Universidade de Aveiro
2021

Bruno Miguel
Pacheco Pereira

Meta-análise de metilação de DNA em invertebrados: tendências evolutivas e indicações para avaliação ambiental

Meta-analysis of DNA methylation in invertebrates: evolutionary trends and indications for environmental assessment

Universidade de Aveiro
2021

Bruno Miguel Pacheco Pereira

Meta-análise de metilação de DNA em invertebrados: tendências evolutivas e indicações para avaliação ambiental

Meta-analysis of DNA methylation in invertebrates: evolutionary trends and indications for environmental assessment

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Biologia Aplicada, realizada sob a orientação científica da Doutora Joana Luísa Pereira, Investigadora do Departamento de Biologia e do Centro de Estudos do Ambiente e do Mar da Universidade de Aveiro e da Doutora Jana Asselman, Professora do Departamento de Ciências Animais e do Blue Growth Research Lab da Universidade de Ghent.

Dedico este trabalho à minha família e amigos pelo seu constante apoio.

o júri

presidente	Prof. Doutora Maria Adelaide de Pinho Almeida professora catedrática da Universidade de Aveiro
arguente	Doutora Joana Isabel do Vale Lourenço investigadora auxiliar da Universidade de Aveiro
orientadora	Doutora Joana Luísa Lourenço Estevinho Pereira investigadora auxiliar da Universidade de Aveiro

agradecimentos

Eu em primeiro lugar gostaria de agradecer à Doutora Joana Luísa Pereira, que durante todo este período tumultuoso, esteve sempre disponível a ajudar. Dando a sua assistência sempre que necessário, e guiando o meu trabalho na direção correta, permitindo sempre que eu aprendesse e crescesse com os meus erros

À Doutora Jana Asselman que igualmente esteve sempre disposta a ajudar-me e a orientar o meu trabalho de maneira a elevar a qualidade científica deste mesmo.

Em particular, gostaria de agradecer o aluno de Doutoramento Guilherme de Abreu Jeremias que mais que ninguém seguiu o meu trabalho de perto. Incentivando-me sempre a fazer mais e melhor e apoiando-me sempre que necessário.

Ao todos os membros do grupo “Aquatic Toxicology & Risk Assessment”, que apesar de apenas lá estar pouco tempo me ajudaram sempre que necessitei e pelo bom ambiente no qual fui recebido.

Ao meu amigo, João que sempre me apoiou e onde pude encontrar alguém de confiança e igualmente “cultured”.

Aos meus pais, a minha mãe que sempre me incentivou e encorajou a perseguir os meus sonhos, ao meu pai que sempre me apoiou por todas a dificuldades nesse caminho e especialmente a minha avó Amelia Norberta que toda a minha vida me apoiou, qualquer que fosse a situação ou dificuldade.

A todos que nestes últimos anos estiveram na minha vida e a fizeram melhor pela sua presença.

Obrigado.

palavras-chave

Regiões genéticas codificantes, Evolução, Adaptação, Epigenética, Meta-analise, CpG O/E.

resumo

O impacto humano no ambiente e nos ecossistemas tem aumentado, especialmente nas décadas mais recentes. Esta pressão antropogénica é traduzida em fenómenos que são globais, como fragmentação de habitat, extinção em massa ou declínio da biodiversidade, que ameaçam os serviços dos ecossistemas. A mitigação destes impactos negativos em diferentes ecossistemas é largamente dependente da nossa capacidade para avaliar os efeitos das pressões nos sistemas biológicos com resolução suficiente, permitindo estabelecer relações fidedignas de causa-efeito.

Ferramentas de elevada resolução para avaliar estes efeitos têm emergido recentemente, por exemplo, as baseadas na expressão génica e no epigenoma. Este é fortemente influenciado pelas condições ambientais, com consequências na expressão génica. Assim, o epigenoma constitui uma nova oportunidade para a compreensão mecanicista dos efeitos tóxicos dos contaminantes ambientais e para o desenvolvimento de biomarcadores de exposição e efeito. Neste contexto, os objetivos específicos da presente Dissertação foram: (i) recolher dados genómicos relativos a várias espécies de invertebrados representantes de diferentes compartimentos ambientais; (ii) quantificar a distribuição do potencial para metilação de regiões genéticas codificantes nos genomas destes organismos, com base na prevalência de dinucleótidos citosina-guanina (CpG); (iii) elaborar sobre a capacidade diferencial das espécies de invertebrados para responderem ou potencialmente se adaptarem a pressões ou flutuações ambientais, com base na plasticidade fenotípica mediada pela metilação do DNA e/ou mutações relacionadas; (iv) concluir sobre as espécies-modelo mais favoráveis para abordar a metilação de regiões genéticas codificantes num contexto ecotoxicológico.

De entre o conjunto de 27 espécies selecionadas para meta-análise com dados genómicos suficientes disponíveis, 11 apresentaram distribuições do rácio entre frequência de CpG observada e esperada (CpG O/E) que sugerem a existência de grupos de genes muito suscetíveis a metilação, e 16 apresentaram distribuições que sugerem que todo o genoma é pouco ou muito pouco suscetível a metilação. Padrões de redução de suscetibilidade à metilação do ADN foram identificados entre espécies filogeneticamente próximas, que apoiam a teoria de que a metilação do DNA poderá ter tido um papel relevante na adaptação das espécies a novos nichos ecológicos, p.ex. durante a colonização de ambientes terrestres ou dulçaquícolas a partir de ambientes primitivos marinhos. A concentração de dinucleótidos CG nos genomas analisados permitiu ainda discutir sobre a adequação de diferentes organismos-modelo usados em ecotoxicologia para o desenvolvimento de biomarcadores baseados na metilação do ADN. Neste contexto, destacam-se as seguintes espécies, como modelos com uma maior quantidade relativa de dinucleótidos CG: *N. vectensis* e *T.* californic*us* no ambiente marinho; *D. magna* e *D. polymorpha* em ecossistemas dulçaquícolas; *I. scapularis* e *H. saltator* no compartimento solo; *A. albimanus*, *A. albopictus* e *N. vitripennis* para representar espécies aéreas. Complementarmente a esta evidência, a escolha de organismos-modelo neste contexto deve ser guiada pela informação existente sobre as suas respostas ecotoxicológicas ao nível fenotípico.

keywords

Gene body, evolution, adaptation, epigenetics, meta-analysis, CpG O / E.

abstract

The anthropogenic impact in a variety of ecosystems has been increasing significantly in recent decades. This anthropogenic pressure translates into globally spreading phenomena like habitat fragmentation, mass extinction or biodiversity decline, that threaten ecosystem services. The mitigation of such detrimental ecological impacts is largely dependent on our ability to properly assess effects of the pressures on biological systems, i.e., with sufficient resolution levels that allows establishing reliable cause-effect relationships.

High-resolution tools to assess these effects have recently been emerging, e.g., tools based in gene expression and the epigenome. The latter is strongly influenced by environmental conditions, with consequences in gene expression. Thus, the epigenome constitutes a new opportunity for the mechanistic understanding of toxic effects of environmental contaminants and for the development of biomarkers of exposure and effect. In this context, the specific objectives of the present dissertation were as follows: (i) to collect reliable genomic data of various invertebrate species representing different environmental compartments; (ii) quantify the distribution of the potential for gene body methylation based on the prevalence of CpG dinucleotides in the genome of each species; (iii) elaborate on the differential ability of invertebrate species to respond or potentially adapt to challenging and/or highly fluctuating environments based on phenotypic plasticity mediated by DNA methylation and/or related mutations; (iv) conclude on the most favourable models to address gene body methylation within an ecotoxicological context.

Among the set of 27 species with reliable genomic data publicly available selected for meta-analysis, 11 recorded an observed-to-expected CpG ratio distribution (CpG O/E) that suggest the existence of gene groups with high susceptibility to methylation, and 16 have distributions suggesting that the entire genome is little or very little susceptible to methylation. Patterns of reduced susceptibility to DNA methylation have been identified among phylogenetically closely related species, which supports the theory that DNA methylation may have played a relevant role in species adaptation to new ecological niches, e.g., during the colonization of terrestrial or freshwater environments from primitive marine environments. The concentration of CG dinucleotides in the analysed genomes also allowed to discuss the suitability of different model organisms used in ecotoxicology for the development of biomarkers of exposure or effect based on DNA methylation. In this context, the following species are noteworthy as those presenting higher relative levels of CG dinucleotides in their genome: *N. vectensis* and *T. californicus* in the marine environment; *D. magna* and *D. polymorpha* in freshwater ecosystems; *I. scapularis* and *H. saltator* in the soil compartment; *A. albimanus*, *A. albopictus* and *N. vitripennis* representing aerial species.

