UNIVERSITÀ POLITECNICA DELLE MARCHE

STAZIONE ZOOLOGICA ANTON DOHRN
SZN

Department of Life and Environmental Sciences
PhD course in Life and Environment Sciences
Curriculum Marine Biology and Ecology

DNA Methylation Machinery in *Mytilus galloprovincialis* and its Crucial Role during Embryogenesis

PhD candidate

Claudia La Vecchia

Claudia La Vecchia

Academic Tutor:

Prof. Francesco Regoli

External Tutor:

Dr. Enrico D'Aniello

XXXIV cycle
AA 2020-2021

Summary

ABSTRACT **4**

ABBREVIATIONS **6**

INDEX OF FIGURES **8**

INDEX OF TABLES **10**

STRUCTURE OF THE THESIS **11**

CHAPTER 1. INTRODUCTION **12**

1.1 The model system: The mollusk *Mytilus galloprovincialis* **12**

1.1.1 Phylogeny and evolution of mollusca 12

1.1.2 Phylogeny of bivalves 15

1.1.3 *Mytilus galloprovincialis* taxonomy 15

1.1.4 Morphology of mussels: external and internal structure 16

1.1.5 *Mytilus* reproduction 18

1.1.6 *Mytilus* early development and life cycle 18

1.1.7 *Mytilus galloprovincialis* genome 22

1.2 A brief introduction to epigenetics **23**

1.2.1 DNA methylation mechanism and its function 26

1.2.2 Writing DNA methylation: DNA methyltransferase family 28

1.2.3 Erasing DNA methylation: Tet eleven translocation proteins 29

1.2.4 Reading DNA methylation: Methyl Binding Proteins 30

1.2.5 DNA methylation role in metazoans 31

1.2.6 DNA methylation in mollusks 33

1.3 Environmental epigenetics: The role of DNA methylation in marine organisms **36**

1.3.1 Anthropogenic pollutants: Contaminants of emerging concern 37

1.3.2 Pharmaceuticals at sea: fate, impacts, and epigenetic involvement? 38

CHAPTER 2. AIMS OF STUDY **40**

CHAPTER 3. MATERIALS AND METHODS **42**

3.1 *Mytilus galloprovincialis* spawning induction, *in vitro* fertilization and larvae rearing 42

3.2 Identification of *Mytilus galloprovincialis* sequences 42

3.3 Comparative genome analyses 43

3.4 *In silico* gene expression 44

3.5 RNA extraction and RNA quality detection 44

3.6 cDNA synthesis 45

3.7 Gene expression by quantitative Real time PCR 45

3.8 Ribonucleic probes preparation 47

3.8.1 Amplification of genes of interest 47

3.8.2 TOPO-TA cloning of PCR products 48

3.8.3 Chemical bacterial transformation 49

3.8.4 Isolation of plasmid DNA from *E. coli* 49

3.8.5 DNA sequencing 49

3.8.6 Probe synthesis 49

3.9 Whole-mount in situ hybridization (WISH) 50

3.9.1 Embryos preparation 50

3.9.2 Colorimetric and fluorescent WISH protocol 50

3.10 Embryo whole mount immunochemistry assay 52

3.11 Genomic DNA extraction 52

3.12 Quantification of global DNA methylation levels 53

3.13 Preparation of inhibitors: 5-azacytidine and Bobcat 339 53

3.14 Statistical analysis 53

3.15 Scanning electron microscopy 53

3.16 Propidium iodide fluorescence assay and fluorescein diacetate stain 54

3.17 Preparation of pharmaceuticals: Carbamazepine and Valsartan 54

3.18 Methyl-RAD library preparation and sequencing 55

CHAPTER 4. RESULTS **57**

4.1 Evolutionary history of DNA methylation related genes in bivalvia: New insights from *Mytilus galloprovincialis* **57**

4.1.1 Identification of *M. galloprovincialis* DNA-methylation machinery genes 57

4.1.2 Mussel DNA methylation related gene structure 58

4.1.3 Domain architecture of *M. galloprovincialis* DNA methylation toolkit related proteins 59

4.1.4 The DNA-methylation machinery is conserved across all bivalves 61

4.1.5 Evolutionary analysis of bivalve DNA methylation machinery genes 62

4.1.6 Gene expression analysis of DNA methylation-related genes in *M. galloprovincialis* adult tissue and in bivalve adult tissue 64

4.2 Characterization of DNA methylation during *M. galloprovincialis* embryogenesis 66

4.2.1 Spawning induction and embryo rearing in the lab 66

4.2.2 Stages of embryonic development of *Mytilus galloprovincialis* 68

4.2.3 DNA methylation related genes expression during embryonic and larva stages 71

4.2.4 Temporal and spatial expression of different methylation-related genes 72

4.2.5 Decrease of 5mC correlated with an increase of 5hmC during development 76

4.2.6 Global DNA methylation changes during *M. galloprovincialis* early embryogenesis 78

4.2.7 Inhibition of DNA methylation induces severe alteration of *M. galloprovincialis* development 79

4.2.8 Inhibition of DNA demethylation induces aberrant phenotype during *M. galloprovincialis* development 86

4.3 Effects of emerging contaminants on *Mytilus galloprovincialis* embryos: the case of carbamazepine (CBZ) and valsartan (VAL) 92

4.3.1 The effects of emerging contaminants in mediating DNA methylation changes during *M. galloprovincialis* development 92

4.3.3 Preliminary sequence analyses 96

CHAPTER 5. DISCUSSION AND CONCLUSION 100

REFERENCES 105

PUBLICATIONS 118

ABSTRACT

Epigenetic modifications play a key role in regulating gene expression in all organisms. Even if every cell in an organism is genetically identical, differences between tissues and organs are controlled by epigenetic modifications. DNA methylation is the most stable epigenetic modification influencing gene expression and is necessary for crucial biological processes such as embryonic development in both vertebrates and invertebrates. As a consequence, differentiated cells develop a stable and unique DNA methylation pattern that regulates tissue-specific gene transcription: a critical mechanism for development. While the mechanism of DNA methylation is well described in mammals and during development, such regulation and its key mechanisms are still far to be elucidated in phylogenetically distant organisms, especially in aquatic ecosystems, where embryos develop directly exposed to outer environmental parameters. Therefore, in order to fill current knowledge gaps, I decided to investigate the mechanisms underlying this process in the marine bivalve *Mytilus galloprovincialis*. To shed light on this process, I first screened its genome, detecting two major groups of DNMT enzymes, three MBD members, and one TET protein, which were subject to phylogenetic studies and protein domain characterization followed by gene expression analysis in different adult tissues. Then, to clarify the role of DNA methylation during development, I investigated the expression of these genes together with the dynamics of global DNA methylation levels during embryogenesis. In addition, the role of DNA methylation during embryogenesis was further assessed by the use of specific inhibitors for DNMT1 and TET. Interestingly, we found similar to other vertebrates and invertebrates that embryos treated with these inhibitors display delayed or arrested development: a likely consequence of altered DNA methylation. Moreover, very recently many studies underline the role of epigenetic mechanisms in ecological and evolutionary processes by inducing phenotypic plasticity in marine metazoan, which often enables the rapid reaction of an organism to environmental changes. Therefore, as DNA methylation functions also promote phenotypic plasticity, I started also to investigate if exposure to specific emerging contaminants (ECs) can influence DNA methylation patterns in *Mytilus galloprovincialis* embryos by taking advantage of an innovative technique, the Methyl-RAD sequencing. In conclusion, my findings provide the first insight into DNA methylation in *Mytilus galloprovincialis* and its role in proper embryo development

providing fundamental new pieces of information necessary to understand the complex role played by this mechanism in regulating genome activity in bivalves. Ultimately, this study offers a unique opportunity for further in-depth investigations that could help to improve the efficiency of *Mytilus galloprovincialis* husbandry, one of the most important aquaculture resources worldwide, and more generally of other Lophotrochozoans.

ABBREVIATIONS

List of the abbreviations used in the thesis, in alphabetical order:

4mC: N-4 cytosine methylation
5-aza: 5-azacytidine
5caC: 5-carboxylcytosine
5fC: 5-formylcytosine
5hmC: 5-hydroxymethylcytosine
5mC: 5-methylcytosine
6mA: N-6 adenine methylation
BCIP: 5-bromo-4chloro-3-indolyl-phosphate
BER: base excision repair
bp: base pair
CBZ: carbamazepine
CDS: protein coding sequence
CECs: contaminants of emerging concerns
CpG: cytosine followed by guanine
CxxC: Zinc finger domain
DMPA1: DNMT1-associated protein 1
DMSO: dimethyl sulfoxide
DNMTs: DNA methyltransferases
DUI: double uniparental inheritance
ESCs: embryonic stem cells
fsw: filtered sea water
GBM: gene body methylation
gDNA: genomic DNA
hpf: hours post fertilization
LCMA: last common ancestor of all mollusk
MBD: methyl binding domain
MeCP2: methyl-CpG-binding protein 2
mESCs: mouse embryonic stem cells
mtDNA: mitochondrial DNA

NBT: nitroblue tetrazolium

ncRNA: non-coding RNA

NSAIDs: non-steroidal anti-inflammatory drugs

PAV: presence-absence variation

PGCs: primordial germ cells

PI: propidium iodide

PP: pharmaceutical products

PWWP: pro-trp-trp-pro

REs: repetitive elements

RFTS: replication foci targeting sequence

RGMa: Repulsive Guidance Molecule

SAM: s-adenyl methionine

SE: standard error

SEM: scanning electron microscopy

TDG: thymine DNA glycosylase

TRD: transcription repression domain

TEs: transposable elements

TET: tet eleven translocation

TPM: transcript per million

USGS: united states geological survey

VAL: valsartan

INDEX OF FIGURES

Figure 1: Molluscan phylogenetic tree based on transcriptomic and genomic data. 14
Figure 2: Improved molluscan phylogenetic tree based on molecular clock analysis. 14
Figure 3: Phylogenetic relationship between the major clades of the molluscan class Bivalvia 15
Figure 4: External shell morphology of *Mytilus galloprovincialis* 16
Figure 5: Internal anatomy of *M. galloprovincialis*. 17
Figure 6: Cartoon showing three different developmental stages of *Dreissena rostriformis* a freshwater mussel. 19
Figure 7: Typical spiral cleavage patterns in oyster development. 20
Figure 8: Scheme illustrating the overall morphology of the trochophore larva. 21
Figure 9: Generic life cycle of *Perna canaliculus* mussel. 21
Figure 10: Cartoon of the epigenetic landscape described by Waddington. 24
Figure 11: Major mechanisms of epigenetic modifications controlling gene expression. 25
Figure 12: Structure of cytosine and 5-methylcytosine. 26
Figure 13: DNA methylation and demethylation. 27
Figure 14: A biochemical pathway of dynamic modification of Cytosine. 30
Figure 15: Cartoon showing the two phases of DNA methylation reprogramming during development in mice. 35
Figure 16: Epigenetic modifications during an environmental stress episode. 37
Figure 17: Schematic protocol of MethylRAD library preparation. 56
Figure 18: Exon-Intron structure of Dnmt1, Dnmt3, Tet, Mbd2/3, Mbd4 genes. 59
Figure 19: Domain architecture of *M. galloprovincialis* DNMTs. 60
Figure 20: Domain architecture of *M. galloprovincialis* TET and MBDs. 61
Figure 21: Summary of the genes involved in the DNA methylation machinery found in the 23 bivalve species analyzed in this study. 62
Figure 22: Evolutionary analysis of bivalve DNA methylation machinery genes. 63
Figure 23: Heat map of DNA methylation-related genes expression in adult tissue of six bivalve species from transcriptome database. 65
Figure 24: Expression profile of Dnmt1, Dnmt3, Tet, Mbd2/3a-b, and Mbd4 in different *M. galloprovincialis* tissue. 65
Figure 25: Mussel spawning process. 67
Figure 26: Representation of a successful spawning day. 67
Figure 27: Stage of *Mytilus galloprovincialis* embryonic development. 70
Figure 28: Expression profile of *dnmt1*, *dnmt3*, *tet* in different *Mytilus galloprovincialis* developmental stages. 71
Figure 29: Representative microscopy images of in situ hybridization of Actin expression in Trochophore and D-Larva stage of *Mytilus galloprovincialis*. 73
Figure 30: Representative confocal microscopy images of fluorescent *in situ* hybridization of Dnmt1, Dnmt3, Tet, with antisense probe during different developmental stages from oocyte to 48 hpf. 75

Figure 31: Representative confocal immunofluorescence images showing the distribution of 5mC and 5hmC during *Mytilus galloprovincialis* embryogenesis. 77
Figure 32: Levels of global methylation during early development in *M. galloprovincialis*. 78
Figure 33: Influence of 5-azacitydine treatment after 1.5 hpf on *M. galloprovincialis* embryos at different doses and stages. 80
Figure 34: Influence of 5-azacitydine on *M. galloprovincialis* at 7 and 48 hpf. 81
Figure 35: Analyses to determine methylation levels in embryos treated with 5-azacitydine. 82
Figure 36: Representative confocal fluorescence images of embryos in presence of Propidium Iodide (PI). 84
Figure 37: Influence of different doses of 5-azacitydine on *M. galloprovincialis* embryos with treatment from 24 to 48 hpf. 85
Figure 38: Influence of Bobcat 339 treatment after 1.5 hpf on *M. galloprovincialis* embryos at different doses. 87
Figure 39: Influence of Bobcat 339 treatment after 1.5 hpf on *M. galloprovincialis* embryos at different doses. 88
Figure 40: Influence of Bobcat 339 treatment after 1.5 hpf on *M. galloprovincialis* embryos at different doses. 89
Figure 41: Representative confocal fluorescence images of embryos in presence of Propidium Iodide (PI). 91
Figure 42: Molecular structure of pharmaceuticals. 92
Figure 43: Embryos after exposure at 48 hpf with the ECs. 94
Figure 44: Example of a gDNA extraction. 95
Figure 45: Example of restriction digestion analyses with LpnPI. 96
Figure 46: Length distribution after trimming. 97
Figure 47: Quality distribution after trimming. 97
Figure 48: Nucleotide contributions after trimming. 98
Figure 49: GC-content after trimming. 98

INDEX OF TABLES

Table 1: Taxonomic classification of *Mytilus galloprovincialis* Lamarck, 1819. 16
Table 2: List of primers used for qPCR. 47
Table 3: List of primers used for probe synthesis 48
Table 4: Developmental timetable of *M. galloprovincialis.* 68
Table 5: Detailed trim results. 96

STRUCTURE OF THE THESIS

The entire work of this thesis is composed of six chapters. It begins in **Chapter 1** with a general introduction organized into three parts. In detail, the introduction is characterized by an overview of mollusks, with a particular focus on *M. galloprovincialis*, the organism used in this study. The second part introduces the biological background on which this thesis is based and provides an in-depth description of epigenetic regulation with a detailed description of DNA methylation function and regulation in invertebrates and vertebrates. The third part describes the presence of pharmaceuticals in marine ecosystems and the opportunity to use epigenetic tools to study this crucial environmental issue. **Chapter 2** introduces the main objectives of this PhD thesis. Then, **Chapter 3** is the section Materials and Methods with a description of all the methodologies used. The experimental work and the results obtained were presented in **Chapter 4**. This thesis work ends with a discussion and general conclusions in **Chapter 5**.

Part of the work presented in this PhD thesis has been published as a manuscript, listed below:

Gerdol M., **La Vecchia C.**, Strazzullo M., De Luca P., Gorbi S., Regoli F., Pallavicini A., D'Aniello E. (2021). "Evolutionary History of DNA Methylation Related Genes in Bivalvia: New Insights From *Mytilus galloprovincialis*". Frontiers in Ecology and Evolution, Volume 9, Article 698561.

https://doi.org/10.3389/fevo.2021.698561

CHAPTER 1. INTRODUCTION

1.1 The model system: The mollusk *Mytilus galloprovincialis*

The Mediterranean mussel, *Mytilus galloprovincialis* is an edible cosmopolitan bivalve mollusk and it is considered one of the world's most successful invading species, thanks to its quick rate of spread and its ability to displace native mussels (Bonham, 2017). Similarly, to other marine invertebrates, mussels have separate sexes and they reproduce by spreading their gametes during the spawning process into the open water. External fertilization produces planktonic larvae which in open water can travel for very long distances thanks to oceanic currents (Gosling, 2003; Pechenik, 2009). Generally, after 1 or 2 months strictly depending on food and temperature availability, the metamorphosis process takes place and leads to the end of planktonic life. The settlement of juveniles finally coincides with the start of the sessile adult life (Gosling, 2003). Generally, *Mytilus galloprovincialis* lives on hard substrates and is found along the coast and rocky shores (Fisheries). This species has a relatively wide geographical distribution: the native range is the Mediterranean and eastern Atlantic from southern Spain to northern France, Ireland, the Arctic region, and North Africa. It is also present on the Pacific coast of North America, Japan, Hong Kong, South Africa, Chile, and Australia (https://www.cabi.org/isc/datasheet/73756).

Due to its widespread distribution and impressive filter-feeding and sedentary lifestyle behavior, *Mytilus galloprovincialis* has been employed as a sentinel species for monitoring coastal environments. Indeed, mussels are constantly exposed to a wide range of potentially anthropogenic pollutants, biotoxins, and pathogenic microorganisms, and they are capable of uptake several contaminants, thus reflecting the impact of different environmental disturbances (Beyer et al., 2017). Furthermore, these organisms have the extraordinary capability of environmental adaptation. Recent studies on the *Mytilus galloprovincialis* genome have indeed demonstrated the importance of a standing genetic variation for a rapid adaptation in environmental gradients in marine systems (Bitter et al., 2019).

1.1.1 Phylogeny and evolution of mollusca

Mollusks are considered the second largest phylum with over 100,000 described extant species most of which are marine. Many mollusks have important ecological roles in

marine and terrestrial environments; in addition, some of them are considered commercially and economically important as sources of food or jewelry.

Interestingly, mollusks are highly diverse animals, not just in size and anatomical structure, but also in behavior and in habitat (Haszprunar and Wanninger, 2012; Gosling, 2015). The phylum has a very ancient and long evolutionary history with a rich fossil record rooted in the early Cambrian, about 540 million years ago. This implies a dating of the last common ancestor of all mollusks (LCAM) to the Ediacaran (Wanninger and Wollesen, 2018). However, due to their extremely different morphology, the origins and evolutionary relationships between the major molluscan lineages still remain unresolved. The classical methods based on morphological and traditional molecular phylogenetic (nuclear ribosomal gene sequences) approaches unfortunately often suffer from poor resolution. However, nowadays more robust phylogeny methods based on the most recent sequencing technologies can provide an enormous help in better answering unresolved questions. The use of more than 300 orthologous genes was used for the phylogenetic reconstruction of 42 taxa from all major lineages of the mollusk, except the Monoplacophera class (Kocot et al., 2011). These analyses allowed to reconstruct a well-supported tree topology for Mollusca (Figure 1). Although several conflicting phylogenetic hypotheses have been proposed, Mollusca is divided into eight major lineages: Aplacophora (contain two classes: Neomeniomorpha and Chaetormomorpha) and Polyplacophora share a common ancestor and are grouped in Aculifera (Sigwart and Sutton, 2007). Cephalopoda and Scaphopoda belong to the subphylum of Conchifera and these classes represent the sister taxon of Pleistomollusca. The relationship between Gasteropoda and Bivalvia has been recognized not only with molecular studies but has also been recognized thanks to the strong morphological similarity at a precise larval stage, the veliger larvae. These two classes are included in a proposed clade Pleistomollusca (Kocot et al., 2011).

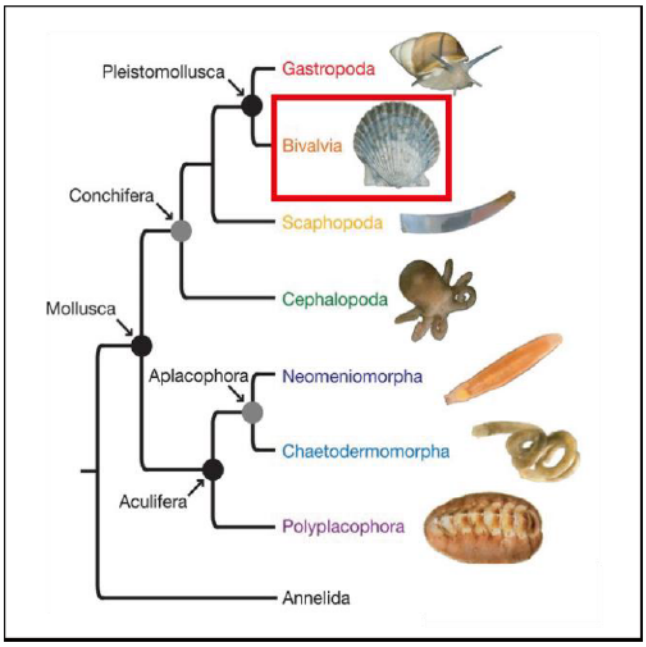

Figure 1: Molluscan phylogenetic tree based on transcriptomic and genomic data. *Mytilus* belongs to the Bivalvia class (red rectangle). Adapted from (Kocot et al., 2011).

However, very recently new studies allowed the recovery of Monoplacophora as the sister taxon of a clade made from the rest of Conchifera. Moreover, interestingly Gastropoda was also recovered as the sister taxon of Scaphopoda in most analyses (Smith et al., 2011; Kocot et al., 2020) (Figure 2). Therefore, even though molluscan origin and evolution remain unresolved, significant progress has been made in recent years to better resolve the relationship among organisms belonging to this phylum.

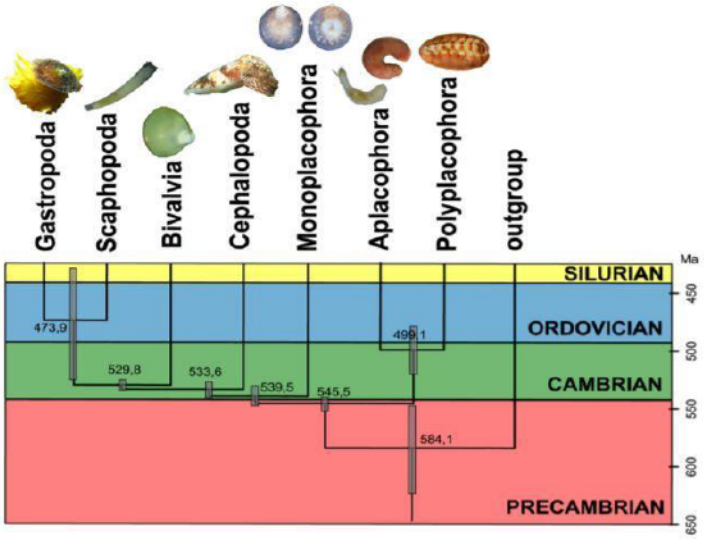

Figure 2: Improved molluscan phylogenetic tree based on molecular clock analysis. Adapted from (Kocot et al., 2020).

1.1.2 Phylogeny of bivalves

Mytilus galloprovincialis belongs to the class of Bivalvia, which is the second-largest class within the phylum of Mollusca, with more than 80% of species living in marine environments (Gosling, 2015). Among them, mussels, clams, and oysters are those with the most significant ecological and economical value. The bivalve classes are the most modified of all mollusks, since they have lost the radula and developed gills, which represent an efficient system for respiration and ciliary feeding (Gosling, 2015). Bivalves are divided traditionally into five subclasses, Protobranchia and Autobranchia, organized into superorders Nepiomorphia, Pteriomorphia, and Heteroconchia (Bieler and Mikkelsen, 2006; Bieler et al., 2010). To improve and resolve the evolutionary relationship between Bivalvia a new classification was proposed by using molecular analysis, anatomical analysis, and also geographical information. This new classification divided bivalves into the primitive and relatively small groups of Protobranchia, the large group of Pteriomorphia, Palaeoheterodonta (comprising mostly freshwater clams and mussels), Imparidentia, and the two additional small groups of Archiheterodonta and Anomalodesmata (Figure 3). Mussel belongs to a group of Pteriomorphia, in the order Mytiloida (Bieler et al., 2010, 2014).

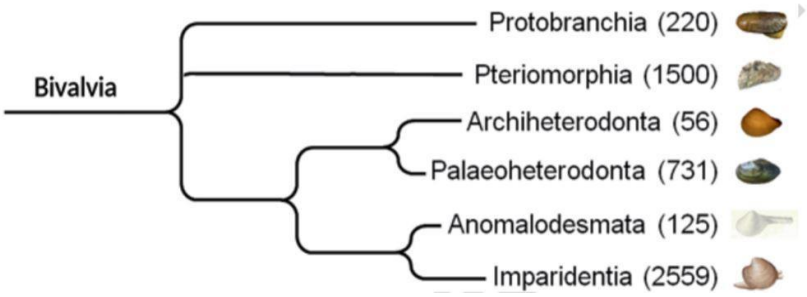

Figure 3: Phylogenetic relationship between the major clades of the molluscan class Bivalvia. Adapted from (Gerdol et al., 2018).

1.1.3 *Mytilus galloprovincialis* taxonomy

This thesis is focused on *Mytilus galloprovincialis*, a species belonging to the genus *Mytilus* and to the family Mytilidae: *M. galloprovincialis* (Mediterranean mussel) is a commercial species, which was described for the first time by Lamarck (1819) (Table 1). This species is spread from the Mediterranean Sea to the Atlantic coast of Europe. The presence of this species has been also documented in Ireland, southwest England, and

North America, where it showed adaptation to other environments (Gardner, 1992).

Taxonomic classification	
Kingdom	Animalia
Phylum	Mollusca
Class	Bivalvia
Subclass	Pteriomorphia
Order	Mytilida
Superfamily	Mytiloidea
Family	Mytilidae
Subfamily	Mytilinae
Genus	*Mytilus*
Species	*Mytilus galloprovincialis*

Table 1: Taxonomic classification of *Mytilus galloprovincialis* Lamarck, 1819.

1.1.4 Morphology of mussels: external and internal structure

External structure. Mussels are shelled animals and have two shell valves of calcium carbonate of similar size, representing a hard exoskeleton, secreted by mantle hinged dorsally by a calcified elastic ligament, containing the soft tissue (Figure 4). The shell has an extended form and right and the left shells size are usually identical. Shells are composed of calcium carbonate and a small number of organic matrices. Moreover, shell structure is determined by environmental conditions in the living area (Suzuki and Nagasawa, 2013).

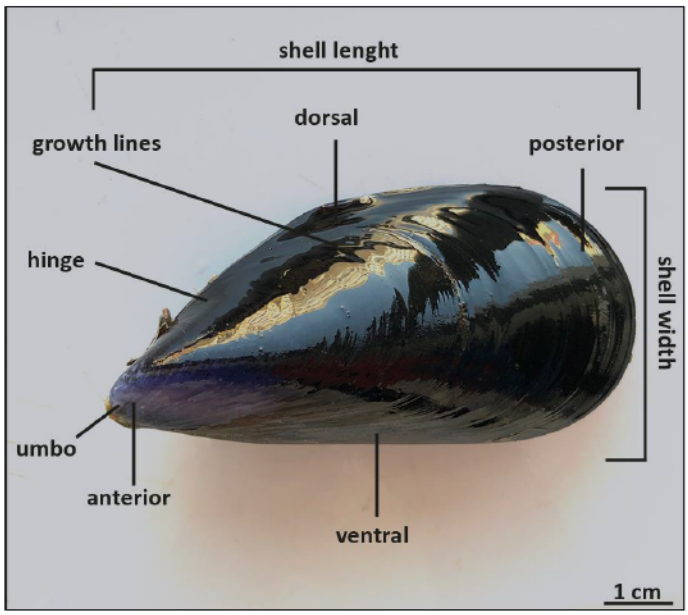

Figure 4: External shell morphology of *Mytilus galloprovincialis.*

Internal structure. The mussel's body contains many organs (Figure 5). The trunk of mussels is covered with a mantle. This structure represents a highly modified epidermis consisting of two lobes of tissue that wraps the animal and gives protection to the soft body other than secreting the shell's valve. The digestive system, blood circulatory system, adductors muscle, heart, nervous system, urogenital system, foot and gills are located in the mantle cavity (Eggermont et al., 2020) (Figure 5). Generally, the mantle margins are fused between the inhalant and exhalant openings called siphons. The mantle contains most of the gonads and its inner face is covered by cilia necessary to direct particles to gills. Mussels are filter-feeding animals: the gills replace the palps as feeding organs and also have a respiratory role, containing cilia responsible for drawing water into the mantle cavity and carrying it through gill filaments. The double function of the gill makes it also the main target organ in the bioaccumulation of different pollutants, heavy metals, and hydrocarbons. In order to prevent gills saturation, the function of labial palps, situated on either side of the mouth, is to remove the food track from the gills (Ruppert et al., 2004) (Figure5). The mouth is located on the front end of the trunk and the anus on the opposite side. The mouth is ciliated and ciliary movement drives particles in the stomach. The digestive system of mussels is composed of the esophagus, an oval-shaped stomach, a digestive gland, and an intestine (Gosling, 2015) (Figure 5).

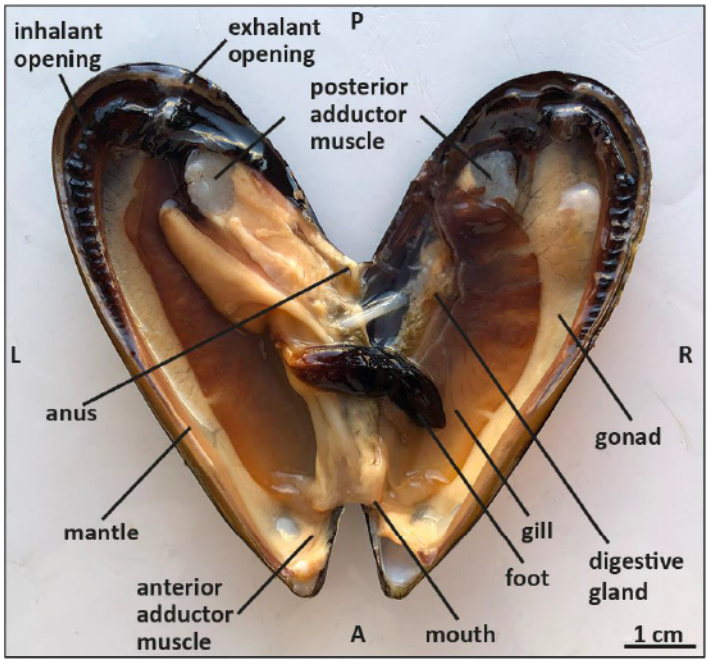

Figure 5: Internal anatomy of *M. galloprovincialis*.

1.1.5 *Mytilus* reproduction

Mussels are dioecious broader spawner animals and there are no external characteristics to discriminate between male and female. The only way to discriminate between sex is to direct observation of gonadal tissues under the microscope (Gosling, 2015). Very interestingly, *Mytilus* species show an unusual phenomenon called Doubly uniparental inheritance (DUI) of mtDNA. Initially described in *M. edulis*, DUI has been detected also in other species of *Mytilus*. Uniparental inheritance of mtDNA is a general rule in animal species, where mtDNA is exclusively maternal. Instead, *Mytilus* species contain two types of mitochondrial DNA (mtDNA) "F genome" and "M genome". This system has changed the paradigm of strict maternal inheritance and could be controlled by sex-determining genes or genes linked to sex-determining (Skibinski et al., 1994; Breton et al., 2007). *Mytilus* reproductive cycle consists of three main phases: gonad growth, spawning, and gonad redevelopment (Gosling, 2015). Despite the influence of genetic and neuroendocrine factors on the reproductive cycle, this process is also affected by environmental factors influencing the start and the duration of the cycle. The most important environmental factors are temperature, salinity, food availability, and light. Among these, the temperature is a key environmental factor influencing both gametogenesis and spawning in *M. galloprovincialis* (Gosling, 2015). For example, some studies have determined that high temperatures can induce a significant delay in gametogenesis, showing a negative correlation between the rate of gametogenesis and temperature (Fearman and Moltschaniwskyj, 2010). Considering the distribution of the genus *Mytilus*, there is an enormous variation in the spawning season. For example, in *Mytilus edulis* the gonads are usually ripe in February and partial spawning occurs in the spring. Moreover, the second period of gametogenesis takes place in the summer months, which culminates in spawning in early autumn (Duinker et al., 2008). Meanwhile, in *M. galloprovincialis* several spawning events occur with a peak at the end of winter (January-February), but the release of gametes is possible even in summer (Da Ros et al., 1985).

1.1.6 *Mytilus* early development and life cycle

During the spawning period, most bivalves, including those belonging to the *Mytilus* genus, release their gametes in the column of water, where external fertilization takes place. In most mollusks, meiosis is completed after fertilization with the extrusion of polar

bodies at the animal pole of the zygote and the point of fusion of the two gametes defines the orientation of the first cleavage. Mollusca belongs to the multiphylum group Spiralia (or Trochozoa), which includes phyla which display a very different adult body plan but share some conserved features. Mollusks are protostome animals and exhibit a pattern of early development called spiral cleavage similar to most groups except for certain taxa with great yolk-rich eggs, such as, for example, cephalopods, which exhibit discoidal cleavage (Haszprunar and Wanninger, 2012; Wanninger, 2015; Wu et al., 2020). Mollusks have bilateral symmetry, like the majority of coelomate invertebrates. In contrast with deuterostomes, protostomes are characterized by the formation of the organism's mouth at the blastopore near the opening to the gut, which is produced during gastrulation. Another major difference between the two groups is represented by the timing of the formation of the anus, which occurs after the mouth opening in protostomes and before this event in deuterostomes (Grubb, 2006a) (Figure 6). The fertilized egg rapidly divides and becomes a ball of cells ready to swim after a few hours thanks to cilia. Moreover, bivalves have an indirect development via trochophore and veliger-type larvae. (Fisheries; Wanninger, 2015) (Figure 6).

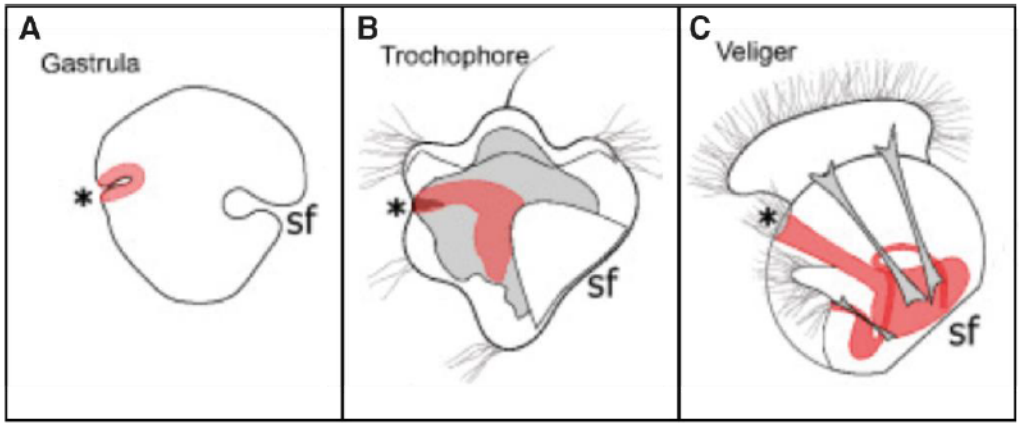

Figure 6: Cartoon showing three different developmental stages of *Dreissena rostriformis* a freshwater mussel. (A-C) Different developmental stages: Gastrula, Trochophore, and Veliger. Germ layers and major derivatives in cartoon schemes are depicted in grey (mesoderm), red (endoderm), and white (ectoderm), respectively. Asterisks mark the blastopore/mouth; sf indicates the shell field. Adapted from (Salamanca-Díaz et al., 2021).

The mollusks have isolecithal eggs, in which yolk is equally sparse. In these species cleavage extends through the entire egg and this type of cleavage is called holoblastic (Grubb, 2006b). The peculiarity of this cleavage pattern is that cleavage planes are not parallel or perpendicular to the animal-vegetative axis of the egg but are at oblique angles. Embryos produced by this type of cleavage is a stereoblastula, characterized by an

absence of blastocoel. Briefly, the first two cleavages produce four large macromeres A, B, C, and D. Subsequently in the successive cleavage, each macromere gives rise to a bud of a small micromere. Despite some differences between different species, this conserved spiral cleavage program has a cell lineage with a high degree of conservation (Grubb, 2006a; Wanninger and Wollesen, 2015) (Figure 7). The mussel life cycle includes several free-swimming larvae stages before the larvae undergo the metamorphosis process and eventually attach themselves to a suitable substrate. Larval life begins with a non-feeding trochophore-like stage and is characterized by its name-giving prototroch and apical sensory organ with a tuft of long cilia (Figure 8). Following this larval stage, many bivalves develop a secondary larva, the veliger. The velum is the structure that gives the name to this larva and it is considered an elaborate prototech used for swimming and feeding (Wanninger, 2015).

