi.org/10.1126/sciadv.adc8786

Wang, X., & Bhandari, R. K. (2019). DNA methylation dynamics during epigenetic reprogramming of medaka embryo. *Epigenetics*, *14*(6), 611–622. https://doi.org/10.1080/15592294.2019.1605816

Wang, X., & Bhandari, R. K. (2020). The dynamics of DNA methylation during epigenetic reprogramming of primordial germ cells in medaka (*Oryzias Latipes*). *Epigenetics*, *15*(5), 483–498. https://doi.org/10.1080/15592294.2019.1695341

Wang, X., Ma, X., Wei, G., Ma, W., Zhang, Z., Chen, X., Gao, L., Liu, Z., Yuan, Y., Yi, L., Wang, J., Tokumoto, T., Huang, J., Chen, D., Zhang, J., & Liu, J. (2021). The role of DNA methylation reprogramming during sex determination and transition in zebrafish. *Genomics, Proteomics & Bioinformatics*, *19*(1), 48–63. https://doi.org/10.1016/j.gpb.2020.10.004

Wu, S. F., Zhang, H., & Cairns, B. R. (2011). Genes for embryo development are packaged in blocks of multivalent chromatin in zebrafish sperm. *Genome Research*, *21*(4), 578–589. https://doi.org/10.1101/gr.113167.110

Wu, X., & Zhang, Y. (2017). TET-mediated active DNA demethylation: Mechanism, function and beyond. *Nature Reviews Genetics*, *18*(9), 517–534. https://doi.org/10.1038/nrg.2017.33

Yamaguchi, S., Shen, L., Liu, Y., Sendler, D., & Zhang, Y. (2013). Role of Tet1 in erasure of genomic imprinting. *Nature*, *504*(7480), 460–464. https://doi.org/10.1038/nature12805

Zeng, Y., & Chen, T. (2019). DNA methylation reprogramming during mammalian development. *Genes*, *10*(4), 257. https://doi.org/10.3390/genes10040257

Zhang, B., Wu, X., Zhang, W., Shen, W., Sun, Q., Liu, K., Zhang, Y., Wang, Q., Li, Y., Meng, A., & Xie, W. (2018). Widespread enhancer dememorization and promoter priming during parental-to-zygotic transition. *Molecular Cell*, *72*(4), 673–686.e6. https://doi.org/10.1016/j.molcel.2018.10.017

How to cite this article: Matlosz, S., Franzdóttir, S. R., Pálsson, A., & Jónsson, Z. O. (2024). DNA methylation reprogramming in teleosts. *Evolution & Development*; e12486. https://doi.org/10.1111/ede.12486