INDEX

1. INTRODUCTION 2
1.1. Epigenetics 8
1.1.1. Epigenetic mechanisms 9
1.2. Epigenetic responses to environmental stressors 21
1.2.1. Transgenerational Epigenetic Inheritance 29
1.2.2. Methylation in Vertebrates vs Invertebrates 30
1.3. Assessment of epigenetic modifications – most common methods 31
1.4. Structure and objectives of the Dissertation 34
2. Material and methods 36
2.1. Species selection and justification 36
2.1.1. Validation species 40
2.1.2. Non-validation terrestrial species 42
2.1.3. Non- validation aerial species 46
2.1.4. Non-validation freshwater species 49
2.1.5. Non-validation marine species 51
2.2. Data Processing 54
2.2.1. Preparation of files sourcing bioinformatics analysis 54
2.2.2. Bioinformatic analysis 57
2.2.3. Data analysis 58
3. Results and Discussion 63
3.1. Likelihood of gene body methylation of invertebrate genomes 64
3.1.1. Validation Species 67
3.1.2. Non-Validation Species 69
3.2. Differences in gene body methylation in closely related species 78
3.3. Species appropriateness for the development of epigenetic biomarkers supporting environmental assessment 82
4. Conclusions and future perspectives 86
References 88

1. INTRODUCTION

Humans have always had an impact on the environment and on the ecosystems they interact with but in more recent decades this impact has increased, and the consequent detrimental effects have been reasoned for long (Dudgeon et al. 2006a, b; DURANCE and ORMEROD 2009; Fenoglio et al. 2010; Pereira et al. 2010; Carpenter et al. 2011a; Madin et al. 2016). Such anthropogenic pressure ultimately translates into globally spreading phenomena like habitat fragmentation, mass extinction, biodiversity decline, spread of invasive species, and overall depletion of ecosystem services, these being of particular importance to humanity (May 1988; Ehrenfeld 2010; Dirzo et al. 2014; Wilson et al. 2016; Otto 2018).

The benefits that ecosystem services provide are reflected in the many and varied benefits humans take advantage off as provided by the natural environment and its ecosystems, e.g., agroecosystems, forest ecosystems, grassland ecosystems and aquatic ecosystems. Examples of some ecosystem's services are the pollination of crops, weather mitigation, clean air and human well-being in both mental and physical aspects. These services are also often integral for many processes on which we are reliant upon, like the provisioning of clean drinking water, waste decomposition and the overall resilience and productivity of food supplying ecosystems. Ecosystem services can be categorised into 4 different types: provisioning, like the production of food and water; regulating, such as the control of the climate and disease; supporting, for instance nutrient cycles and oxygen production; and cultural, for example spiritual and recreational benefits (Millennium Ecosystem Assessment 2005).

A majority of ecosystem services are either being degraded or used in an unsustainable way including the provisioning of food, the regulation of clean air and water and also the regulation of regional and local climate. Because of efforts done to boost the availability of given services, such as food provisioning, many ecosystem functions have been degraded. These costs are often shifted between groups of people: taking the above example, the people benefiting from the increase supply of food may not be the same people that face the cost of that said increase (Millennium Ecosystem Assessment 2005; Field et al. 2014).

Over the past several decades, humans have altered ecosystems to a degree that cannot be compared to any other period of time in human history. This alteration has been usually in the pursuit of satisfying our ever-increasing demand for various resources like for e.g., food, fresh water and fuel. The amount of land used for cultivation (e.g., there are areas where at least 30% of the landscape is used for crops), shifting cultivation, livestock production or freshwater aquaculture, has increased so much that it accounts for one quarter of Earth's terrestrial surface (Millennium Ecosystem Assessment 2005; Dubois 2011). In the latter few decades of the twentieth century, around 20% of the world's coral reefs were lost, and another 20% were degraded; also, 35% of mangrove land was lost during this time (Cesar et al. 2003; Goldberg et al. 2020) The quantity of water that is being sequestered by dams quadrupled since 1960 and water held in reservoirs is three to six times more than that held in rivers. The amount of water being extracted from rivers and lakes has doubled since 1960 and most of it is used (70% worldwide) for agriculture. Since the beginning of the Industrial Age the concentration of carbon dioxide in the atmosphere has increased by about 47% (from about 280 to 412 parts per million in 2019), mainly due to utilization of fossil fuels and land use changes (Alan Buis 2019).

Across a wide variety of taxonomic groups, the majority of these changes reflect a loss of biodiversity. The majority of species' population sizes, ranges, or both are currently dropping, and the distribution of species on Earth is becoming increasingly homogeneous (Millennium Ecosystem Assessment 2005; Pecl et al. 2017). Indeed, the number of species on the planet is declining, over the past few centuries, the rate of species extinction has increased by as much as 1,000 times above the natural background rates, and this increase can be attributed to humans (Millennium Ecosystem Assessment 2005; de Vos et al. 2015). Some 10–30% of mammals, birds and amphibian species are currently threatened with extinction. Freshwater ecosystems tend to have the highest proportion of species threatened with extinction and genetic diversity has declined globally. These phenomena can either be a direct outcome of human activity or an indirect effect, for example when linked to climate change.

Among ecosystem provisioning services, freshwater supply is remarkably concerning, not only due to the fact that humans need it to survive but because of its noteworthy, limited

supply. Only 2.5% of the volume of all water on Earth is freshwater, and then 68.7% of it is permanently frozen and 29.9% is ground water; only 0.26% of freshwater is available in lakes, rivers and reservoirs (Carpenter et al. 2011b; Reid et al. 2019). Freshwater is utilized in various human activities besides direct consumption, such as agriculture, various industrial processes, and energy production. The problem of freshwater scarcity is an ever-increasing problem for several reasons anchored first in the physical scarcity of water available due to the increased frequency and severity of drought events and the overutilization of aquatic resources, and second in the scarcity of water of good quality resulting from the pollution of aquatic resources (Reid et al. 2019; Santos et al. 2021). Nowadays, there are millions of people who do not have an easy access to freshwater and this scarcity will only increase with the expected increase in human population, as well as with the consequent need for an increase in energy, industrial and agricultural production. This will raise the pressure in an already stressed system (Carpenter et al. 2011b; Reid et al. 2019). Remarkable direct anthropogenic pressures over freshwater ecosystems nowadays are land-use, human alteration of water flows, and chemical inputs that often come from agricultural and urban areas, as well as the discharge of contaminated effluents that can be produced in e.g., mining and industrial activities that have been leading to an increase in the quantity of exogenous chemicals such as organic compounds, acids and heavy metals alkalis released and spread throughout freshwater ecosystems (Carpenter et al. 2011b; Reid et al. 2019; Santos et al. 2021).

Marine ecosystems provide various services that are of importance to humanity, and among them are: the regulation of the global climate, contribution to the water cycle, maintenance of biodiversity and perhaps the one that affects humanity in a more direct way, i.e., the provision of food through fisheries. Fish contributed around 17% of total animal protein and 7% of all proteins consumed globally in 2017, making food provision particularly significant; fish provides 20% of the average per capita intake of animal protein for about 3.3 billion people and it provides 50% or more of total animal protein intake in some countries (Food and Agriculture Organization of the United Nations 2020). These ecosystems are currently being pressured by a variety of different factors, from climate change that induces the acidification of ocean due to carbon sequestering or rising sea

temperatures, which accelerates the acidification of the oceans and impacts temperature-sensitive ecosystems like coral reefs, to factors more directly related to human actions like overfishing, pollution with metals, toxins, plastics and more recently micro- and nano-plastics. These more direct impacts are related to the human proximity to the coast, with nearly 40% of the people in the world living within 100 km of the coast (Joyner and Frew 1991; Reish et al. 2002; Aronson et al. 2011; Mearns et al. 2012; Naser 2013; Gaw et al. 2014; Madin et al. 2016; Peng et al. 2020; Falkenberg et al. 2020).

Terrestrial ecosystems have also been severely affected by the destruction of habitats, overhunting, pollution and climate change. These pressures may have severe impacts on the provisioning services that terrestrial environments are able to provide and that humans need, particularly food provision. Because soil provides the natural habitat for plant growth and development, as well as serve as water and nutrient reservoirs, it is essential for the proper development of plants and microorganisms. An overarching problem is soil-degradation by the means of soil overexploitation that lowers the current and/or future capacity of the soil to support life (Bridges and Oldeman 1999). One of the main specific pressures that act on terrestrial ecosystems is soil pollution. The main and most studied aspects of soil pollution is that by heavy metals and by pesticides. Soils have the capacity of retaining various pollutants like heavy metals or pesticides and thus act as a pollution absorbers (Wołejko et al. 2020). Heavy metals accumulate in soils primarily as a result of dry and wet atmospheric deposition from a variety of sources; the primary and most significant source of heavy metals is industrial emissions (e.g., chemistry, mining, iron and steel industry, metallurgy, building and electronics industry), fuel combustion, waste management, and transportation (automobile traffic, fuel composition, road types, types of the engine and driving mode are important constraining variables) (Weissmannová and Pavlovský 2017). Pesticides are the most common contaminants among all xenobiotics in soil. Due to the fast human population growth in the last 50 years and a growing demand for high-quality food, agricultural producers frequently implement pesticide-intensive processes, resulting in increased pesticide accumulation in the surface layers of soil that often persists for long periods of time. Metal and pesticide contamination in soils often

translates into bioaccumulation and biomagnification phenomena across the food chain, which can potentially threaten human and environmental health (Wołejko et al. 2020).

The limitation of such detrimental impacts in different ecosystems, hence promoting the services they can provide, is largely dependent on our ability to properly assess effects of the pressures. Assessment allows linking pressing factors with their effects, which is critical for example to define restrictions to pollutants' emission within a regulatory context or to establish appropriate measures for recovery management. An example of this context is the Water Framework Directive (WFD) that guides towards reaching the common goal of a 'good status' for all water bodies in the EU (Shea and Thorsen 2012; Santos et al. 2021). Furthermore, ecological risk assessment (ERA), which assesses the likelihood that adverse ecological effects will occur or are occurring as a result of exposure to one or more stressors, is perhaps the most comprehensive and integrated framework for assessing the responses of ecosystems to anthropogenic pressure, either prospectively or retrospectively (Shea and Thorsen 2012). How likely an organism or ecological component responds to a specific stressor indicates on how sensitive they are to that stressor. Responses of biological systems can be measured in number of different ways considering several endpoints and using a number of different tools, while ecotoxicological evidence more directly addresses cause-effect relationships (Shea and Thorsen 2012). Common tools to assess responses at the individual level are, for example, acute toxicity tests that generally measure the lethality of stressors (e.g., through median lethal concentration, LC50, benchmarks) or chronic toxicity tests that generally measure sub-lethal impacts in reproductive rates or growth (through estimated benchmarks such as effect concentrations, EC*x*, or data-based benchmarks such as the no-observed-effect-concentration, NOEC, or the lowest-observed-effects-concentration, LOEC) (Akhila et al. 2007; Newman 2019). Responses at the sub-individual level, although needing further extrapolation for an appraisal of the potential higher-level outcomes that more closely relate to ecosystem impairment, are very useful tools for assessing environmental impacts at a mechanistic level, and often as early warning signals of effects at the individual level (Strimbu and Tavel 2010). In this context, biochemical biomarkers are often used. For example, cholinesterases are enzymes that have an important role in neurotransmission and can respond to neurotoxic stressors,

therefore their measured activity can reflect the effects of these stressors (Walker 1992; Colovic et al. 2013).