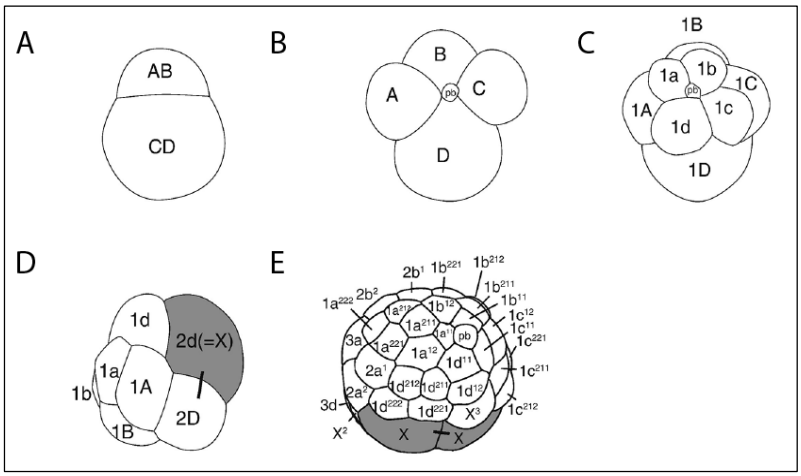

Figure 7: Typical spiral cleavage patterns in oyster development. (A) Animal view of two-cell stage *Saccostrea kegaki* embryos. (B) Animal view of four-cell stage *S. kegaki* embryos. After fertilization, the embryo forms a first polar body (pb) prior to the establishment of four macromeres that correspond to the left, ventral, right, and dorsal sides of the future larval body. (C) Animal view of eight-cell stage *S. kegaki* embryos. The following unequally blastomeres cleavage produces the four micromeres designated by lowercase letters. (D) Lateral view of nine-cell stage *S. kegaki* embryos. The blastomere 2d (=X) is colored in grey. (F) Animal view of 48-cell stage *S. kegaki* embryos. Adapted from (Kin et al., 2009).

Larvae continue to grow and swim until metamorphosis, during which the animal changes from a planktonic to benthic existence. The planktonic phase gives the chance to colonize other areas and is important for the dispersion of the species. Before the larval settlement, the larva actively detects the substrate with an apical tuft or the foot. When a suitable substrate is chosen metamorphosis takes place and this is the start point of the benthic life (Wanninger, 2015) (Figure 9).

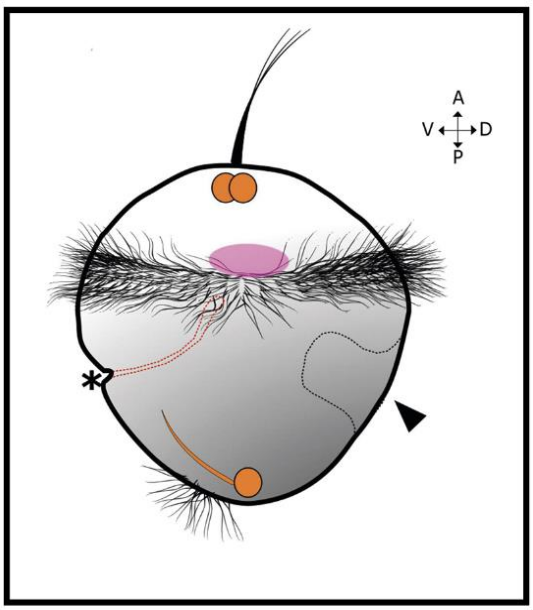

Figure 8: Scheme illustrating the overall morphology of the trochophore larva. An asterisk on the larvae indicates the mouth opening. Arrowhead points to the shell field on the right side. Orange marked areas denote the developing nervous system based on data from (Pavlicek et al., 2018) Pavlicek et al. (2018). The pink marked area denotes the developing muscular system. Dotted black lines highlight shell field cells. Dotted red lines highlight the developing digestive system. Adapted from (Salamanca-Díaz et al., 2021).

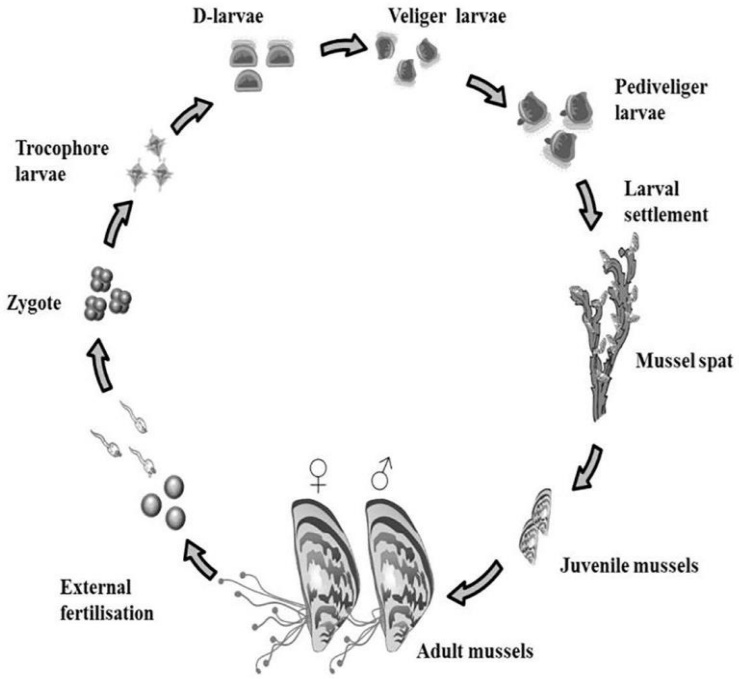

Figure 9: Generic life cycle of *Perna canaliculus* mussel. After external fertilization, fertilized eggs develop and become a ball of cells that begin to swim once cilia appear. A ciliated trochophore larva is reached after 24 hpf, followed by a veliger larva. The pediveliger stage of development is when the foot first appears. This stage precedes settlement and metamorphosis. When a substrate is suitable, the foot begins the secretion of byssus soon after metamorphosis, which is the end of a pelagic larval phase and the start of benthic life. Adapted from (Rusk, 2012).

1.1.7 *Mytilus galloprovincialis* genome

Despite the ecological and economical relevance of mussels, the first genome sequencing study in *M. galloprovincialis* was performed very recently, in 2016. This low-coverage whole-genome sequencing provided some important information about important features such as gene content, functional element, and repetitive elements (REs) (Murgarella et al., 2016). In this manuscript, the authors reported that *Mytilus galloprovincialis* genome is estimated to be around 1.6 Gb and is characterized by a high level of heterozygosis and numerous mobile elements. This information is not surprising considering the high genetic diversity within the genus *Mytilus*. *M. galloprovincialis* genome contains many REs, a large portion of which was identified as transposable elements (TEs). The authors postulated that such elements probably had an important role in adaptation and evolution. However, given the very big size of the genome and the difficulty linked with its assembly, the identification of REs was not an easy task (Murgarella et al., 2016). Very recently, Gerdol et al. provided very new insights on the *Mytilus galloprovincialis* genome. After the assembly of a 1.28-Gb reference genome and the resequencing of 14 individuals coming from two independent populations, the authors were able to demonstrate that *Mytilus galloprovincialis* has a complex pan-genomic structure consisting of a basic set of ~45,000 genes plus ~20,000 dispensable genes (Gerdol et al., 2020). A pan-genome consists of two types of genes: a set of "core genes" present in all individuals of the same species and "dispensable genes" which are found only in some individuals of the population. While core genes are related to housekeeping function, dispensable genes are not essential but contribute to species diversity and might encode for genes that may confer selective advantage (Medini et al., 2005). *M. galloprovincialis* has a very high number of dispensable genes (20,000), which might be the key to explaining the resilience and invasiveness of this species (Gerdol et al., 2020). Moreover, another important feature is that three different species of *Mytilus* are widespread in Europe and hybridize when they occur in sympatry and produce a hybrid population (Vendrami et al., 2020).

1.2 A brief introduction to epigenetics

Initially, epigenetics was defined in 1942 by Conrad Waddington as the causal interactions between genes and their product that bring the phenotype into being, so the change in phenotype without a change in genotype. Waddington used the term epigenesis derived from the Aristotelian word to connect the process of gene control during the developmental process (Waddington, 2011). During the last century, the meaning of the term "epigenetics" has evolved by the increasing knowledge of DNA, molecular biology, gene expression, and many related aspects. In 1994 Holliday offered an important contribution to the definition of epigenetics because for the first time he included in the definition the concept that the stable mitotic inheritance of given patterns of gene activity is a key feature of epigenetic control (Holliday, 1994). The implication of Holliday's redefinition was significant. Nowadays, the most accepted definition of epigenetics combines the concept of changes in gene expression and the implication of mitotic inheritance. Wu and Morris simplified Holliday's definition, giving rise to the definition: "Epigenetic is the study of changes in gene function that are mitotically and/or meiotically heritable and that do not entail a change in the DNA sequence" (Wu Ct and Morris, 2001). Epigenetic regulation is deeply involved in mediating different gene expression profiles in different cell types. This concept was proposed by Conrad Waddington in 1957 with a metaphor known as "epigenetic landscape". This model shows a ball rolling by valleys separated by ridges. The valley represents different stable cellular states, and as development proceeds, the ball travels down different valleys representing a metaphor for cell lineage specification (Waddington, 2014) (Figure 10). Epigenetic modifications play a key role in regulating gene expression. They are involved in many cellular processes, such as cellular differentiation, development, tumorigenesis, germ cell pluripotency, and inheritable genome imprinting. Accumulating evidence established that epigenetic alterations consist of a complex regulatory transcriptional pathway: DNA methylation, histone modifications, and micro-RNA mediated gene silencing (Ho and Burggren, 2010; Becker and Workman, 2013; Greally, 2018; Skvortsova et al., 2018). These modifications reflect the high plasticity of the genome and contribute to modifying gene expression independent of DNA sequence alterations (Peaston and Whitelaw, 2006). Additionally, these modifications can be passed from generation to generation in a phenomenon known as a transgenerational inheritance (Bošković and Rando, 2018). These epigenetic

processes ensure that the expression of all genes is tightly regulated to achieve a correct pattern of differentiated cells, even though every cell in an organism shares the same genetic information. These critical mechanisms play a fundamental role in important biological processes necessary to control the genetic information during the course of life (Moore et al., 2013).

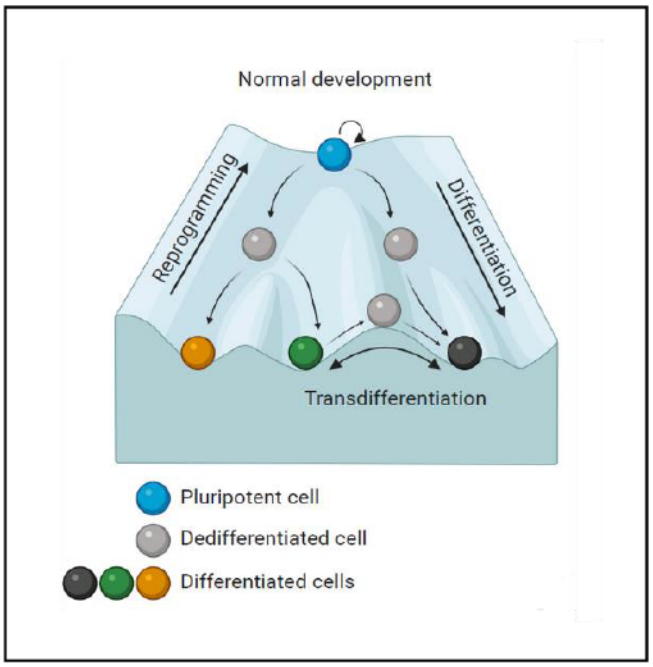

Figure 10: Cartoon of the epigenetic landscape described by Waddington. Pluripotent cells can differentiate into cells from different lineages. The reverse process from a somatic cell to a pluripotent cell is known as reprogramming. In addition, differentiated cells can switch between lineages, a process known as transdifferentiation. Adapted from (Granados et al., 2020).

Major types of epigenetic processes influencing gene function in Metazoan are shown in Figure 11.

- **DNA methylation**. This mechanism will be discussed more in detail in the following section.
- **Histone modifications.** Chromatin is a complex of DNA packaged with proteins in the nucleus. The fundamental unit of chromatin is the nucleosome composed of four proteins called core histones H2A, H2B, H3, H4, and two linker histones H1 and H5, which are known as core histones. An octamer formed by these proteins is capable of binding and organizing approximately 146 base pairs (bp) of DNA (Smith, 1991). The N-terminals “tails" are unstructured, differently from the core that is predominantly globular, and play an important role in nucleosome dynamics. A specific feature of tails and histones themselves is the large number of post-translational modifications that they undergo. Among these modifications of acetylation, methylation, phosphorylation, and ubiquitination, are the most well

studied. These modifications often result in disrupting histone-DNA interactions, which determine the nucleosome to become unwound. This state of open chromatin, called euchromatin, can facilitate the accessibility of DNA to transcription factors and ultimately mediate gene activation (Lawrence et al., 2016). These modifications, therefore, generate synergistic or antagonistic interaction with specific proteins, called readers, which interfere with chromatin function and manipulate gene expression and transcription (Jenuwein and Allis, 2001; Kouzarides, 2007).

- **Non-coding RNA (ncRNA).** Non-coding RNAs have very recently emerged (in 2002) as regulators in gene expression and in interaction with chromatin-modifiers activities (Allis and Jenuwein, 2016). Non-coding RNAs affect the epigenetic landscape through different mechanisms, which include guiding chromatin-modifying enzymes to their sites of action and also the direct binding with DNA and the formation of different complexes such as RNA:DNA: DNA triplex, RNA: DNA binding protein, RNA: DNA hybrid, RNA: RNA hybrid of lncRNA with a nascent transcript (Rinn and Chang, 2012).

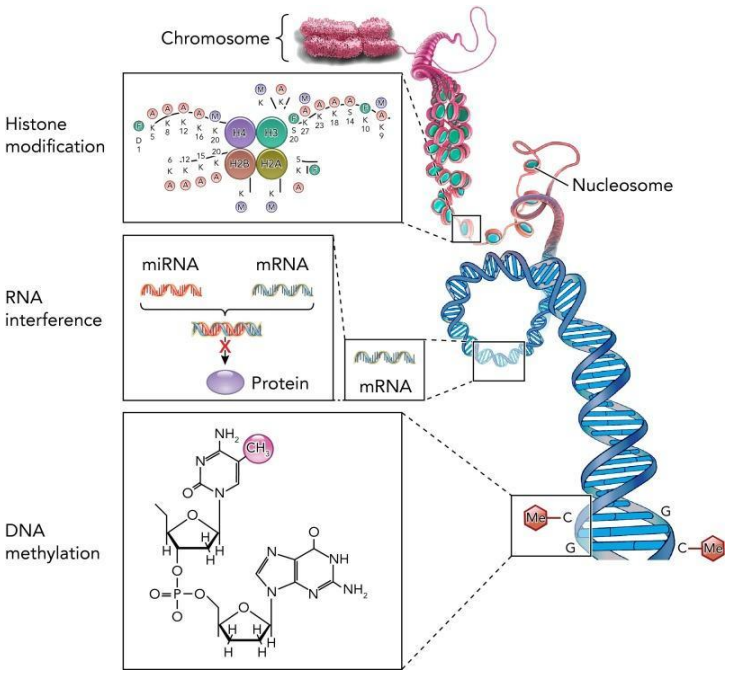

Figure 11: Major mechanisms of epigenetic modifications controlling gene expression. Histone modification, RNA interference, DNA methylation. Adapted from (Hagood, 2014).

1.2.1 DNA methylation mechanism and its function

DNA methylation is a comprehensively studied and well-characterized epigenetic modification that involves a chemical modification to DNA and has been subject to intense study in the last few years. Thanks to Holliday's studies, 30 years after the discovery of DNA methylation, it was possible to demonstrate that this process can alter gene expression by influencing the binding affinities of proteins or transcription factors to DNA. DNA methylation is a heritable epigenetic mark involving the covalent transfer of a methyl group from a methyl donor, S-adenosyl-methionine (SAM), to the C-5 position of the cytosine ring of DNA by DNA methyltransferases (DNMTs) (Figure 12) (Jin et al., 2011; Moore et al., 2013).

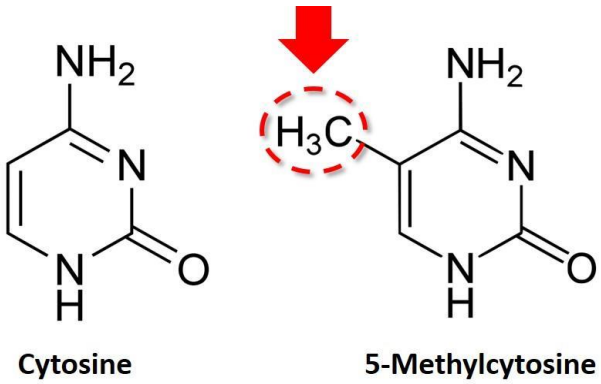

Figure 12: Structure of cytosine and 5-methylcytosine. The methyl group is indicated by the red arrow in a red dashed circle.

This mechanism is necessary to modulate gene expression both in specialized tissue-specific cell types and in pluripotent cells undergoing lineage commitment. Furthermore, DNA methylation is an ancient mechanism shared by all living organisms, from prokaryotes to humans (Willbanks et al., 2016). The human genome contains approximately 1% of 5-methylcytosines (5mC) that are found mainly in the context of cytosines followed by guanine or CpG sites. Due to the increase of spontaneous deamination of 5mC to thymidine, CpG sites are sparsely scattered and underrepresented in the genome, but the majority (70-80%) are methylated (Ehrlich et al., 1982; Bird, 2002). Moreover, other studies have recently revealed that methylation occurs also in non-CpG sites but the role of non-CpG methylation is still unclear. DNA methylation in CpG sites plays a key role in controlling the machinery of gene transcription as it can interfere with the binding capabilities of transcription factors. In fact, commonly, areas of high CpG density overlap with promoter or enhancer regions and it is generally

associated with gene silencing (Schübeler, 2015). However, DNA methylation is more represented within gene bodies, where it mediates on the contrary a positive correlation with gene expression levels (Jjingo et al., 2012). Furthermore, this process also regulates alternative splicing, indeed, it has been shown that loss of DNA methylation alters splicing patterns in a genome-wide manner (Shayevitch et al., 2018). Among many roles, DNA methylation is used by cells to maintain a correct expression pattern and to regulate various biological processes (Liu et al., 2016) as genomic imprinting, X-chromosome inactivation, and repression of transposable elements (Bommarito and Fry, 2019). Lastly, deregulation of DNA methylation patterns is directly involved in human diseases, such as imprinting disorders or cancer (Jones and Liang, 2009; Moore et al., 2013). DNA methylation is modulated by epigenetic modifications, which are provided by specific enzymes (known as "writers") that are later recognized by effector proteins ("readers"). Most, if not all, epigenetic marks are reversible, and various enzymes ("erasers") remove these marks (Nicholson et al., 2015). The correct interplay of the complex governed by these three classes of proteins regulates the gene activation or repression (Figure 13). Defects in this system contribute to dysregulation of gene expression, potential defects during embryonic development, and also cancer initiation (Rasmussen and Helin, 2016; Bommarito and Fry, 2019).

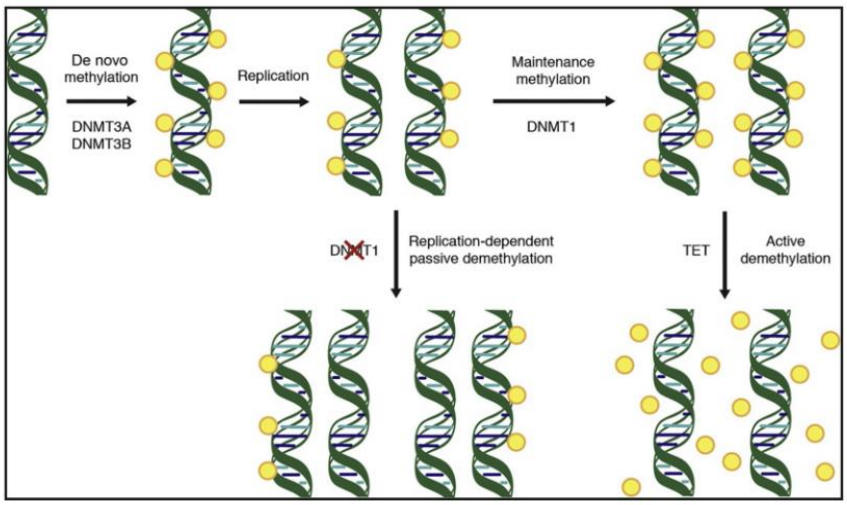

Figure 13: DNA methylation and demethylation. DNA methylation *de novo* is written by DNMT3, whilst DNMT1 is the enzyme responsible for DNA methylation of maintenance. In absence of DNMT1 cells undergo DNA passive demethylation. Active demethylation is also catalyzed by the ten-eleven translocation (TET) proteins. The yellow circle indicates methylated DNA. Adapted from (Bommarito and Fry, 2019).

1.2.2 Writing DNA methylation: DNA methyltransferase family

DNA methylation is catalyzed by the DNA methyltransferases family (DNMTs) that transfer a methyl group from S-adenyl methionine (SAM) to the fifth carbon of a cytosine residue to form 5mC (Lyko, 2018). Hollyday and Pugh proposed the existence of enzymes that preferentially catalyze the methylation of hemimethylated DNA generated during DNA replication, necessary to replicate the pattern of methylation. This prediction brings to the idea of DNMT maintenance enzymes. On the other hand, *de novo* DNMTs or switch enzymes also exist (Holliday and Pugh, 1975). These enzymes have a central role in epigenetic gene regulation and the human genome encodes five DNMTs: DNMT1, DNMT2, DNMT3A, DNMT3B, and DNMT3L (Lyko, 2018). Although these enzymes have a similar structure, they have different functions and expression patterns. The canonical DNMTs catalyzing the addition of 5mC to CpG are DNMT1, DNMT3A, and DNMT3B. DNMT1 is called DNA methyltransferase of maintenance, as it is necessary to fully methylate hemimethylated DNA generated during mitosis. DNMT1 localizes to the replication fork and provides a mechanism for the inheritance of DNA methylation patterns therefore this mechanism is associated with cellular memory as originally hypothesized by Holliday and Pugh (Moore et al., 2013). Recently, Dnmt1 has been described not only as of the canonical maintenance methyltransferases but unexpectedly it also displays *de novo* methylation with a specific retrotransposon target (Haggerty et al., 2021). Differentially from DNMT1, DNMT3A and DNMT3B are called *de novo* methyltransferases and catalyze the addition of 5mC to unmodified CpG sites. DNMT3A-B enzymes have a ubiquitous and non-selective activity and play a crucial role in somatic differentiation. One additional member of DNMT3 is DNMT3L, which does not have methyltransferase activity (Jones and Liang, 2009; Lyko, 2018). This protein shares some conserved domains with DNMT3A-B but lacks the catalytic domain and only weakly binds to DNA. DNMT3L seems instead to play a key role in stimulating the activity of DNMT3A-B (Suetake et al., 2004). Regarding DNMT2, this enzyme does not play any role in DNA methylation in mammals, being instead involved in aspartic acid tRNA methylation (Goll et al., 2006). Its possible role in DNA methylation in other organisms, where it is the only present DNA methyltransferase, is still poorly understood (Ashapkin et al., 2016).

1.2.3 Erasing DNA methylation: Tet eleven translocation proteins

DNA methylation has an essential role in genome stability. More importantly, DNA methylation is reversible and this active process of deposition and removal of 5mC has been observed during several key stages of vertebrate embryogenesis (Wu and Zhang, 2014). Despite 5mC being both chemically and genetically stable, it can still be reversed in several ways. Indeed, the loss of DNA methylation has been observed in different biological contexts through active or passive processes. Passive demethylation refers to a dilution of 5mC during replication in absence of functional DNA machinery (Messerschmidt et al., 2014). Furthermore, active demethylation is possible thanks to an enzymatic process that removes and modifies 5mC (Wu and Zhang, 2017). The mechanism of active DNA demethylation was recently discovered, in 2009, with the identification of Ten-eleven translocation (TET) proteins. Tahiliani et al. (Tahiliani et al., 2009) demonstrated that TET protein oxidizes 5mC in 5-hydroxymethylcytosine (5hmC) *in vivo* and in mouse embryonic stem cells (mESCs), in which 5hmC constitutes about 0.03% of total nucleic acid. These proteins are homologs of the Trypanosoma proteins JBP1 and JBP2 involved in the oxidation of the 5-methyl group of thymine (Tahiliani et al., 2009). In the same year, another group identified the presence of high 5hmC in mouse Purkinje Neurons and granule cells (Kriaucionis and Heintz, 2009). Afterwards, it was robustly demonstrated that TET enzymes mediate the serial conversion of 5mC to 5hmC, 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC). To complete the active demethylation process, thymine DNA glycosylase (TDG) and the base excision repair pathway (BER) mediate the excision of the 5fC and 5caC, which act as a substrate for these enzymes. Finally, the TDG pathway removes the modified cytosine and replaces it with an unmodified cytosine completing the DNA demethylation pathway (Kohli and Zhang, 2013) (Figure 14). The TET family is conserved in the animal kingdom. Mammalian genomes encode three TET genes: TET1, TET2, and TET3. TET proteins are very large, ranging from 180 to 239 kDa (Wu and Zhang, 2011). Active DNA methylation is involved in numerous biological processes such as pre-implantation and primordial germ cells (PGCs) development, embryonic stem cells (ESCs) maintenance and differentiation, and neuronal function. Moreover, TET is implicated also in DNA repair and in genomic instability. Mutation of TET genes and the consequent aberrant

demethylation is very often observed in cancer (Wu and Zhang, 2017).

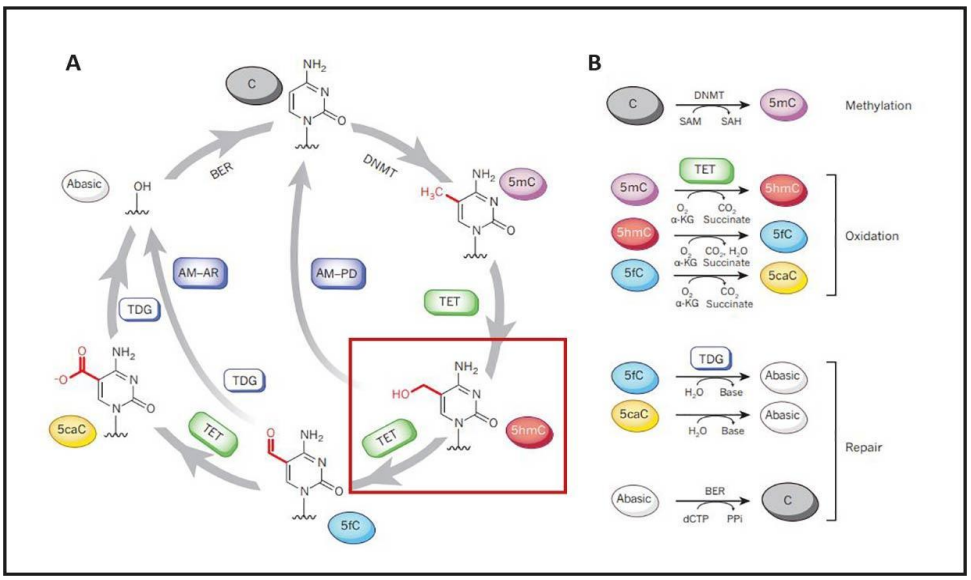

Figure 14: A biochemical pathway of dynamic modification of Cytosine. (A) Cytosine was modified by DNMTs enzymes at 5mC. 5mC can be oxidized by TET enzymes to 5hmC (bordered in red), 5fC, 5caC. Active modification followed by active restoration, indeed 5fC and 5caC are excised by TDG generating an abasic site that will be used by the base excision repair (BER) process that regenerates unmodified C. (B) Individual reactions of pathways are shown. Adapted from Kohli and Zhang, 2013.

1.2.4 Reading DNA methylation: Methyl Binding Proteins

Methyl binding proteins (MBD) usually bind to methylated DNA sequence and thereby influence transcription by association with other factors and by regulating repression or formation of heterochromatin (Fatemi and Wade, 2006). DNA methylation is recognized by three separate families of proteins: the MBD proteins, the UHRF proteins and the zinc finger proteins (Moore et al., 2013). These proteins show a conserved methyl-CpG-binding domain that has the specificity to bind single methylated CpG sites (Nan et al., 1993). This family includes MeCP2, and MBD1-6 (Meehan et al., 1989; Du et al., 2015). The methyl-CpG-binding protein (MeCP2) is the first identified protein. MBD family members contain several different domains reflecting different roles. Thanks to this specific MBD domain, methyl binding proteins recognize 5mC in CpG nucleotides. These proteins act as transcriptional repressors and play a critical role in determining the transcriptional state of the genome. Given their features, they function in coordinating crosstalk between DNA methylation, histone modifications, and chromatin organization that ultimately determine the correct and coherent transcriptional program in a given cell (Bogdanović and Veenstra, 2009; Fournier et al., 2012).

1.2.5 DNA methylation role in metazoans

DNA methylation is an important epigenetic modification involved in the regulation of many biological processes. In prokaryotes, DNA methylation is considered the major epigenetic regulator since these organisms lack key elements such as histones and nucleosomes. N-6 adenine methylation (6mA) and N-4 cytosine methylation (4mC) are mainly found in bacteria (Bheemanaik et al., 2006). DNA methylation in prokaryotes is involved not only in defence mechanisms against bacteriophage but is also involved in cellular events such as the regulation of cell cycle, gene expression, and virulence (Seong et al., 2021). In contrast, 5mC is prevalent in eukaryotes and significant differences exist between kingdoms. Unlike animal genomes that are methylated mostly on CpG island, plants and fungi show extensive methylation also in CG, CNG (where N denotes any generic nucleotide), and CHH (where H denotes A, C, or T) (Henderson and Jacobsen, 2007). Moreover, DNA methylation in plants and fungi is predominant in repetitive DNA sequences and transposons (Henderson and Jacobsen, 2007; Nai et al., 2020). Furthermore, DNA methylation patterns are different between vertebrates and invertebrates. Indeed, vertebrates and invertebrates show a highly divergent pattern of DNA methylation, with vertebrate genomes being highly and globally methylated (Suzuki and Bird, 2008; Zemach et al., 2010). Invertebrate genomes on the contrary are generally only slightly methylated and CpG methylation is mostly concentrated in the exons and introns of certain genes (Feng et al., 2010; Zemach et al., 2010). *Ciona intestinalis*, which is one of the closest living relatives to vertebrates (Delsuc et al., 2006), shows 60% of DNA methylation in gene bodies (Suzuki et al., 2007). These data were confirmed also in the genome of various invertebrates by the use of bisulfite sequencing confirming the prominent gene body methylation as a common feature of invertebrate genomes (Feng et al., 2010; Xiang et al., 2010; Zemach et al., 2010; Nanty et al., 2011; Sarda et al., 2012). Bird and Taggart divided genomes into three different types based on their methylation patterns back in 1980: “insect-type”, which shows little or no methylation, “echinoderm type”, where both methylated and unmethylated fractions are present and “vertebrate-type”, where heavy methylation is observed (Bird and Taggart, 1980). Generally invertebrate animals display a “mosaic pattern” of DNA methylation, similarly to some plants, such as *Arabidopsis thaliana*, and their genomes are slightly methylated (Suzuki and Bird, 2008). Interestingly, among invertebrates, DNA

methylation patterns can significantly differ. For example, DNA methylation is extremely different in insects and even if initially *Drosophila melanogaster* was considered to lack DNA methylation, very recently new data suggest the existence of DNA methylation (<1% of cytosine methylated) both in adult tissues and during development. However, there is still controversial debate on this topic in literature (Dunwell et al., 2013). Recently, the presence of *Dnmt1* and *Dnmt3* genes at a detectable level and a specific pattern of 5mC has been reported in other insects (Dunwell et al., 2013). Interestingly, studies in *Apis mellifera* showed the importance of DNA methylation in the generation of different colony phenotypes (Lyko et al., 2010). DNA methylation can be equally variable in other invertebrates. In nematodes, *Caenorhabditis elegans* completely lack DNA methylation and have no conventional DNMT enzymes in its genome (Head, 2014). However, on the contrary, the parasitic nematode *Trichella spiralis* shows a methylated genome, demonstrating the great variability that can exist in the same phylum (Gao et al., 2012). In addition, another very interesting case is represented by *Platynereis dumerilii*, which shows a very high level of CpG methylation during embryonic/larval development, comparable with that of mammalian somatic cells (Planques et al., 2021). Despite this general difference between invertebrates and vertebrates, there is an interesting exception motivated by convergent evolution: the sponge *Amphimedon* has a hypermethylated genome (80%) and reveals striking similarities with vertebrates, indeed, most CpG islands are highly methylated (de Mendoza et al., 2019). In vertebrate animals, DNA is globally methylated and occurs throughout the entire genome such as in repeat elements, transposons, intergenic regions, and gene bodies (Suzuki and Bird, 2008; Zemach et al., 2010). Interestingly, vertebrates show two main kinds of methylation: a higher methylation pattern exhibited by fishes and amphibians and a lower methylation pattern exhibited by mammals and birds (Head, 2014). Another difference in DNA methylation is the meaning and the function of this process through evolution. Although the role played by DNA methylation in regulating gene expression is clear, the function of DNA methylation varies according to the genome target. Gene body methylation (GBM) is associated with actively transcribed genes and the regulation of alternative splicing, whereas DNA methylation in promoters and enhancers (regulatory regions), and in transposable elements, are typically linked with gene repression (Jones, 2012; Keller et al., 2015). Interestingly, a direct correlation between gene body methylation and expression has been found in plants, corals, and mollusks up to humans, further

suggesting that this regulation is widely conserved throughout the animal kingdom (Dixon et al., 2016).

1.2.6 DNA methylation in mollusks

DNA methylation has been intensively studied in vertebrates. However, these studies are still in their infancy in invertebrates and even more in mollusks, notwithstanding the growing research interest placed on in DNA methylation. Recently, Fallet *et al.* (Fallet et al., 2020) have collected in a review all the studies performed on DNA methylation in a diverse range of mollusks clades including 8 gastropods, 7 bivalves and 1 cephalopod species. Mollusks global cytosine methylation was found to range from 5% to 15% (Fallet et al., 2020), and was considered lower than what is typically observed in plants, fish and mammals (Law and Jacobsen, 2010; Metzger and Schulte, 2016) but higher than in insects (Bewick et al., 2017). Usually, mollusks have been reported to show a range of methylated cytosine between 0.4% and 4.5% of 5mC in their genome (Wang et al., 2014; Müller et al., 2016; García-Fernández et al., 2017; Geyer et al., 2017; Riviere et al., 2017) and interestingly compared to plants they usually have a much higher level of methylation in the context of CpG (Zhang et al., 2020). Most of the DNA methylation studies in mollusks are currently based on *Crassostrea gigas* and *Biomphalaria glabrata,* where it has been widely shown that a mosaic DNA methylation type, consisting of a large portion of fully methylated DNA separated by a large block of unmethylated DNA, is present (Fallet et al., 2020). Furthermore, similar to other invertebrates (Sarda et al., 2012), bivalves show a higher level of methylation mainly in gene bodies, while intergenic regions are less methylated (Wang et al., 2014; Venkataraman et al., 2020). Currently, the role of DNA methylation in promoter and enhancer region is still far to be elucidated (Fallet et al., 2020). In oysters, *in silico* analysis revealed a correlation between gene function and methylation levels in the intragenic region (Gavery and Roberts, 2010). Then, to better characterize DNA methylation patterns in *C. gigas*, Gavery and Roberts adopted a strategy using methylation enrichment and high-throughput bisulfite sequencing to describe the genome-wide distribution of DNA methylation in the gill tissue. Results indicate that methylated genes were associated with high transcript abundance (Gavery and Roberts, 2013). Generally, those genes showing low methylation contents may be subject to specific transcriptional opportunities such as alternative transcription start sites, increasing sequence mutations, and exon skipping. Hence, DNA

methylation contributes to increased phenotypic plasticity in highly fluctuating environments facilitating a variety of transcriptional opportunities (Roberts and Gavery, 2012; Gavery and Roberts, 2014). Other studies in *Crassostrea gigas* (Olson and Roberts, 2014; Wang et al., 2014) and in *Octopus vulgaris* stressed this concept and showed that methylation is focused in the CpG context (Díaz-Freije et al., 2014). Moreover, very recently new analyses on a differentially methylated region in two yesso scallop *Patinopecten yessoensis* revealed that DNA methylation regulates different genes and pathways involved in melanin and porphyrin formation and is responsible for the different shells colorations (Yuan et al., 2021).

1.2.7 DNA methylation reprogramming during development

In mammals, DNA methylation plays a crucial role in determining cell fate specification. This process is based on two reprogramming events. The first takes place just after fertilization from the formation of the zygote and the second one in primordial germ cells (PGCs) (Figure 15). Indeed, during the development of complex organisms, cells must coordinate correctly timed and spatial events to guide cell differentiation (Wu and Zhang, 2017). Changes in DNA methylation in the early embryo are crucial for stochastic expression of lineage-determining transcription factors and allow cells to acquire their differentiated state through a gradual loss of potency. During the early stages of life, transcription is strictly modulated over time to prepare for the transition from the developmental to the adult organism (Hemberger et al., 2009). The first evidence that the methylation mechanism is essential for embryonic development was in 1992 when a knockout experiment of *Dnmt1* in mice showed that this gene was necessary for the correct development as its ablation resulted in lethal embryos between E8.0 and E10.5 (Li et al., 1992). In a similar way, a few years later, the knockout of *Dnmt3b* in mice resulted also in embryonic lethality while the knockout of *Dnmt3a* resulted in the survival of mice until 4 weeks (Okano et al., 1999).