DNA and RNA can be understood as the biological levels bearing the highest resolution regarding the assessment of effects of environmental pollutants. Several tools are available that use DNA to detect and measure effects of environmental stressors, based for example on structural alterations like damage. DNA damage can be caused by either endogenous or exogenous factors. Exogenous DNA damage occurs when environmental, physical, and chemical factors damage the DNA. Endogenous DNA damage occurs when chemically active DNA interacts with water and reactive oxygen species in hydrolytic and oxidative processes (ROS). Failure of DNA repair systems to repair this damage can lead to a variety of illnesses, including neurological disorders and sporadic malignancies, not to mention overall degradation of biological functions (Chatterjee and Walker 2017). Also, gene expression, i.e., the transcription of DNA to RNA, is a process that can be specifically affected by stressors, constraining phenotypic responses and thus defining the response of the exposed organisms to e.g., contaminants (Snell et al. 2003; Schirmer et al. 2010). Through the analysis of gene expression profiles, one can clarify the mechanistic behind the effects of a given contaminant, but it is also possible in some particular cases to identify the contaminant(s) that are impacting the studied individuals; this is because toxicant-induced genes are only a small part of genome and define specific pathways activated by the exposure, which can be unique signatures of the stressors involved (Snell et al. 2003; Schirmer et al. 2010; Louis et al. 2017).

At this high-resolution level, the study of the epigenome, that consists of heritable phenotypic changes that do not involve alterations in the DNA sequence, has also been emerging. It is known that the epigenome is largely influenced by environmental conditions, and epigenetic changes may alter gene expression, and thus the way the genome responds to environmental stress. Therefore, epigenetics provides a whole new venue for the mechanistic understanding of toxic effects of environmental chemicals, as well as to the development of early-warning biomarkers of exposure, both highly relevant features within ERA frameworks. In spite of its potential in ecotoxicology, epigenetics is a relatively new field, and a lot of research still needs to be done to fill many gaps, which

nowadays prevent its wide incorporation in ERA (Vandegehuchte and Janssen 2011; Suarez-Ulloa et al. 2015; Jeremias et al. 2018a, 2020; Šrut 2021).

1.1. Epigenetics

Epigenetics is defined as the study of potentially heritable phenotype changes that do not involve alterations in the DNA sequence. These changes occur through several different mechanisms but the most common are DNA methylation, histone modification and non-coding RNAs (Vandegehuchte and Janssen 2011), hence these are developed further in section 1.1.2. These epigenetic alterations can be a response to environmental cues, and they are generally reversible. So, they may contribute to the development of abnormal phenotypes and also the modulation of normal physiological responses.

The epigenetic mediation of physiological responses to environmental stimulation has been demonstrated widely. An example among many is the vernalization reaction of plants growing at high altitudes (Jaenisch and Bird 2003). Vernalization is a quantitative reaction to low temperature that causes flowering to occur sooner. The FLOWERING LOCUS C (FLC) gene is assumed to play a key part in the vernalization response, allowing the vernalized state to be passed down in a mitotically stable manner. It has been demonstrated that DNA methylation regulates FLC activity, and that genome-wide demethylation, induced by 5-azacytidine or a METHYLTRANSFERASEI (METI) antisense construct, increases blooming in vernalization-responsive *Arabidopsis* ecotypes and mutants (Sheldon et al. 2000).

In addition, generational and transgenerational inheritance of epigenetic marks has been (more scarcely) demonstrated as a new scope for the understanding of local adaptation processes, long-term responses and resilience to environmental changes by organisms in the wild. For example, Jeremias et al. (2018b) found that specific methylation patterns were transferred from a parental generation exposed to salinity stress to a non-exposed generation (F3) of *Daphnia magna*. In particular, hypomethylation of six protein-coding genes with important roles in the organisms' response to environmental change like DNA damage repair, cytoskeleton organization, and protein synthesis was recorded, which may increase the potential for future generations to better deal with the same stressor. But

environmental stress can also induce negative effects at the phenotypic level, driven by epigenetic mechanisms, that reduce organisms' fitness. For example, mice exposed to social defeat stress develop long-term depression-like symptoms. These mice were observed to have a decrease in expression of brain-derived neurotrophic factor (BDNF) transcripts linked to an increase in methylation at histone 3 lysine 27, corresponding to the promoter regions of the same transcripts (Tsankova et al. 2006); offspring whose parents had been exposed to social defeat stress also exhibited depressive and anxiety-like behaviours (Dietz and Nestler 2012).

1.1.1. Epigenetic mechanisms

Epigenetic mechanisms have distinct ways of influencing gene expression and therefore their own ways of altering phenotypes. The most common epigenetic mechanisms are histone modifications, non-coding RNAs and DNA methylation, the latter being the most comprehensively studied so far.

1.1.1.1. Histone modifications

Histones are proteins abundant in lysine and arginine residues, which are found in the nuclei of eukaryotic cells. They act as spools around which DNA winds creating structural units called nucleosomes, thus preventing DNA from becoming tangled and protecting it from DNA damage. H1/H5 (linker histones), H2, H3, and H4 (core histones) are the five families of histones. The nucleosome core is formed of two H2A-H2B dimers and a H3-H4 tetramer (McGinty and Tan 2015). The structure of histones is not static and post-translational modifications of their protruding amino-terminal tails, which they all have, can change their mutual affinity as well as their affinity towards DNA and other chromatin-associated proteins, hence the role of these post-translational changes in controlling gene expression. These modifications include (but are not limited to) methylation, citrullination, acetylation and phosphorylation (Lennartsson and Ekwall 2009) (see Figure 1 for a summary). The various post-translational modifications of histones that are herein

described directly affect the structure of chromatin, regulating the level of chromatin condensation and thus transcription, affecting gene expression (Gibney and Nolan 2010).

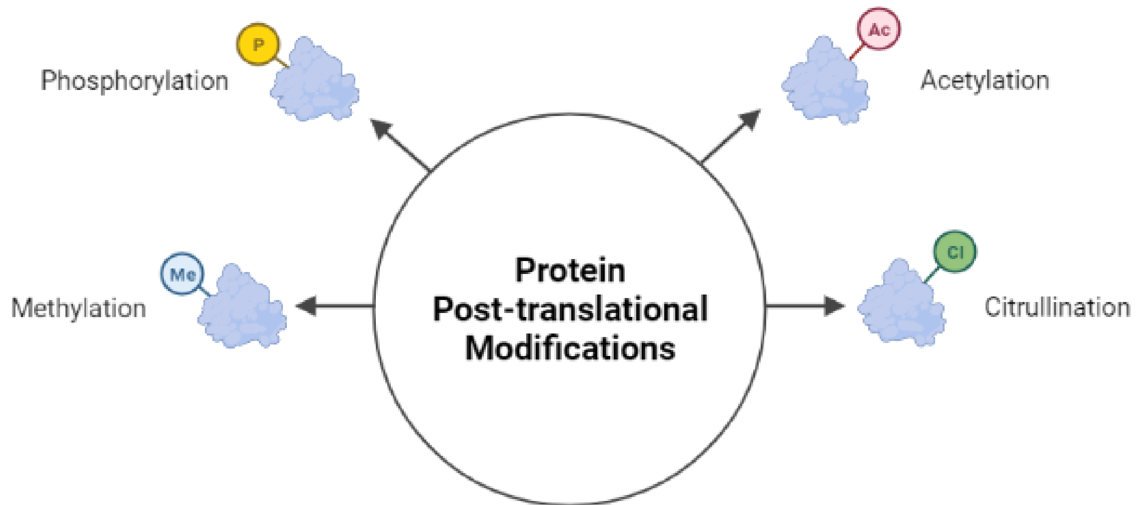

Figure 1. Summary on the most common and best known post-translational modifications of histones considered in the epigenetics arena. Adapted from "Protein Post-Translational Modifications", by BioRender.com (2021). Retrieved from https://app.biorender.com/biorender-templates

Histone methylation usually occurs in lysine residues through the addition of one, two, or many methyl groups, which can be recognized by proteins containing certain domains like Tudor, chromo or PHD that in turn can differentiate between mono, di and tri-methyl lysine. Histone lysine methylations impart active or repressive transcription based on their locations and methylation states, unlike other histone modifications that merely identify active or repressed chromatin states. H3K4, H3K36, and H3K79 methylations are assumed to indicate active transcription, whereas H3K9, H3K27, and H4K20 methylations are thought to indicate silenced chromatin states (Hyun et al. 2017).

Citrullination converts arginine into citrulline in a process driven by enzymes like peptidyl-arginine deiminases (PADs) that hydrolyse the imine group of arginine and attach a keto group. This process removes a positive charge from the amino acid residue. The activation of gene expression has been found to be an outcome of this process as it renders histones

less bound to DNA and by consequence makes the chromatin more accessible (Christophorou et al. 2014). PADs can also counteract the effect of arginine methylation (less common than lysine methylation; see above) by eliminating or blocking mono-methylation of arginine residues on histones; through this process the effect of arginine methylation on transcriptional activity can be antagonized (Cuthbert et al. 2004), configuring an indirect effect of the histone citrullination machinery.

The acetylation of histones typically occurs in lysine residues. This happens through the addition of an acetyl group and has a major chemical effect on the lysine as it neutralises the positive charge it possesses. This makes the histone and the negatively charged DNA less electrostatically attracted to each other, resulting in the relaxing of the chromatin structure. Histones that are highly acetylated induce a higher accessibility to chromatin, usually associated with active transcription. Histone acetyltransferases tend to act on more than one lysine in order to have a substantial effect on chromatin structure, hence lysine acetylation appears to be less precise than methylation (Grunstein 1997).

Histone phosphorylation can occur on the amino acids serine, threonine or tyrosine. This modification adds a negatively charged phosphate group, which results in a repulsive force between the negatively charged phosphates of the DNA and the modified amino acid. This alters the structure of the chromatin and changes DNA accessibility. This modification is generally associated with transcriptional activation (Jaenisch and Bird 2003; Gibney and Nolan 2010; Rossetto et al. 2012).

Environmental stress has been shown to influence these modifications. For example, Lee et al. (2018) demonstrated that alterations in global histone modification status, particularly the hyper-methylation of H3K36, occurred in *C. riparius* following exposure to bisphenol-A, as well as that arsenic may affect histone methylation primarily by competing with the methyl group S-adenosylmethionine (SAM) during its own methylation metabolism. Gilmour et al. (2003) showed that treatment with PM_{10} and H_2O_2 increased the activity of histone acetyltransferases as well as the level of acetylated histone 4 in human alveolar cells, and Burkhart et al. (2007) showed that phosphorylation of histone H3 serine 10 and 28 increased under osmotic stress.

1.1.2.2. Non-coding RNAs

Non-coding RNAs (ncRNA) are RNA molecules that do not translate into a protein. Among ncRNAs the main types are microRNAs, small interfering RNAs, and long non-coding RNAs. They act by being involved in the translation and splicing through sequence specific recognition of RNA substrates and thus are involved in the regulation of gene expression (Gibney and Nolan 2010).

MicroRNAs (miRNAs) are small non-coding RNA molecules of around 22 nucleotides that interact with complementary sequences in mRNA molecules by base-pairing. The genes encoding for miRNA are usually transcribed by RNA polymerase II, then they form a double-stranded RNA hairpin structure, which is cut into a smaller precursor miRNA by the RNase III Drosha, partnered by the microprocessor complex subunit DGCR8. This miRNA precursor is then exported from the nucleus by the nucleocytoplasmic shuttler Exportin-5. In the cytoplasm, the pre-miRNA hairpin is cleaved by the RNase III enzyme Dicer. Despite the fact that both strands of the duplex are capable of acting as functional miRNA, only one of them is usually incorporated into the RNA-induced silencing complex (RISC) where the miRNA and its target interact. miRNAs don't need to have 100% complementary in order to pair with their target and are able to use as few as 6–8 nucleotides; thus, miRNAs can have hundreds of different mRNA targets. The mature miRNA is part of an active RISC containing Dicer and many associated proteins including members of the Argonaute (Ago) protein family. This complex then interacts with mRNA and can interfere with translation in different ways. For example, if there is near 100% complementary the mRNA molecule can be cleaved by the RISC complex, which leads to its degradation through natural mechanisms (removal of the polyA tail and then degraded thought nucleases). Alternatively, the RISC can recruit proteins that have silencing domains to inhibit the binding of eukaryotic translation initiation factor and the 60S ribosomal subunit (Bushati and Cohen 2007) (Figure 2a).

Small interfering RNA (siRNA) is a type of double-stranded RNA with a length of 20 to 27 base pairs. By degrading mRNA after transcription and blocking translation through the same mechanisms as miRNA, they interfere with the expression of certain genes with complementary nucleotide sequences. Contrarily to miRNAs, siRNAs have perfect

complementary with their targets, thus having only one target (Mack 2007; Nambudiri and Widlund 2013).

Long non-coding RNAs (lncRNAs) bear lengths exceeding 200 nucleotides. In many instances, mRNAs and lncRNAs are similar in terms of their biogenesis and RNA polymerase II (Pol II) transcribes lncRNAs from genomic loci with comparable chromatin states to mRNAs; they are frequently 5'-capped, spliced, and polyadenylated. Thus, in general terms, they lack any biochemical distinction from mRNAs. However, some specific features distinguish lncRNA from mRNA, namely: (i) lncRNAs do not have a translated open reading frame (ORF); (ii) lncRNAs tend to be shorter than mRNAs, have fewer but longer exons, and are expressed at relatively low levels, exhibiting poorer primary sequence conservation. lncRNAs interact primarily with mRNA, DNA, proteins, and miRNA; as a result, they influence gene expression in a number of ways at the epigenetic, transcriptional, post-transcriptional, translational, and post-translational levels. By regulating histone methyltransferases or demethylases or by recruiting chromatin-modifying complexes, lncRNAs can affect various histone modifications such as methylation and acetylation. They can also act as molecular scaffolds binding two or more protein molecules that then perform specific biological functions. They can recruit DNA methyltransferases to regulate target gene transcription, act as co-factors to modulate transcription factor activity, inhibit the transcriptional regulation of genes, cooperate with transcription factors and complexes to repress gene transcription and be involved in the mechanisms of alternative splicing. They can act as endogenous target mimics (eTMs), which regulate gene expression by competing with miRNAs, mediate the decay of mRNA by preventing the interaction of RNA-binding proteins and mRNAs or they can bind to mRNA transcripts and help to recruit RNA-binding proteins to stabilize mRNA (Ernst and Morton 2013; Quinn and Chang 2016; Zhang et al. 2019) (Figure 2 b, c, d, e, f, g).

Non-coding RNAs have been shown to be affected by environmental conditions. For example, the association between microalbuminuria, elements identified in urine (Hg, Pb, As, and Cd), and miR-21, miR-126, miR-155, and miR-221 levels was investigated by Kong et al. (2012). They found that the levels of miR- 21 and miR-221 were negatively associated with As and Pb levels, and that the levels of miR-21 were associated with microalbuminuria.

As a result, in the context of elemental exposure, miR-21 and miR-221 have been proposed as biomarkers of renal function. In another study, pregnant mothers and their infants were examined for the potential association of the presence of As in drinking water and maternal urine with miRNAs expression in cord blood. A number of cord blood miRNAs (let-7a, miR-107, miR-126, miR-16, miR-17, miR-195, miR-20a, miR-20b, miR-454, miR-96, and miR-98) have significant relationships with arsenic levels in urine, and these miRNAs have been linked to cancer and diabetes (Rager et al. 2014). The formation of siRNA was already associated to salt stress, mediated by other genes like P5CDH and SRO5 as show in Borsani et al. (2005), and the lncRNAs MALAT1 and SOX2OT were shown to be induced in human cells exposed to ionizing radiation (Chaudhry 2013).