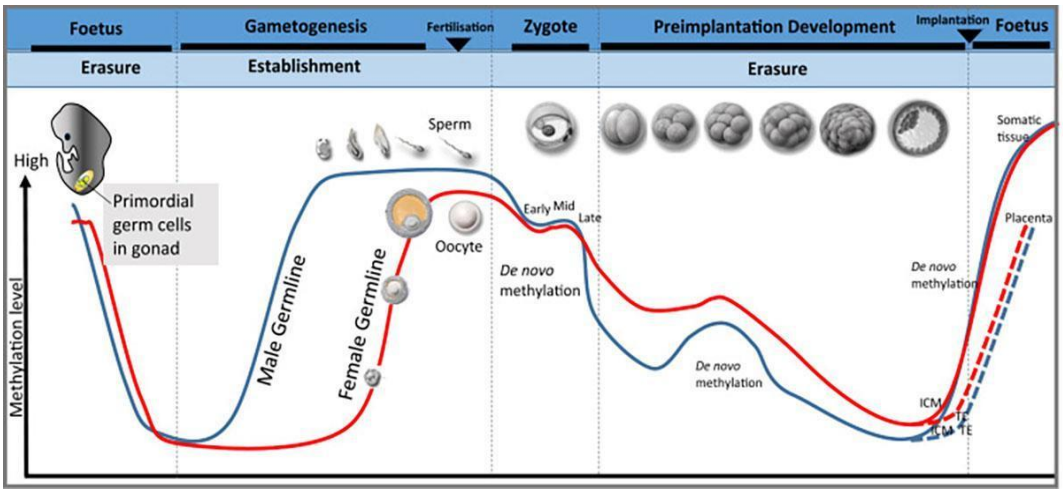

Figure 15: Cartoon showing the two phases of DNA methylation reprogramming during development in mice. Adapted from (Huntriss, 2021).

Regarding invertebrates, methylome reprogramming seems to not occur in pre-bilaterians, such as cnidarians, nor in protostomes, such as insects. Things change in deuterostomes that show DNA methylation reprogramming both in invertebrates and still more evident in vertebrates (Xu et al., 2019). Interestingly, recent data showed that DNA methylation plays a complex role in regulating genome activity and is involved in the regulation of proper development processes and gametogenesis in various lophotrochozoa including mollusks and bivalves. Many studies indeed describe that DNA methylation levels change during specific stages of development or function for proper gonads development as shown in *Crassostrea gigas*, *Patinopecten yessoensis, Pinctada fucata, Octopus vulgaris, Biomphalaria glabrata, Platynereis dumerlii* (Gavery and Roberts, 2010; Riviere et al., 2013, 2017; Díaz-Freije et al., 2014; Geyer et al., 2017; Li et al., 2019; Luviano et al., 2021; Planques et al., 2021). These studies are based on the recent advance of the last generation of methylome sequencing in combination with the use of specific inhibitors of the main enzyme involved in DNA methylation, the DNMT1.

1.3 Environmental epigenetics: The role of DNA methylation in marine organisms

Genetic, environmental or stochastic factors influence epigenetic mechanisms. Several epigenetic mechanisms could change under the exogenous influence. An important aspect of epigenetic patterns, in fact, is that they are able to change in response to environmental stimuli. The branch of biological studies investigating the cause-effect relationship between environmental factors and epigenetic mechanisms is called environmental epigenetics (Figure 16) (Bollati and Baccarelli, 2010). Anthropogenic pressure on the oceans in combination with multiple ecosystem stressors and enhanced biological invasions negatively affected the abundance of marine taxa and their functionality (Lotze et al., 2006). These considerations incorporate epigenetics into modern evolutionary theory because epigenetic variations are transmitted intergenerational and make direct contributions to the evolutionary process modulating the effects of natural selection (Ashe et al., 2021). Several studies are building convincing evidence that epigenetic mechanisms can be environmentally induced and can affect and enhance phenotypic plasticity and phenotypic variance. There is always more evidence that there is a complex relationship between epigenetic changes and phenotypic plasticity that could have evolutionary consequences (Kronholm and Collins, 2016; Ashe et al., 2021). Many studies suggest that DNA methylation is required for an adaptive strategy in response to environmental stressors exposure through the regulation of gene expression (Flores et al., 2013). The contribution of epigenetic mechanisms to phenotypic plasticity and acclimatization during adaptive responses has emerged in the last decade as an important area of study. Environmental epigenetic analyses are becoming always more useful to understand the environmental resistance and resilience of marine organisms (Torda et al., 2017). Moreover, recently, new evidence indicates that contaminants of emerging concern (CECs) could play a key role as epigenetic modulators (Alam et al., 2021; Šrut, 2021). In this context, environmental epigenetic analyses are becoming promising tools for pollution biomonitoring (Cherblanc et al., 2013; Suarez-Ulloa et al., 2015).

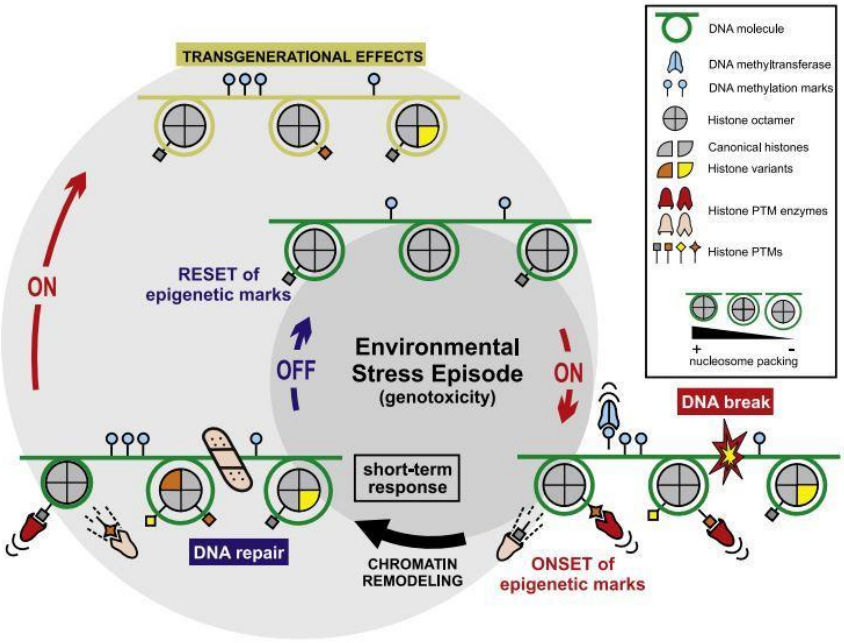

Figure 16: Epigenetic modifications during an environmental stress episode. The onset of epigenetic marks induces a short-term response to DNA damage. Many epigenetic marks will remain throughout generations in the case that stress persists, establishing the base of adaptation. Adapted from Suarez-Ulloa et al., 2015.

Recently a growing series of studies are investigating how environmental stressors can mediate epigenetic changes in aquatic mollusks as reported by Fallet et al. 2020. Both abiotic or biotic factors (including metals, pharmaceutical compounds, herbicides, temperature, salinity, toxic algae, or ocean acidification) can mediate changes in DNA methylation or impact the DNA methylation machinery genes in mollusks and bivalves (Ittiprasert et al., 2015; Suarez-Ulloa et al., 2015, 2019; Sánchez-Argüello et al., 2016; Ardura et al., 2017; Farias et al., 2017; Geyer et al., 2017; Gonzalez-Romero et al., 2017; Nica et al., 2017; Rondon et al., 2017a; Thorson et al., 2017; Sussarellu et al., 2018; Bogan et al., 2020; Downey-Wall et al., 2020; Johnson et al., 2020; Venkataraman et al., 2020; Lim et al., 2021).

1.3.1 Anthropogenic pollutants: Contaminants of emerging concern

"Contaminants of Emerging Concern" (CECs) were initially defined by the United States Geological Survey (USGS) as "any synthetic or naturally occurring chemical or any microorganism that is not commonly monitored in the environment but has the potential to enter the environment and cause known or suspected adverse ecological and/or human health effects". CECs, hence, are contaminants that have been recently introduced into

the environment or contaminants that have been present in the environment for longer periods but only recently detected (Houtman, 2010). The last recent definition of "emerging contaminants" refers even more to the potentially harmful effects and to the environmental hazard related to human health and the environment (Khan et al., 2022). Basically, CECs are substances that are essential for the contemporary society and we use them in everyday life for all kinds of purposes, such as pharmaceuticals, personal care products, pesticides, insecticides, plasticizers, detergents, flame retardants, steroid hormones, and so on (Houtman et al., 2010). Since some of these chemicals are persistent in the environment and are highly resistant to degradation they represent a growing "emergence" because of the potentially harmful effect they can pose on organisms and human health (Da Ros et al., 1985; D'Aniello, 2009; University of Arizona. Water Resources Research Center et al., 2013; Fabbri et al., 2014; Balbi et al., 2016; Dai et al., 2016; Franzellitti et al., 2019; Rodil et al., 2019; Miglioli et al., 2021).

1.3.2 Pharmaceuticals at sea: fate, impacts, and epigenetic involvement?

Pharmaceutical compounds represent a category of great concern because of the harmful effects they could exert on organisms. Pharmaceutical Products (PP) include antibiotics, illicit drugs, steroids, analgesics, and beta-blockers among other drugs and medications (Gogoi et al., 2018). These compounds are used in human and veterinary medicine and have been identified as a potential risk in our aquatic environment since the late 1990s (Daughton and Ternes, 1999). During the last year, pharmaceuticals are considered contaminants of concern on the global scale and have been detected in seawater with variables ranging between ng/L to µg/L and ng/g (Mezzelani et al., 2020). The numbers of increased studies on PP reflect the concern of this emergent class of pollutants. Indeed, due to the extensive demand for pharmaceuticals in human and veterinary medicine, their discharge into the environment is increasing (Mompelat et al., 2009). Furthermore, pharmaceuticals are considered pseudo-persistent pollutants since they enter the environment continually at very low concentrations. In addition, another important aspect is that these compounds are biologically active molecules and they could be converted into more toxic or pharmacologically active metabolites, which are largely unstudied, and could cause a significant impact on terrestrial and aquatic life (Gogoi et al., 2018). Much research has been done on steroid hormones, antibiotics, non-steroidal anti-inflammatory drugs (NSAIDs), psychiatry, and cardiovascular drugs, because these compounds are the

most frequently detected in marine environments (Ebele et al., 2017; Desbiolles et al., 2018; Mezzelani et al., 2018). These compounds pose a risk to aquatic organisms and several articles have been published on the ecotoxicological effects of various classes of pharmaceuticals with deleterious effects on marine organisms from molecular up to physiological levels representing a crucial environmental issue (Fabbri et al., 2014; Balbi et al., 2016; Franzellitti et al., 2019; Rodil et al., 2019; Miglioli et al., 2021). Biological disruption and/or dysfunction have been reported in lower marine organisms exposed to trace/levels of emerging contaminants and a growing series of evidence suggests that these alterations in lower organisms exposed to particular ECs may not be exclusively the result of traditional toxic responses but rather to interference with epigenetic regulations (Wilkinson et al., 2016).

CHAPTER 2. AIMS OF STUDY

Phenotypic plasticity has emerged as an important area of study involving those processes that might be involved in supporting organismal resistance to environmental changes. Among many mechanisms responsible to mediate phenotypic plasticity in metazoan epigenetic alterations affecting gene expression often induces a rapid reaction of organisms to environmental changes. Epigenetic mechanisms have attracted strong attention over the past years. Accumulating evidence clearly established that epigenetic modifications consist of a complex regulatory transcriptional pathway (DNA methylation, histone modification, and micro-RNA mediated gene silencing) that determine gene expression changes independent of DNA sequence alterations. Furthermore, these alterations including heritable changes play a key role in many developmental and physiological processes. In relation to DNA methylation in mollusks, a growing series of studies have recently shown that DNA methylation contributes to reproduction and development in many species, whilst others showed how DNA methylation could be sensitive to biotic and abiotic stresses. Indeed, many studies have recently reported how epigenetic changes can mediate genetic adaptation during a biological invasion or trigger those processes by which marine organisms respond to environmental changes. However, the molecular nature and the role of epigenetic regulations (DNA methylation, histone modification, and non-coding RNA) remains poorly described in species of ecological and economical value as mollusk and are still in their infancy. Therefore, in this context, my PhD studies have been focused to determine the role of DNA methylation during the developmental and adult stages of *Mytilus galloprovincialis* and investigating if exposure to specific ECs has persistent epigenetic effects on the phenotype of this ecological and economically relevant species. Understanding the role of DNA methylation and the potential epigenetic alterations following exposure to specific pharmaceutical contaminants could have important implications to predict and better understand the effects of toxicants exposure at the individual, population, and ecological levels.

Specifically, my research activity was addressed:

1. At performing a comprehensive investigation of the key genes involved in DNA

methylation machinery in the marine bivalve *Mytilus galloprovincialis* and to characterize their expression patterns in different adult tissues in comparison with other bivalves.

2. At determining the role of DNA methylation during *Mytilus galloprovincialis* embryogenesis by using different approaches.
3. At investigating the effects of exposure to specific Emerging Contaminants in mediating DNA methylation changes during *Mytilus galloprovincialis* development, from fertilization to the D-larva stage.

CHAPTER 3. MATERIALS AND METHODS

3.1 *Mytilus galloprovincialis* spawning induction, *in vitro* fertilization and larvae rearing

Sexually mature specimens of *M. galloprovincialis* were obtained from Irsvem Srl, a commercial shellfish farm (Bacoli, Napoli, Italy). Dry animals were transported to the lab and on the same day roughly cleaned from barnacles or algae growth on their shells. The cleaning helps to provide mechanical stimulation for the spawning. At this point, around 20-30 mussels were transferred in a tank with filtered seawater (FSW) at 18°C and spread to easily monitor the spawning. When mussels begin to spontaneously spawn, each individual was first washed with FSW and then located in a Becker containing 200 ml of FSW to avoid fertilization among males and females. Once various mussels spawned, gametes were recovered with a glass pipette and filtered through 40 µm (sperms) and 100 µm (oocytes) meshes to remove impurities. At this point, mature eggs were fertilized with an egg: sperm ratio of 1:15 in a volume of 50 ml. To increase the variability each fertilization was performed between a mix of oocytes and sperms coming from 6 animals (3 males and 3 females). After 30 minutes of contact time oocytes were washed a couple of times to eliminate most of the sperm and observed under the microscope to verify fertilization success (>80). The fertilized embryos were then finally transferred in a 50 ml cell culture flask or petri dish at a density of 250 larvae ml^{-1} FSW and allowed to grow until the stage of interest (D-larva at 48 hours post-fertilization). Embryos were grown in an incubator at 18°C with a 12h:12h light-dark photoperiod (Adapted from Fabbri et al., 2014).

3.2 Identification of *Mytilus galloprovincialis* sequences

The human sequences of DNMT1 (P26358.2), DNMT2 (O14717.1), DNMT3a (Q9Y6K1.4), DNMT3b (Q9UBC3.1), DNMT3L (Q9UJW3.3), MBD1 (Q9UIS9.2), MBD2 (Q9UBB5.1), MBD3 (O95983.1), MBD3L1 (Q8WWY6.2), MBD3L2 (Q8NHZ7.3), MBD3L3 (A6NE82.1), MBD3L5 (A6NJ08.1), MBD3L4 (A6NE82.1), MBD4 (O95243.1), MeCp2 (P51608.1), and TET1 (Q8NFU7.2) were retrieved from UniprotKB. These were used as queries for tBLASTn searches against the *M. galloprovincialis* genome assembly, version mg10 (accession ID: GCA_900618805.1)

(Gerdol et al., 2020). Missing hits were re-checked against previous versions of the same genome assembly (mg3 and mg5, more redundant, but slightly more complete) to rule out the possibility that the genes of interest were erroneously discarded during the assembly phases aimed at removing uncollapsed heterozygous genomic regions. All analyses were initially carried out using an e-value threshold of 1E-5, and overlapping gene annotations were translated to protein sequences to verify their reliability and completeness, based on the pairwise comparison with human queries and on the evaluation of the presence of the expected conserved domains with InterProScan 5. Predicted protein-coding sequences (CDS) were further validated with available assembled transcriptomes from the same species, which allowed for detection and correct erroneous exon annotations. Exon and intron boundaries were therefore refined, whenever needed, through the alignment between genomic DNA and cDNA sequences with NCBI Splign.

3.3 Comparative genome analyses

The validated translated amino acid sequences of *M. galloprovincialis* were used as queries for BLASTp searches against the virtually-translated proteomes derived from all the available genomes of species belonging to the class Bivalvia. In addition, tBLASTn searches were carried out whenever no ortholog could be identified in the proteomes, to detect the presence of non-annotated partial genes. In detail, the following bivalve genomes were screened: *Argopecten irradians*, *Argopecten purpuratus*, *Bathymodiolus platifrons*, *Crassostrea gigas*, *Crassostrea virginica*, *Limnoperna fortunei*, *Magallana hongkongensis*, *Mizuhopecten yessoensis*, *Modiolus philippinarum*, *Mytilus coruscus*, *Pecten maximus*, *Pinctada fucata*, *Saccostrea glomerata*, and *Scapharca broughtonii* among Pteriomorphia; *Archivesica marissinica*, *Cyclina sinensis*, *Dreissena rostriformis*, *Lutraria rhynchaena*, *Panopea generosa*, *Ruditapes philippinarum*, and *Sinonovacula constricta* among Euheterodonta; *Megalonaias nervosa* among Palaeoheterodonta. In addition, the presence of DNA methylation machinery transcripts was also assessed in the basal bivalve subclade Protobranchia: in this case, due to the lack of genomic data, tBLASTn searches were carried out in the transcriptomes of three representative species, obtained from a previously published study: *Solemya velum* (order Solemiyda), *Ennucula tenuis* (order Nuculida), and *Yoldia limatula* (order Nuculanida). The best matching sequences for each species were retrieved to build a list of ortholog sequences.

The correct identification of the sequences was evaluated by the assessment of primary sequence homology and the presence of the expected conserved domains based on literature data. Partial sequences unusable for downstream phylogenetic analyses were replaced with assembled transcripts recovered from the NCBI SRA database (https://www.ncbi.nlm.nih.gov/sra), whenever available. Highly fragmented sequences with unequivocal homology with the queries were marked as present, but they were not used for downstream analyses as they were not deemed to be phylogenetically informative.

3.4 *In silico* gene expression

The expression levels of the genes of interest were calculated *in silico* in *M. galloprovincialis*, *Modiolus philippinarum*, *Crassostrea gigas*, *Argopecten purpuratus*, and *Mizuhopecten yessoensis* as representative species for Pteriomorphia, and *Archivesica marissinica* as a representative species for Heterodonta. All these species had a fully sequenced genome and tissue-specific RNA-seq sequencing libraries available. In detail, raw reads, downloaded from the NCBI SRA database, were mapped to the reference genomes, with high stringency (length fraction = 0.75, similarity fraction = 0.98), using the CLC Genomics Workbench v. 20 (Qiagen, Hilden, Germany). Gene expression levels were calculated as Transcripts Per Million (TPM), to allow a more reliable comparison both within and between samples, and represented as a heat map based on Log (TPM + 1) values.

3.5 RNA extraction and RNA quality detection

Adult *M. galloprovincialis* were obtained from a commercial shellfish farm (Bacoli, Napoli, Italy). Three living adult animals were used for RNA extraction from different tissues. From each mussel, hemolymph was withdrawn from the adductor muscle of each mussel with a 0.5 mm diameter (25G) disposable needle. The hemolymph sampled at time zero was centrifuged at 4°C a 3000 g for 10 min, and the pellet containing the hemocytes was resuspended in 350 µl of RLT buffer included in the RNeasy Micro Kit (Qiagen) and stored at −80°C until RNA isolation. Tissue samples from muscle, gills, digestive gland, mantle edge, and foot were isolated from the same three animals that were used to isolate hemocytes and processed with the same kit. Before RNA extraction adult tissues have been homogenized with Branson Ultrasonics™ Sonifier™ SFX450

Cell Disruptors. Samples were sonicated 3 times for 15 sec using a frequency of 20 Hz, with an interval of 30 sec between each sonication step. Then, total RNA isolation was conducted using the RNeasy Micro kit (Qiagen) for RNA purification, following the manufacturer's protocol. Furthermore, total extraction from different larval stage embryos of *M. galloprovincialis* was carried out. *M. galloprovincialis* embryos were sampled throughout embryonic development including oocyte, 3hpf, 6 hpf, 10 hpf, 16 hpf, 24 hpf, 48 hpf. Four independent *in vitro* fertilization were performed, representing 4 biological replicates. Then, defined stages of *M. galloprovincialis* embryos were centrifuged (maximum speed, 2 min) and the pellet was quickly frozen in liquid nitrogen, and stored at -80°C, to obtain approximately 7500 embryos/replicate for each defined stage (Balbi et al., 2016). Total RNA isolation was conducted using the RNeasy Micro kit (Qiagen) for RNA purification, following the manufacturer's protocol. Next, the concentration and quality of the RNA were verified using a NanoDrop (ND-1000 Spectrophotometer; NanoDrop Technologies, Wilmington, DE, USA) and electrophoresis using 1% agarose gel to verify the integrity of 28s and 18s ribosomal RNA.

3.6 cDNA synthesis

500 ng of total RNA, from adult tissues samples, and 1 µg of total RNA, from each developmental stage was retro-transcribed using QuantiTect Reverse Transcription Kit (Qiagen) following the manufacturer's instruction. Briefly, total RNA was added to 2 µl of gDNA Wipeout Buffer 7x to obtain a final volume of 14 µl. This mixture was incubated for 5 min at 42°C and cooled on ice. This step is crucial to removing contamination of genomic DNA. Subsequently, 6 µl of the solution containing 1 µl of RT Primer mix, 4 µl of Quantiscript RT buffer, 5x and 1µl of Quantiscript Reverse Transcriptase, were added. The samples were incubated for 30 min at 42°C, and finally for 3 min at 95°C to inactivate Quantiscript Reverse Transcriptase. The resulting cDNA was used for the Real-Time PCR (qPCR).

3.7 Gene expression by quantitative Real time PCR

Real-time qPCR has been carried out in order to analyze the expression of DNA methylation-related genes in different tissues and in different development stages of *M. galloprovinicalis.* The quantitative PCR allows for quantification of the nucleic acid

sample in relation to the amount of DNA produced in a PCR reaction, measured by evaluating the fluorescence of an intercalating dye (in this case SYBR green dye), which is monitored at each cycle during the amplification. The absence of genomic DNA contamination was ensured thanks to two DNAse treatments carried out during RNA extraction and subsequently during cDNA retro-transcription. Real-time qPCR reactions were performed with three biological replicates for tissue analysis and four biological replicates for analysis during developmental stages and each of them with technical triplicates. Specific pairs of primers were designed for each gene using the Primer3 program. These primer pairs were chosen to amplify products of 100-250 bp in length. All the forward and reverse primers used for Real-Time PCR are listed in Table 3. The primers were specifically designed either on different exons or exon/exon junctions to prevent the amplification of contaminant genomic DNA. The amplification efficiency of each primer was calculated using a dilution series of cDNA. Five serial dilutions were set up to determine the efficiency of the reaction of all primers. Dissociation curves were carried out to ensure that primer pairs amplified single products with no detectable primer dimers. In addition, the expected size of each amplicon was verified by 1.5% agarose gel electrophoresis with GeneRuler 100 bp ladder (ThermoFisher) as a size marker. qPCR was carried out with Applied Biosystems (Life Technologies) ViiA7 384 wells format. Each reaction was performed in a final volume of 9 µL containing 5µL of Fast SYBR Green Master Mix, 1 µL of diluted cDNA (dilution 1:5 for adult tissue, dilution 1:10 for analysis during the developmental stage), and 0.7 pmol/µL specific primers. PCR amplification was performed using the following thermal profile: 95°C for 10 min, 95°C for 1 sec, and 60°C for 20 sec, 39 cycles of amplification. qPCR raw data were analyzed with Applied Biosystems ViiA™ 7 Real-Time PCR System software. Expression levels for each gene, in adult tissues and different embryonic stages, were standardized to EF1α as a reference gene, and all the data were analyzed using the $2^{-\Delta\Delta Ct}$ Livak Method. For the analysis of adult tissues, we decided to use the expression of Dnmt1 in hemocytes as a baseline of expression level, so all the ΔΔCt values were calculated using the Dnmt1 ΔCt in hemocytes as a reference for each individual. The data shown are the average of the relative expression levels in all three animals. For the analysis of different developmental stages, we decided to use the expression of each gene in oocytes as a baseline of expression level, so all the ΔΔCt values were calculated using the analyzed gene ΔCt in

oocytes as a reference for each individual. The data shown are the average of the relative expression levels in all four replicates.

Target	Primer	Sequence (5'→ 3')	Amplicon (bp)
ef1a	F	CGTTTTGCTGTCCGAGACATG	135 bp
	R	CCACGCCTCACATCATTTCTTG	
dnmt1	F	AGGGGAAAGGCAAATCATCT	173 bp
	R	TGCTGCAGGTTCCTCTTTTT	
dnmt3	F	CAGGGATGTTTTGCATTGTG	152 bp
	R	AAGCTGATCCATGGAAATGC	
tet	F	ACAATGGCAGCACAGTGGTA	222 bp
	R	TCCGTGAAGGTGTAAGTGGAG	
mbd	F	CCACCTGGTTGGAGAAGAGA	168 bp
	R	ACCAGTGCGAAAATCAAAGG	
mbd4	F	GTTCAACGATTGGCTGGTAAA	196 bp
	R	TACGGTTCACACCACCTTGA	

Table 2: List of primers used for qPCR. Mbd2/3 amplifies both variants.

3.8 Ribonucleic probes preparation

3.8.1 Amplification of genes of interest

Specific oligonucleotides for *actin*, *dnmt1*, *dnmt3*, *tet*, were designed in order to amplify by PCR the corresponding DNA fragment for each gene (Table 4). PCR amplification was performed using cDNA at 7, 16, and 24 hpf. All primers except those for *dnmt3* were designed including Sp6 or T7 promoter sequences for labeling purposes (Table 4). The presence of these promoter sequences allows to avoid the cloning step, therefore, *actin*, *dnmt1,* and *tet* sense and antisense probes were retro transcribed directly from the PCR product.

Target	Primer	Sequence (5'→ 3')	Amplicon (bp)	Probe
actin	F	**ACGATTTAGGTGACACTATAG AA**ACCGCCGCTTCTTCATCTTC	222 bp	sense Sp6
actin	R	**AGTTAATACGACTCACTATAG GGA**ACCACCAGACAAGACGG		antisense T7
dnmt1	F	**ACGATTTAGGTGACACTATAG AA**CCCTCTGAGGATGTGGAGGA	1129 bp	sense Sp6
dnmt1	R	**AGTTAATACGACTCACTATAG GGA**GAGGTAAACACCAGGGGAC G		antisense T7
dnmt3	F	CAGGGATGTTTTGCATTGTG	590 bp	sense Sp6 TOPO-TA cloning
dnmt3	R	GGCTGTGCTGCTGAAATCAG		Antisense T7 TOPO-TA cloning
tet	F	**ACGATTTAGGTGACACTATAG AA**TGAAGGCACAGAGC	1187 bp	sense Sp6
tet	R	**AGTTAATACGACTCACTATAG GG**TTCTTCTTTCCGGGACCTTT		antisense T7

Table 3: List of primers used for probe synthesis. Primer sequences are used to synthesize probes. Highlighted bold nucleotides are the Sp6 and T7 promoters.

The PCR reaction was performed in a total volume of 50 µl, using about 100 ng of cDNA as a template, 0.2 mM of dNTP mix (dATP, dTTP, dCTP, dGTP), 1mM of $MgCl_2$ solution, 1x PCR buffer, 0,025U/µl of Taq DNA polymerase, and 2 pmol/µl of each forward and reverse suitable oligonucleotide. The amplification cycles were conducted by means of a C1000 Touch Thermal Cycler. After denaturation at 95°C for 5 min, 35 cycles were structured as follows: denaturation at 94°C for 45 seconds, annealing at 46-60°C for 1-1.5 minutes, and extension at 72°C for 30 seconds. Finally, an extension cycle was carried out at 72°C for 5 minutes to complete all the strands. PCR fragments or digested DNA obtained for probe preparation were purified with Macherey-Nagel™ NucleoSpin™ Gel and PCR Clean-up Kit. Their concentration was verified using a NanoDrop (ND-1000 Spectrophotometer; NanoDrop Technologies, Wilmington, DE, USA).

3.8.2 TOPO-TA cloning of PCR products

Only *dnmt3* amplicon obtained by PCR was cloned in the TOPO® TA Cloning® Kit

(Invitrogen), following the manufacturer's indications. The TOPO TA cloning kit allows direct insertion of a PCR amplification product into the pCR® II-TOPO® plasmid vector. The reaction was performed in a total volume of 6 µl, using 4,5 µl of insert (50 ng), 1 µl of saline solution (dilution 1:4), and 0.5 µl of pCR® II-TOPO® plasmid vector (10ng/µl). The reaction mix was incubated at RT for 1 hour and then directly used for chemical transformation in *E. coli* bacterial cells.

3.8.3 Chemical bacterial transformation

This approach allows the transformation of bacterial cells (TOP10, Escherichia coli) with a plasmid containing DNA of interest. For transformation, 40 µl of competent cells, prepared by the Molecular Biology Service of the Stazione Zoologica Anton Dohrn in Naples, were thawed on ice and mixed with 2 µl of ligation mixture. Tubes were swirled gently and incubated on ice for 30 min. After incubation, cells were heat-shocked in a 42°C water bath for 30sec and then immediately transferred on ice for 2 min. Then, 300 µl of LB was added to the cells and incubated at 37°C for 1h with shaking at 230 rpm. For positive selection, two different aliquots (50, 150 µl) of transformation mixtures were plated on LB agar plates containing the appropriate antibiotic and incubated at 37°C.

3.8.4 Isolation of plasmid DNA from *E. coli*

A single bacterial colony containing the plasmid DNA of interest was grown in a suitable volume of LB in the presence of the appropriate antibiotic, shaking at 37°C overnight. The GenElute HP Plasmid DNA Miniprep Kit, based on the alkaline method, was used to isolate the plasmid DNA from the cells according to the manufacturer's instructions.

3.8.5 DNA sequencing

The DNA sequences were obtained using the Automated Capillary Electrophoresis Sequencer 3730 DNA Analyzer (Applied Biosystem, Foster City, CA) by the Molecular Biology Service at the Stazione Zoologica Anton Dohrn in Naples.

3.8.6 Probe synthesis

1 µg of linearized plasmid DNA, or PCR purification, was used as a template for the ribonucleic probe synthesis. This template was added to the following reaction mix:

transcription buffer (1/10; Roche); Digoxigenin or Fluorescein labeling mix, containing 0.65 mM of ATP, CTP, and GTP and 0.35 mM of DIG-11-UTP or fluorescein-12-UTP (Roche); Sp6 or T7 RNA polymerase (2U/ µl; Roche); Protector RNase inhibitor (1U/ µl). The reaction was performed in a total volume of 20 µl (in H2O DEPC-treated). Synthesis reaction was conducted at 37°C for 2 hours, then 2 µl of 10U/µl DNase I RNase free were added and incubated at 37°C for 20 minutes to remove the DNA template. The reaction was stopped by adding 2 µl of 0.2 M EDTA pH 8. Then, in order to purify probes, samples were precipitated adding 75 µl of ethanol 100% and 2.5 µl of LiCl. The samples were mixed and stored overnight at 20°C. Then, it has been centrifuged at 13000 rpm for 30 minutes, at 4°C. The precipitated probes were washed with ethanol 70%, centrifuging at 13000 rpm for 15 minutes. Then, the ethanol was removed and the samples were air-dried at R.T. Finally, probes were diluted in a suitable volume of H_2O.

3.9 Whole-mount in situ hybridization (WISH)

The ISH technique on the whole embryos allows to highlight and analyze the expression of a given gene, using a sequence-specific RNA antisense probe labelled with digoxigenin or fluorescein.

3.9.1 Embryos preparation

In situ hybridization was carried out on *M. galloprovincialis* whole embryos at different developmental stages fixed at 4 °C in MEM-PFA (4% paraformaldehyde, 0.1 M MOPS pH 7.4, 0.5 NaCl, 1mM EGTA, 2mM $MgSO_4$, 0.05 % Tween 20) overnight. Embryos were progressively dehydrated in methanol 100% with 10 minutes of washes with a gradual series of methanol and conserved at -20°C until processing.

3.9.2 Colorimetric and fluorescent WISH protocol

Day 1. The dehydrated stored embryos were progressively rehydrated in a graduate scale of 100% - 70% - 50% - 30% methanol in TBS-T (136 mM NaCl, 25 mM Tris pH 7.5, 2.7mM KCl, 0.1% Tween 20): one wash of 10 minutes for each methanol solution. After that, the samples were washed 3 times for 10 minutes in TBS-T and then treated with proteinase K (100ng/ml) in TBS-T for 20 minutes at 37°C to increase the permeability to the cells and the accessibility to mRNA target. After digestion, embryos were post-fixed

for 20 minutes in MEM-PFA and then they were washed 3 times for 10 minutes in TBS-T. Successively embryos were incubated in 0.1 M triethanolamine (TEA) for 10 minutes, 0.1 M TEA with 2.5µl acetic anhydride/ml for 10 minutes, and 0.1 M TEA with 5µl acetic anhydride/ml for 10 minutes, followed by 3 wash for 10 minutes in TBS-T. At this point, embryos were re-fixed in MEM-PFA for 20 minutes and then washed 3 times for 10 minutes in TBS-T. Embryos were incubated for 10 minutes with 50% TBS-T and 50% Hybridization Buffer (HB) (50% formamide, 1x Denhardt's solution, 6x SSC, 100µg/ml Yeast tRNA, 75mM NaCitrate, 0.005% heparin, 0.1% Tween 20). Prehybridization was done in the hybridization buffer for 3-4 hours at 70°C. Digoxigenin/fluorescein-labelled probes were added in 0.5ng/µl and hybridization was performed for 16 hours overnight at 70°C.

Day 2. The following day, any excess of the probe was removed. A series of washes were carried out by varying the temperature and salinity conditions washes. Embryos were washed at 65°C in 50% hybridization buffer/ 50% 2x SSC for 20 minutes, 25% hybridization buffer/ 75% 2x SSC for 20 minutes, 2x SSC for 30 minutes at room temperature, two washes in 0.2x SSC with 0.3% CHAPS for 30 minutes at room temperature. Then embryos were washed twice in MaNaT (100mM Maleic acid, 150 mM NaCl, 0.1% Tween 20) for 10 minutes and incubated in a Blocking Solution (MaNaT containing 1% Blocking Reagent in MAB) for at least 1 hour. At this point, antibody incubation was performed for 18 hours in a fresh Blocking Solution containing 1:2000 of the anti-digoxigenin or 1:400 of anti-fluorescein antibodies overnight at 4°C (Roche Molecular Biochemicals).

Day 3. The samples were extensively washed in MaNaT two times for 5 minutes and six times for 15 minutes. At this point, for the colorimetric *in situ* protocol, to identify the localization of the RNA of interest with labelled DIG and recognized by DIG alkaline phosphatase-conjugated, are provided with the appropriate substrates. This substrate will be converted by the phosphatase into a blue precipitate. Therefore, embryos are incubated in 1 ml of MaNaT containing 4.5µl of NBT (nitroblue tetrazolium) and 3.5 µl of BCIP (5-bromo-4chloro-3-indolyl-phophate). The time of formation of the precipitate is conditioned by the amount of antibody bound and indirectly by the probe. When the signal is visible, the color reaction is stopped using 1x TBS (Adapted from (Nederbragt et al., 2002). Subsequently, larvae were mounted for imaging for further observation with a light microscope (ZEISS Axio Imager 2).

For *in situ* fluorescence, a protocol previously used in the lab for *Ciona robusta* was used and adapted for *Mytilus* (D'Aniello, 2009)). Specifically, during the third day, after extensively washing in MaNaT, embryos were incubated, in the 1x Plus Amplification Diluent (Perkin-Elmer) for 15 minutes at RT and then 1:400 Cy3 or Cy5 diluted in the same solution for 2 hours. After the reaction, embryos were washed several times in TBS. Finally, embryos were stained with DAPI (1:10000) in MaNaT for 10 min followed by three washes in MaNaT for 5 minutes. Afterwards, larvae were mounted for imaging for further observation with a confocal microscope (Zeiss LSM700).

3.10 Embryo whole mount immunochemistry assay

Larvae at different developmental stages were fixed in 4% PFA for 15 minutes at room temperature and washed five times in phosphate-buffered saline with 0.1% Tween-20 (PBST). Embryos were then permeabilized in PBS with 1% of TRITON 100 (PBS-TX) for 15 min at room temperature and washed several times in PBS. Successively, embryos were incubated with 2M HCl for 30 min at 37°C. Treatment with 2M HCl is needed to optimize the 5mC and the 5hmC signal. At this point, embryos were washed several times in PBST, and subsequently, they were equilibrated in blocking solution (1 mg/ml BSA, 4% sheep serum in PBST) for 1 hour at room temperature. Then, embryos were incubated overnight at 4°C with both the 5-Methylcytosine Monoclonal Antibody (Epigentek 33D3) (1:500) and the 5-Hydroxymethylcytosine Polyclonal Antibody (Active Motif 39770) (1:1000) in blocking solution. The day after, larvae were then washed five times, 10 minutes each, with PBST and incubated for 1 h at room temperature with the secondary antibodies anti-mouse AlexaFluor 555 (Invitrogen) and anti-rabbit AlexaFluor 488 (Invitrogen) diluted in blocking solution (Riviere et al., 2013). After being washed in PBST several times, larvae were mounted for imaging with a confocal microscope (Zeiss LSM700).