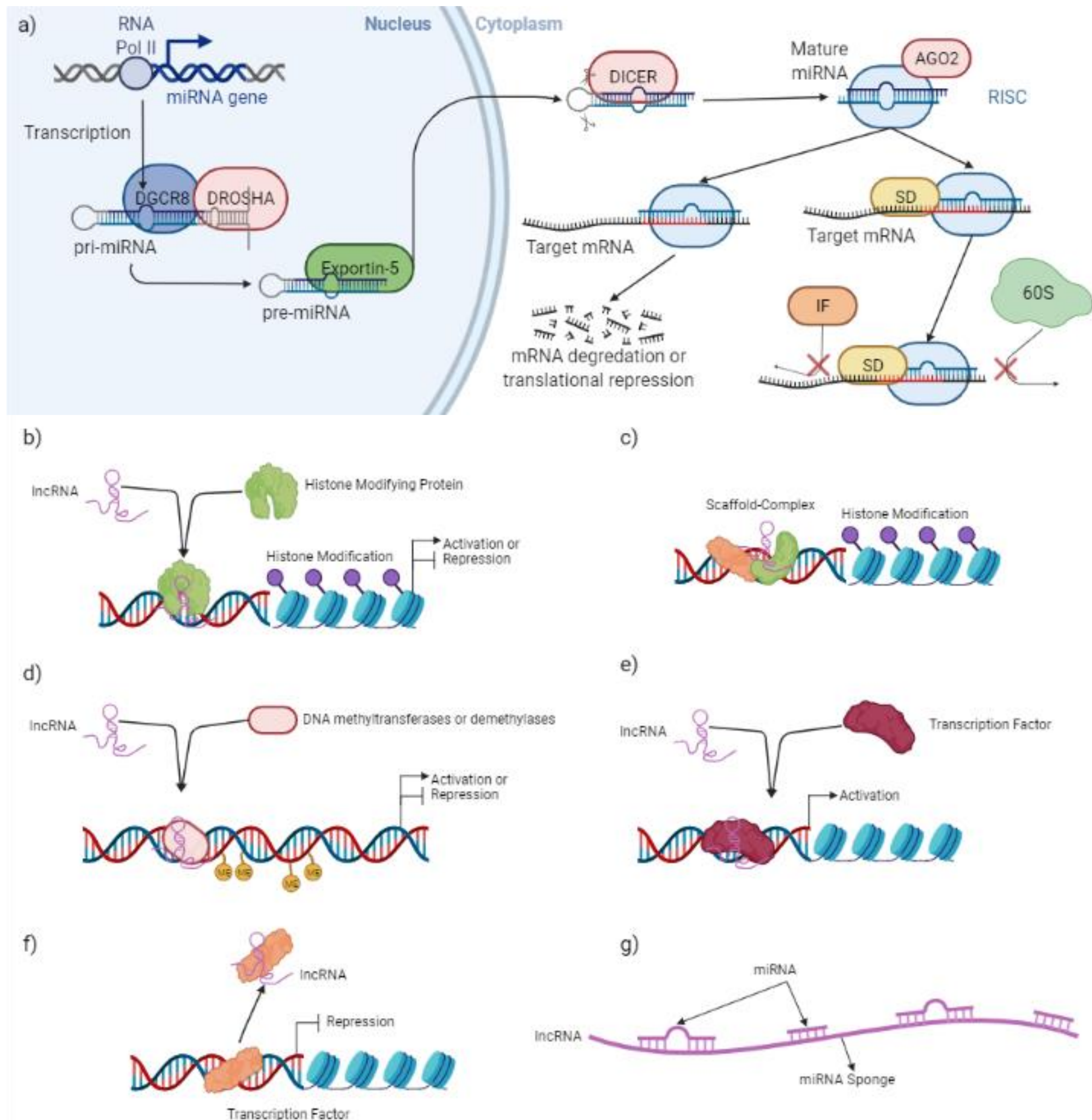

Figure 2. Representation of clarifying selected aspects modulating the role of ncRNAs as players in epigenetic mechanisms. In panel (a), a summary on the biogenesis and general mode of action of miRNAs is given. In the remaining panels, multiple aspects of the activity of lncRNAs are presented, namely in recruiting chromatin-modifying complexes (b); as molecular scaffolds (c); in recruiting DNA methyltransferases (d); in modulating transcription factor activity (Activation of gene transcription) (e); in modulating transcription factor activity (Repression of gene transcription) (f); in acting as a miRNA sponge (g). Created with BioRender.com.

1.1.2.3. DNA methylation

DNA methylation is the epigenetic mechanism better understood and the one that has been more explored within an ecotoxicological context. DNA methylation is a biological process that consists in the addition of a methyl group (CH3) from S-adenyl methionine (SAM) to the 5' position of a cytosine ring, creating a 5-methylcytosine (m5C). DNA methylation occurs in cytosines (also, but less prominently in adenines) in both eukaryotes and prokaryotes, especially in cytosines followed by guanines (CpG dinucleotides). When a CpG-rich region appears in the genome sequence, it is usually called a CpG island. CpG islands are defined as regions with a length greater than 200bp, a G+C content greater than 50% and a ratio of observed to expected CpG greater than 0.6. DNA methylation is catalysed by specific enzymes that are named methyltransferases (DNMTs) – see Figure 3 for an overview of the mechanisms and actors involved in DNA methylation. The maintenance methyltransferase DNMT1 is thought to be important for copying DNA methylation patterns to daughter strands during DNA replication. Early in development, the *de novo* methyltransferases DNMT3a and DNMT3b establish DNA methylation patterns. DNMT3L is a homologous protein to the other DNMT3, but it lacks catalytic activity. DNMT3L helps the methyltransferases DNMT3a and DNMT3b by boosting their ability to bind to DNA, hence promoting their activity.

DNA methylation has a reverse process, i.e., DNA demethylation (Figure 3), which can be passive or active. Passive DNA demethylation occurs in dividing cells as the result of the inhibition or dysfunction of DNMT1. The process itself is not well understood but it is known, for example, that primordial germ cells are characterised by the downregulation of pathways associated with DNA methylation maintenance like Dnmt1, UHRF1 (Ubiquitin Like With PHD And Ring Finger Domains 1) and BER (Base excision repair) pathways (Kurimoto et al. 2008). Mulholland et al. (2020) discovered that DPPA3 (Developmental Pluripotency Associated protein 3) causes large-scale passive demethylation by binding to UHRF1 and displacing it from chromatin, thus preventing DNA methylation maintenance. Because DNMT1 is responsible for actively maintaining DNA methylation during cell duplication, when this machinery stops working, newly incorporated cytosines remain

unmethylated, and there is a consequent reduction of the overall methylation level following each cell division (Dean 2016).

Both dividing and non-dividing cells can undergo active DNA demethylation, although the process requires enzyme activities to convert the 5mC back to a naked cytosine (Figure 3). Demethylation is accomplished through a sequence of chemical events, usually TET (dioxygenase) oxidation of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), then subsequent changes, as depicted in figure 3, result in a product that is identified by the BER (Base excision repair) pathway, which replaces the modified base with naked cytosine. Although the BER route is widely accepted as the ultimate step in DNA demethylation, the specific enzymes and chemical intermediates generated during DNA demethylation are still a source of controversy (Jaenisch and Bird 2003; Gibney and Nolan 2010; Vandegehuchte and Janssen 2011; Kohli and Zhang 2013; Wu and Zhang 2017).

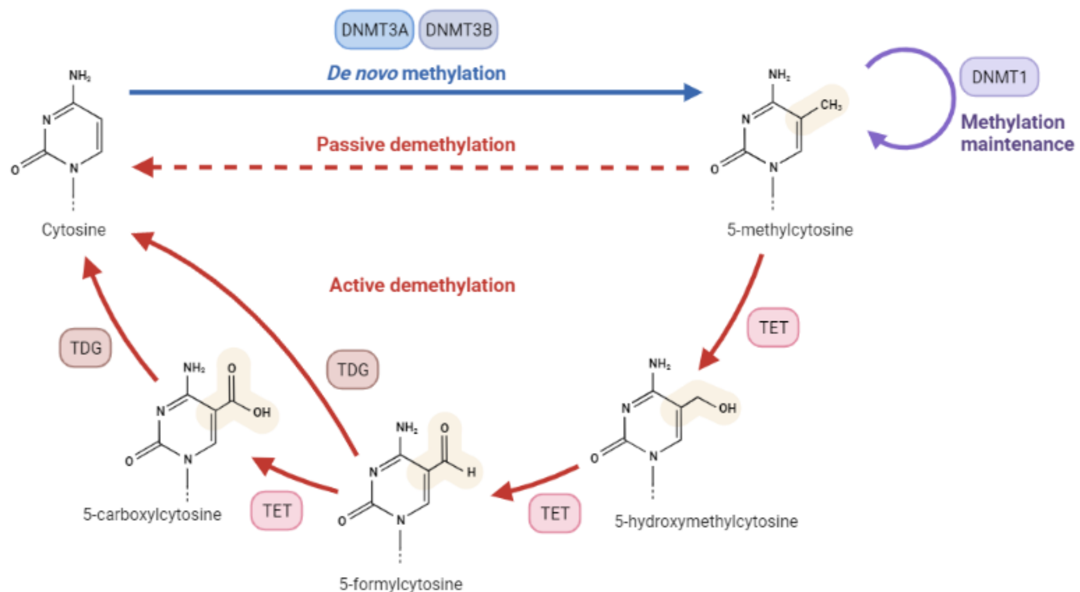

Figure 3. DNA methylation and demethylation pathways and its components. Reprinted from "DNA Methylation", by BioRender.com (2021). Retrieved from https://app.biorender.com/biorender-templates.

In animals, DNA methylation is reprogrammed in two stages: one during the formation of the germ line, and the other shortly after fertilization, as part of an ongoing cycle that ensures genetic material is passed down through generations (Cedar and Bergman 2012) –

see a graphical summary in Figure 4. At the E7.5 stage of embryo development, primordial germ cells (PGCs) develop, and DNA methylation is completely erased as they proliferate and migrate towards the genital ridge. In male and female embryos, new DNA-methylation landscapes are generated in germ-cell precursors in an asymmetrical manner following sex determination. In the male embryo, *de novo* methylation takes place before meiosis in mitotically arrested cells and is completed before birth. In the female embryo, primary oocytes enter meiosis and arrest in Prophase-I; the DNA methylation landscape is established only after birth during the follicular/oocyte growth phase. Following fertilization, a fresh wave of DNA demethylation occurs that differs from that seen in the parental genomes. The paternal genome's DNA methylation is promptly erased by an active mechanism in the zygote. Demethylation of the maternal genome is slower and is dependent on DNA replication (passive demethylation). These post-fertilisation demethylation events do not include imprinted germline differentially methylated regions, resulting in parental-allele-specific methylation of these elements in early embryos and consequent parental-allele-specific expression of associated imprinted genes. New methylation landscapes emerge in tandem with blastocyst implantation and cell-lineage determination and are linked to cellular differentiation (Smallwood and Kelsey 2012).

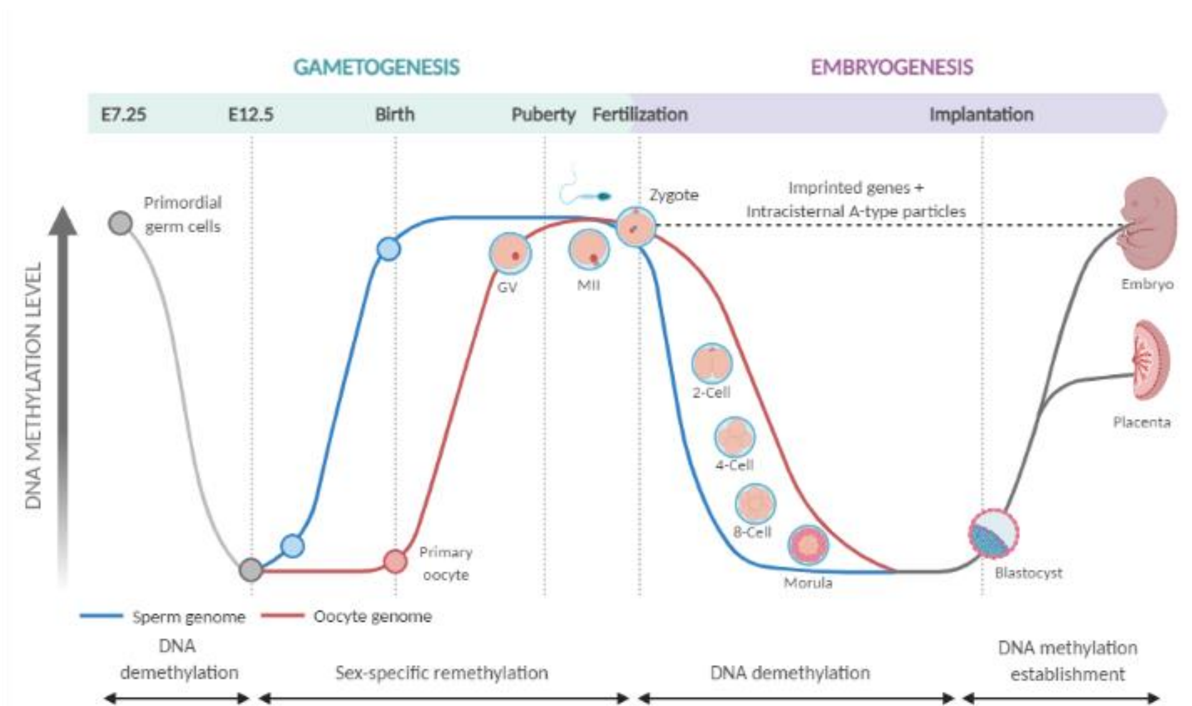

Figure 4. DNA methylation levels during mammalian development. Reprinted from "DNA Methylation Levels During Mammalian Development", by BioRender.com (2021). Retrieved from https://app.biorender.com/biorender-templates

Through the addition of the methyl group, DNA methylation alters the biophysical characteristics of the DNA, bearing two effects that are usually associated with gene repression. First, it inhibits the recognition of DNA by blocking transcriptional activators from binding to the DNA sequence (Figure 5a). In this context, Watt and Molloy (1988) demonstrated that methylation at a CpG site centrally located within the binding site strongly inhibited complex formation and the binding of transcription factor MLTF. Second, they facilitate the binding to the DNA of other proteins like methyl-binding proteins (MBPs), which recognize methylated DNA and recruit co-repressors to silence gene expression directly (Figure 5b). For example, among the MBD family of proteins, MBD1, MBD2, MBD3 are transcriptional repressors through various mechanisms, resulting in the recruitment of co-repressors and histone deacetylases; these histone deacetylases promote the compaction of DNA leading to the characteristic remodelling of chromatin. MECP2 is a member of the MDB family and can bind to methylated CpG via its MBD domain, as well as exert repressive effects on transcription over hundreds of base pairs via its second

functional domain, the transcriptional repression domain. The co-repressor Sin3 complex, which contains histone deacetylase 1 and 2, or other co-repressor complexes, is recruited by this repressor domain. MECP2 can also change chromatin compaction by binding to linker DNA and nucleosomes, which creates a physical barrier to the transcriptional machinery (Gibney and Nolan 2010). More mechanisms that alter gene expression within the general scope of DNA methylation are those involving DNMTs specifically interacting with transcription factors, which results in site-specific methylation at promoter regions. This site-specific methylation is then responsible for the assembly of proteins that recognize methylated DNA at these places, similar to the one described above. These assemblies then directly influence further action of the transcriptional machinery or cause alterations in chromatin structure, which in turn affects normal gene expression mechanisms (Gibney and Nolan 2010).

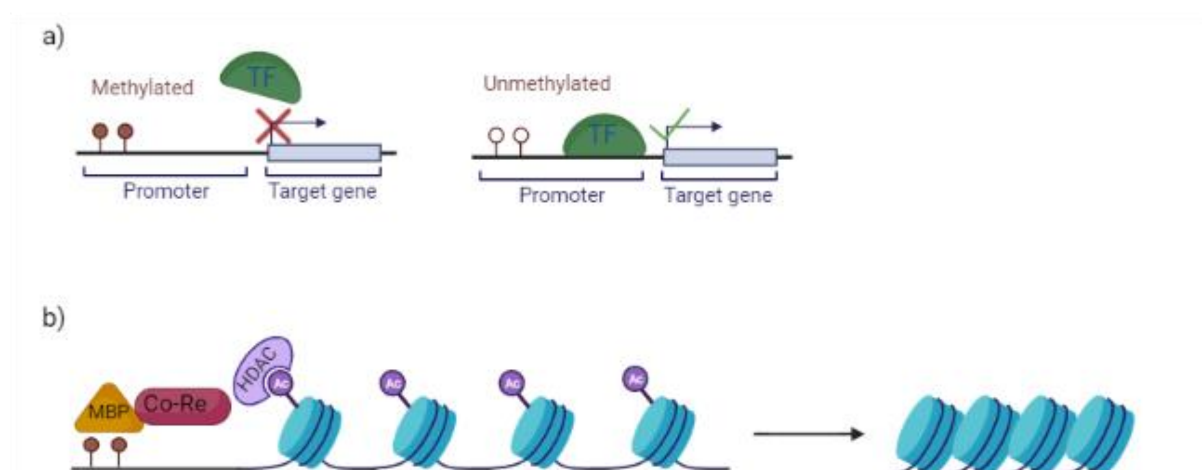

Figure 5. Methods by which DNA methylation can repress gene expression. a) Methylation prevents the binding of transcription factors thus inhibiting transcription. b) Methylation recruits methyl-binding proteins (MBPs), which in turn recruit co-repressors and histone deacetylases that remove acetyl groups from histones leading to a more tightly bound chromatin inducing transcription repression. Created with BioRender.com.

Although the most recognised effect of DNA methylation is the repression of gene expression, DNA methylation can also promote it by being specifically identified by transcription activators. Studies in humans and mice, for example, already showed the expression of dozens of transcription factors with binding preferences for specific methylated sequences. This can happen when these transcription factors that bind DNA

methylation site recruit proteins that catalyse DNA demethylation and through that way gene expression is rather activated. C/EBP and KLF4 transcription factors, for example, attract the methylcytosine dioxygenase TET2 to enhancers for demethylation during cell-type reprogramming, as shown by Sardina et al. (2018). DNA methylation can also play a role in gene expression activation by displacing other repressive modifications. For example, the genomic distribution of DNA methylation is largely mutually exclusive from the genomic distribution of histone H3 Lys27 trimethylation (H3K27me3), a gene-repressive histone modification catalysed by Polycomb repressive complex 2 (PRC2). When PRC2 is displaced by DNA methylation and hence loses H3K27me3, gene activation is occasionally observed (Greenberg and Bourc'his 2019).

1.2. Epigenetic responses to environmental stressors

Over the past few years, several studies looking at the effects of various stressors on epigenetics marks have been conducted. In the context of the present Dissertation, this section reviews what has been done in this field, particularly focusing on organisms other than humans, thus effects on those organisms rather than on human health. A summary of the studies reviewed is given in table 1.

Most studies regarding epigenetic effects of environmental stressors study DNA methylation rather than other epigenetic mechanisms. Metal elements are perhaps the most studied environmental stressors in this context. Exposure to Cd apparently induces both hypo- and hypermethylation depending on the organism and on the exposure routine. Cribiu et al. (2018) found global DNA hypomethylation in *Gammarus fossarum* after 14 days of exposure, but then hypermethylation after 1 month of exposure. Global DNA hypermethylation caused by Cd was shown in *Cantareus asperses* (Nica et al. 2017a, b) and also in *Lumbricus terrestris* (Šrut et al. 2017), while in *Drosophila melanogaster* there was both increased and decreased methylation in several genes (Guan et al. 2019); for *Daphnia magna* (Vandegehuchte et al. 2009a) and *Artemia franciscana* (Pestana et al. 2016), the methylation status did not respond to Cd. Various other metals induce variable DNA methylation patterns. For example, the F1 progeny of *D. magna* exposed to Zn showed a

global reduction in DNA methylation (Vandegehuchte et al. 2009b). In *Enchytraeus crypticus*, Cu exposure in adult animals resulted in no effects regarding global methylation level (Noordhoek et al. 2018) but another study performed on juveniles reported a tendency for the increase global DNA methylation (Bicho et al. 2020a). Cu caused DNA methylation changes in some specific genes of *Crassotrea gigas* (Sussarellu et al. 2018). Studies are available examining the effect of environmental exposure of earthworms collected from mining areas associated exposure to the prevailing elements (Au, Ag, As and Cu) and particular DNA methylation patterns (Kille et al. 2013); and in a study using multigenerational exposure of *Chironomus riparius* to sediments contaminated with metals, global DNA hypermethylation in F0 and hypomethylation in F2 and F4 was found (Im et al. 2019). The mussels *Mytilus galloprovincialis* and *Xenostrobus secures* were shown to experience a decrease in genome-wide DNA methylation when inhabiting ports with metal pollution (Ardura et al. 2018).