3.11 Genomic DNA extraction

Genomic DNA (gDNA) from different early developmental stages was extracted using E.Z.N.A Mollusc DNA Kit (Omega D3373) according to the manufacturer's instructions. DNA concentration and quality were verified using Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific) and electrophoresis using a 0.7% agarose gel.

3.12 Quantification of global DNA methylation levels

Global cytosine methylation levels were determined using the 5mC DNA Elisa kit (Zymo Research D5325) following the manufacturer's protocol. The 5mC DNA ELISA Kit is a powerful tool that allows to accurately quantitate 5mC in a genomic DNA sample in less than 3 hours. The kit features an Anti-5-Methylcytosine monoclonal antibody that is both sensitive and specific for 5mC. Briefly, purified and denatured, single-stranded DNA samples (30 ng) are coated on the well surface in 5mC Coating Buffer. Then, after several washes, the Anti-5-methylcytosine monoclonal antibody and the secondary antibody (that is horseradish peroxidase (HRP conjugate)) were added to the wells. Detection of 5mC occurs after the addition of the HRP Developer. The percentage of 5mC in a DNA sample can be accurately quantified thanks to specially designed controls included with the kit. Each sample was assayed in two technical duplicates. The data shown are the average of two biological replicates.

3.13 Preparation of inhibitors: 5-azacytidine and Bobcat 339

5-azacytidine (purity, more than 98%, Sigma A2385) was dissolved in dimethyl sulfoxide (DMSO) initially at 50 mM as a stock solution and stored at -80°C. Bobcat 339 (purity, more than 98%, Selleckchem S6682) was dissolved in dimethyl sulfoxide (DMSO) at 3mM as a stock solution and stored at -80°C. The working solutions were prepared by diluting the stock solution. DMSO was also used as a control during treatments.

3.14 Statistical analysis

Data are reported as the standard error of the mean (SE) of the given number of biological replicates (n). Data sets were compared by use of an unpaired Student's t-test, two-way ANOVA followed by Bonferroni's post hoc or Chi-square test as reported in each figure legend with Graph Pad Prism 8 software. Statistically significant differences were accepted when the p-value was at least <0.05.

3.15 Scanning electron microscopy

To obtain high-magnification images we employed Scanning Electron Microscopy (SEM) analysis. The samples were fixed with 3% glutaraldehyde in filtered seawater overnight. Then, samples were washed five times in filtered seawater for 10 min and

progressively dehydrated in ethanol 100% with 15 min washes (twice for each point) with a gradual series of ethanol/filtered seawater (30%, 50%, 70%, 90%, 100%), followed by three wash in 100% ethanol. In this way, samples were stored at 4°C until further processing: critical point dryer and sputter coating. These two steps were performed by the Advanced Microscopy Center (AMC) in Stazione Zoologica Anton Dohrn. Briefly, drying was performed by using a critical point dryer Leica EM CPD300. Samples then were mounted on specific stubs and subsequently, they were sputtered with gold and observed with a scanning electron microscope (JEOL JSM 6700F).

3.16 Propidium iodide fluorescence assay and fluorescein diacetate stain

Propidium iodide assay and fluorescein diacetate stain were used to determine cell viability. Propidium iodide (PI) solution (P4864) was obtained from Sigma Aldrich (Milan, Italy). PI fluorescence assay was used for cell viability analysis. Indeed, PI can only pass through disordered areas of the membrane of dead cells and intercalates with the DNA of the nuclei, emitting red fluorescence light.
Fluorescein Diacetate (FDA) (F7378) was obtained from Sigma Aldric (Milan, Italy). FDA is an esterase substrate that can permeate cell membranes. Only living cells convert the non-fluorescent FDA into the green fluorescent compound fluorescein following the enzymatic activity of esterases. Fluorescein accumulates in viable cells, causing them to fluoresce green. Embryos treated with different concentrations of inhibitors were used to test the toxicity of the treatment. After 48 hours of treatment with inhibitors, a positive control for PI and FDA staining was prepared by fixing embryos with 4% PFA for 1 hour, followed by 3 washing with PBS. Then, positive control, DMSO treated and embryos treated with different doses of inhibitors were incubated in filtered seawater containing PI (10 µg/ml) and FDA (7.5 mM) for 15min. After incubation, embryos were imaged using a confocal microscope (Zeiss LSM700).

3.17 Preparation of pharmaceuticals: Carbamazepine and Valsartan

Carbamazepine (purity, more than 98%, Sigma C0244) was dissolved in dimethyl sulfoxide (DMSO) at 42.35 mM as a stock solution and stored at -20°C. Valsartan (purity, more than 98%, Sigma SML014) was dissolved in dimethyl sulfoxide (DMSO) at 31.65 mM as a stock solution and stored at -20°C.

The working solutions were prepared by diluting the stock solution and were used for exposure treatment purposes.

3.18 Methyl-RAD library preparation and sequencing

Methyl-RAD sequencing is a rapid approach to study DNA methylation profiles of different genomic regions that takes advantage of Mrr-like enzymes. The peculiarity of these enzymes is that they can discriminate between CG and non-CG methylation, since the methylation status of each site is interrogated independently. In our approach, for the production of our libraries, we choose LpnPI (NEB) enzyme that generates 32-base-long fragments with the methylated site in the middle. Moreover, the generated fragments have arbitrary four-base 5' overhangs at each end since we use adaptors with fully degenerate ends (5'-NNNN-3'). Methyl-RAD preparation began with the digestion of 100 ng genomic DNA in a 15µl reaction containing 5U LpnPI at 37°C for 4 hours. Then, the enzyme was inactivated at 50°C for 20 minutes and samples were transferred to ice. Then digestion was verified on a 1% agarose gel. During this enzymatic step, genomic DNA is digested to produce 32-base-long restriction fragments around $C^{m}CDG$ sites with 4 bases 5' protruding ends (Figure 17). These conditions are optimized to efficiently generate 32 bp fragments from fully modified CpG sites in genomic DNA. Under this condition star activity may occur, however, the star activity has minimal effect on generating the correct 32 bp fragments. In the next step, these fragments were ligated with adaptors. Ligation efficiency depends on the total number of adapters and the number of fragment ends in a digested sample. For this second step, a 9 µl ligation master mix containing 0.8 µM each of two adaptors, 1 mM ATP (NEB), and 800 U of T4 DNA ligase (NEB) was added to the digestion solution. The total volume of 30 µl was incubated for 6–12 h at 4°C (Figure 17). Then ligation products were amplified. In this step, ligation constructs are amplified using a set of primers that introduce sample-specific barcodes and the Illumina-specific sequencing and cluster amplification binding sites. Ligation products were amplified in 20 µl reactions containing 9 µl ligated DNA. PCR was conducted using a thermal cycler with 20 cycles of 98°C for 5s, 60°C for 60s, and 72°C for 10s. The target band (170 bp) was excised from a 2% agarose gel. Each sample has to be assigned a unique combination

of “BC” and “HT” oligos. BC oligos introduce index 1 (“i7”) and HT oligos introduce index 2 (“i5”) (Figure 17).

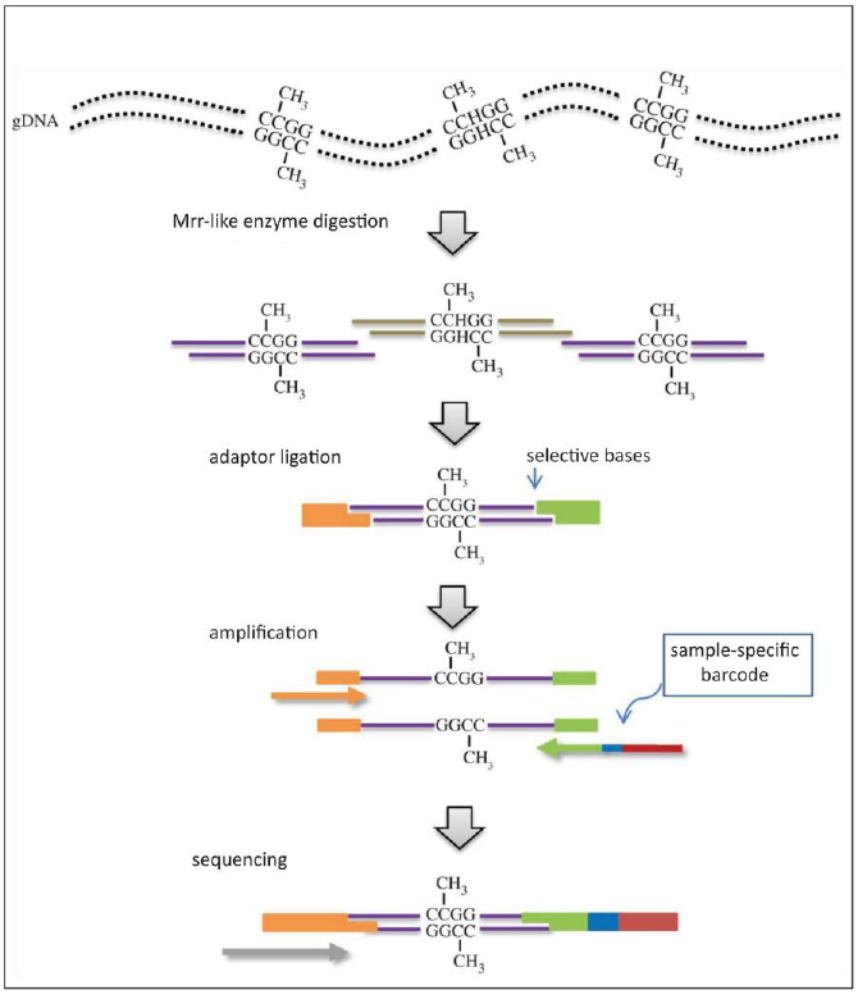

Figure 17: Schematic protocol of MethylRAD library preparation. Genomic DNA is digested with Mrr-like enzyme-producing 32-bp fragments with four-base 3’ overhangs at each end. Adaptors with compatible overhang (NNNN) are ligated of these fragments and the constructs are amplified and purified by gel extraction. After the incorporation of sample-specific barcodes, the products are ready for sequencing. Adapted from (Wang et al., 2015).

CHAPTER 4. RESULTS

4.1 Evolutionary history of DNA methylation related genes in bivalvia: New insights from *Mytilus galloprovincialis*

Currently, DNA methylation remains poorly understood in mollusks and completely unexplored in *Mytilus galloprovincialis*. Furthermore, a detailed analysis of the key genes involved in DNA methylation machinery in *M. galloprovincialis* is still missing. Therefore, my PhD study provides the first insights into DNA methylation mechanism in *M. galloprovincialis* by identification and characterization of genes involved in this process through a genome-wide analysis, protein domain characterization, and analyses of their expression patterns (*in vitro* and *in silico*) in different adult tissues. Furthermore, comparative analyses with other bivalves and their phylogenetic relationship have been performed. This part of the work has been recently published in the journal Frontiers in Ecology and Evolution.

4.1.1 Identification of *M. galloprovincialis* DNA-methylation machinery genes

In order to identify DNA-methylation-related genes in *M. galloprovincialis* our lab started a collaboration with Prof. Pallavicini Alberto and Prof. Marco Gerdol from the University of Trieste in Italy. Therefore, initially, I employed a survey analysis by either extracting annotation from the mg10 version assembly or by recovering full-length CDS from alternative versions of the assembly and/or transcriptome data. These survey analyses revealed that *M. galloprovincialis*, compared with the reference human sequences, displays:

I) A single DNMT1 ortholog. The identified sequence displayed a 53% pairwise amino acid sequence of homology with human DNMT1 and the preservation of all expected conserved domains.

II) A single DNMT3 ortholog. In this case, the mussel sequence showed remarkable structural differences compared with human sequences and displayed a longer N-terminal region. However, the alignable portion of the mussel sequence has different homology with human DNMT3 paralogs, with a slightly superior homology with 35% of identity with DNMT3B than 33% with DNMT3A.

III) A single TET ortholog. The identified sequence displayed a similar degree of similarity with the three human TET paralogs (50.91% identity with TET1, 53.19% identity with TET2, 56.48% identity with TET3), with significant homologies detected in the conserved C-terminal domain.

IV) A single MBD2/MBD3 ortholog with two variants, originated by alternative splicing called MBD2/3a and MBD2/3b. The identified sequence displayed a similar degree of similarity with the two human paralogs, with 48.81% identity with MBD2 and 50.00% identity with MBD3. However, *M. galloprovincialis* orthologs display an organization more similar to MBD3 since human MBD2 is characterized by an N-terminal glycine-rich low complexity region, which was not detected in mussels.

V) A single MBD4 ortholog. The alignable portion has 24% of homology with human sequence. The orthology relationship was made clear by the presence of two distinct regions of homology, corresponding to the MBD (57% sequence identity) and to the DNA glycosylase (71% sequence identity) domains.

No reliable ortholog sequence to the other DNA methylation-related proteins could be identified in *M.galloprovincialis*. In particular, we did not find any orthologs for the MeCP2 considered the founding member of the MBD family of vertebrates. Moreover, our survey allowed the identification of two DNMT2 orthologs, but these sequences have not been subject to these analyses since they are involved in the methylation of aspartic acid tRNA.

4.1.2 Mussel DNA methylation related gene structure

Then, in order to determine the exon-intron structure of these identified genes, I took advantage of the SPLIGN NCBI software and by an accurate alignment of the cDNA sequence against their genomic counterparts, I was able to define the number of the exons of each gene and their organization (Figure 18):

- Dnmt1 gene contains 38 exons, 3 less than the human ortholog.
- Dnmt3 gene contains 22 exons, very similar to the human Dnmt3b gene (24 exons) and dissimilar to Dnmt3a (9 exons).
- Tet gene contains 10 exons, very similar to those found in the three human paralogs Tet 1-2-3, which are respectively 12-11-11 exons (NCBI).

- Mbd2/3 contains 6 exons, just one fewer Mbd2, and Mbd3 human genes.
- Mbd4 contains 21 exons, a much higher number of exons compared to the human orthologs showing only 8.

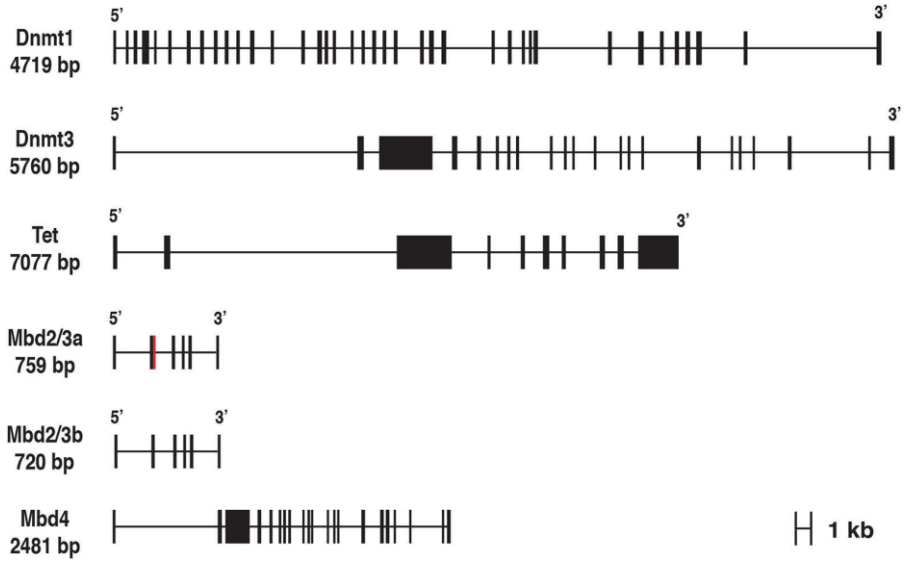

Figure 18: Exon-Intron structure of Dnmt1, Dnmt3, Tet, Mbd2/3, Mbd4 genes. The red square in Mbd2/3a gene indicates the optional 39 bp region alternatively spliced in the exon 2 of the Mbd2/3 gene. This region discriminates between Mbd2/3a and Mbd2/3b variants.

4.1.3 Domain architecture of *M. galloprovincialis* DNA methylation toolkit related proteins

At this point I was interested in determining the level of conservation of *Mytilus* methylation toolkit protein domains with those of human and other representative lophotrochozoan species. In order to determine the protein structures, I employed the Pfam and InterPro databases. This analysis reveals that *M. galloprovincialis* DNMT1 contains the same highly conserved domains of human protein (Jeltsch and Jurkowska, 2016; Ren et al., 2018). These include a binding domain at the N-terminal for the interaction with transcriptional repressor the DNMT1-associated protein 1 (DMPA1) (Rountree et al., 2000), a replication foci domain (RFD) necessary for mediating anchoring to the replication fork (Leonhardt et al., 1992), a zing finger CXXC motif able to bind unmethylated CpG and two Bromo-adjacent homology domain necessary for mediating gene expression silencing (Pradhan et al., 2008). Last, the C-terminal region contains a methylase domain (C5 cytosine methyltransferase), responsible for the catalytic activities on a cytosine residue in a DNA sequence (Razin, 1989). All these domains contribute together to the proper DNMT1 activity on maintenance methylation

(Jeltsch and Jurkowska, 2016). Regarding *Mytilus* DNMT3 the domain architecture is also very similar to the human DNMT3a and DNMT3b despite the presence of a very long N-terminal region devoid of recognizable domains, which is encoded by exon that is not conserved across Lophotrochozoa. The sequence consists of three main domains: The N-terminal domain, the Pro-Trp-Trp-Pro (PWWP) domain required for DNMT3 chromatin association via recognition of Histone H3 trimethylated at lysine 36. This binding is necessary and responsible for directing the *de novo* methylation activity at the pericentric heterochromatin (Chen et al., 2004) the central ADD domain bind to histone H3 unmethylated tails at lysine 4 (Zhang et al., 2010); the C-terminal domain, similarly to DNMT1 is the C-5 cytosine-specific DNA methylase (Figure 19).

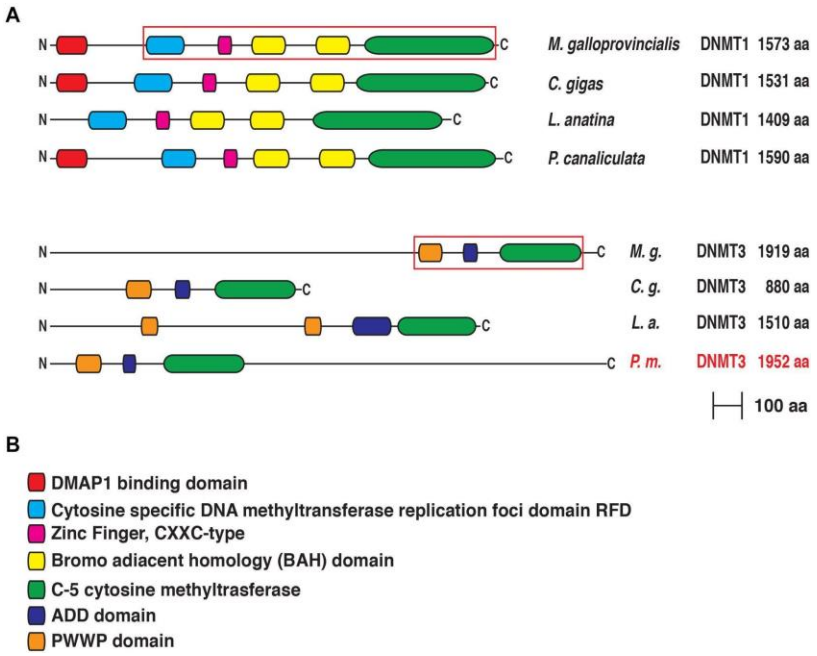

Figure 19: Domain architecture of *M. galloprovincialis* DNMTs. (A) Conserved domain organization of DNMT1 and DNMT3 of *M. galloprovincialis* in comparison with those of *Crassostrea gigas*, *Lingula anatina*, and *Pomacea canaliculata*. (B) List of functional domains.

As we reported from the survey analyses, *M galloprovincialis* only has a single TET protein. My analyses showed that like DNMT3, the mussel TET protein shows a long N-terminal region with no detectable domains. TET protein has only a single well defined structural domain, the double-stranded β-helix 2-oxoglutarate and Fe (II)-dependent dioxygenase domain functioning for the conversion of 5-methylcytosine to hydroxymethylcytosine and necessary to induce demethylation. Similarly to other lophotrochozoans species, *mytilus* lacks the CXXC motif and cysteine-rich domain that

characterizes two of three mammalian TET (Nakajima and Kunimoto, 2014) (Figure 19). I then went on and analyzed the methyl CpG binding proteins. In *M galloprovincialis*, both alternative spliced isoforms of MBD2/3 generate similar proteins that consist of three typical MBD domains. In vertebrates, MBD2 includes a central transcriptional repressor domain (TRD) that I did not find. However, while its presence in the mussel sequence cannot be ruled out, this would require additional characterization. Finally, MBD4 displayed a structural architecture similar to other metazoan orthologs, showing the characterizing MBD domain followed by a larger glycosylase domain at the C-terminal end (Kondo et al., 2005) (Figure 20).

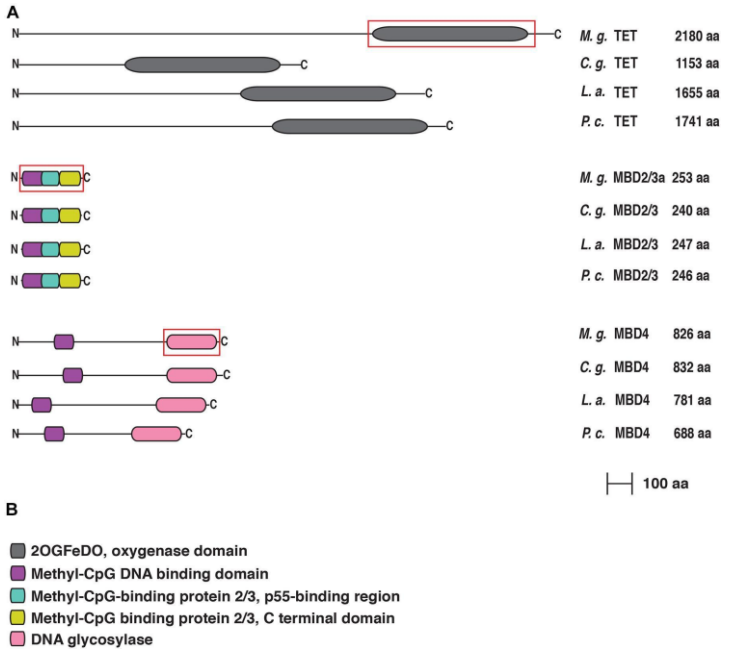

Figure 20: Domain architecture of *M. galloprovincialis* TET and MBDs. (A) Conserved domain organization of TET, MBD2/3, and MBD4 *of M. galloprovincialis* in comparison with those of *Crassostrea gigas*, *Lingula anatina* and *Pomacea canaliculata*. (B) List of functional domains**.**

4.1.4 The DNA-methylation machinery is conserved across all bivalves

At this point in collaboration with Prof. Pallavicini and Prof. Gerdol from the University of Trieste, I decided to perform a comparative analysis supporting the existence or not of these species in the other 23 species belonging to the class Bivalvia. This study was made using genome assemblies with different quality and completeness, ranging from chromosome-scale to highly fragmented and partial draft. The results showed that all the genes involved in DNA methylation machinery appear to be nearly invariably present with a single copy in bivalves. Moreover, Dnmt1, Dnmt3, Mbd2/3, Mbd4 orthologs, full

CDS or partial sequences, are present in all species analyzed. Meanwhile, Tet could not be retrieved from *M. nervosa* and *P. generosa* (Figure 21). The lack of detection of some of the genes involved in the DNA methylation machinery in some of the target species is probably due to several technical factors; these two species have a relatively low assembly and annotation quality, which also lacked transcriptome data to fill in such gaps.

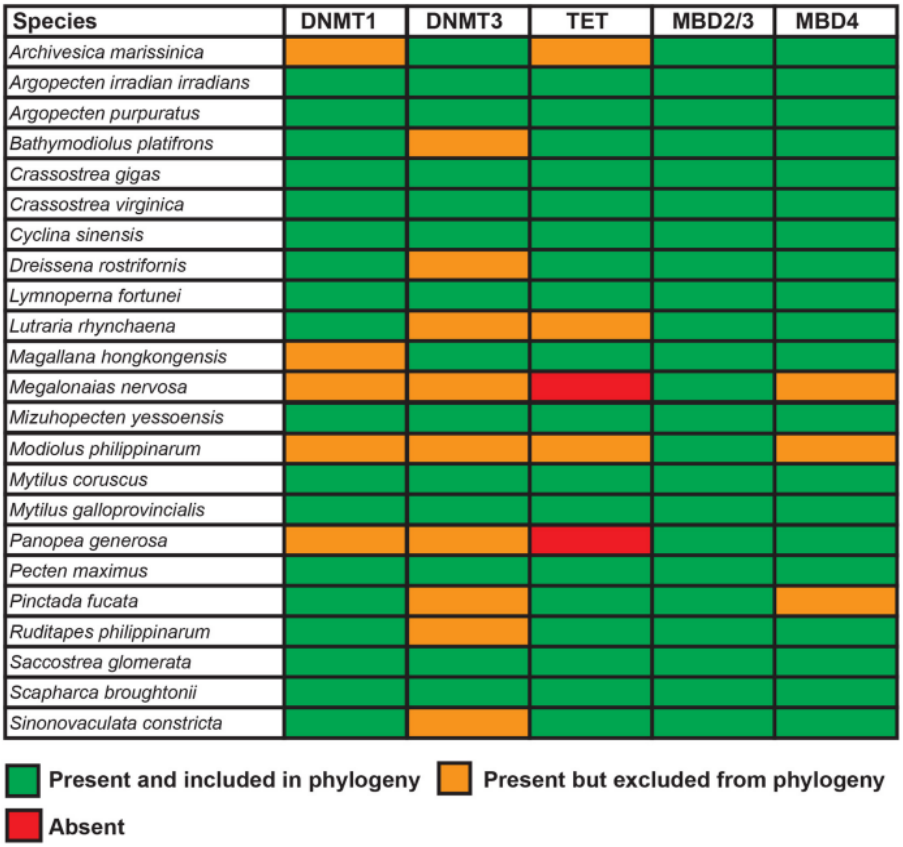

Figure 21: Summary of the genes involved in the DNA methylation machinery found in the 23 bivalve species analyzed in this study. Regarding the Dnmt3, these analyses allowed us to determine that the previously reported (Männer et al., 2021) apparent lack of Dnmt3 in the genome of some bivalve species was most likely a technical artifact linked with misassembles or incorrect gene annotation.

4.1.5 Evolutionary analysis of bivalve DNA methylation machinery genes

To investigate more in-depth, the evolution of DNA methylation-related genes, thanks to the expertise with phylogenesis of Prof. Marco Gerdol from the University of Trieste, phylogenetic analyses have been performed on DNA methylation related genes. Prof. Gerdol's results showed that most genes involved in DNA methylation machinery displayed sequence phylogenies nearly identical to the currently accepted species

phylogeny (Figure 22).

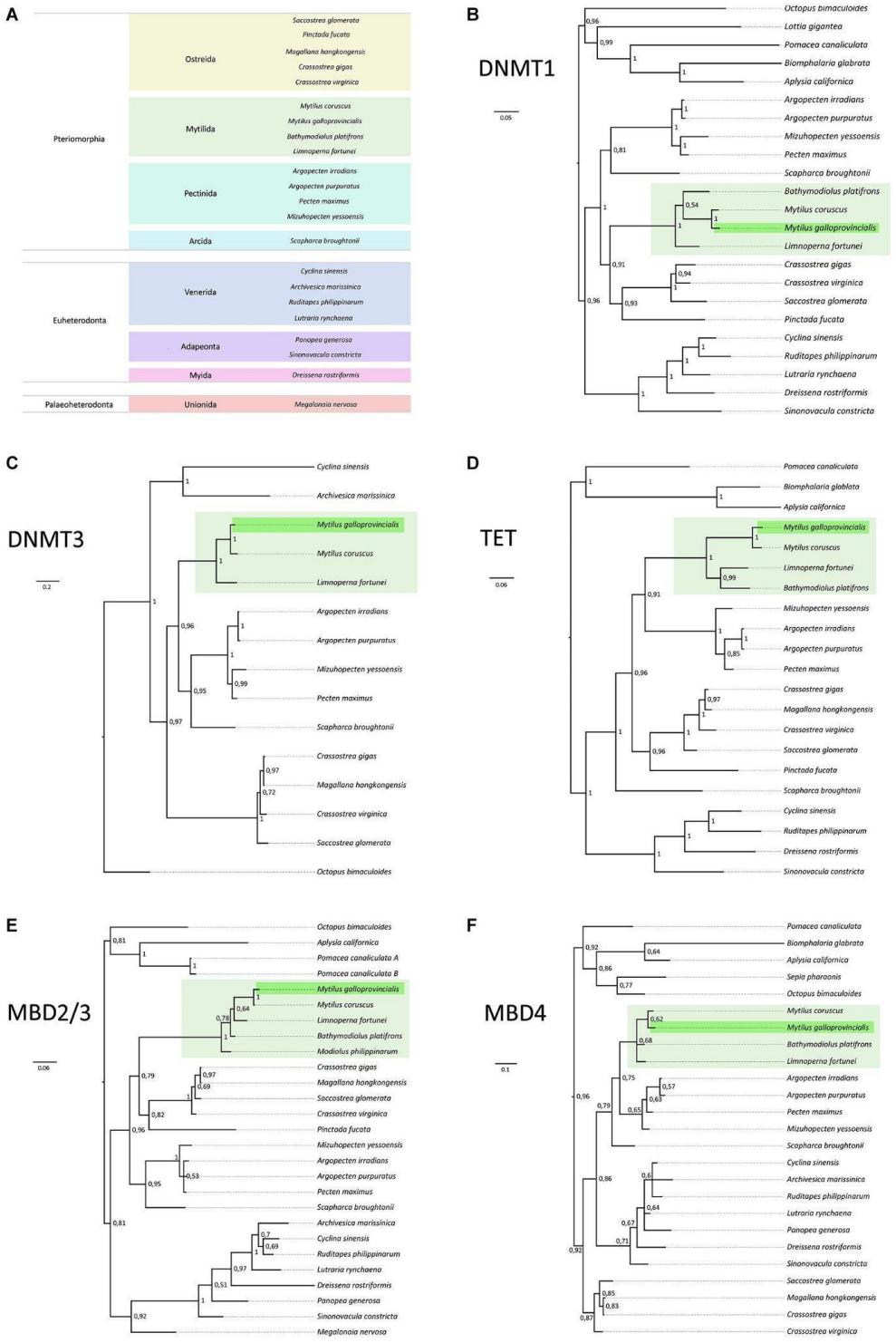

Figure 22: Evolutionary analysis of bivalve DNA methylation machinery genes. (A) Taxonomic classification of the bivalve species considered in this study. (B-F) Bayesian phylogeny of bivalve DNMT1, DNMT3, TET, MBD2/3, MBD4 orthologs. Sequences from *M. galloprovincialis* and mytilid bivalves are highlighted with a green and light green background, respectively.

4.1.6 Gene expression analysis of DNA methylation-related genes in *M. galloprovincialis* adult tissue and in bivalve adult tissue

At this point to investigate the expression of DNA methylation-related genes in adult tissues of various bivalve species in collaboration with Prof. Pallavicini and Prof. Marco Gerdol I performed *in silico* analyses to assess their transcriptional activity. Results revealed a poor transcriptional activity, seldom reaching values higher than 10 Transcript Per Million (TPM) and most often displaying values lower than 3 TPM (Figure 23). The only exception is Mbd2/3 which displayed higher expression in all tissues probably due to the involvement of this gene in other processes. Among the genes investigated, Dnmt3 was the one displaying the lowest expression levels, in all adult tissue, and all species, with the exception of *C. gigas* displayed the highest levels of Dnmt3 expression in all tissues analyzed. Moreover, although reproductive tissues were only available for two species (*C. gigas* and *M. yessoensis*) RNA-seq data suggests that Dnmt1, Tet, Mbd2/3 and Mbd4 are expressed in both male and female gonads at levels comparable to other tissue. On the other hand, Dnmt3 maintained a poor expression in the gonads of *M. yessoensis* (TPM < 1) and a moderate expression in the gonads of *Crassotrea gigas*. *In silico* analysis shows that *M. galloprovincialis* was characterized by gene expression patterns in line with the other species analyzed for the available tissues. At this point, to confirm the *in silico* analyses I performed a qPCR experiment in order to analyze the expression of Dnmt1, Dnmt3, Tet, Mbd2/3a-b, and Mbd4 in different *M. galloprovincialis* adult tissues: mantle edge, foot, gills, digestive gland, adductor muscle and hemocytes (Figure 24). The results obtained were consistent with RNA-seq data, confirming the very low expression of Dnmt1, Dnmt3, Tet, and the moderate expression of Mbd2/3a-b in all the analyzed adult tissues. The only discrepancy I found was for Mbd4, which appeared to be strongly expressed in all tissues by qPCR, in contrast with RNA-seq data, which only identified gills and adductor muscle as the preferential sites of expression in *Mytilus galloprovincialis*. This inconsistent result may be due to a number of different reasons such as the genetic variability between the animals, their different age, diverse geographical backgrounds, and ecological parameters of the sampling site.

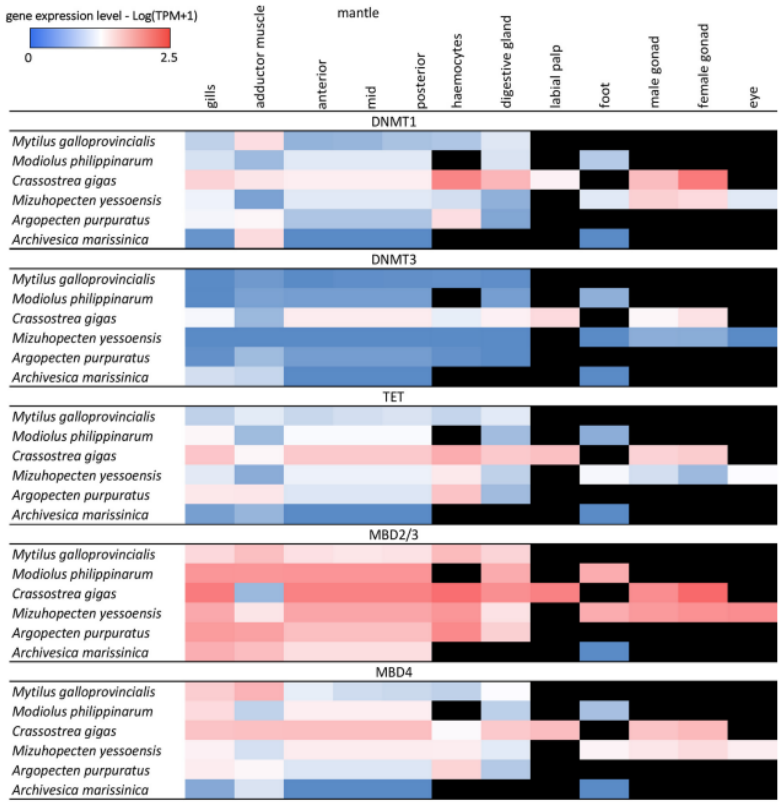

Figure 23: Heat map of DNA methylation-related genes expression in adult tissue of six bivalve species from transcriptome database. Black cells indicate missing data.

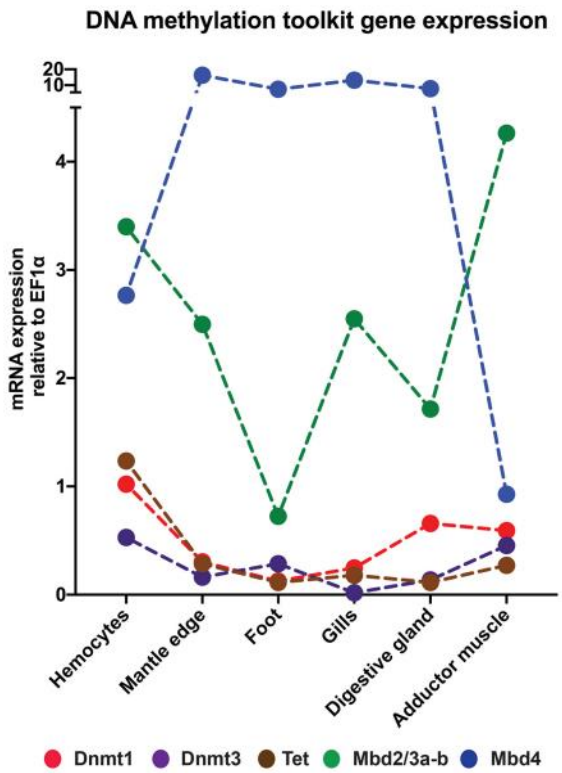

Figure 24: Expression profile of Dnmt1, Dnmt3, Tet, Mbd2/3a-b, and Mbd4 in different *M. galloprovincialis* tissue. The relative expression has been analyzed in comparison to the Dnmt1 expression in hemocytes.

4.2 Characterization of DNA methylation during *M. galloprovincialis* embryogenesis

Once provided first insights on DNA methylation machinery genes in *M. galloprovincialis*, the next step has been to characterize the role of DNA methylation during the first stage of development of *Mytilus galloprovincialis* from fertilization to D-larva stage at 48 hpf.

4.2.1 Spawning induction and embryo rearing in the lab

When I started my PhD, Dr. D'Aniello, my co-supervisor at the Stazione Zoologica, where I spent most of my lab experience, had just introduced *Mytilus galloprovincialis* in his lab in the Biology and Evolution of Marine Organism department at the Stazione Zoologica. Until that time, *Mytilus* embryos had never been used at Stazione Zoologica for Developmental Biology studies and neither for embryotoxicity studies. Although Dr. D'Aniello has a strong experience in developmental biology with aquatic animals we started to learn and apply different protocols for *Mytilus* spawning and larval rearing. Therefore, during the first year of PhD to achieve this objective, I firstly approached several different spawning methods based on thermal shock or chemical stimulation with an injection of KCl or H_2O_2 (Wong and Arshad, 2013; Bantoto-Kinamot, 2016; Resgalla, 2016). Among different methods I attempted, I found that the most useful was just to use animals obtained during the same day of the spawning from the commercial shellfish farm, keep them in dry conditions for a couple of hours, roughly cleaned from barnacles, algae, and byssus and later immersing them in fresh seawater (FSW) at 18°C and wait until they start to spawn. Later, once identified animals that started to spawn, I washed them to avoid self-fertilization and located them singularly in a cylinder to allow them to spawn most of their gametes (Fabbri et al., 2014) (Figure 25).

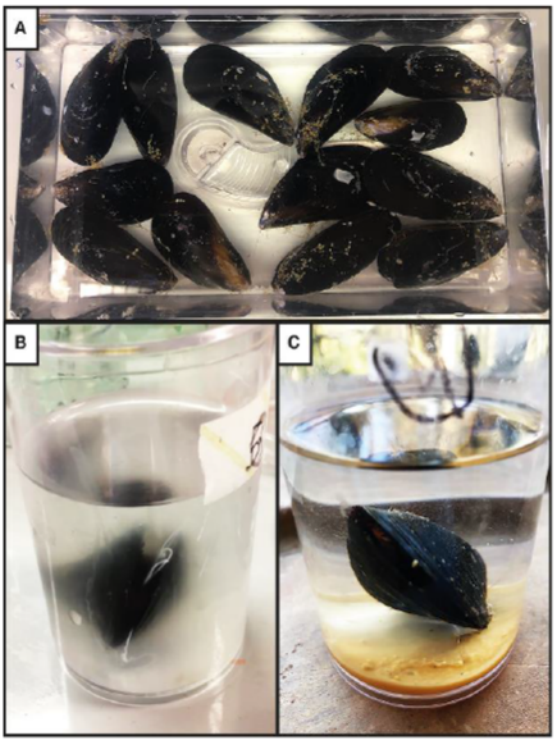

Figure 25: Mussel spawning process. (A) A tank containing cleaned mussels in fresh seawater (FSW) at 18°C. (B) Cylinder with FSW containing a male mussel that is spawning his sperm. (C) Cylinder with FSW containing a female mussel that is spawning her oocytes.

Usually, females spawn white/pink/orange masses that are groups of hundreds of eggs stuck together. On the other hand, males spawn a white cloud that dissolves in the volume of water (Figure 26). Spawning can last for hours and before *in vitro* fertilization I usually check gamete quality. The detailed and full procedure applied can be found in the material and methods section.

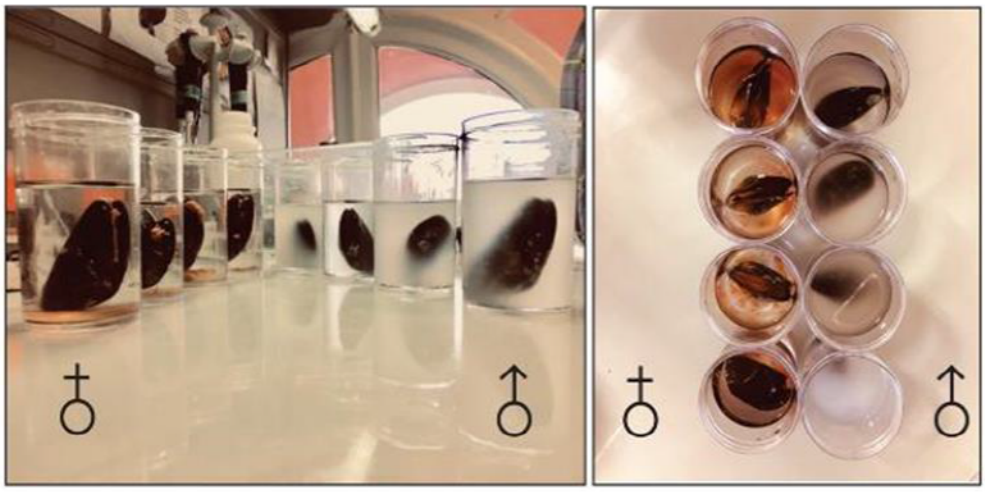

Figure 26: Representation of a successful spawning day.

4.2.2 Stages of embryonic development of *Mytilus galloprovincialis*

At this point, I learned how to grow embryos from fertilization to the D-larva stage at 96 hours post-fertilization. Many protocols are available in the literature on mussel growing and embryo development. Specifically, I followed the protocol applied by Fabbri et al., 2014 with just some modifications. Then, given that one of my main PhD goals was to understand the role of DNA methylation during the development I first decided to follow *Mytilus galloprovincialis* embryogenesis in detail during the main developmental stages. This study allowed me to build a full developmental panel (currently missing in literature for *M. galloprovincialis*) that we thought will be useful for the future use of this organism not only for embryotoxicity studies but also for basic developmental biology studies with the main stages reported in Table 5 and represented in Figure 27. Mature sperm showed a typical structure formed by a conical head with an acrosome of elongated shape (Figure 27 A). The unfertilized eggs (ca. 70 µM) have a peculiar shape, they appear elongated, similar to an olive (Figure 27 B). When fertilization takes place, the perivitelline space appears around 20 minutes later. Moreover, since meiosis is only completed after fertilization, the oocyte extrudes also the first polar body and changes shape becoming oval (Figure 27 C). Mussels develop by a "spiral cleavage" pattern.

Developmental stages	Time after fertilization
Extrusion of the first *polar* body	20 minutes
First cleavage	1.2 hours
Second cleavage	2 hours
Morula stage	3.5 hours
Stereoblastula stage	5.5-6 hours
Rotatory blastula	6.5-7 hours
Gastrula stage	8 hours
Early trochophore stage	16 hours
Middle-late trochophore stage	24 hours
D-larva stage	48 hours

Table 4: Developmental timetable of *M. galloprovincialis.*

The fertilized eggs undergo a series of cleavages and the first asymmetric cleavage happens after 1.2 hpf producing two blastomeres: AB and CD (Figure 27 D). After 2 hours, the second cleavage occurs orthogonal to the first cleavage, creating four

macromeres (Figure 27 E). These macromeres divide unequally to generate the first quartet of micromeres (not shown in the panel). Then, after the generation of these first quarter of micromeres, the macromeres undergo several rounds of unequal cleavages that should give rise to successive generations of animal micromere quarters (not shown here). Later as development proceeds, around 3.5 hpf embryos reached the morula stage (Figure 27 F). Embryogenesis then proceeds through the blastula stage produced by cleavage that has no blastocoel and is called stereoblastula. From 6.5-7 hpf all the embryos that were previously sitting on the bottom become competent to swim in the water column thanks to the bundle of well-developed cilia visible also in the picture. Our observations are in agreement with what has been previously reported by (Ruiz et al., 2008) and therefore these embryos are also called rotatory blastula (Figure 27 G). In figure 27 H-I are instead represented embryos that enter the gastrula stage. As development proceeds at 16 hpf, the first trochophore appears: these larvae have a prototroch consisting of a ciliated band or ring of cilia and an apical tuft of cilia which allow them to have active and coordinated swimming. The ciliated early-trochophore shows the blastopore on the ventral side of the embryos and the shell field on the dorsal side (Figure 27 J). As development proceeds, valves appear as the smallest structures on the dorsal area of the middle trochophore (Figura 27 K). At 24 hpf middle-late trochophore larva form the shell rudiment (Figure 27 L). Production of normal D-shape larvae or veliger from fertilized eggs was completed after 48 hpf. The body of the larva now possesses a distinct, disk-shaped velum that D-stage larvae use to swim. Moreover, at this stage, the larva has rudimental gills while the mouth, stomach, and digestive gland are well developed (Figure 27 M). As the veliger stage grows three pairs of striated muscle velum retractors and shell anterior adductors developed (Figure 27 N-O).

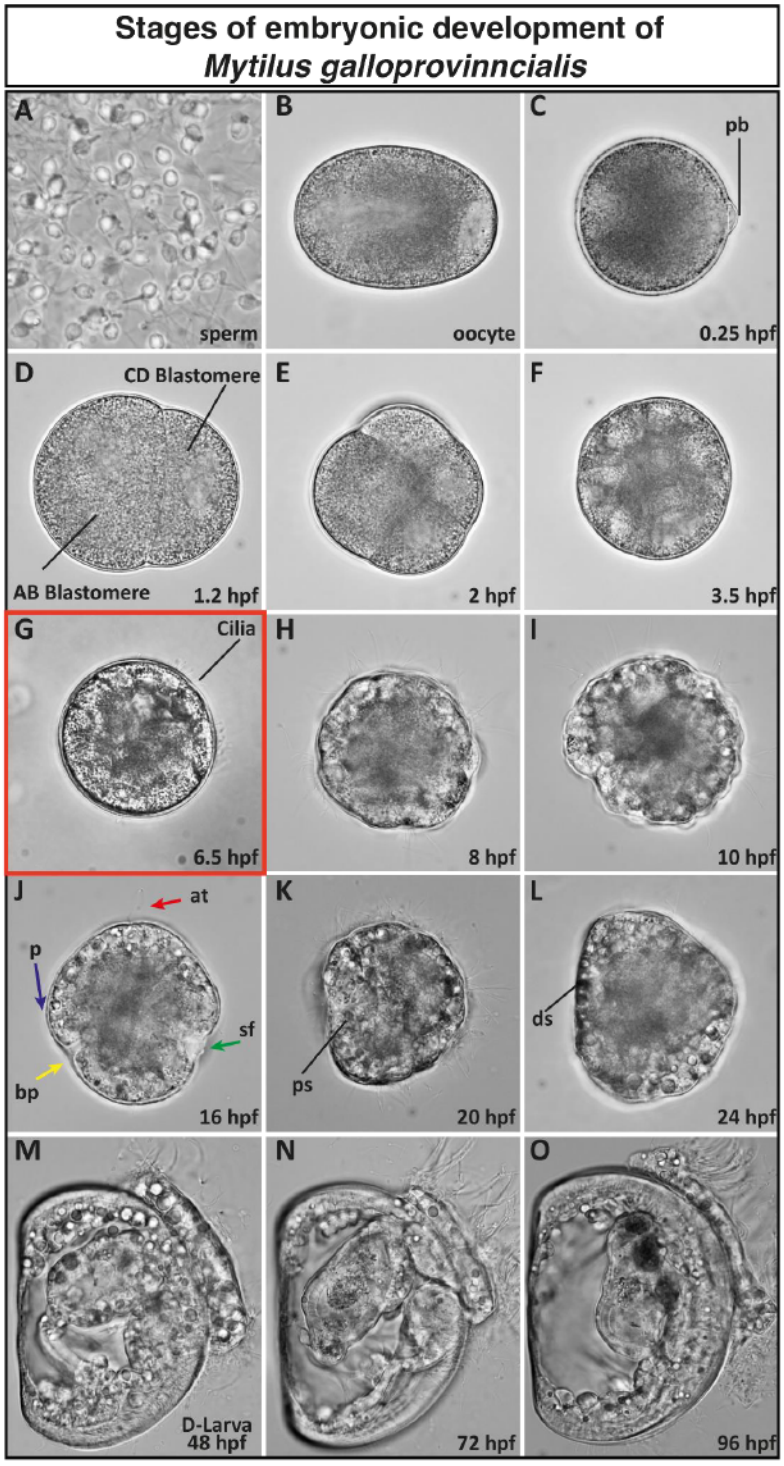

Figure 27: Stage of *Mytilus galloprovincialis* embryonic development. Embryology and larval development in *M. galloprovincialis* species. (A) Sperm morphology. (B) Unfertilized oocyte. (C) Fertilized oocyte showing extrusion of the polar body. (D) Two-cell stage. (E) Four-cell stage. (F) Morula stage. (G) The red square represents the rotatory blastula stage. (H-I) Gastrula stage. (J-L) Early-Middle-Late trochophore stage. (M-O) D-larva stage. pb, polar globe; ps, primordial shell, ds, developing shell; red arrow indicates apical tuft, blue arrow indicates prototroch, yellow arrow indicates blastopore, green arrows indicate shell field.

4.2.3 DNA methylation related genes expression during embryonic and larva stages

DNA methylation role during development has never been explored before in *Mytilus galloprovincialis*. Therefore, once I established how to grow *Mytilus* embryos, initially to answer this question I decided to analyze expression levels of the main DNA methylation-related genes, from oocytes to the D-Larva stage at 48 hpf employing qPCR. Gene expression was examined in four replicate experiments. Specifically, I analyzed Dnmt1 expression whose function is for the maintenance of DNA methylation, Dnmt3 involved instead in the methylation *de novo,* and finally Tet, the key enzyme involved in the active demethylation process through the oxidation of 5-methylcytosine in 5-hydroxymethylcytosine. The results obtained and shown in Figure 28 reveal a decrease in Dnmt1 expression from the initial stages of development until D-larva at 48 hpf.

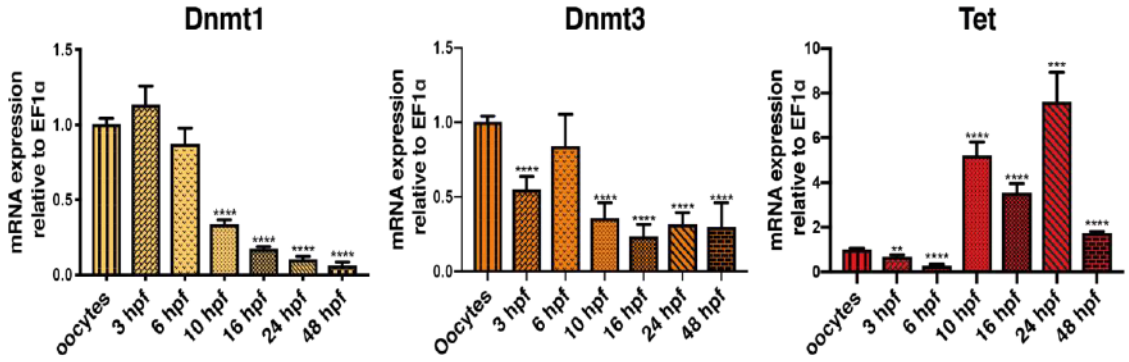

Figure 28: Expression profile of *dnmt1*, *dnmt3*, *tet* in different *Mytilus galloprovincialis* developmental stages. Statistical analysis was performed with the unpaired Student's t-test of each stage versus oocytes by using the GraphPad Prism 8 software. Data are expressed as mean SEM, (n=4). ± *p <0.05. The relative expression has been analyzed in comparison to the expression of each gene in oocytes.

Similarly, I found that the Dnmt3 expression shows a similar expression trend. Interestingly, my data are in agreement with what has been previously found during development in other mollusk species such as *Crassostrea gigas* and *Patinopecten yessoensis* (Riviere et al., 2013; Xu et al., 2019). Moreover, these data do also recapitulate at least in part the expression of these genes in vertebrates (Firmino et al., 2017; Uysal et al., 2017). Interestingly, when I analyzed the expression of Tet, I found an opposite trend in comparison to Dnmt1 expression. Indeed, as visible from Figure 28, its expression initially decreases until 6 hpf, and then from 10 hpf during the gastrula stage it increases dramatically with the maximum expression at 24 hpf. A comparison with other molluscan

species cannot be performed, since at the present time we are the first group to have ever analyzed the expression of this gene in this phylum. However, our data seems to agree with what has been found also during mouse development at least for Tet1 and Tet2 (Tan and Shi, 2012). Furthermore, our results, showing a peak of Tet expression during development are consistent with those recently reported in other species from fish to sea urchin and lancelet (Skvortsova et al., 2021). Finally, our expression analyses revealed also that Dnmt3 is expressed at very low levels compared to Dnmt1 and Tet during *Mytilus galloprovincialis* development (comparison of target Ct with that of Ef1α, data not shown). In conclusion, my data on Dnmt1, Dnmt3, and Tet in *Mytilus galloprovincialis* show that these genes are expressed and regulated in a similar way to what was previously demonstrated in vertebrates and invertebrates.

4.2.4 Temporal and spatial expression of different methylation-related genes

At this point, I was interested in determining the expression pattern of these genes in the whole embryos during various developmental stages. To this goal, I first set up the *in situ* experiment on *Mytilu*s embryos with both the chromogenic and fluorescence methods by taking advantage of a strongly expressed muscle marker: the actin gene, whose expression was already reported at trochophore and D-Larva stages (Balseiro et al., 2013). Initially, I performed chromogenic *in situ* protocol and the results showed expression of the actin gene in a bilateral population of muscle cells in the trochophore embryos and in a larger population of muscles on the interior of the shell at the D-larva stage (Figure 29 A-B). Such expression patterns are in agreement with previous studies (Balseiro et al., 2013). At this point, by using the same marker, I decided to attempt a protocol for fluorescent *in situ* hybridization that had never been applied in *Mytilus galloprovincialis* embryos before. As visible in Figure 29 C-D, the results are the same as in the chromogenic protocol. The main advantage of using a fluorescence-based protocol is that gene expression can be revealed faster. Indeed, since the staining in the chromogenic protocol is the result of an enzymatic reaction that produces an insoluble precipitate of purple color, several hours may be required to obtain a signal for poorly expressed genes. Furthermore, both experiments were also performed with the actin sense probe and I

never observed any staining or fluorescent signal, further confirming the specificity of the expression pattern.

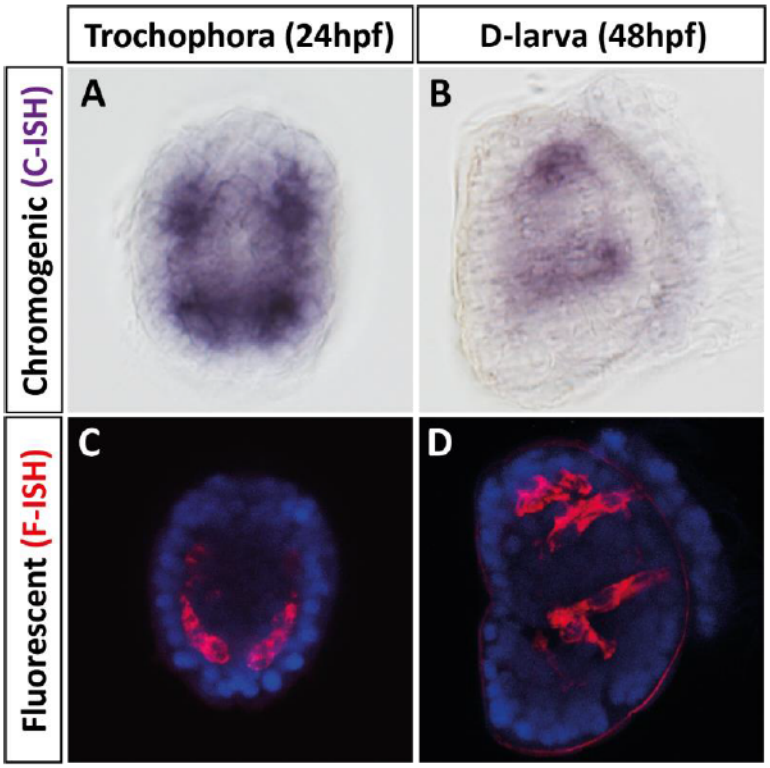

Figure 29: Representative microscopy images of in situ hybridization of Actin expression in Trochophore and D-Larva stage of *Mytilus galloprovincialis*. (A-B) Chromogenic whole-mount *in situ* protocol. (C-D) Fluorescent whole-mount *in situ* protocol.

Once this technique was established in the lab, I was interested in determining the spatial expression of DNA methylation toolkit genes in the whole embryos as it has never been reported in any bivalves before. Therefore, after I prepared the sense and antisense probes for Dnmt1, Dnmt3 and Tet, their expression pattern was examined by fluorescent *in situ* hybridization (FISH) (Figure 30). Specifically, Dnmt1 seems to have a quite ubiquitous expression in all analyzed stages. Moreover, it is very interesting to note that although FISH cannot be considered a quantitative technique, the results are mostly consistent with qPCR data (Figure 28). Indeed, as visible in Figure 30 A1-A7, the fluorescent signal strongly decreases from 10 hpf to 48 hpf. Regarding Dnmt3, the *in situ* experiments showed that also this gene is ubiquitously expressed throughout the whole embryos (Figure 30 B1-B7) and again with an expression that seems to decrease from oocyte to D-larva (reduction from oocyte to D-Larva). Moreover, the Dnmt3 signal appears to be weaker than Dnmt1, further confirming our observation made with the qPCR data (Figure 28). At this point, I analyzed the expression of Tet, the main enzyme involved in the

demethylation process. This analysis also reveals a ubiquitous expression pattern for this gene. Interestingly, as visible in Figure 30, Tet expression, at least from 10 hpf up to 48 hpf, displayed an opposite trend compared with Dnmt1 (Figure 30 A4-A7; Figure 30 C4-C7). This is exactly what I found with qPCR data (Figure 28). However, I did also find some inconsistent results regarding the expression of Tet during the first stages of development, where qPCR data showed an initial low expression, whilst the *in situ* data seemed to not show the same trend (Fig. 28; Figure 30 C1-C3). However, it is not always obvious to perform this kind of correlation, since *in situ* experiments are not quantitative experiments. The specificity of the signal for all the genes was confirmed from the absence of any signal in the embryos hybridized with a specific sense probe.

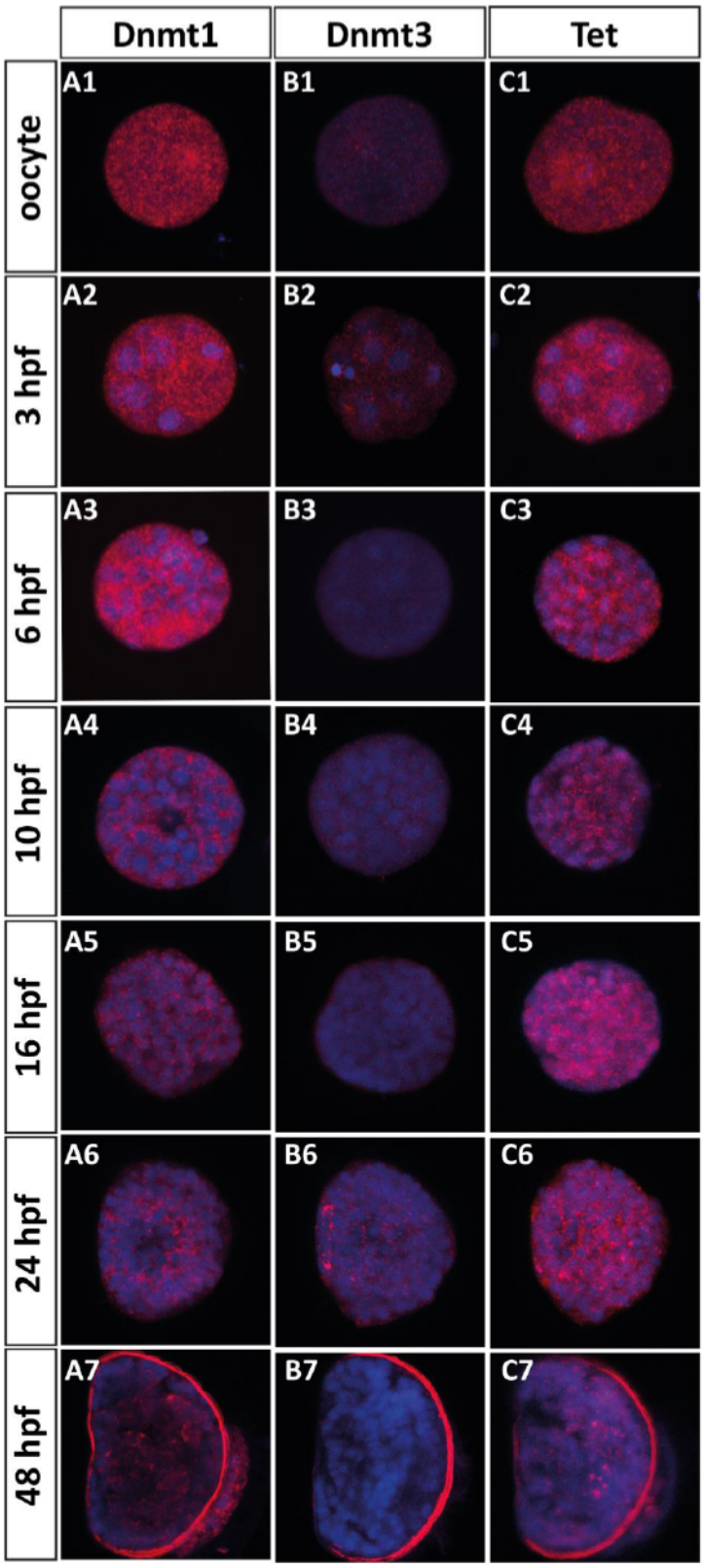

Figure 30: Representative confocal microscopy images of fluorescent *in situ* hybridization of Dnmt1, Dnmt3, Tet, with antisense probe during different developmental stages from oocyte to 48 hpf. (A1-A7) Dnmt1; (B1-B7) Dnmt3; (C1-C7) Tet.

4.2.5 Decrease of 5mC correlated with an increase of 5hmC during development

To better understand if *Mytilus* embryos undergo a reprogramming of their methylome during development I decided to examine the levels of genomic methylation (5mC) and hydroxymethylation (5hmC) during *M. galloprovincialis* development, by taking advantage of specific 5mC and 5hmC antibody. The result showed in all the analyzed stages a positive signal for both 5mC and 5hmC (Figure 31). Specifically, oocytes showed the sign of methylation and demethylation (Figure 31, A2 and A3). However, very interestingly, as development proceeds at 3, 6 and 10hpf embryos showed a quite evident methylation (5-methylcytosine) status with a lower demethylation status (5-hydroxymethylcitosine) (Figure 31 B2, C2, D2; B3, C3, D3). Later, at 16, 24 and 48 hpf it is quite easy instead to observe a demethylation process as shown by a decrease in the fluorescence intensity of 5mC immunofluorescence (Figure 31 E2, F2, G2) with a quite evident parallel increase of the 5hmC signal (Figure 31 E3, F3, G3). These results are somehow in agreement with the expression we found with the qPCR analyses that showed an opposite trend respectively for Dnmt1 (methylation) with stronger expression from oocytes to 6 hpf and with Tet (demethylation) showing an increased expression from 10 to 24 hpf (Figure 28). However, it is important to take into account that our qPCR experiment revealed mRNA expression whilst the respective proteins can also act at later stages based on their half-life. To attain this goal, I also attempted some immunofluorescence experiments using human DNMT1 antibody but unfortunately, the results did not show any specific and repetitive signal given to some differences in the sequence of the epitope. These results seem to suggest that the *Mytilus galloprovincialis* genome is not constantly methylated during larval life and is reminiscent of what has been found in vertebrates, which undergo a reprogramming of their DNA after fertilization (Huntriss, 2021).

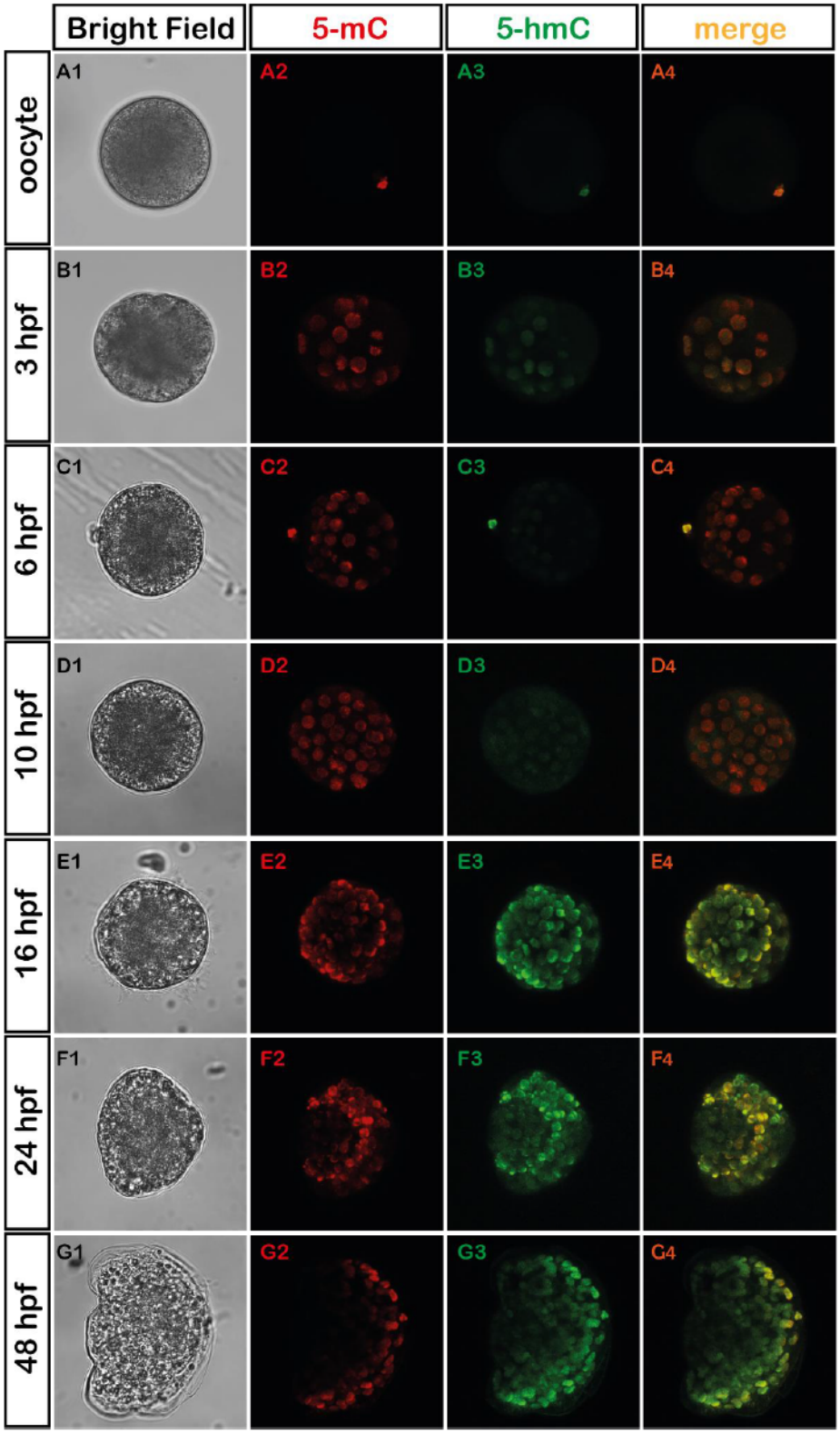

Figure 31: Representative confocal immunofluorescence images showing the distribution of 5mC and 5hmC during *Mytilus galloprovincialis* embryogenesis. In red embryos were incubated with 5mC antibody, in green embryos were incubated with 5hmC antibody. In orange, the two-channel have been merged.

4.2.6 Global DNA methylation changes during *M. galloprovincialis* early embryogenesis

To further demonstrate potential changes of DNA methylation during the developmental stage, I decided to perform global DNA methylation analyses of genomic DNA during seven different embryonic stages: oocytes, 3 hpf, 6 hpf, 10 hpf, 16 hpf, 24 hpf, 48 hpf. Analyses have been performed taking advantage of an Elisa Kit for the detection of relative amounts of global 5mC in *M. galloprovincialis* genomic DNA by using a specific monoclonal antibody. The results showed that DNA methylation was not constant and there were some changes in 5mC levels at different developmental stages, although these variations were not always statistically significant (Figure 32).

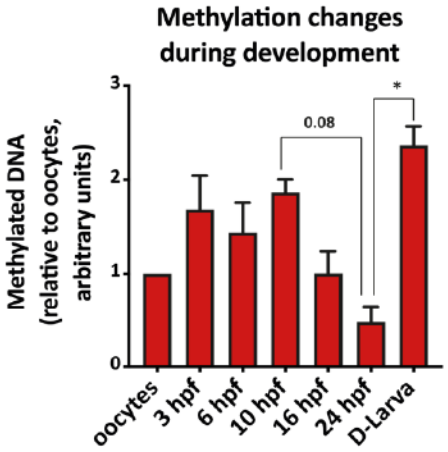

Figure 32: Levels of global methylation during early development in *M. galloprovincialis*. CpG DNA methylation during early development. Amount of CpG methylation during the early development stages using fluorescent ELISAs. The results were normalized regarding oocytes and are given for 10 ng genomic DNA coming from two biological replicates. * $p<0.05$, one-way Anova followed by Bonferroni's post hoc test.

However, the most interesting data I found is the quite evident reduction in the methylation levels that can be observed at 24 hpf in Figure 32. This data completely corroborates my previous immunofluorescent experiment, where I observed strong increases in the staining of 5hmC intermediate of demethylation and a reduction of 5mC staining (Figure 31 F2; F3). Furthermore, a strong reduction in methylation levels at 24 hpf is also in agreement with my qPCR experiment showing a strong reduction in Dnmt1 expression with a parallel increase in Tet expression (Figure 28). Moreover, interestingly, our Elisa data also completely recapitulate what has been found in *Patinopecten yessoensis* (Li et al., 2019) and at least in part with what has been found in *Crassostrea*

gigas (Riviere et al., 2013). Some of the inconsistencies between Elisa (Figure 32, a quantitative method) and immunofluorescence (Figure 31, not a quantitative method) could also be explained by the 3D structure of the chromatin that prevents in some areas the proper binding of the antibody.

4.2.7 Inhibition of DNA methylation induces severe alteration of *M. galloprovincialis* development

At this point to gain more insight into the potential role of DNA methylation during *M. galloprovincialis* development, we decided to use a well-known hypomethylating agent: 5-azacytidine. 5-azacytidine is incorporated in DNA and binds DNMT1 irreversibly leading to a progressive loss of DNA methylation through cell divisions. This drug has been already used in many organisms including mollusks, such as *Crassostrea gigas* and *Biomphalaria glabrata* (Riviere et al., 2013; Geyer et al., 2017). However, since this drug has never been used in *M. galloprovincialis*, I first tested its activity using different doses from 0.01 µM to 10 µM with the treatment starting from 1.5 hpf. During these analyses, I found that 10 µM concentration appears to be much more harmful to embryos and therefore this concentration was excluded from the analyses. At this point, I went on by treating embryos from 1.5 hpf with four different concentrations of 5-azacytidine (0.01µM, 0.1µM, 0.5µM, 1µM) until the D-larva stage at 48 hpf and I observed embryos phenotype during two different time points: after 7 hpf (Figure 33 A-D) and at the end of the treatment at 48 hpf (Figure 33 E-I). The 0.01µM did not show any defects with most of the embryos being D-Larva at 48 hpf similarly to those treated with the vehicle (Figure 33 I). Interestingly, my initial observation with the microscope at 7 hpf revealed no major abnormalities at this stage although it is necessary to take into account that the microscope magnification used did not allow me to easily discern the morphology of the embryo (Figure 33 A-D). However, during repetitive phenotype observation, I found that embryos treated with the different doses of 5-azacytidine showed a decrease in the swimming behaviour that was proportional to the increases in the used doses. The apparent alteration in the swimming behaviour could be the result of a delay in development. At this point to better understand how 5-azacytidine and therefore DNA methylation inhibition can impair *Mytilus galloprovinicalis* proper development I decided to analyze the percentage of embryos that were able to reach the D-Larva stages. As visible from Figure 33, all the embryos treated with different doses of 5-azacitidine from 1.5 hpf were unable to reach

the D-Larva stage. All the embryos were aberrant and the phenotype became stronger in proportion to dose increase.