Another type of contaminants commonly addressed are pesticides and endocrine disruptors. The fungicide vinclozolin (VZ), an endocrine disruptor, has the potential to change DNA methylation in several invertebrate species. VZ induces a reduction in global DNA methylation in both *D. magna* and *Physella acuta* (Vandegehuchte et al. 2010; Sanchez-Arguello P et al. 2016), but no intergenerational effects can be observed (Müller et al. 2016). Exposure to VZ and the phytoestrogen genistein induced changes in global DNA methylation in *Aedes albopictus,* which correlated with reduced insecticide sensitivity until the F2 generation (Oppold et al. 2015). The herbicide diuron induced no effects regarding DNA methylation in whole tissue of *C. gigas* adults exposed during gametogenesis, but higher DNA methylation was observed in the sperm of exposed adults, as well as in the spat (Bachère et al. 2017) and in the digestive gland, followed by a decrease in DNMT activity (Akcha et al. 2021). Both Rondon et al. (2017) and Bachère et al. (2017) found that parental exposure to diuron during gametogenesis caused global DNA hypermethylation, altered transcriptomic profiles and reduced growth and survival in the first year of life regarding the unexposed mussel spat. Bisphenol A (BPA) is an endocrine disrupting chemical that has been shown to cause disruptions in epigenetic pathways in *Eisenia fetida* through the repression of the expression of DNMT1 and DNMT3b genes in

its male organs (Novo et al. 2018). All of the studies that addressed environmental exposure to 5-azacytidine reported a consequent decrease in DNA methylation, both at the global and subcellular levels (Vandegehuchte et al. 2010; Lindeman et al. 2019), as well as in specific target genes (Athanasio et al. 2018), with intergenerational effects being confirmed in unexposed F1 and F2 generations (Vandegehuchte et al. 2010) and in F0 and F1 generations (Lindeman et al. 2019).

Several studies use *D. magna* as a model organism to address DNA methylation upon exposure to various environmental factors. *D. magna* showed global DNA hypermethylation in multiple generations (F0–F2) when exposed to water contaminated with faecal coliforms (Chatterjee et al. 2019); additionally, gamma irradiation or the toxic cyanobacterium *Microcystis aeruginosa* were shown to induce significant methylation changes at specific CpG sites that could persist until the F3 generation post-exposure (Asselman et al. 2017) . The epigenome of *D. magna* is also responsive to other stressors like raised salinity levels, which induces global decrease in DNA methylation as well as hypomethylation of genes involved in general responses to environmental changes (Asselman et al. 2015; Jeremias et al. 2018b).

Other stressors that have been tested include valproic acid (VPA), curcumin, prednisolone (PDS) and perfluorooctane sulfonate (PFOS). VPA and curcumin have altered the expression of genes that encode epigenetic regulators like HDACs, HDAC subunits, HATs and methylation enzymes in *T. castaneum* (Bingsohn et al. 2016). PDS exposure in *P. acuta* resulted in a linear decrease of 5 mC content in exposed adults, which was consistent over three exposed generations (F0–F2) (Bal et al. 2017), and PFOS exposure of *Glyptocidaris crenularis* resulted in an increase in genome-wide DNA methylation, methylation polymorphisms and also demethylation rates (Ding et al. 2015). Studies have also been conducted with nanomaterials in *E. crypticus*, and an increase in global DNA methylation that persisted in the subsequent unexposed generations was generally found (Bicho et al. 2020a, b).

Histone modifications have also been analysed in respect to the effect of environmental stressors, albeit to much less extent than DNA methylation. In Camacho et al. (2018), a transgenerational decrease in the germline levels of H3K9me3 and H3K27me3, lasting up

to 5 generations, was found in *C. elegans* exposed to BPA; and in Rudgalvyte et al. (2017), modifications to H3K4me3 induced transcription of 1467 genes and repressed it for 508 genes following chronic exposure of *C. elegans* to methylmercury. Wamucho et al. (2020) found a transgenerational increase in the methylation levels of H3K4me2 in response to pristine Ag-NP exposure, and decreased H3K4me2 and H3K9me3 levels following exposure to sulfidized Ag-NPs in *C. elegans*. Decreased methylation of histone H3K9 was found in the parental generation exposed to heavy crude oil in *C. elegans* (Yang et al. 2018a) and H3K27me3 enrichment was observed in perinatal *D. magna* exposed to 5-azacytidine (Lindeman et al. 2019).

Table 1. Overview of the studies examining epigenetic changes using invertebrates exposed to environmental stressors.

Reference	Organism	Stressor	Effect
(Wamucho et al. 2020)	*C. elegans*	Silver nanoparticles (Ag-NP)	Transgenerational increase in histone methylation levels at H3K4me2 in response to pristine Ag-NP exposure and decreased H3K4me2 and H3K9me3 levels exposure to transformed sAg-NPs
(Camacho et al. 2018)	*C. elegans*	Bisphenol A	Transgenerational decrease in the germline levels of H3K9me3 and H3K27me3 lasting up to 5 generations with modest impacts on reproduction extending until the F7
(Yang et al. 2018a)	*C. elegans*	Iranian heavy crude oil	Decreased methylation of histone H3 (H3K9) was found in the IHC-exposed parental generation
(Rudgalvyte et al. 2017)	*C. elegans*	Methylmercury	Modifications of histone H3K4me3 induced enhanced signal in 1467 genes and reduced in 508.
(Nica et al. 2017a)	*Cantareus aspersus*	Cd	Hypermethylation
(Nica et al. 2017b)	*Cantareus aspersus*	Cd	Increase in 5hmC content in 1 mg/L Cd treatment group
(Santoyo et al. 2011)	*Earthworms*	Silver and gold mine	Inverse correlation between the percentage of methylated DNA and total tissue As, As + Hg, As + Hg + Se + Sb, and inorganic As + Hg
(Novo et al. 2018)	*Eisenia fetida*	Bisphenol A	Lower expression at higher concentrations
(Bicho et al. 2021)	*Enchytraeus crypticus*	CuO NMs and CuCl2	Ago1, H3-dimethyl and 5mC were affected by both Cu forms, concentrations and across generations. Effects evident also in unexposed F6 generation
(Bicho et al. 2020a)	*Enchytraeus crypticus*	Copper oxide nanomaterials (CuO NMs) and copper salt (CuCl2)	Significant difference in F3, F4, F6, F7 in comparison to F1 in global methylation after exposure to CuO NMs, no difference in EF1 genespecific DNA methylation, changes in epigenetic gene targets
(Bicho et al. 2020b)	*E. crypticus*	Nanostructured tungsten carbide cobalt (WCCo NMs) and cobalt (CoCl2)	WCCo nanomaterials caused increase in global DNA methylation, which continued in the unexposed generations. Increase in methylation was associated with increase in reproduction

(Noordhoek et al. 2018)	*Enchytraeus crypticus, Folsomia candida*	Copper (Cu)	No methylation recorded in *F. candida*, 1.4% methylation in *E. crypticus*, no effect of Cu on DNA methylation in E. crypticus
(Kille et al. 2013)	*Lumbricus rubellus*	As and Cu mine	Association of methylation patterns with soil As concentrations in one earthworm lineage
(Šrut et al. 2017)	*Lumbricus terrestris*	Cd	Hypermethylation after exposure to Cd
(Newbold et al. 2019)	*Octolasion lacteum*	Ionising radiation within the Chernobyl exclusion zone	No change in methylation comparing to the control groups
(Bingsohn et al. 2016)	*Tribolium castaneum*	Valproic acid (VPA), curcumin	Curcumin and VPA induced the expression of epigenetic regulatory genes in the F0 and untreated F1 generation
(Pestana et al. 2016)	*Artemia franciscana*	Cd, Zinc (Zn)	No epigenetic alterations
(Akcha et al. 2021)	*Crassostrea gigas*	Diuron	No change in global DNA methylation level in whole tissue increased DNA methylation in the digestive gland, decrease in DNMT activity upon exposure to diuron
(Sussarellu et al. 2018)	*Crassostrea gigas*	Cu	No change in global DNA methylation, decrease in 5-hydroxymethylcytosine, changes in methylation of some gene regions (HoxA1, Hox2 and Notochord), decrease in DNMT3 expression
(Rondon et al. 2017)	*Crassostrea gigas*	Diuron	Parental diuron exposure had an impact on the DNA methylation pattern of progeny, most of the differentially methylated regions occurred within coding sequences
(Bachère et al. 2017)	*Crassostrea gigas*	Diuron	Exp 1: No difference in DNA methylation in exposed adults' whole tissue, higher DNA methylation in spat coming from exposed adults; Exp 2: higher DNA methylation in sperm of exposed adults, higher DNA methylation in spat coming from exposed adults
(Ding et al. 2015)	*Glyptocidaris crenularis*	Perfluorooctane sulfonate	Increase in DNA methylation polymorphism, methylation and demethylation rate
(Ardura et al. 2018)	*Mytilus galloprovincialis, Xenostrobus securis*	Non-essential metal polluted ports	Demethylation in stressed environments