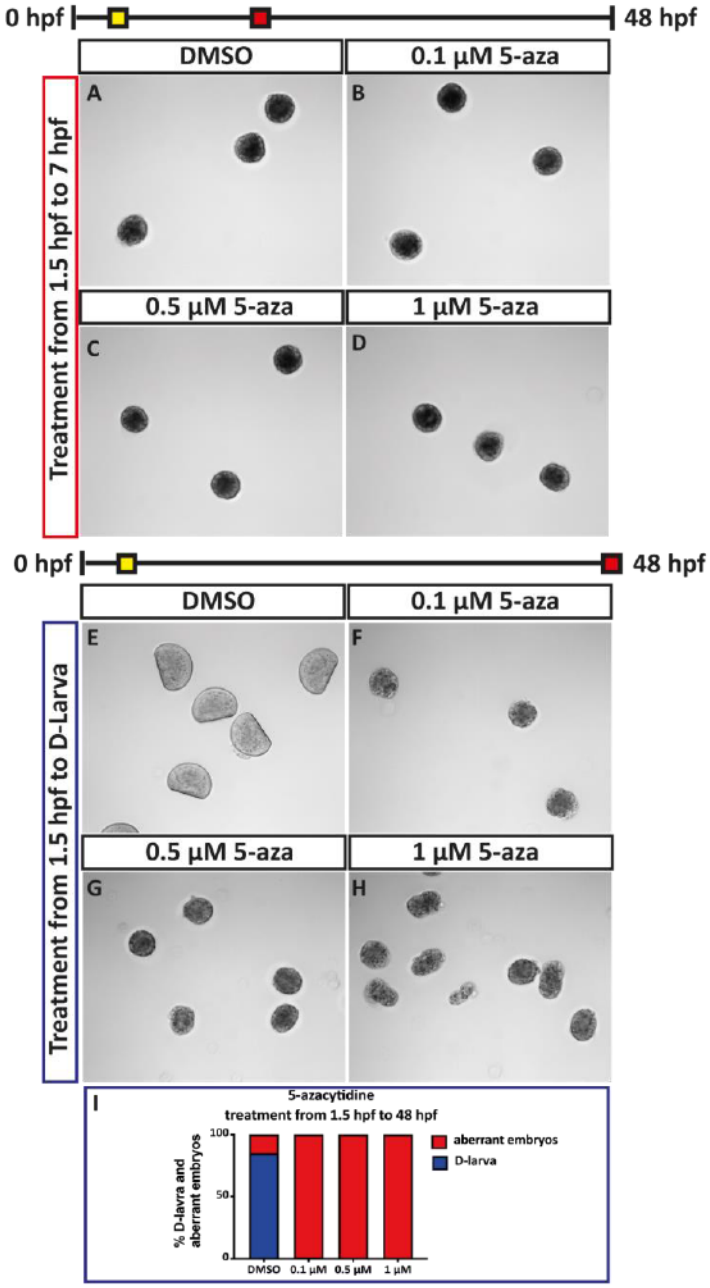

Figure 33: Influence of 5-azacitydine treatment after 1.5 hpf on *M. galloprovincialis* embryos at different doses and stages. (A-D) Embryos treated with different doses of 5-azacitydine from 1 hpf to 7 has been used as a statistical method. *$p<0.0001$ represent statistical significance between DMSO versus 0.1µM, 0.5µM, 1µM 5-azacitydine. (E-H) Embryos treated with different doses of 5-azacitydine from 1 hpf to 48 hpf. (I) Frequency distribution at 48 hpf of D-Larva stages after treatment with 5-azacitydine from 1.5 hpf. The black bar indicates the developmental time from fertilization to 48 hpf. The yellow square indicates the moment (1.5 hpf) when the treatment started. The red square indicates the developmental time when embryos have been observed under an optical microscope.

Subsequently, to better understand the potential effects of this compound on *Mytilus galloprovincialis* development and to better discern embryos morphology, I decided to perform specific analyses with the Scanning Electron Microscopy (SEM) by using just 0.1 µM dose of 5-azacytidine able to impair proper D-Larva development. Interestingly, SEM analyses confirmed my previous observation performed with a normal optic microscope. To this aim, I performed double analyses at 7 and 48 hpf. At 7 hpf embryos treated only with the vehicle showed a normal blastula rotatory with well-developed cilia, on the contrary, instead, embryos treated with the DNMT1 inhibitor clearly showed a strong delay in the development as visible also by the very low development of cilia, which confirms my previous observation on swimming behaviour (Figure 34 A-B). Observation at a later stage revealed instead that whilst DMSO treated embryos at 48 hours give rise to a well-defined D-larva, those treated with inhibitor remain mostly arrested during development with embryos consisting in a cluster of cell non properly aggregated (Figure 34 C-D), as they were probably blocked during the gastrulation. This phenotype was reminiscent of what has been previously found also in *Crassostrea gigas* (Riviere et al., 2013).

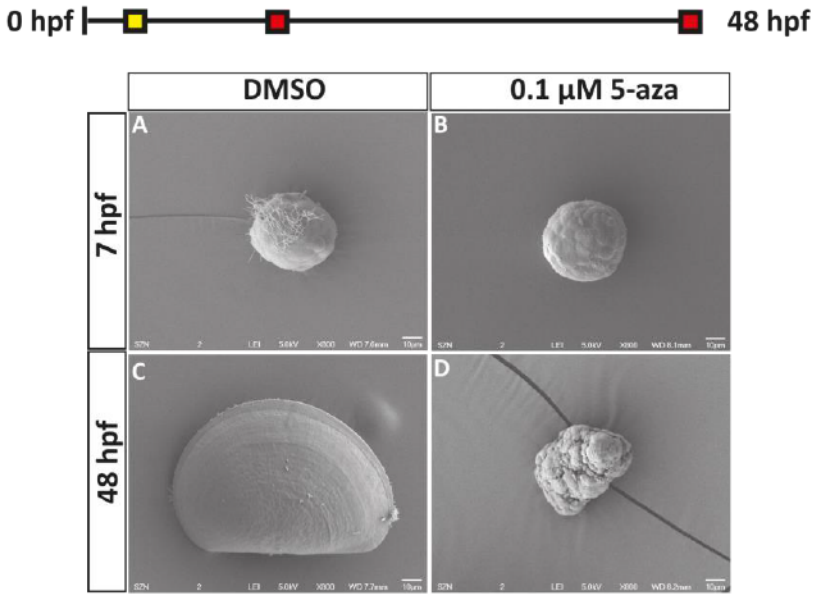

Figure 34: Influence of 5-azacitydine on *M. galloprovincialis* at 7 and 48 hpf. Scanning electronic microphotographs of *Mytilus* embryos treated with DMSO (A, C) or 5-azacitydine (B, D). The black bar indicates the developmental time from fertilization to 48 hpf. The yellow square indicates the moment (1.5 hpf) when the treatment started. The red square indicates the developmental time when embryos have been observed at SEM.

At this point, to show that these phenotypes were due to inappropriate DNA methylation during the early stage of development rather than the result of unspecific teratogenic effects of 5-azacytidine I first decided to analyze the ability of 5-azacytidine to determine a reduction in the methylation in *Mytilus galloprovincialis* embryos. To this goal, I used two approaches and interestingly both of them showed the ability of 5-azacytidine to determine a reduction in the methylation levels (Figure 35). However, the reduction I found when I performed the Elisa experiment did not result in any statistical significance, although this could be the result of a low number of biological replicates used (n=2).

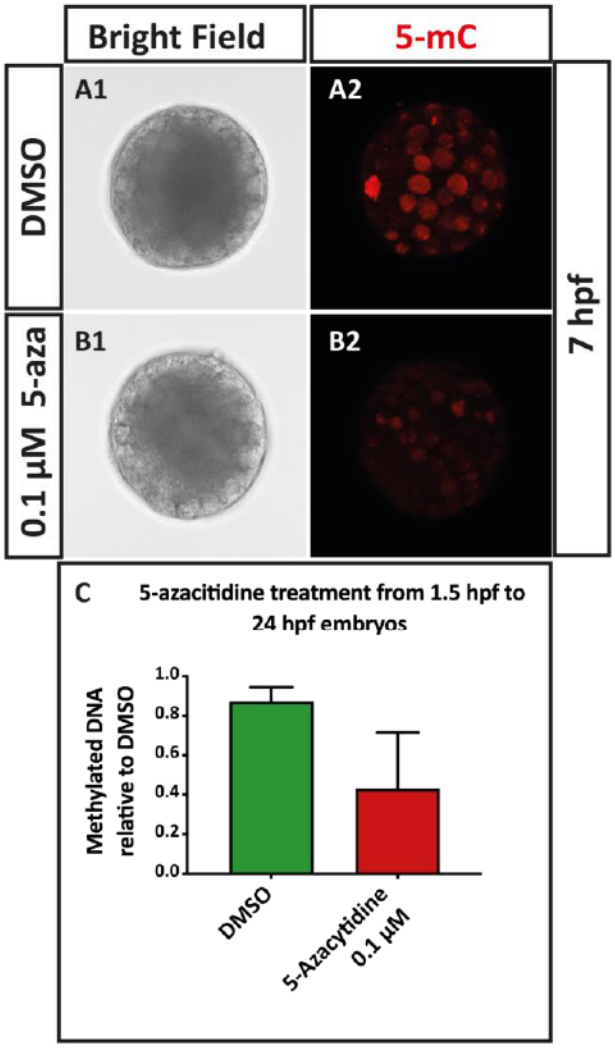

Figure 35: Analyses to determine methylation levels in embryos treated with 5-azacitydine. (A1-B2) Immunofluorescence analysis showing the distribution of 5mC on *M. galloprovincialis* at 7hpf in presence of vehicle and 0.1µM of 5-azacytidine. Treatment started at 1.5 hpf. (C) Elisa analysis to determine the level of global methylation during early development in *M. galloprovincialis* at 24 hpf in presence of vehicle and 0.1µM of 5-azacytidine. Treatment started at 1.5 hpf.

Next, to rule out the possibility that the reduction in the methylation levels in embryos treated with 5-azacitydine was the result of nonspecific toxicity, I decided to take advantage of Propidium Iodide assay (PI) and fluorescein diacetate staining (FDA). PI is a vital dye that can be used to determine if the observed cells are viable or dead, since PI is a membrane-impermeant dye excluded from viable cells. FDA on the other side is a cell-permeant esterase substrate producing free fluorescent fluorescein into viable cells. Enzymatic activity and cell-membrane integrity are required for the fluorescent products. This type of double essay (PI-FDA) is commonly used for identifying dead or live cells in a population or whole embryos. For these experiments, we used a double control. Specifically, one control was based on the use of D-larvae or aberrant embryos from the batch just treated with DMSO. The second control was represented by the use of the same DMSO treated embryos, but treated for 1 hour in paraformaldehyde 4% (PFA), inducing cell death (Sharp et al., 2013). In parallel, embryos treated with different doses of 5-azacytidine were also tested (Figure 36). The results obtained were convincing in excluding nonspecific embryo toxicity. Indeed, in embryos treated only with DMSO (Figure 36 B3-C3), the PI did not penetrate cells because they are alive and viable embryos appeared with green fluorescence (Figure 36 C3). Meanwhile, in fixed DMSO embryos, as expected, red fluorescence was observed (Figure 36 B2, B3), with the absence of green fluorescence further supporting the presence of dead cells (Figure 36 C1, C2). Interestingly, when I analyzed aberrant embryos treated with different concentrations of 5-azacytidine from 1.5 hpf to 48 hpf I did not find any specific toxicity. Indeed, I never observed any positive staining to PI, suggesting that at least at these doses the Dnmt1 inhibitors are not toxic (Figure 36 B4-B6; These results were further confirmed also by the positive staining with FDA (Figure 36 C4-C6).

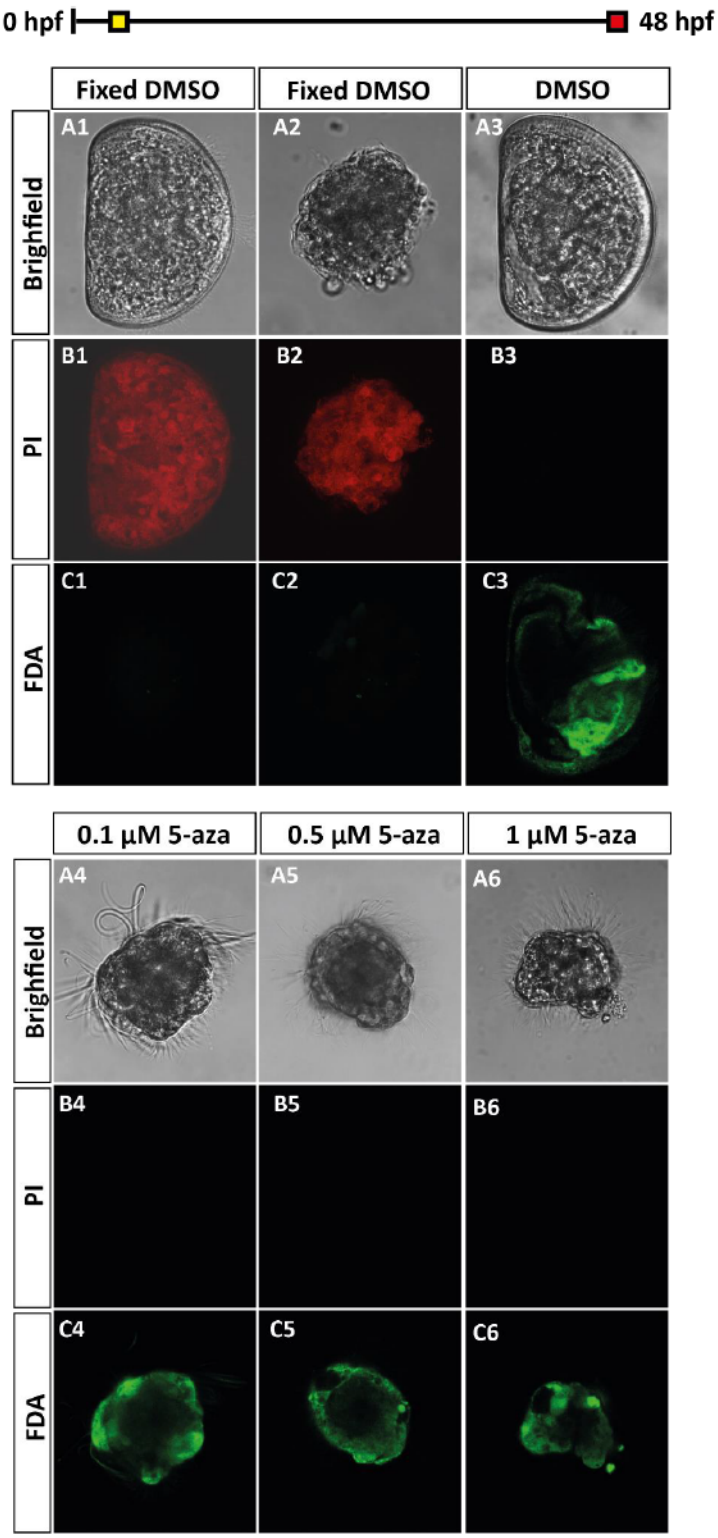

Figure 36: Representative confocal fluorescence images of embryos in presence of Propidium Iodide (PI) and Fluorescein diacetate (FDA). (A1, A2, B1, B2, C1, C2) Embryos treated with DMSO and fixed with 4% PFA. (A3, B3) Embryos treated only with DMSO. (A4-A6; B4-B6; C4-C6), Embryos treated with different doses of 5-azacytidine. All the treatments have been performed from 1 hpf to 48 hpf.

The lack of specific teratogenic effect of the used doses of 5-azacytidine was confirmed also when I decided to treat embryos with the same doses, but with the treatment that started at 24 hpf (Figure 37 A-E). Results clearly showed that embryos develop mostly normally in all the analyzed conditions, differently from what I observed when the treatment started earlier (Figure 33). This data further suggests the specificity of the DNMT1 inhibitor phenotype that is due more likely to alteration of *Mytilus* embryonic DNA methylome rather than to unspecific teratogenic effects. The lack of malformations I found when treatment started at 24 hpf could also be explained by our previous qPCR experiments showing a strong reduction of Dnmt1 at 24 hpf. (Figure 28).

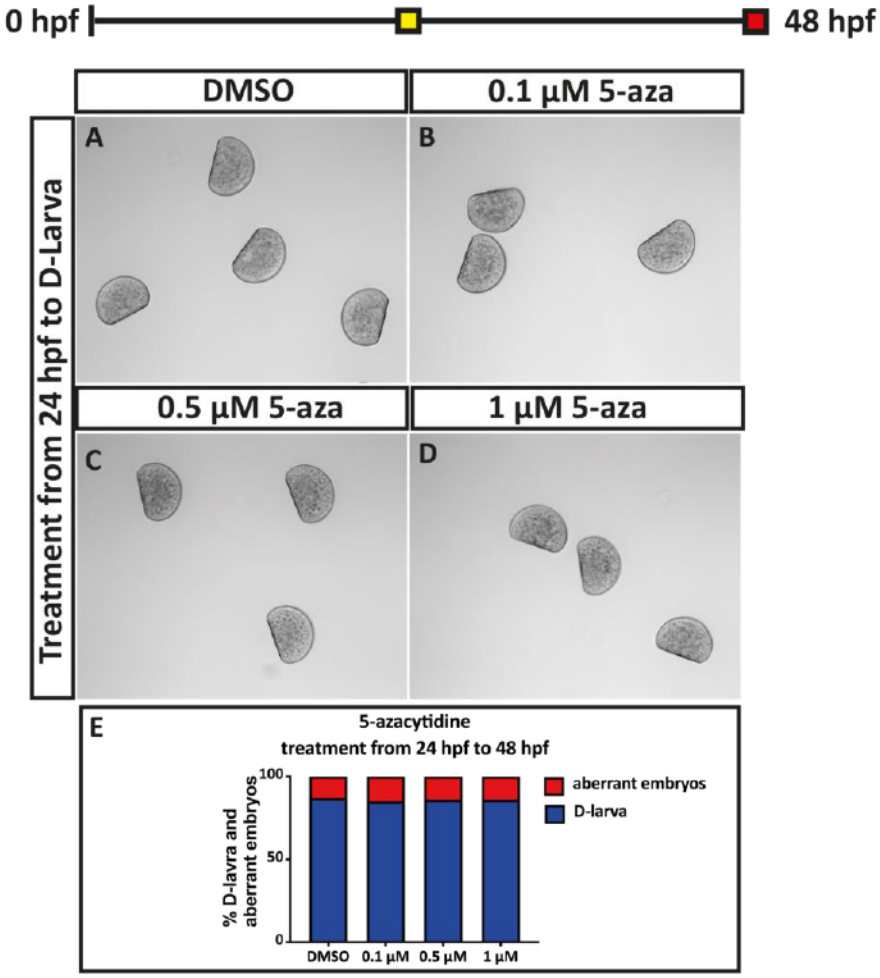

Figure 37: Influence of different doses of 5-azacitydine on *M. galloprovincialis* embryos with treatment from 24 to 48 hpf. (A-D) Embryos treated with 0.1µM, 0.5µM, 1µM 5-azacitydine from 24 to 48 hpf. (E) Frequency distribution at 48 hpf of D-Larva stages after treatment with 5-azacitydine from 24 hpf. Chi square test has been used as a statistical method. No significance was found between DMSO and treated embryos. The black bar indicates the developmental time from fertilization to 48 hpf. The yellow square indicates the moment (24 hpf) when the treatment started. The red square indicates the developmental time when embryos have been observed under an optical microscope.

4.2.8 Inhibition of DNA demethylation induces aberrant phenotype during *M. galloprovincialis* development

At this point, I decided also to investigate the role of DNA demethylation as currently there are no functional studies on this topic in non-vertebrate lineage. Therefore, to shed light on the role of the TET enzyme in the active removal of 5mC through the formation of 5hmC (intermediate of DNA demethylation) I decided to employ a specific inhibitor recently discovered Bobcat 339 (Chua et al., 2019). Differently from 5-azacitydine, which is a well-known inhibitor of DNA methylation both in vertebrates and invertebrate species, there was not much data available on the used doses of this compound in invertebrates. Chua et al., 2019 reported a significantly reduced global 5hmC levels by using Bobcat 339 at 10 μM on immortalized mouse hippocampal neuronal cell line. Therefore, starting with this information I first tested Bobcat 339 activity using different doses from 0.1 μM to 30 μM and by treating embryos from 1.5 to 48 hpf. The result clearly showed that 10-20-30 μM concentrations appear to be too harmful to embryos. Indeed, after only a few hours after drug application embryos were completely disaggregated into cells with no recognizable shape of any embryos and died soon. Therefore, these doses were excluded from the analyses. At this point, based on the same approaches I used for 5-azacytidine treatment, embryos have been treated initially from 1.5 hpf to 48 hpf with four different concentrations of Bobcat 339 (0.1μM, 1μM, 3 μM, 5μM) and observed first at 7 hpf (Figure 38 A-F). The results showed that most of the treated embryos from 1.5 hpf have defects starting at 1 μM with extroflection of some blastomeres. This phenotype becomes more severe at 3 μM (Figure 38 D) whilst at 5 μM (Figure 38 E) treated embryos showed also huge defects in blastomere adhesion, migration, and polarity as shown by the magnification in Figure 38 F. To better understand these phenotypes, I then performed SEM analyses on embryos fixed at this stage (Figure 38 G-J). Results showed a delay in the development for the 1 μM and confirmed our observation on blastomere extroflection with strong problems in proper development probably due to alteration in cell adhesion and migration.

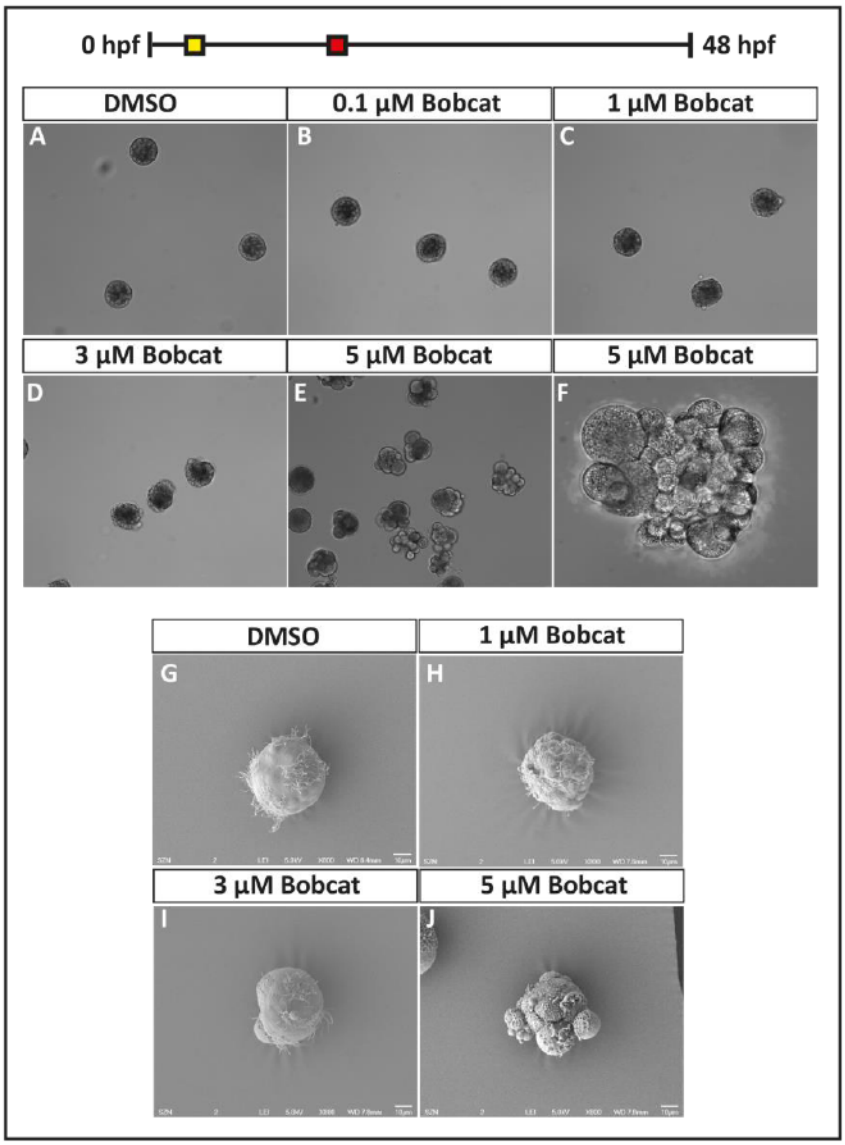

Figure 38: Influence of Bobcat 339 treatment after 1.5 hpf on *M. galloprovincialis* embryos at different doses. (A-E) Embryos treated with different doses of Bobcat 339 from 1 hpf to 7 hpf. (F) Magnification of aberrant embryos treated with 5 µM of inhibitor. (G, J) Scanning electronic microphotographs of *Mytilus* embryos treated with DMSO (G) or Bobcat 339 (H-J). The black bar indicates developmental time from fertilization to 48 hpf. The yellow square indicates the moment (1.5 hpf) when the treatment started. The red square indicates the developmental time when embryos have been observed under an optical microscope and SEM.

I then proceeded and analyzed the effects of the treatment at 48 hpf (Figure 39). As expected I found that most of the embryos were unable to develop proper D-Larva. Specifically, embryos treated with 0.1 µM showed a similar number of D-Larva as control. The 1 µM dose instead results in a big impairment in the development of proper D-Larva as only a little percentage of embryos reached this stage whilst the remaining embryos showed aberrant embryos or partial D-Larva as visible also by the SEM analysis

(Figure 39 C, G, I). Regarding embryos treated with 3 µM it looks that inhibition of TET at this dose mediates an arrest in the development between gastrula and late gastrula stage (Figure 39 D, H, I). Moreover, as expected, the severe phenotype we observed at 5 µM resulted in dead embryos that were fully disaggregated in separated blastomeres (Figure 39 E, I).

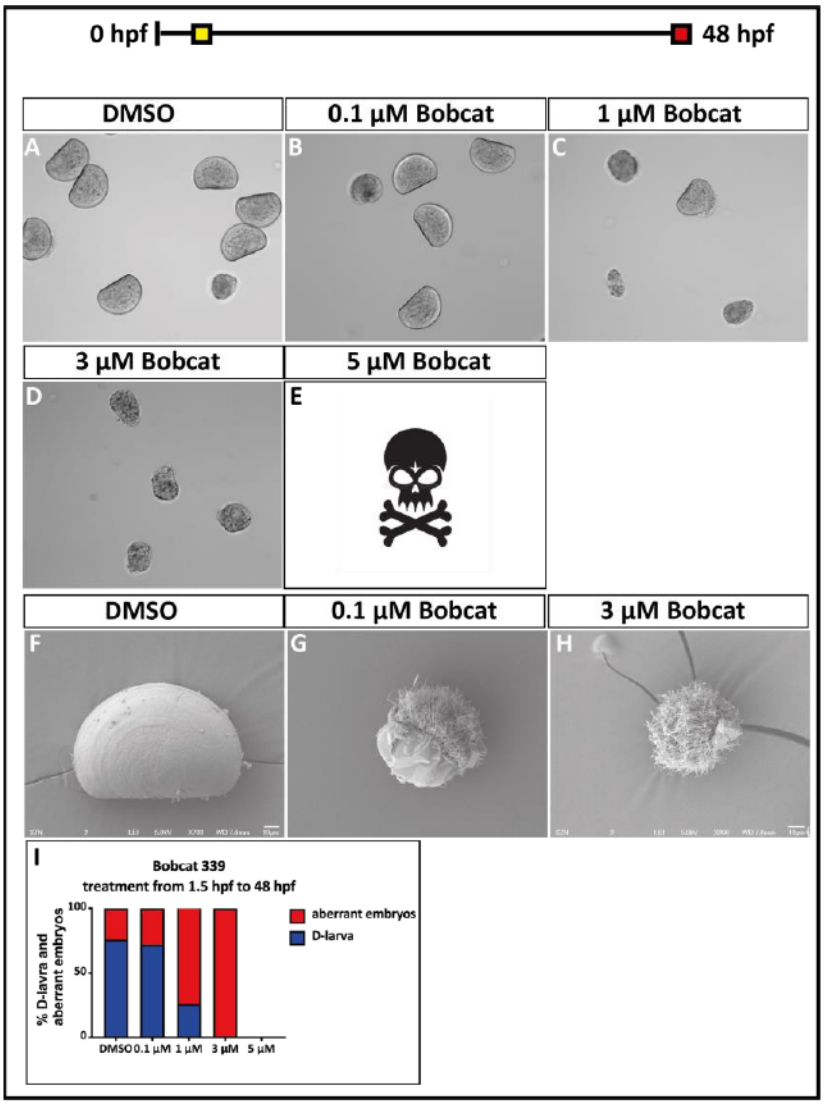

Figure 39: Influence of Bobcat 339 treatment after 1.5 hpf on *M. galloprovincialis* embryos at different doses. (A-E) Embryos treated with different doses of Bobcat 339 from 1 hpf to 48 hpf. (F, H) Scanning electronic microphotographs of *Mytilus* embryos treated with DMSO (F) or Bobcat 339 (G-H). The black bar indicates the developmental time from fertilization to 48 hpf. (I) Frequency distribution of embryonic stage of *M. galloprovincialis in vitro* fecundation assay in the absence (Control) or in presence of different doses of Bobcat 339 at 48 hpf. Chi-square test has been as a statistical method. *p<0.0001 represents statistical significance between DMSO versus 1µM, 3µM 5µM of Bobcat 339. The yellow square indicates the moment (1.5 hpf) when the treatment started. The red square indicates the developmental time when embryos have been observed under an optical microscope and SEM.

At this point, I was interested to determine the potential effect of Bobcat 339 when the treatment started at 24 hpf. These data are very interesting because differently from 5-azacytidine the treatment is still able to mediate a significant impairment in the development (Figure 40 A-F). Specifically, the 1 µM dose did not result in embryo defects and the 3 µM dose resulted mostly in partial D-Larva phenotype. Finally, regarding embryos treated with the 5 µM dose, although they did not develop in any D-Larva, the arrested aberrant embryos were not disaggregated as observed during the earlier treatment (Figure 40 A-F).

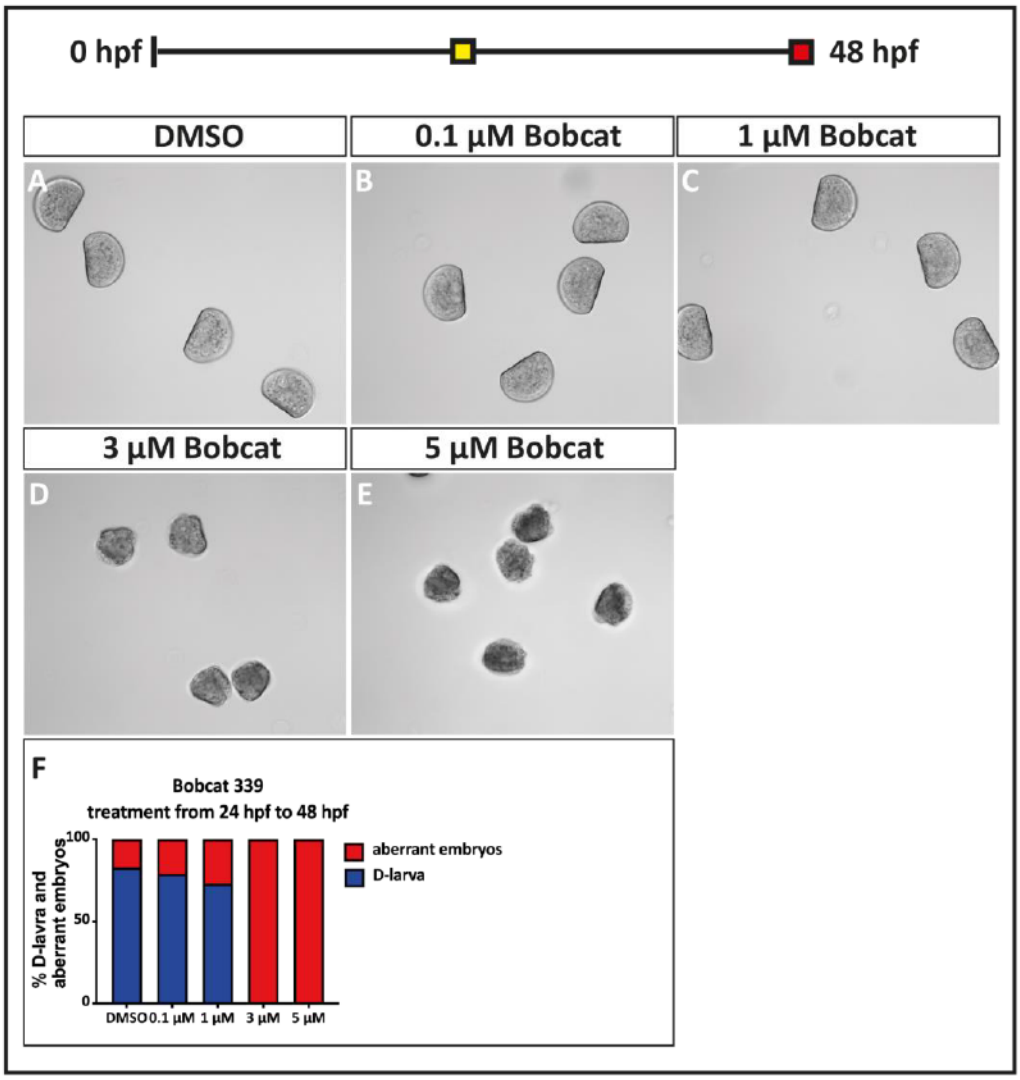

Figure 40: Influence of Bobcat 339 treatment after 1.5 hpf on *M. galloprovincialis* embryos at different doses. (A-E) Embryos treated with different doses of Bobcat 339 from 24 hpf to 48 hpf. (F) Frequency distribution of embryonic stage of *M. galloprovincialis in vitro* fecundation assay in the absence (Control) or in presence of different doses of Bobcat 339 at 48 hpf. Chi square test has been as a statistical method. *$p<0.0001$ represents statistical significance between DMSO versus 3µM and 5µM of Bobcat 339. The yellow square indicates the moment (1.5 hpf) when the treatment started. The red square indicates the developmental time when embryos have been observed under an optical microscope.

These results could be explained by my previous result on gene expression where I found that Tet, differently from Dnmt1, shows an opposite trend in its expression that increases as development proceeds and displays the highest expression at 24 hpf (Figure 28). These results in association with the Propidium Iodide and Fluorescein diacetate (Figure 41) assay I performed as I did for the 5-azacytidine seem to suggest that Bobcat 339 is not toxic for embryos and therefore any alteration in embryos development is the result of proper DNA demethylation mechanisms. Further studies will be necessary to confirm the ability of this compound to mediate alteration in demethylation. First of all, a specific Elisa Kit for detecting 5hmC will be used to determine the effect on demethylation with a low dose of Bobcat 339. Furthermore, immunofluorescence studies will be also employed to determine the effects of DNA demethylation in the whole embryos.

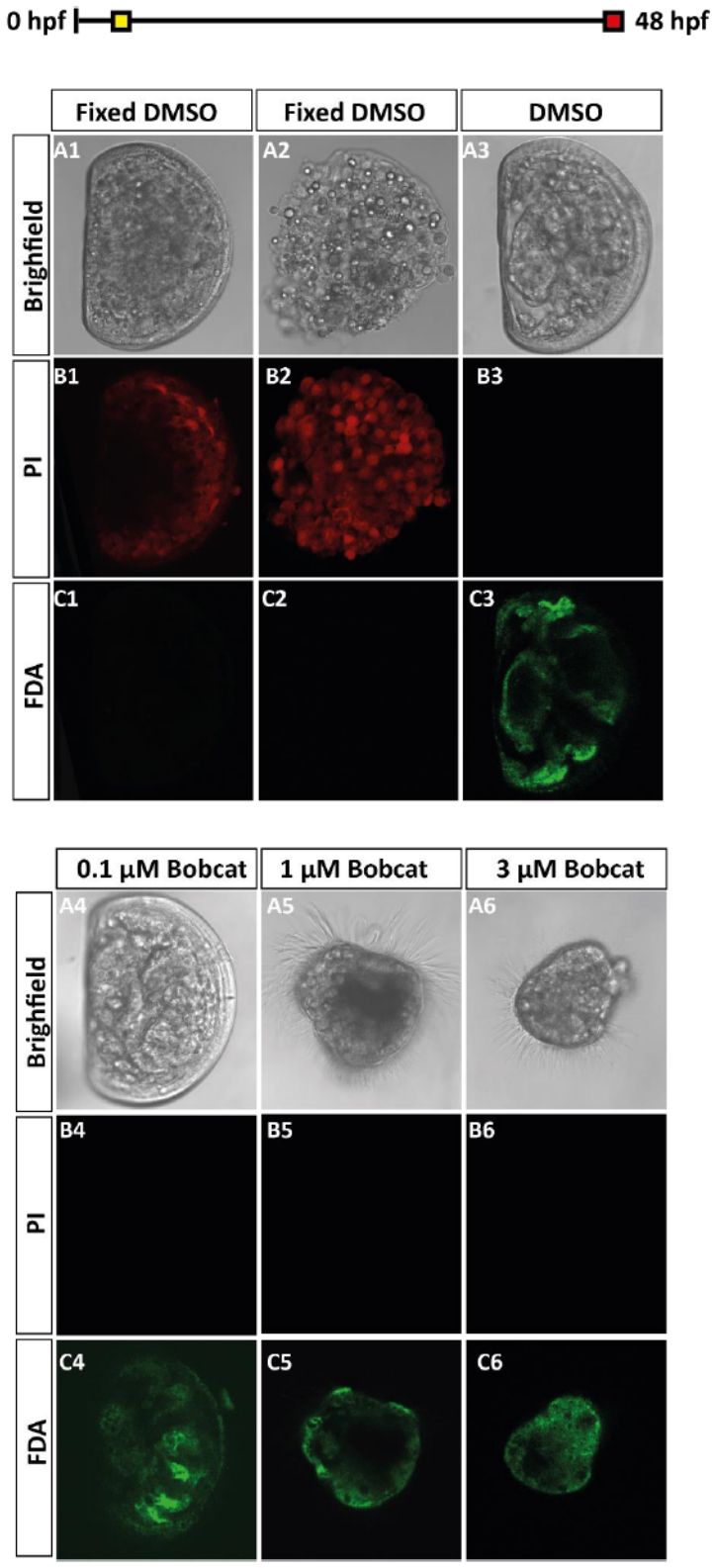

Figure 41: Representative confocal fluorescence images of embryos in presence of Propidium Iodide (PI) and Fluorescein diacetate (FDA). (A1, A2, B1, B2) Embryos treated with DMSO and fixed with 4% PFA. (A3, B3) Embryos treated only with DMSO. (A4-A6; B4-B6), Embryos treated with different doses of Bobcat 339. All the treatments have been performed from 1 hpf to 48 hpf.

4.3 Effects of emerging contaminants on *Mytilus galloprovincialis* embryos: the case of carbamazepine (CBZ) and valsartan (VAL)

Once the critical role of DNA methylation during *M. galloprovincialis* development was established, I was interested during my studies to determine if specific environmental stressor could interfere with DNA methylation mechanisms during mussel development. DNA methylation, indeed, has been already shown to be an important mechanism in gene regulation in response to environmental stresses and accumulating evidence clearly established the connection between DNA methylation and phenotypic plasticity in marine metazoan (Suarez-Ulloa et al., 2015; Rondon et al., 2017b). Based on this, the last part of my PhD project was focused on the study of the potential effects of Emerging Contaminants in mediating DNA methylation changes during *Mytilus galloprovincialis* development. A lack of information exists regarding the fate of pharmaceuticals in the marine ecosystem. Among such compounds, I decided to investigate the effects of two pharmaceuticals with widespread contamination in marine ecosystems on *Mytilus* embryogenesis and its potential effects on DNA methylation. To this aim we chose the antiepileptic Carbamazepine (CBZ), and the antihypertensive Valsartan (VAL) (Figure 42).

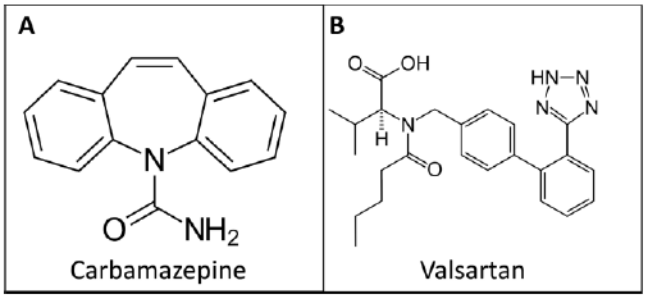

Figure 42: Molecular structure of pharmaceuticals. (A) Carbamazepine and (B) Valsartan.

4.3.1 The effects of emerging contaminants in mediating DNA methylation changes during *M. galloprovincialis* development

The potential impacts of Carbamazepine and Valsartan on *Mytilus galloprovincialis* organisms have been already reported in many studies including those of my tutor Prof. Regoli. (Franzellitti et al., 2019; Mezzelani et al., 2020, 2021). Interestingly, recently, carbamazepine has been also tested at different concentrations during embryogenesis

(Franzellitti et al., 2019). The results of these studies showed a strong impact of this pharmaceutical on *Mytilus* embryogenesis, as also revealed by changes in the transcriptional levels of many genes. On the other hand, although Valsartan has been previously reported as one of the most abundant Emerging Contaminants in *Mytilus* tissues, together with carbamazepine (Mezzelani et al., 2020), its effects on embryogenesis have never been analyzed. Therefore, given the interesting results of Franzellitti et al. 2019, which showed a strong impact of Carbamazepine on gene expression during embryogenesis, I decided to investigate the effects of this compound in mediating alteration of many transcriptional pathways could also affect the levels of DNA methylation. In this context, I decided also to test in parallel the effects of Valsartan and the combination of both compounds together. Regarding the analyses of potential DNA methylation, we decided to use a new approach based on Methyl-Rad sequencing. This method is based on methylation-dependent restriction enzymes that can discriminate between CG and non-CG methylation as the methylation status of each site is independently interrogated. However, given the high cost of the sequencing, experiments have been performed using a single dose of both compounds (0.5 µg/L) and their combination in quadruplicates. The decision to use these specific doses relies on the fact that they are considered to be in the range of environmental concentrations (Mezzelani et al., 2020; Almeida et al., 2021) and on the previously published data showing alteration in gene expression at these doses (Franzellitti et al., 2019). In detail, the experimental design includes 3 different conditions with organisms exposed to 0.5µg/L of CBZ, 0.5µg/L of VAL, a mixture of 0.5µg/L of CBZ and 0.5µg/L of VAL together. As control has been used the same vehicle where the compounds have been dissolved (DMSO). I next proceeded by treating fertilized embryos after 1 hpf with the compounds of interest. The embryos were then allowed to grow until 48 hpf at D-Larva stage. Later the embryos have been counted and phenotypic analyses were performed to discriminate the potential defects compared to the control. Our analyses did not show any specific defect in both compounds, differently from what has been previously shown for Carbamazepine (Franzellitti et al., 2019) (Figure 43). The discrepancy of these results could be the result of different components. First, different animals could behave slightly differently based on their genetic background or another possible explanation could be linked to the different timing of the start of the incubation with the compounds (0.5 hpf, previous analyses -1 hpf, our analyses). However, given the strong effects that have been

previously found on changes in gene expression for Carbamazepine, I decided to proceed and extract genomic DNA that was subsequently used for the Methyl-Rad sequencing after the opportune digestion and library preparation.

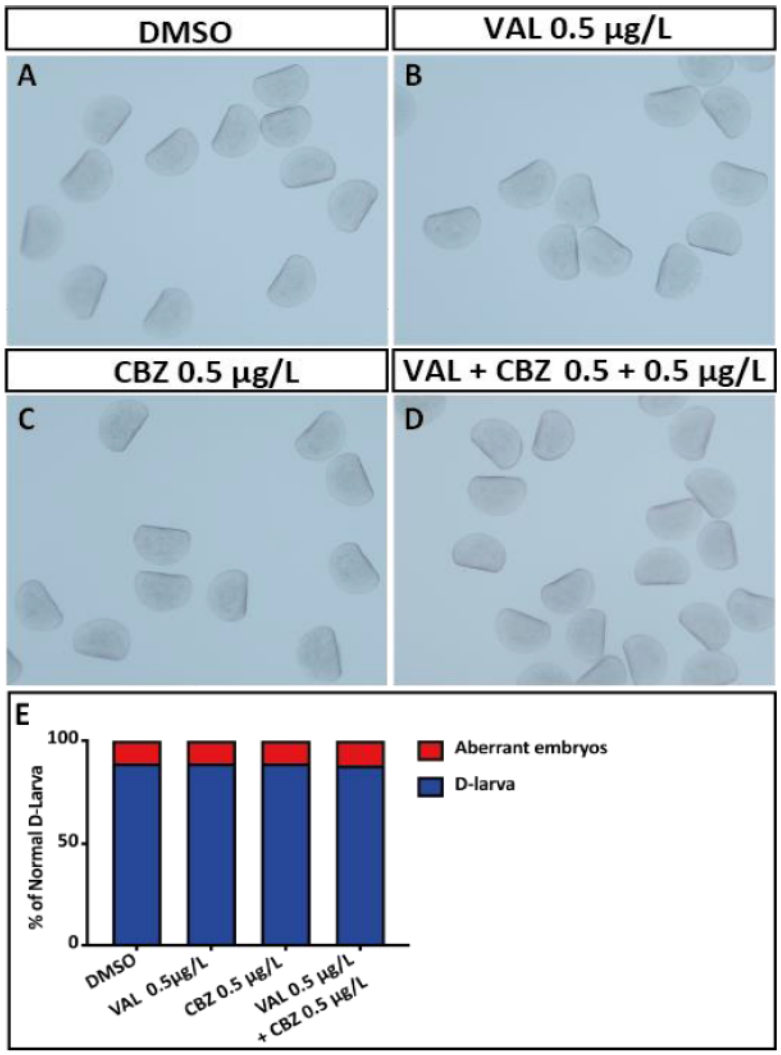

Figure 43: Embryos after exposure at 48 hpf with the ECs. (A-D) Representative image showing the phenotype of D-Larva in presence: (A) DMSO, (B) Valsartan, (C) Carbamazepine, (D) Valsartan and Carbamazepine together. (E) Graphic showing the percentage of D-Larva (blue) and aberrant embryos (red) in the four different conditions. Frequency distribution at 48 hpf of D-Larva stages after treatment with Carbamazepine and Valsartan from 1 hpf to 48 hpf. Chi-square test has been used as a statistical method. No significance has been found between DMSO and treated embryos.

4.3.2 Methyl-RAD library preparation

To determine the potential effects on DNA methylation of two pharmaceutical compounds object of the study I took advantage of a recent sequencing analysis called Methyl-RAD, which is a useful approach to investigate epigenetic changes in response to environmental stress. Genomic DNA (gDNA) was extracted from samples (Figure 44) and gDNA quality and integrity were then evaluated on agarose gel (0.7% of agarose).

High-quality and high-molecular-weight DNA is needed to fully benefit from Methyl-RAD sequencing.

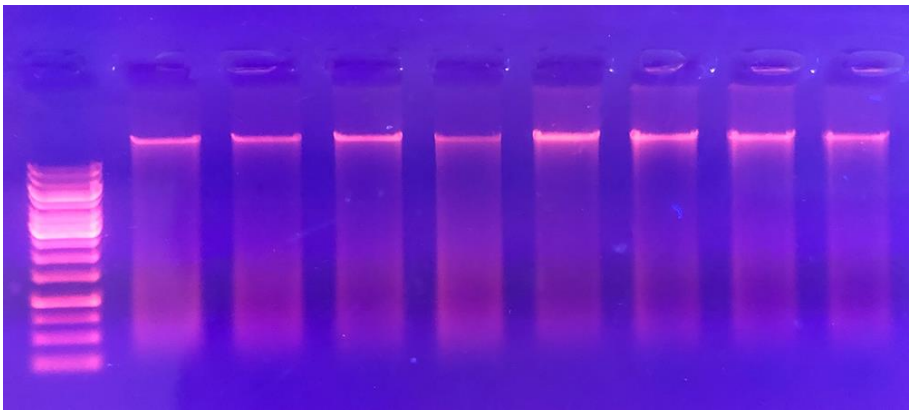

Figure 44: Example of a gDNA extraction. The first 4 lanes show the gDNA coming from the first elution, while the second 5 lanes showed gDNA extracted with a second elution.

At this point, I sent gDNA samples (five for each condition) to the Lab of Prof. Pallavicini Alberto in Trieste for the library preparation. Following the protocol, gDNA was digested with LpnPI enzyme to produce 32-base-restriction-fragments that were later ligated with specific adapters necessary to provide a unique barcoding sequence for each condition. In column 1 of Figure 44, a blue arrow shows gDNA digested with LpnPI, ligated with adaptors, and amplified, showing a fragment of the expected size (170 bp). In column 2, the yellow arrow shows gDNA not digested, ligated with adaptors and amplified. In this case, we did not find a fragment of the expected size (Figure 45). Then, the libraries were pooled at equimolar concentration and sent for sequencing with 1X75 Illumina technology at BMR genomics, Padova.

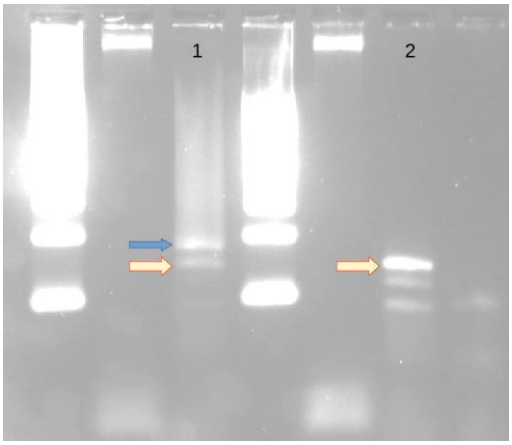

Figure 45: Example of restriction digestion analyses with LpnPI. Agarose gel electrophoresis (2% agarose) of amplification products of (column 1 blue arrow) gDNA digested and ligated with adaptors and (column 2 yellow arrow) gDNA not digested with LpnPI and ligated with adaptors.

4.3.3 Preliminary sequence analyses

Just before I was about to complete the writing of my thesis, we received the sequences, which at the present time have been only subjected to trimming analyses. Each sequencing library produced more than 10 million reads. Based on the result of the quality check (QC), a pre-processing step of trimming was applied using QIAGEN CLC Genomics Workbench. In this step, raw reads were first trimmed to remove low-quality and adapter sequences. Reads containing ambiguous nucleotides were also removed. Moreover, we decided to maintain only sequences with a length comprised between a minimum of 32 and a maximum of 33 nucleotides. Around 16 billion nucleotides were trimmed altogether, and approximately 100 million reads were completely removed during trimming (Table 6). High quality (HQ) reads will be used for subsequent analysis. Around 90% of these HQ reads could be mapped to the *M. galloprovincialis* genome assembly, version mg10 (Gerdol et al., 2020).

Trim	Input reads	No trim	Trimmed	Nothing left or Discarded
Trim on quality	309,234,567	309,192,106	42,461	0
Ambiguity trim	309,234,567	309,234,387	180	0
Adapter trimming	309,234,567	3,419,687	305,642,967	171,913
Filter on length	309,062,654	209,951,071	0	99,111,583

Table 5: Detailed trimming results.

After trimming, the program generates a graphical report containing plots of various QC metrics. This step assists in the quality control process by assessing and visualizing statistics related to:

1. Sequence read lengths and base coverage
2. Nucleotide contributions and base ambiguities
3. Quality scores
4. Over-represented sequences, which may suggest contamination events

The graph in figure 46 shows the percentage of reads that were observed for each individual sequence length after trimming. Specifically, the result of this analysis shows that our sequences are all 32-33 bases long, as we expected from the enzymatic cut.

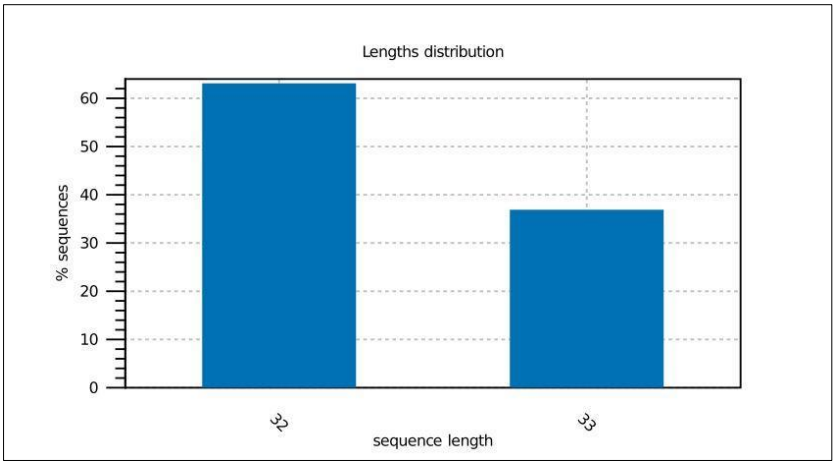

Figure 46: Length distribution after trimming.

Then, the quality of the sequences was evaluated by using a Phred quality score and most of the reads in our samples showed a score of 35 (Figure 47). The quality of a sequence is calculated as the arithmetic mean of its base qualities. PHRED-scores of 30 and above are considered an indication of high quality.

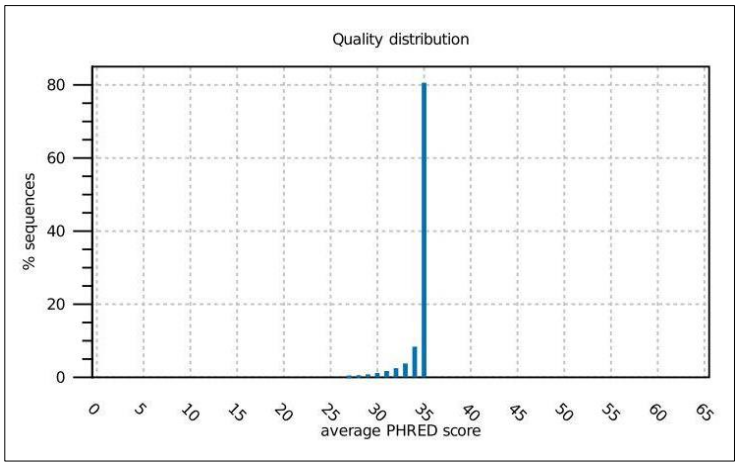

Figure 47: Quality distribution after trimming.

Moreover, nucleotide contributions were analyzed (Figure 48). These analyses calculate absolute coverages for the four DNA nucleotides (A, C, G, or T) for each base position in the sequences. A necessary premise is that LpnPI generates a double-stranded DNA break on the 3' side of the modified cytosine at a fixed distance (N_{12}/N_{16}). If the target sites are symmetrically methylated, LpnPI cleaves bidirectionally to generate 32-base-long fragments with the methylated site in the middle. This is visible in the plot of nucleotide contribution. The proportion of bases before and after the recognition site remains relatively constant with a predominance of T and A (as expected from the GC content of the mussel genome). Moreover, as we expected, there is a clear difference

between the four bases in the middle of the sequences with a predominance of C followed by G representing the recognition site.

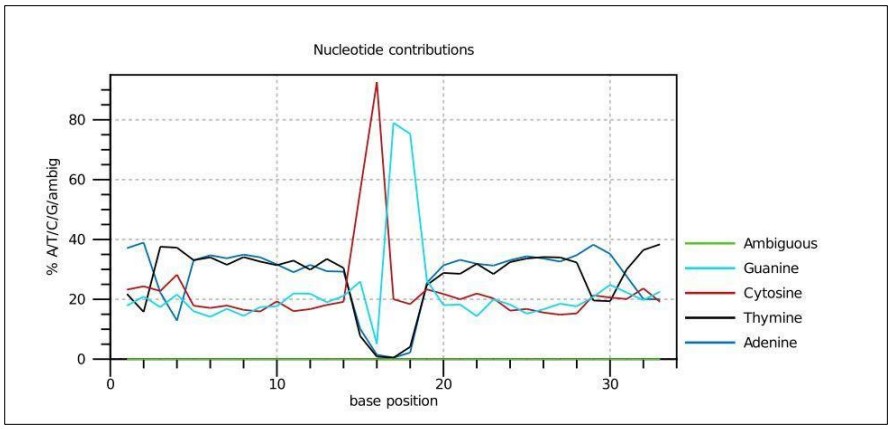

Figure 48: Nucleotide contributions after trimming.

This data is confirmed also by the GC-content plot that calculates absolute coverage of C's + G's for each base position in the sequences.Again, there is a peak of GC in the putative recognition site of LpnPI enzyme (Figure 49).

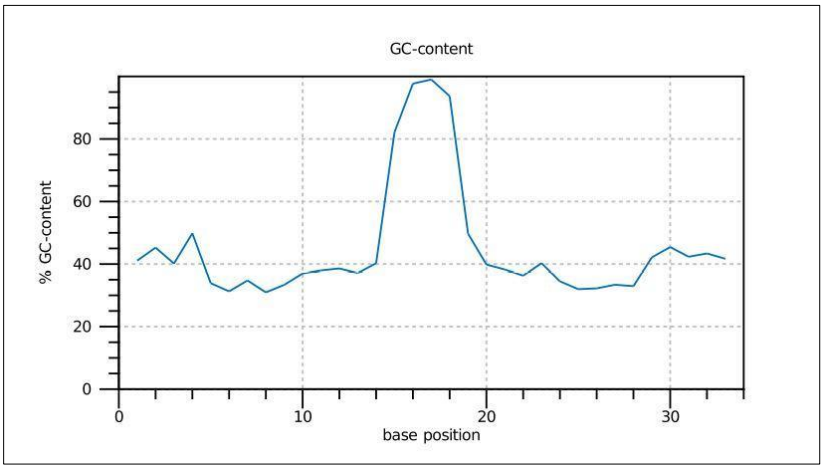

Figure 49: GC-content after trimming.

This is the first output of the experiment. In the following steps, high-quality reads will be analyzed, probably using two different approaches. In the first approach, mapped reads will be used to assign read count scores to each sample LpnPI site in the *Mytilus galloprovincialis* genome. Furthermore, obtained 32bp sequences from the control samples will be used to produce a virtual genome that will be used to identify differences with data coming from reads of the other conditions. In this study, only CpG sites will be considered. The use of different algorithms will allow the identification of the genomic

location of differentially methylated sites (DMSs) in all the different conditions. These regions will span from the transcription start site, exon, intron, and intergenic to the untranslated regions covering 5' UTR and 3' UTR. Later, the effects of Carbamazepine and Valsartan to mediate potential methylation changes in specific genomic regions (hypermethylation or hypomethylation in enhancer or gene body regions) will also be correlated to potential differential gene expression with further RNAseq o qPCR data. Ultimately, the results I will obtain from this Methyl-Rad sequencing will also help to better investigate the features of *Mytilus galloprovincialis* methylome that are currently completely unexplored and to integrate recent findings on the function of DNA methylation of marine organisms to these environmental threats.

CHAPTER 5. DISCUSSION AND CONCLUSION

DNA methylation is an important epigenetic modification, evolutionary conserved in all domains of life (Bird, 2002). The regulatory function of DNA methylation in the vertebrate genome has been deeply investigated and shown to be involved in many biological processes such as regulation of gene expression, DNA structure, and control of transposable elements (Law and Jacobsen, 2010; Schübeler, 2015). Overall, vertebrate genomes show widespread and "global" methylation, in contrast, most invertebrate genomes are "mosaic" and sparsely methylated, with some that have completely lost DNA methylation (Xu et al., 2019). Additionally, DNA methylation mechanism is a critical regulator of mammalian embryonic development. Mammalian genomes undergo two waves of reprogramming of CpG methylation, following fertilization and after germline cell specification. During these processes, genome-wide demethylation occurs through passive, but also active processes mediated by TET enzymes (Eckersley-Maslin et al., 2018; Greenberg and Bourc'his, 2019). Although most of the studies on this subject have so far targeted only a few invertebrate models, very recently we assisted in an increase in the number of laboratories that are investigating these processes in invertebrates including Lophotrochozoa. Furthermore, very recently a growing series of publications are reporting the role played by DNA methylation in various marine organisms and how epigenetic changes can impact their adaptability to a new environment. In this context, the goal of my PhD project was to explore the role of cytosine DNA methylation both in adults and embryos of an emerging developmental and evolutionary biology non-model species, the bivalve *Mytilus galloprovincialis*. Therefore, during my PhD, I attempted to answer these questions using different approaches.

- ***Mytilus galloprovincialis*** **possesses an ancestral-like repertoire of DNA methylation toolkit genes**

First, by taking advantage of collaboration with Prof. Pallavicini and Prof. Gerdol from the University of Trieste I decided to explore the evolutionary history of DNA methylation related genes in Bivalvia and more in detail in *Mytilus galloprovincialis* by using different bioinformatics software. These analyses revealed for the first time the existence of a full complement of DNA methylation-related genes in *Mytilus* genome,

Dnmt1, Dnmt3, Tet, Mbd2/3 and Mbd4 (Figure 18). These genes display a relatively good conservation of the key conserved domain with other bivalves and human orthologs suggesting that *Mytilus* methylation genes can exert a similar function to vertebrate counterparts (Figure 19-20). Phylogenetic analyses revealed a monophyletic origin of all five genes object of the study (Figure 22). *In silico* analyses using available transcriptome (RNA-Seq) revealed for the first time the presence of these genes in 23 bivalve species (Figure 21) and their poor expression in 6 analyzed bivalve species including *Mytilus galloprovincialis* (Figure 23). These data were also confirmed when I performed qPCR analyses of Dnmt1-3, Tet and Mbds on various tissues including hemocytes, mantle edge, foot, gills, digestive gland and adductor muscle (Figure 24). The expression of DNA methylation toolkit genes into the analyzed adult tissue hints at a possible function during adult life and potential response to environmental changes as in other mollusks (Fallet et al., 2020).

- **DNA methylation role during *Mytilus galloprovincialis* embryogenesis**

Currently, very little is known about the potential role of DNA methylation during embryogenesis in invertebrates and even less in mollusks. However, very recently, the dynamics of methylation changes occurring during development were observed in the early embryos of echinoderms, tunicates, polychaetes, lancelets, and also in some bivalve species, in contrast with sea anemones and honey bees (Riviere et al., 2013, 2017; Li et al., 2019; Xu et al., 2019; Skvortsova et al., 2021). Therefore, to gather the first glimpse of this process during *Mytilus galloprovincialis* embryogenesis, the expression of DNA methylation machinery toolkit genes was analyzed during different developmental stages. Concerning Dnmt1 (maintenance DNA methylation) and Dnmt3 (de-novo DNA methylation), my data are somehow in agreement with what has been previously found both in invertebrate and vertebrate species (Riviere et al., 2013; Firmino et al., 2017; Uysal et al., 2017; Li et al., 2019), with a peak in their expression in oocytes and at 3 hpf, that then dramatically decreases as development proceeds (Figure 28; 30). Interestingly, the initial maximum Dnmt1 expression observed during the first stage of development reflects also the stronger levels of DNA methylation I found with both Elisa and immunofluorescence analyses (Figure 31-32). Indeed, the peak of 5mC content observed around the 10 hpf could be the result of the relatively high level of Dnmt1 expression during the first stage of development (Figure 28; 30). Moreover, the dramatic decrease in

DNA methylation (Figure 31-32) at 24 hpf also correlates with the reduction in Dnmt1 expression at this stage (Figure 28, 30). These data, therefore, seem to demonstrate a potential reprogramming of DNA methylation during *Mytilus galloprovincialis*. Therefore, to gain more insights into this process our group decided for the first time in mollusks to investigate in parallel the demethylation mechanisms. Therefore, I analyzed the expression of Tet, the main enzyme of DNA demethylation and necessary to convert 5mC in 5hmC. Interestingly, gene expression analyses (Figure 28; 30) during development showed that the expression of this gene increased as development proceeded. These data are fully in agreement with a demethylation process I observed with global DNA methylation analyses and immunofluorescence. The presence of a strong peak in Tet expression at 24 hpf (Figure 28) correlates with the very low level of DNA methylation at this stage and increased content of 5hmC (Figure 31-32). Concomitantly to my PhD studies, a very interesting manuscript reporting TET demethylation enzyme expression during development in sea urchin, lancelet and zebrafish has been recently published in BioRxiv (Skvortsova et al., 2021). Again, the data reported in this manuscript showed a peak of Tet expression specifically during mid-development stages, which is very similar to my observations (Figure 28). The potential ability of *Mytilus galloprovincialis* to go through a DNA methylation reprogramming were also shown when I employed two inhibitors of the main methylation and demethylation gene, the 5-azacytidine for DNMT1 and Bobcat 339 for TET. This approach has been used because currently the classical methods based on the use of CRISPR/Cas9 or morpholino injection are not easy to apply in *Mytilus* . The treatment with drugs however represents a good alternative tool to answer these questions, although it is hard sometimes to discriminate between nonspecific toxicity or off-target inhibition. However, whilst Bobcat 339 inhibitor has never been used before in any invertebrate due to its recent discovery (Chua et al., 2019), 5-azacitidine has been instead widely used in many vertebrate and invertebrate organisms for understanding DNA methylation function but also as a therapeutic treatment for cancer. Previous studies have reported that 5-azacytidine induces significant cytotoxicity. To avoid this problem, I performed experiments during the development with very low doses in comparison to what I found in the literature. DNMT1 inhibition with 5-azacytidine revealed a strong impairment of *Mytilus galloprovincialis* development (Figure 33-34), with embryos showing strong defects that become more evident as development proceeded. Specifically, embryos

observed at 7 hours when the treatment started at 1.5 hpf do not show huge defects other than some developmental delay (Figure 33 A-I and 34 A-B). Differently, observation of embryos after a long time of treatment showed huge developmental defects with embryos that were arrested and showed evident alterations (Figure 33 E-I and 34 C-D). Interestingly, specific Elisa and immunofluorescence analyses confirmed a reduction in DNA methylation in embryos treated with 5-azacytidine (Figure 35). The multidisciplinary approach we used seems to support that aberrant embryos are the results of specific Dnmt1 inhibition at the very first stage of development being its expression higher (Figure 28). On the contrary, the absence of defects in embryos treated with 5-azacytidine from 24 hpf (Figure 37 A-E) could be explained by the dramatic reduction of Dnmt1 expression from this stage (Figure 28). Currently, in *Mytilus galloprovincialis* there is no available data on the timing when zygotic transcriptional activation starts. In this context, the lack of big defects in 5-azacytidine treated embryos at 7 hpf could be explained by the fact that embryo development still relies on maternal transcripts, as they probably did not enter the Zygotic Genomic Activation (ZGA) yet. Regarding 5-azacytidine, potential methylation independent cytotoxicity or off-target effects cannot be excluded. However, my experiments with Propidium Iodide and Fluorescein diacetate (Figure 36), together with the lack of defects when treatment started at 24 hpf, seem to exclude this possibility (Figure 37). However, to confirm the specificity of our results on the inhibition of DNMT1, future analyses will be performed also with alternative inhibitors (Decitabine and Flavanone) reported to be less cytotoxic (Luviano et al., 2021; Planques et al., 2021). Furthermore, my results with TET inhibitor revealed developmental defects that highlight the importance of the demethylation process as in vertebrates. Moreover, differently from the 5-azacytidine, TET inhibitor, the Bobcat 339 was able to mediate defects not only when the treatment started at 1.5 hpf but also when it started at 24 hpf (Figure 38-39-40-41). These results somehow confirm the data coming from mice and zebrafish showing that TET knockouts in these species mediate defects mainly during later developmental embryonic stages (Bogdanović et al., 2016; Dai et al., 2016) corresponding to the peak of Tet expression, that I also found during my analyses (Figure 28). Therefore, my findings on the demethylation role in *Mytilus* seems tare consistent with those reported in vertebrates, opening a new scenario for future studies on this process in mollusks. However, in *Mytilus* I also observed a very strong phenotype at earlier stages in particular when embryos were treated with higher doses of Bobcat 339

(Figure 38 E-F). Specifically, embryos showed strong defects with blastomere extroflection, and alteration in cell adhesion, migration, polarity and differentiation. Similar defects to this phenotype have been reported also in vertebrates embryos with altered function of a specific protein called Repulsive Guidance Molecule (RGMa) with a phenotype at early stages resembling those found by me (Figure 38 E-F; J) (Key and Lah, 2012). This protein selectively mediates chemorepulsion, cell adhesion and cell migration at different times and places in the embryo. Future experiments will be performed to analyze the level of expression of the *Mytilus galloprioviancialis* orthologs in control and embryos treated with TET inhibitors. Furthermore, I will investigate if potential changes in the expression of this gene could be correlated to methylation changes in its regulatory regions or gene body.

In summary, my PhD data provide the first evidence of the evolutionary conservation of DNA methylation machinery toolkit genes in *Mytilus galloprovincialis.* Furthermore, the use of different approaches allowed me to show that the *Mytilus galloprovincialis* genome is methylated and undergoes reprogramming during development. Alteration of this process similar to what happened in vertebrates results in embryos showing many developmental defects. Furthermore, the results I will obtain from the very recent Methyl-Rad sequencing will offer a unique opportunity to shed light on the effect of Emerging Contaminants on DNA methylation during *Mytilus* development. Ultimately, I also predict that the results I will obtain from these sequencing analyses will help also to understand if *Mytilus galloprovincialis,* like other mollusks, possesses a mosaic organization of DNA methylation and which areas of the genome are preferentially methylated (intragenic versus intergenic region). Finally, my characterization of developmental stages will certainly help those laboratories interested to use *Mytilus galloprovincialis* as a model for evolutionary and developmental studies (EVO-DEVO).

REFERENCES

Alam, M. N., Shapla, U. M., Shen, H., and Huang, Q. (2021). Linking emerging contaminants exposure to adverse health effects: Crosstalk between epigenome and environment. *J. Appl. Toxicol.* 41, 878–897.

Allis, C. D., and Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* 17, 487–500.

Almeida, Â., Soares, A. M. V. M., Esteves, V. I., and Freitas, R. (2021). Occurrence of the antiepileptic carbamazepine in water and bivalves from marine environments: A review. *Environ. Toxicol. Pharmacol.* 86, 103661.

Ardura, A., Zaiko, A., Morán, P., Planes, S., and Garcia-Vazquez, E. (2017). Epigenetic signatures of invasive status in populations of marine invertebrates. *Sci. Rep.* 7, 42193.

Ashapkin, V. V., Kutueva, L. I., and Vanyushin, B. F. (2016). Dnmt2 is the most evolutionary conserved and enigmatic cytosine DNA methyltransferase in eukaryotes. *Russ. J. Genet.* 52, 237–248.

Ashe, A., Colot, V., and Oldroyd, B. P. (2021). How does epigenetics influence the course of evolution? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 376, 20200111.

Balbi, T., Franzellitti, S., Fabbri, R., Montagna, M., Fabbri, E., and Canesi, L. (2016). Impact of bisphenol A (BPA) on early embryo development in the marine mussel *Mytilus galloprovincialis*: Effects on gene transcription. *Environ. Pollut.* 218, 996–1004.

Balseiro, P., Moreira, R., Chamorro, R., Figueras, A., and Novoa, B. (2013). Immune responses during the larval stages of Mytilus galloprovincialis: metamorphosis alters immunocompetence, body shape and behavior. *Fish Shellfish Immunol.* 35, 438–447.

Bantoto-Kinamot, V. (2016). Effects of induced spawning on early development in snout otter clam, Lutraria philippinarum (Deshayes, 1854) (Bivalvia: Mactridae). *Iran. J. Fish. Sci.* 15, 1230–1236.

Becker, P. B., and Workman, J. L. (2013). Nucleosome remodeling and epigenetics. *Cold Spring Harb. Perspect. Biol.* 5. doi:10.1101/cshperspect.a017905.

Bewick, A. J., Vogel, K. J., Moore, A. J., and Schmitz, R. J. (2017). Evolution of DNA Methylation across Insects. *Mol. Biol. Evol.* 34, 654–665.

Beyer, J., Green, N. W., Brooks, S., Allan, I. J., Ruus, A., Gomes, T., et al. (2017). Blue mussels (Mytilus edulis spp.) as sentinel organisms in coastal pollution monitoring: A review. *Mar. Environ. Res.* 130, 338–365.

Bheemanaik, S., Reddy, Y. V. R., and Rao, D. N. (2006). Structure, function and mechanism of exocyclic DNA methyltransferases. *Biochem. J* 399, 177–190.

Bieler, R., Carter, J. G., and Coan, E. V. (2010). Classification of bivalve families. *Malacologia* 52, 113–133.

Bieler, R., and Mikkelsen, P. M. (2006). Bivalvia - a look at the Branches. *Zool. J. Linn. Soc.* 148, 223–235.

Bieler, R., Mikkelsen, P. M., Collins, T. M., Glover, E. A., González, V. L., Graf, D. L., et al.

(2014). Investigating the Bivalve Tree of Life--an exemplar-based approach combining molecular and novel morphological characters. *Invertebr. Syst.* 28, 32–115.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev.* 16, 6–21.

Bird, A. P., and Taggart, M. H. (1980). Variable patterns of total DNA and rDNA methylation in animals. *Nucleic Acids Res.* 8, 1485–1497.

Bitter, M. C., Kapsenberg, L., Gattuso, J.-P., and Pfister, C. A. (2019). Standing genetic variation fuels rapid adaptation to ocean acidification. *Nat. Commun.* 10, 5821.

Bogan, S. N., Johnson, K. M., and Hofmann, G. E. (2020). Changes in Genome-Wide Methylation and Gene Expression in Response to Future pCO2 Extremes in the Antarctic Pteropod Limacina helicina antarctica. *Frontiers in Marine Science* 6, 788.

Bogdanović, O., Smits, A. H., de la Calle Mustienes, E., Tena, J. J., Ford, E., Williams, R., et al. (2016). Active DNA demethylation at enhancers during the vertebrate phylotypic period. *Nat. Genet.* 48, 417–426.

Bogdanović, O., and Veenstra, G. J. C. (2009). DNA methylation and methyl-CpG binding proteins: developmental requirements and function. *Chromosoma* 118, 549–565.

Bollati, V., and Baccarelli, A. (2010). Environmental epigenetics. *Heredity* 105, 105–112.

Bommarito, P. A., and Fry, R. C. (2019). "The role of DNA methylation in gene regulation," in *Toxicoepigenetics* (Elsevier), 127–151.

Bonham, V. (2017). *Mytilus galloprovincialis*. *Invasive Species Compendium (CAB, Wallingford, UK)*.

Bošković, A., and Rando, O. J. (2018). Transgenerational Epigenetic Inheritance. *Annu. Rev. Genet.* 52, 21–41.

Breton, S., Beaupré, H. D., Stewart, D. T., Hoeh, W. R., and Blier, P. U. (2007). The unusual system of doubly uniparental inheritance of mtDNA: isn't one enough? *Trends Genet.* 23, 465–474.

Chen, T., Tsujimoto, N., and Li, E. (2004). The PWWP domain of Dnmt3a and Dnmt3b is required for directing DNA methylation to the major satellite repeats at pericentric heterochromatin. *Mol. Cell. Biol.* 24, 9048–9058.

Cherblanc, F. L., Davidson, R. W. M., Di Fruscia, P., Srimongkolpithak, N., and Fuchter, M. J. (2013). Perspectives on natural product epigenetic modulators in chemical biology and medicine. *Nat. Prod. Rep.* 30, 605–624.

Chua, G. N. L., Wassarman, K. L., Sun, H., Alp, J. A., Jarczyk, E. I., Kuzio, N. J., et al. (2019). Cytosine-Based TET Enzyme Inhibitors. *ACS Med. Chem. Lett.* 10, 180–185.

Dai, H.-Q., Wang, B.-A., Yang, L., Chen, J.-J., Zhu, G.-C., Sun, M.-L., et al. (2016). TET-mediated DNA demethylation controls gastrulation by regulating Lefty–Nodal signalling. *Nature* 538, 528–532.

D'Aniello, E. (2009). Studies on RX gene function and its involvement in Ciona intestinalis ocellus differentiation. doi:10.21954/ou.ro.0001004c.

Da Ros, L., Bressan, M., and Marin, M. G. (1985). Reproductive cycle of the mussel (Mytilus

galloprovincialis Lmk) in Venice Lagoon (North Adriatic). *Boll. Zool.* 52, 223–229.

Daughton, C. G., and Ternes, T. A. (1999). Pharmaceuticals and personal care products in the environment: agents of subtle change? *Environ. Health Perspect.* 107 Suppl 6, 907–938.

Delsuc, F., Brinkmann, H., Chourrout, D., and Philippe, H. (2006). Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* 439, 965–968.

de Mendoza, A., Hatleberg, W. L., Pang, K., Leininger, S., Bogdanovic, O., Pflueger, J., et al. (2019). Convergent evolution of a vertebrate-like methylome in a marine sponge. *Nature Ecology & Evolution* 3, 1464–1473.

Desbiolles, F., Malleret, L., Tiliacos, C., Wong-Wah-Chung, P., and Laffont-Schwob, I. (2018). Occurrence and ecotoxicological assessment of pharmaceuticals: Is there a risk for the Mediterranean aquatic environment? *Sci. Total Environ.* 639, 1334–1348.

Díaz-Freije, E., Gestal, C., Castellanos-Martínez, S., and Morán, P. (2014). The role of DNA methylation on Octopus vulgaris development and their perspectives. *Front. Physiol.* 5, 62.

Dixon, G. B., Bay, L. K., and Matz, M. V. (2016). Evolutionary Consequences of DNA Methylation in a Basal Metazoan. *Mol. Biol. Evol.* 33, 2285–2293.

Downey-Wall, A. M., Cameron, L. P., Ford, B. M., McNally, E. M., Venkataraman, Y. R., Roberts, S. B., et al. (2020). Ocean Acidification Induces Subtle Shifts in Gene Expression and DNA Methylation in Mantle Tissue of the Eastern Oyster (Crassostrea virginica). *Frontiers in Marine Science* 7, 828.

Duinker, A., Håland, L., Hovgaard, P., and Mortensen, S. (2008). Gonad development and spawning in one and two year old mussels (Mytilus edulis) from Western Norway. *J. Mar. Biol. Assoc. U. K.* 88, 1465–1473.

Dunwell, T. L., McGuffin, L. J., Dunwell, J. M., and Pfeifer, G. P. (2013). The mysterious presence of a 5-methylcytosine oxidase in the Drosophila genome: possible explanations. *Cell Cycle* 12, 3357–3365.

Du, Q., Luu, P.-L., Stirzaker, C., and Clark, S. J. (2015). Methyl-CpG-binding domain proteins: readers of the epigenome. *Epigenomics* 7, 1051–1073.

Ebele, A. J., Abou-Elwafa Abdallah, M., and Harrad, S. (2017). Pharmaceuticals and personal care products (PPCPs) in the freshwater aquatic environment. *Emerging Contaminants* 3, 1–16.

Eckersley-Maslin, M. A., Alda-Catalinas, C., and Reik, W. (2018). Dynamics of the epigenetic landscape during the maternal-to-zygotic transition. *Nat. Rev. Mol. Cell Biol.* 19, 436–450.

Eggermont, M., Cornillie, P., Dierick, M., Adriaens, D., Nevejan, N., Bossier, P., et al. (2020). The blue mussel inside: 3D visualization and description of the vascular-related anatomy of Mytilus edulis to unravel hemolymph extraction. *Sci. Rep.* 10, 6773.

Ehrlich, M., Gama-Sosa, M. A., Huang, L. H., Midgett, R. M., Kuo, K. C., McCune, R. A., et al. (1982). Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res.* 10, 2709–2721.

Fabbri, R., Montagna, M., Balbi, T., Raffo, E., Palumbo, F., and Canesi, L. (2014). Adaptation of the bivalve embryotoxicity assay for the high throughput screening of emerging contaminants in *Mytilus galloprovincialis*. *Mar. Environ. Res.* 99, 1–8.

Fallet, M., Luquet, E., David, P., and Cosseau, C. (2020). Epigenetic inheritance and intergenerational effects in mollusks. *Gene* 729, 144166.

Farias, N. D., de Oliveira, N. F. P., and da Silva, P. M. (2017). Perkinsus infection is associated with alterations in the level of global DNA methylation of gills and gastrointestinal tract of the oyster Crassostrea gasar. *J. Invertebr. Pathol.* 149, 76–81.

Fatemi, M., and Wade, P. A. (2006). MBD family proteins: reading the epigenetic code. *J. Cell Sci.* 119, 3033–3037.

Fearman, J., and Moltschaniwskyj, N. A. (2010). Warmer temperatures reduce rates of gametogenesis in temperate mussels, *Mytilus galloprovincialis*. *Aquaculture* 305, 20–25.

Feng, S., Cokus, S. J., Zhang, X., Chen, P.-Y., Bostick, M., Goll, M. G., et al. (2010). Conservation and divergence of methylation patterning in plants and animals. *Proc. Natl. Acad. Sci. U. S. A.* 107, 8689–8694.

Firmino, J., Carballo, C., Armesto, P., Campinho, M. A., Power, D. M., and Manchado, M. (2017). Phylogeny, expression patterns and regulation of DNA Methyltransferases in early development of the flatfish, Solea senegalensis. *BMC Dev. Biol.* 17, 11.

Fisheries, F. A. O. Aquaculture Department. Cultured aquatic species information programme. *Mytilus galloprovincialis*.

Flores, K. B., Wolschin, F., and Amdam, G. V. (2013). The role of methylation of DNA in environmental adaptation. *Integr. Comp. Biol.* 53, 359–372.

Fournier, A., Sasai, N., Nakao, M., and Defossez, P.-A. (2012). The role of methyl-binding proteins in chromatin organization and epigenome maintenance. *Brief. Funct. Genomics* 11, 251–264.

Franzellitti, S., Balbi, T., Montagna, M., Fabbri, R., Valbonesi, P., Fabbri, E., et al. (2019). Phenotypical and molecular changes induced by carbamazepine and propranolol on larval stages of *Mytilus galloprovincialis*. *Chemosphere* 234, 962–970.

Gao, F., Liu, X., Wu, X.-P., Wang, X.-L., Gong, D., Lu, H., et al. (2012). Differential DNA methylation in discrete developmental stages of the parasitic nematode Trichinella spiralis. *Genome Biol.* 13, R100.

García-Fernández, P., García-Souto, D., Almansa, E., Morán, P., and Gestal, C. (2017). Epigenetic DNA Methylation Mediating Octopus vulgaris Early Development: Effect of Essential Fatty Acids Enriched Diet. *Front. Physiol.* 8, 292.

Gardner, J. P. A. (1992). *Mytilus Galloprovincialis* (Lmk) (Bivalvia, Mollusca): The Taxonomic Status of the Mediterranean Mussel. *Ophelia* 35, 219–243.

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

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

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

Gerdol, M., Gomez-Chiarri, M., Castillo, M. G., Figueras, A., Fiorito, G., Moreira, R., et al.

(2018). “Immunity in Molluscs: Recognition and Effector Mechanisms, with a Focus on Bivalvia,” in *Advances in Comparative Immunology*, ed. E. L. Cooper (Cham: Springer International Publishing), 225–341.

Gerdol, M., Moreira, R., Cruz, F., Gómez-Garrido, J., Vlasova, A., Rosani, U., et al. (2020). Massive gene presence-absence variation shapes an open pan-genome in the Mediterranean mussel. *Genome Biol.* 21, 275.

Geyer, K. K., Niazi, U. H., Duval, D., Cosseau, C., Tomlinson, C., Chalmers, I. W., et al. (2017). The Biomphalaria glabrata DNA methylation machinery displays spatial tissue expression, is differentially active in distinct snail populations and is modulated by interactions with Schistosoma mansoni. *PLoS Negl. Trop. Dis.* 11, e0005246.

Gogoi, A., Mazumder, P., Tyagi, V. K., Tushara Chaminda, G. G., An, A. K., and Kumar, M. (2018). Occurrence and fate of emerging contaminants in water environment: A review. *Groundwater for Sustainable Development* 6, 169–180.

Goll, M. G., Kirpekar, F., Maggert, K. A., Yoder, J. A., Hsieh, C.-L., Zhang, X., et al. (2006). Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. *Science* 311, 395–398.

Gonzalez-Romero, R., Suarez-Ulloa, V., Rodriguez-Casariego, J., Garcia-Souto, D., Diaz, G., Smith, A., et al. (2017). Effects of Florida Red Tides on histone variant expression and DNA methylation in the Eastern oyster Crassostrea virginica. *Aquat. Toxicol.* 186, 196–204.

Gosling, E. (2003). Bivalve molluscs; Biology and culture.

Gosling, E. (2015). *Marine Bivalve Molluscs*. John Wiley & Sons.

Granados, K., Poelchen, J., Novak, D., and Utikal, J. (2020). Cellular Reprogramming—A Model for Melanoma Cellular Plasticity. *Int. J. Mol. Sci.* 21, 8274.

Greally, J. M. (2018). A user’s guide to the ambiguous word “epigenetics.” *Nat. Rev. Mol. Cell Biol.* 19, 207–208.

Greenberg, M. V. C., and Bourc’his, D. (2019). The diverse roles of DNA methylation in mammalian development and disease. *Nat. Rev. Mol. Cell Biol.* 20, 590–607.

Grubb, B. J. (2006a). Developmental Biology, Eighth Edition. Scott F. Gilbert, editor. *Integrative and Comparative Biology* 46, 652–653. doi:10.1093/icb/icl011.

Grubb, B. J. (2006b). Developmental Biology, Scott F. Gilbert, editor. *Integr. Comp. Biol.* 46, 652–653.

Haggerty, C., Kretzmer, H., Riemenschneider, C., Kumar, A. S., Mattei, A. L., Bailly, N., et al. (2021). Dnmt1 has *de novo* activity targeted to transposable elements. *Nat. Struct. Mol. Biol.* 28, 594–603.

Hagood, J. S. (2014). Beyond the genome: epigenetic mechanisms in lung remodeling. *Physiology* 29, 177–185.

Haszprunar, G., and Wanninger, A. (2012). Molluscs. *Curr. Biol.* 22, R510–4.

Head, J. A. (2014). Patterns of DNA methylation in animals: an ecotoxicological perspective. *Integr. Comp. Biol.* 54, 77–86.

Hemberger, M., Dean, W., and Reik, W. (2009). Epigenetic dynamics of stem cells and cell

lineage commitment: digging Waddington's canal. *Nat. Rev. Mol. Cell Biol.* 10, 526–537.

Henderson, I. R., and Jacobsen, S. E. (2007). Epigenetic inheritance in plants. *Nature* 447, 418–424.

Ho, D. H., and Burggren, W. W. (2010). Epigenetics and transgenerational transfer: a physiological perspective. *J. Exp. Biol.* 213, 3–16.

Holliday, R. (1994). Epigenetics: an overview. *Dev. Genet.* 15, 453–457.

Holliday, R., and Pugh, J. E. (1975). DNA modification mechanisms and gene activity during development. *Science* 187, 226–232.

Houtman, C. J. (2010). Emerging contaminants in surface waters and their relevance for the production of drinking water in Europe. *Journal of Integrative Environmental Sciences* 7, 271–295.

Huntriss, J. (2021). "Epigenetic reprogramming in the embryo," in *Epigenetics and Reproductive Health* (Elsevier), 97–116.

Ittiprasert, W., Knight, M., and Tucker, M. (2015). Evaluation of cytosine DNA methylation of the Biomphalaria glabratahe at shock protein 70 locus after biological and physiological stresses. *J. Parasitol.* Available at: https://academicjournals.org/journal/JPVB/article-abstract/2B2900955743.

Jeltsch, A., and Jurkowska, R. Z. (2016). Allosteric control of mammalian DNA methyltransferases – a new regulatory paradigm. *Nucleic Acids Research* 44, 8556–8575. doi:10.1093/nar/gkw723.

Jenuwein, T., and Allis, C. D. (2001). Translating the histone code. *Science* 293, 1074–1080.

Jin, B., Li, Y., and Robertson, K. D. (2011). DNA Methylation: Superior or Subordinate in the Epigenetic Hierarchy? *Genes & Cancer* 2, 607–617. doi:10.1177/1947601910393957.

Jjingo, D., Conley, A. B., Yi, S. V., Lunyak, V. V., and Jordan, I. K. (2012). On the presence and role of human gene-body DNA methylation. *Oncotarget* 3, 462–474.

Johnson, K. M., Sirovy, K. A., Casas, S. M., La Peyre, J. F., and Kelly, M. W. (2020). Characterizing the Epigenetic and Transcriptomic Responses to Perkinsus marinus Infection in the Eastern Oyster Crassostrea virginica. *Frontiers in Marine Science* 7, 598.

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

Jones, P. A., and Liang, G. (2009). Rethinking how DNA methylation patterns are maintained. *Nat. Rev. Genet.* 10, 805–811.

Keller, T. E., Han, P., and Yi, S. V. (2015). Evolutionary Transition of Promoter and Gene Body DNA Methylation across Invertebrate–Vertebrate Boundary. *Mol. Biol. Evol.* 33, 1019–1028.

Key, B., and Lah, G. J. (2012). Repulsive guidance molecule A (RGMa): a molecule for all seasons. *Cell Adh. Migr.* 6, 85–90.

Khan, S., Naushad, M., Govarthanan, M., Iqbal, J., and Alfadul, S. M. (2022). Emerging contaminants of high concern for the environment: Current trends and future research. *Environ. Res.* 207, 112609.

Kin, K., Kakoi, S., and Wada, H. (2009). A novel role for dpp in the shaping of bivalve shells revealed in a conserved molluscan developmental program. *Dev. Biol.* 329, 152–166.

Kocot, K. M., Cannon, J. T., Todt, C., Citarella, M. R., Kohn, A. B., Meyer, A., et al. (2011). Phylogenomics reveals deep molluscan relationships. *Nature* 477, 452–456.

Kocot, K. M., Poustka, A. J., Stöger, I., Halanych, K. M., and Schrödl, M. (2020). New data from Monoplacophora and a carefully-curated dataset resolve molluscan relationships. *Sci. Rep.* 10, 101.

Kohli, R. M., and Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* 502, 472–479.

Kondo, E., Gu, Z., Horii, A., and Fukushige, S. (2005). The thymine DNA glycosylase MBD4 represses transcription and is associated with methylated p16(INK4a) and hMLH1 genes. *Mol. Cell. Biol.* 25, 4388–4396.

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

Kriaucionis, S., and Heintz, N. (2009). The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain. *Science*. doi:10.1126/science.1169786.

Kronholm, I., and Collins, S. (2016). Epigenetic mutations can both help and hinder adaptive evolution. *Mol. Ecol.* 25, 1856–1868.

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

Lawrence, M., Daujat, S., and Schneider, R. (2016). Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet.* 32, 42–56.

Leonhardt, H., Page, A. W., Weier, H. U., and Bestor, T. H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* 71, 865–873.

Li, E., Bestor, T. H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69, 915–926.

Lim, Y.-K., Cheung, K., Dang, X., Roberts, S. B., Wang, X., and Thiyagarajan, V. (2021). DNA methylation changes in response to ocean acidification at the time of larval metamorphosis in the edible oyster, Crassostrea hongkongensis. *Mar. Environ. Res.* 163, 105217.

Liu, H., Li, S., Wang, X., Zhu, J., Wei, Y., Wang, Y., et al. (2016). DNA methylation dynamics: identification and functional annotation. *Brief. Funct. Genomics* 15, 470–484.

Li, Y., Zhang, L., Li, Y., Li, W., Guo, Z., Li, R., et al. (2019). Dynamics of DNA Methylation and DNMT Expression During Gametogenesis and Early Development of Scallop Patinopecten yessoensis. *Mar. Biotechnol.* 21, 196–205.

Lotze, H. K., Lenihan, H. S., Bourque, B. J., Bradbury, R. H., Cooke, R. G., Kay, M. C., et al. (2006). Depletion, degradation, and recovery potential of estuaries and coastal seas. *Science* 312, 1806–1809.

Luviano, N., Lopez, M., Gawehns, F., Chaparro, C., Arimondo, P. B., Ivanovic, S., et al. (2021). The methylome of Biomphalaria glabrata and other mollusks: enduring modification of epigenetic landscape and phenotypic traits by a new DNA methylation inhibitor. *Epigenetics Chromatin* 14, 48.

Lyko, F. (2018). The DNA methyltransferase family: a versatile toolkit for epigenetic regulation. *Nat. Rev. Genet.* Available at: https://www.nature.com/articles/nrg.2017.80.

Lyko, F., Foret, S., Kucharski, R., Wolf, S., Falckenhayn, C., and Maleszka, R. (2010). The honey bee epigenomes: differential methylation of brain DNA in queens and workers. *PLoS Biol.* 8, e1000506.

Männer, L., Schell, T., Provataris, P., Haase, M., and Greve, C. (2021). Inference of DNA methylation patterns in molluscs. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 376, 20200166.

Medini, D., Donati, C., Tettelin, H., Masignani, V., and Rappuoli, R. (2005). The microbial pan-genome. *Curr. Opin. Genet. Dev.* 15, 589–594.

Meehan, R. R., Lewis, J. D., McKay, S., Kleiner, E. L., and Bird, A. P. (1989). Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell* 58, 499–507.

Messerschmidt, D. M., Knowles, B. B., and Solter, D. (2014). DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes Dev.* 28, 812–828.

Metzger, D. C. H., and Schulte, P. M. (2016). Epigenomics in marine fishes. *Mar. Genomics* 30, 43–54.

Mezzelani, M., Fattorini, D., Gorbi, S., Nigro, M., and Regoli, F. (2020). Human pharmaceuticals in marine mussels: Evidence of sneaky environmental hazard along Italian coasts. *Mar. Environ. Res.* 162, 105137.

Mezzelani, M., Gorbi, S., and Regoli, F. (2018). Pharmaceuticals in the aquatic environments: evidence of emerged threat and future challenges for marine organisms. *Mar. Environ. Res.* Available at: https://www.sciencedirect.com/science/article/pii/S0141113618300722.

Mezzelani, M., Nardi, A., Bernardini, I., Milan, M., Peruzza, L., d'Errico, G., et al. (2021). Environmental pharmaceuticals and climate change: The case study of carbamazepine in *M. galloprovincialis* under ocean acidification scenario. *Environ. Int.* 146, 106269.

Miglioli, A., Balbi, T., Besnardeau, L., Dumollard, R., and Canesi, L. (2021). Bisphenol A interferes with first shell formation and development of the serotoninergic system in early larval stages of *Mytilus galloprovincialis*. *Sci. Total Environ.* 758, 144003.

Mompelat, S., Le Bot, B., and Thomas, O. (2009). Occurrence and fate of pharmaceutical products and by-products, from resource to drinking water. *Environ. Int.* 35, 803–814.

Moore, L. D., Le, T., and Fan, G. (2013). DNA Methylation and Its Basic Function. *Neuropsychopharmacology* 38, 23.

Müller, R., Charaf, S., Scherer, C., Oppold, A., Oehlmann, J., and Wagner, M. (2016). Phenotypic and epigenetic effects of vinclozolin in the gastropod Physella acuta. *J. Molluscan Stud.* 82, 320–327.

Murgarella, M., Puiu, D., Novoa, B., Figueras, A., Posada, D., and Canchaya, C. (2016). Correction: A First Insight into the Genome of the Filter-Feeder Mussel *Mytilus galloprovincialis*. *PLoS One* 11, e0160081.

Nai, Y.-S., Huang, Y.-C., Yen, M.-R., and Chen, P.-Y. (2020). Diversity of Fungal DNA Methyltransferases and Their Association With DNA Methylation Patterns. *Front. Microbiol.* 11,

616922.

Nakajima, H., and Kunimoto, H. (2014). TET2 as an epigenetic master regulator for normal and malignant hematopoiesis. *Cancer Sci.* 105, 1093–1099.

Nanty, L., Carbajosa, G., Heap, G. A., Ratnieks, F., van Heel, D. A., Down, T. A., et al. (2011). Comparative methylomics reveals gene-body H3K36me3 in Drosophila predicts DNA methylation and CpG landscapes in other invertebrates. *Genome Res.* 21, 1841–1850.

Nan, X., Meehan, R. R., and Bird, A. (1993). Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Res.* 21, 4886–4892.

Nederbragt, A. J., van Loon, A. E., and Dictus, W. J. A. G. (2002). Expression of Patella vulgata orthologs of engrailed and dpp-BMP2/4 in adjacent domains during molluscan shell development suggests a conserved compartment boundary mechanism. *Dev. Biol.* 246, 341–355.

Nica, D., Popescu, C., Draghici, G., Privistirescu, I., Suciu, M., and Stöger, R. (2017). Effect of cadmium on cytosine hydroxymethylation in gastropod hepatopancreas. *Environ. Sci. Pollut. Res. Int.* 24, 15187–15195.

Nicholson, T. B., Veland, N., and Chen, T. (2015). Writers, Readers, and Erasers of Epigenetic Marks. *Epigenetic Cancer Therapy*, 31–66. doi:10.1016/b978-0-12-800206-3.00003-3.

Okano, M., Bell, D. W., Haber, D. A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell* 99, 247–257.

Olson, C. E., and Roberts, S. B. (2014). Genome-wide profiling of DNA methylation and gene expression in Crassostrea gigas male gametes. *Front. Physiol.* 5, 224.

Pavlicek, A., Schwaha, T., and Wanninger, A. (2018). Towards a ground pattern reconstruction of bivalve nervous systems: neurogenesis in the zebra mussel Dreissena polymorpha. *Org. Divers. Evol.* 18, 101–114.

Peaston, A. E., and Whitelaw, E. (2006). Epigenetics and phenotypic variation in mammals. *Mamm. Genome* 17, 365–374.

Pechenik, J. A. (2009). *Biology of the invertebrates*. 6th ed. Maidenhead, England: McGraw Hill Higher Education.

Planques, A., Kerner, P., Ferry, L., Grunau, C., Gazave, E., and Vervoort, M. (2021). DNA methylation atlas and machinery in the developing and regenerating annelid Platynereis dumerilii. *BMC Biol.* 19, 148.

Pradhan, M., Estève, P.-O., Chin, H. G., Samaranayke, M., Kim, G.-D., and Pradhan, S. (2008). CXXC domain of human DNMT1 is essential for enzymatic activity. *Biochemistry* 47, 10000–10009.

Rasmussen, K. D., and Helin, K. (2016). Role of TET enzymes in DNA methylation, development, and cancer. *Genes Dev.* 30, 733–750.

Razin, A. (1989). DNA Methylases. *Genetic Engineering*, 1–11. doi:10.1007/978-1-4615-7084-4_1.

Ren, W., Gao, L., and Song, J. (2018). Structural Basis of DNMT1 and DNMT3A-Mediated DNA Methylation. *Genes* 9. doi:10.3390/genes9120620.

Resgalla, C. (2016). Spawning and multiple end points of the embryo-larval bioassay of the blue mussel *Mytilus galloprovincialis* (Lmk). *Ecotoxicology* 25, 1609–1616.

Rinn, J. L., and Chang, H. Y. (2012). Genome Regulation by Long Noncoding RNAs. *Annu. Rev. Biochem.* 81, 145–166.

Riviere, G., He, Y., Tecchio, S., Crowell, E., Gras, M., Sourdaine, P., et al. (2017). Dynamics of DNA methylomes underlie oyster development. *PLoS Genet.* 13, e1006807.

Riviere, G., Wu, G.-C., Fellous, A., Goux, D., Sourdaine, P., and Favrel, P. (2013). DNA methylation is crucial for the early development in the Oyster C. gigas. *Mar. Biotechnol.* 15, 739–753.

Roberts, S. B., and Gavery, M. R. (2012). Is There a Relationship between DNA Methylation and Phenotypic Plasticity in Invertebrates? *Frontiers in Physiology* 2. doi:10.3389/fphys.2011.00116.

Rodil, R., Villaverde-de-Sáa, E., Cobas, J., Quintana, J. B., Cela, R., and Carro, N. (2019). Legacy and emerging pollutants in marine bivalves from the Galician coast (NW Spain). *Environ. Int.* 129, 364–375.

Rondon, R., Grunau, C., Fallet, M., Charlemagne, N., Sussarellu, R., Chaparro, C., et al. (2017a). Effects of a parental exposure to diuron on Pacific oyster spat methylome. *Environmental Epigenetics* 3. doi:10.1093/eep/dvx004.

Rondon, R., Grunau, C., Fallet, M., Charlemagne, N., Sussarellu, R., Chaparro, C., et al. (2017b). Effects of a parental exposure to diuron on Pacific oyster spat methylome. 3. *Article dvx004*.

Rountree, M. R., Bachman, K. E., and Baylin, S. B. (2000). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nat. Genet.* 25, 269–277.

Ruiz, M., Tarifeño, E., and Llanos-Rivera, A. (2008). Efecto de la temperatura en el desarrollo embrionario y larval del mejillón, *Mytilus galloprovincialis* (Lamarck, 1819). *Revista de biología*. Available at: https://scielo.conicyt.cl/scielo.php?pid=S0718-19572008000100006&script=sci_arttext.

Ruppert, E. E., Barnes, R. D., and Fox, R. S. (2004). Invertebrate zoology :a functional evolutionary approach /. Available at: http://www.sidalc.net/cgi-bin/wxis.exe/?IsisScript=FVL.xis&method=post&formato=2&cantidad=1&expresion=mfn=008317 [Accessed February 8, 2022].

Rusk, A. B. (2012). Larval Development of the New Zealand Mussel Perna Canaliculus and Effects of Cryopreservation: A Thesis Submitted to Auckland University of Technology in Partial Fulfilment of the Requirements for the Degree of Master of Applied Science (MAppSc), 2012.

Salamanca-Díaz, D. A., Calcino, A. D., de Oliveira, A. L., and Wanninger, A. (2021). Non-collinear Hox gene expression in bivalves and the evolution of morphological novelties in mollusks. *Sci. Rep.* 11, 3575.

Sánchez-Argüello, P., Aparicio, N., Guevara, M., Díaz, L., Cervera, M., and Fernandez, C. (2016). Effects on reproduction, genotoxicity and DNA methylation pattern after chronic exposure of the freshwater snail Physa acuta (Gastropoda, Pulmonata) to vinclozolin. *Austin J EnvironToxicol* 2, 1008.

Sarda, S., Zeng, J., Hunt, B. G., and Yi, S. V. (2012). The evolution of invertebrate gene body methylation. *Mol. Biol. Evol.* 29, 1907–1916.

Schübeler, D. (2015). Function and information content of DNA methylation. *Nature* 517, 321–326.

Seong, H. J., Han, S.-W., and Sul, W. J. (2021). Prokaryotic DNA methylation and its functional roles. *J. Microbiol.* 59, 242–248.

Sharp, D. M. C., Picken, A., Morris, T. J., Hewitt, C. J., Coopman, K., and Slater, N. K. H. (2013). Amphipathic polymer-mediated uptake of trehalose for dimethyl sulfoxide-free human cell cryopreservation. *Cryobiology* 67, 305–311.

Shayevitch, R., Askayo, D., Keydar, I., and Ast, G. (2018). The importance of DNA methylation of exons on alternative splicing. *RNA* 24, 1351–1362.

Sigwart, J. D., and Sutton, M. D. (2007). Deep molluscan phylogeny: synthesis of palaeontological and neontological data. *Proc. Biol. Sci.* 274, 2413–2419.

Skibinski, D. O., Gallagher, C., and Beynon, C. M. (1994). Sex-limited mitochondrial DNA transmission in the marine mussel Mytilus edulis. *Genetics* 138, 801–809.

Skvortsova, K., Bertrand, S., Voronov, D., Duckett, P. E., Ross, S. E., Magri, M. S., et al. (2021). Active DNA demethylation of developmental cis-regulatory regions predates vertebrate origins. Available at: https://europepmc.org/article/ppr/ppr416609.

Skvortsova, K., Iovino, N., and Bogdanović, O. (2018). Functions and mechanisms of epigenetic inheritance in animals. *Nat. Rev. Mol. Cell Biol.* 19, 774–790.

Smith, M. M. (1991). Histone structure and function. *Curr. Opin. Cell Biol.* 3, 429–437.

Smith, S. A., Wilson, N. G., Goetz, F. E., Feehery, C., Andrade, S. C. S., Rouse, G. W., et al. (2011). Resolving the evolutionary relationships of molluscs with phylogenomic tools. *Nature* 480, 364–367.

Šrut, M. (2021). Ecotoxicological epigenetics in invertebrates: Emerging tool for the evaluation of present and past pollution burden. *Chemosphere* 282, 131026.

Suarez-Ulloa, V., Gonzalez-Romero, R., and Eirin-Lopez, J. M. (2015). Environmental epigenetics: A promising venue for developing next-generation pollution biomonitoring tools in marine invertebrates. *Mar. Pollut. Bull.* 98, 5–13.

Suarez-Ulloa, V., Rivera-Casas, C., Michel, M., and Eirin-Lopez, J. M. (2019). Seasonal DNA Methylation Variation in the Flat Tree Oyster Isognomon Alatus from a Mangrove Ecosystem in North Biscayne Bay, Florida. *shre* 38, 79–88.

Suetake, I., Shinozaki, F., Miyagawa, J., Takeshima, H., and Tajima, S. (2004). DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. *J. Biol. Chem.* 279, 27816–27823.

Sussarellu, R., Lebreton, M., Rouxel, J., Akcha, F., and Rivière, G. (2018). Copper induces expression and methylation changes of early development genes in Crassostrea gigas embryos. *Aquat. Toxicol.* 196, 70–78.

Suzuki, M. M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* 9, 465–476.

Suzuki, M. M., Kerr, A. R. W., De Sousa, D., and Bird, A. (2007). CpG methylation is targeted to transcription units in an invertebrate genome. *Genome Res.* 17, 625–631.

Suzuki, M., and Nagasawa, H. (2013). Mollusk shell structures and their formation mechanism. *Can. J. Zool.* 91, 349–366.

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324, 930–935.

Tan, L., and Shi, Y. G. (2012). Tet family proteins and 5-hydroxymethylcytosine in development and disease. *Development* 139, 1895–1902.

Thorson, J. L. M., Smithson, M., Beck, D., Sadler-Riggleman, I., Nilsson, E., Dybdahl, M., et al. (2017). Epigenetics and adaptive phenotypic variation between habitats in an asexual snail. *Sci. Rep.* 7, 14139.

Torda, G., Donelson, J. M., Aranda, M., Barshis, D. J., Bay, L., Berumen, M. L., et al. (2017). Rapid adaptive responses to climate change in corals. *Nat. Clim. Chang.* 7, 627–636.

University of Arizona. Water Resources Research Center, Raghav, M., Eden, S., Mitchell, K., and Witte, B. (2013). Arroyo 2013. Available at: https://repository.arizona.edu/handle/10150/325905 [Accessed May 2, 2022].

Uysal, F., Ozturk, S., and Akkoyunlu, G. (2017). DNMT1, DNMT3A and DNMT3B proteins are differently expressed in mouse oocytes and early embryos. *J. Mol. Histol.* 48, 417–426.

Vendrami, D. L. J., De Noia, M., Telesca, L., Brodte, E.-M., and Hoffman, J. I. (2020). Genome-wide insights into introgression and its consequences for genome-wide heterozygosity in the Mytilus species complex across Europe. *Evol. Appl.* 13, 2130–2142.

Venkataraman, Y. R., Downey-Wall, A. M., Ries, J., Westfield, I., White, S. J., Roberts, S. B., et al. (2020). General DNA Methylation Patterns and Environmentally-Induced Differential Methylation in the Eastern Oyster (Crassostrea virginica). *Frontiers in Marine Science* 7, 225.

Waddington, C. H. (2011). The Epigenotype. *Int. J. Epidemiol.* 41, 10–13.

Waddington, C. H. (2014). *The Strategy of the Genes*. Routledge.

Wang, S., Lv, J., Zhang, L., Dou, J., Sun, Y., Li, X., et al. (2015). MethylRAD: a simple and scalable method for genome-wide DNA methylation profiling using methylation-dependent restriction enzymes. *Open Biol.* 5. doi:10.1098/rsob.150130.

Wang, X., Li, Q., Lian, J., Li, L., Jin, L., Cai, H., et al. (2014). Genome-wide and single-base resolution DNA methylomes of the Pacific oyster Crassostrea gigas provide insight into the evolution of invertebrate CpG methylation. *BMC Genomics* 15, 1119.

Wanninger, A. (2015). Mollusca: Bivalvia. *Structure and Evolution of Invertebrate Nervous Systems*, 190–195. doi:10.1093/acprof:oso/9780199682201.003.0019.

Wanninger, A., and Wollesen, T. (2015). Mollusca. *Evolutionary Developmental Biology of Invertebrates 2*, 103–153. doi:10.1007/978-3-7091-1871-9_7.

Wanninger, A., and Wollesen, T. (2018). The evolution of molluscs. *Biol. Rev. Camb. Philos. Soc.* doi:10.1111/brv.12439.

Wilkinson, J. L., Hooda, P. S., Barker, J., Barton, S., and Swinden, J. (2016). Ecotoxic pharmaceuticals, personal care products, and other emerging contaminants: A review of environmental, receptor-mediated, developmental, and epigenetic toxicity with discussion of

proposed toxicity to humans. *Crit. Rev. Environ. Sci. Technol.* 46, 336–381.

Willbanks, A., Leary, M., Greenshields, M., Tyminski, C., Heerboth, S., Lapinska, K., et al. (2016). The Evolution of Epigenetics: From Prokaryotes to Humans and Its Biological Consequences. *Genet. Epigenet.* 8, 25–36.

Wong, N. L., and Arshad, A. (2013). Induced spawning and early development of Modiolus philippinarum (Hanley 1843)(Bivalvia: Mytilidae). *Asian J. Anim. Vet. Adv.* 8, 100–107.

Wu Ct, and Morris, J. R. (2001). Genes, genetics, and epigenetics: a correspondence. *Science* 293, 1103–1105.

Wu, H., and Zhang, Y. (2011). Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes Dev.* 25, 2436–2452.

Wu, H., and Zhang, Y. (2014). Reversing DNA methylation: mechanisms, genomics, and biological functions. *Cell* 156, 45–68.

Wu, L., Hiebert, L. S., Klann, M., Passamaneck, Y., Bastin, B. R., Schneider, S. Q., et al. (2020). Genes with spiralian-specific protein motifs are expressed in spiralian ciliary bands. *Nat. Commun.* 11, 4171.

Wu, X., and Zhang, Y. (2017). TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat. Rev. Genet.* 18, 517–534.

Xiang, H., Zhu, J., Chen, Q., Dai, F., Li, X., Li, M., et al. (2010). Single base–resolution methylome of the silkworm reveals a sparse epigenomic map. *Nat. Biotechnol.* 28, 516–520.

Xu, X., Li, G., Li, C., Zhang, J., Wang, Q., Simmons, D. K., et al. (2019). Evolutionary transition between invertebrates and vertebrates via methylation reprogramming in embryogenesis. *Natl Sci Rev* 6, 993–1003.

Yuan, C., Mao, J., Sun, H., Wang, Y., Guo, M., Wang, X., et al. (2021). Genome-wide DNA methylation profile changes associated with shell colouration in the Yesso scallop (Patinopecten yessoensis) as measured by whole-genome bisulfite sequencing. *BMC Genomics* 22, 740.

Zemach, A., McDaniel, I. E., Silva, P., and Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* 328, 916–919.

Zhang, H., Tan, K., Li, S., Ma, H., and Zheng, H. (2020). DNA methylation in molluscs growth and development: An overview. *Aquac. Res.* doi:10.1111/are.14966.

Zhang, Y., Jurkowska, R., Soeroes, S., Rajavelu, A., Dhayalan, A., Bock, I., et al. (2010). Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res.* 38, 4246–4253.

PUBLICATIONS

"FISH for All: A Fast and Efficient Fluorescent In situ Hybridization (FISH) Protocol for Marine Embryos and Larvae"
P Paganos, F Caccavale, **C La Vecchia**, E D'Aniello, S D'Aniello, and MI Arnone.
Journal: Frontiers in Physiology:
doi:10.3389/fphys.2022.878062

"Evolutionary history of DNA methylation related genes in Bivalvia. New insight from *Mytilus galloprovincialis*"
M Gerdol, **C La Vecchia**, M Strazzullo, P De Luca, S Gorbi, F Regoli, A Pallavicini, E D'Aniello.
Journal: Frontiers in Ecology and Evolution:
doi.org/10.3389/fevo.2021.698561

"The natural organic matter, a controversial element for metal speciation and toxicity on marine organisms"
L Pontoni*, **C La Vecchia***, P Boguta*, M Sirakov, E D'Aniello, M Fabbricino, A Locascio.
Journal: Environmental Chemistry Letters:
doi.org/10.1007/s10311-021-01310-y

"Transphyletic conservation of Nitric Oxide synthase regulation in cephalochordates and tunicates"
F Caccavale, U Coppola, QA Vassalli, **C La Vecchia**, A Palumbo, E D'Aniello, A Locascio, F Ristoratore, S D'Aniello.
Journal: Development Genes and Evolution, MDPI:
doi.org/10.1007/s00427-020-00668-3