(Athanasio et al. 2018)	*Daphnia magna*	5-azacytidine	Prenatal exposures caused changes in the methylome of target genes, combination of pre- and postnatal exposures caused the most extreme reduction in DNA methylation
(Lindeman et al. 2019)	*Daphnia magna*	5-azacytidine	Decrease in global DNA methylation in F0, decrease in global DNA methylation and changes in histone modifications in F1. Increase in H3K27me3 enrichment in perinatal exposed *D. magna*(F1).
(Chatterjee et al. 2019)	*Daphnia magna*	Water contaminated with faecal coliforms	Hypermethylation in F0 and F1
(Jeremias et al. 2018b)	*Daphnia magna*	Salinity	Specific methylation patterns in exposed F0 that were transferred to the three consequent nonexposed generations (F1–F3); hypomethylation of six protein-coding genes with roles in the organisms' response to environmental change
(Trijau et al. 2018)	*Daphnia magna*	γ irradiation	Significant methylation changes at specific CpG positions in every generation independent of dose rate, with a majority of hypomethylation
(Asselman et al. 2017)	*Daphnia magna*	Toxic cyanobacterium *Microcystis aeruginosa*	Differential methylation mostly in exonic regions of genes related to protein synthesis, transport and degradation
(Asselman et al. 2015)	*Daphnia magna*	Arsenic (As), Cd, Lead (Pb), Hypoxia, pH, temperature, toxic cyanobacterium *Microcystis aeruginosa*, salinity stress	Significant decrease in methylation upon salinity stress in comparison to the control
(Vandegehuchte et al. 2010)	*Daphnia magna*	Vinclozolin, 5-azacytidine, 2′-deoxy-5- azacytidine, genistein, biochanin A	Reduction in DNA methylation after exposure to vinclozolin and 5-azacytidine, effect visible in F1 and F2 unexposed organisms after parental exposure to 5-azacytidine
(Vandegehuchte et al. 2009a)	*Daphnia magna*	Cadmium (Cd)	No effect

(Vandegehuchte et al. 2009b)	*Daphnia magna*	Zn	Decrease in DNA methylation in F1
(Cribiu et al. 2018)	*Gammarus fossarum*	Cd	Hypomethylation after 14 days, hypermethylation after 1month
(Bal et al. 2017)	*Physella acuta*	Prednisolone	Linear decrease of 5 mC content in adult snails of F1 generation
(Müller et al. 2016)	*Physella acuta*	Vinclozolin	No effect
(Sanchez-Arguello P et al. 2016)	*Physella acuta*	Vinclozolin	Changes in DNA methylation patterns
(Oppold et al. 2015)	*Aedes albopictus*	Vinclozolin, Genistein	Alteration of the global DNA methylation level in F0, effect persisted in F1 and F2; F1 and F2 had increased sensitivity to insecticide which correlated with the epigenetic changes
(Im et al. 2019)	*Chironomus riparius*	Metal-contaminated field sediments, As	Hypermethylation at F0, hypomethylation at F2 and F4
(Guan et al. 2019)	*Drosophila melanogaster*	Cd	39 demethylated and 24 genes with increased methylation in Cd treatment. In most cases, demethylation activated gene expression. Significant decrease in methylation upon salinity stress in comparison to the control

1.2.1. Transgenerational Epigenetic Inheritance

Despite the enormous potential of epigenetic marks as a tool for informing on previous exposure scenarios, research on lasting transgenerational epigenetic effects after the removal of the stress pressure in organisms other than humans is limited to very few studies with model species. Our understanding of this process in most invertebrates is nearly non-existent. This is owing in part to a paucity of genomic data and limited use of sequencing techniques allowing to identify DNA methylation at a nucleotide resolution level. In addition, many organisms have a high generational turnover and a DNA that is sparsely methylated.

Epigenetic alterations are usually cleared and re-establish at each generation, but some of them can be transferred through successive generations, even after the initial stress has disappeared, by escaping the process of erasing the ancestral methylation patterns. It is also believed that maintenance of epigenetic information is often reliant on interactions between at least two distinct information carriers, namely the chromatin state or small RNAs that could re-establish ancestral methylation states (Bošković and Rando 2018). Small RNAs, for example, can guide *de novo* cytosine methylation at homologous genomic sites in both plants and animals (Zilberman et al. 2003; Aravin et al. 2008). Similarly, small RNAs can direct the formation of H3K9-based heterochromatin, whereas long noncoding RNAs play complex roles in recruitment and modulation of both H3K27/Polycomb and H3K4/Trithorax chromatin pathways (Bošković and Rando 2018). Heterochromatin and DNA alteration, in turn, can have an impact on the expression of short RNA–generating loci (Bošković and Rando 2018). As one prominent example, heterochromatin plays a key role in directing proper expression and processing of piRNA precursor transcripts in flies (Klattenhoff et al. 2009). But despite some components of how methylation erasing is overcome have been unveiled, there still a lot to be discovered and/or clarified. True (i.e., transgenerational) epigenetic inheritance was already found in several species, such as humans, mice, *Arabidopsis*, *Drosophila* or *Daphnia magna* (Lang-Mladek et al. 2010; Jeremias et al. 2018b; Bošković and Rando 2018).

The ability to transmit epigenetic alterations over several generations even after the ceasing of the driving pressure is a ground for microevolution forces which are independent

of changes in the DNA sequence and that can be critical to a rapid adaptation to environmental stress challenges.

1.2.2. Methylation in Vertebrates vs Invertebrates

Considering that DNA methylation is the best-known epigenetic mechanism responding to environmental stressors, there is room for reasoning on whether methylation patterns differ between different groups of organisms such as vertebrates and invertebrates. Indeed, evidence so far generated allows the identification of some differential trends (Table 2). A first distinguishing feature regards the overall level of genomic methylation. While vertebrates present abundant and widespread methylation, with about 70–80% of cytosines in CpG dinucleotides being regularly methylated (Bird and Taggart 1980; Suzuki and Bird 2008; de Mendoza et al. 2019), invertebrates have a sparsely methylated genome, bearing either no methylation at all or predominant non-methylation (Tweedie et al. 1997; Sarda et al. 2012; de Mendoza et al. 2019). The second distinction refers to the genomic regions where methylation may occur. In vertebrates, methylation occurs across the genome, with the exception of CpG islands found at or near the transcription start site of genes, notably housekeeping genes. In invertebrates, methylation is focused on transcription units occurring mostly at CpG sites within exons and introns of gene bodies (Suzuki and Bird 2008; Sarda et al. 2012; de Mendoza et al. 2019). Third, the gene body methylation (most common in invertebrates) is usually associated with increased gene transcription; this specific association is conserved in the most diverse group of vertebrate and invertebrate species. In vertebrates, methylation is also associated with transcription repression and suppression of transposable elements (Suzuki and Bird 2008; Sarda et al. 2012; de Mendoza et al. 2019).

Table 2. Major differences between vertebrates and invertebrates regarding DNA methylation.

Vertebrates	Invertebrates:
Abundant and widespread methylation	Sparsely methylated genomes
Methylation occurs in across all of the genome with the exception of CpG islands	Methylation occurs predominately at CpG islands within exon and introns of gene bodies.
Methylation is associated with transcription repression	Gene body methylation is often associated with increased gene transcription

Gene body methylation in invertebrates seems to target genes that have housekeeping functions, like those involved in transcription and translation, while the loss of methylation occurs in genes in genes involved in cellular signalling and reproductive processes (Suzuki and Bird 2008; Sarda et al. 2012; de Mendoza et al. 2019). Furthermore, in invertebrates, highly methylated genes are more conserved at the sequence level than genes with low levels of methylation and present almost no variation between related species (Sarda et al. 2012; Asselman et al. 2016). On the other hand, poorly methylated genes are reportedly responsive to experimental changes in what concerns gene expression and display plastic and adaptive variations (Gatzmann et al. 2018). This suggests that gene body methylation regulates phenotypic plasticity through the patterns of methylation in important genes, which limits their transcriptional variation.

1.3. Assessment of epigenetic modifications – most common methods

The study of histone modifications and non-coding RNAs is still very limited and so are the methods by which they are studied. One of the most used methods to assess histone modifications is through high-performance liquid chromatography (HPLC), where each component in the sample interacts slightly differently with the adsorbent material, causing different flow rates and leading to the separation of the proteins as they flow out of the column by molecular weight and conformation. This method is usually combined with mass spectrometry (MS) so that after the physical separation of histones, their post-translational

modifications can be identified (Kumar and Wigge 2010; Norouzitallab et al. 2014; Minshull et al. 2016). Chromatin immunoprecipitation (ChIP) is another method for studying histone modifications, that involves the immunoprecipitation of nucleosomes after chromatin fragmentation using an appropriate protein-specific antibody; qPCR is then used to determine the relative enrichment of changed histones with associated DNA fragments (Park 2009; Kumar and Wigge 2010; Norouzitallab et al. 2014).

Non-coding RNAs can be identified through several methods with an experimental and/or a bioinformatics basis. A common experimental approach is the identification of ncRNAs by chemical or enzymatic sequencing of extracted abundant RNAs. After ncRNAs are isolated from cells or organisms, they are size-separated by polyacrylamide gel electrophoresis (PAGE), eluted from the gel, and then identified by 2D RNA fingerprinting or by enzymatic or chemical sequencing of ncRNAs (Huttenhofer 2006). Assessment of ncRNAs can also be achieved by cDNA cloning and sequencing with the generation of cDNA libraries. RNAs are either isolated by size separation of total RNA via denaturing PAGE or alternatively, by employing an antibody against an RNA-binding protein of interest, thus promoting isolation through immunoprecipitation. Reverse transcription of ncRNAs with poly(A) tails into cDNAs can be accomplished using one of three ways: C-tailing, C-tailing and linker addition, or just linker addition, then RT–RCR. cDNA fragments are then cloned into conventional vector systems and sequenced using cycle sequencing. Bioinformatic approaches are frequently performed after sequencing and usually include locating the ncRNA gene to a specific genomic region and identifying structural or sequence motifs that may aid in determining the function of the ncRNA species of interest (Huttenhofer 2006). The identification of ncRNAs can also be done via microarray analysis. Fluorescently labelled samples produced from cellular RNA are hybridized to DNA oligonucleotides, which are spotted onto glass slides (microarrays). RNA is extracted from the samples, then converted to cDNA, or cRNA; in any case, fluorescent dyes are employed to label the probes, such as Cy3 or Cy5. The prepared sample is then hybridized with the microarray. The fluorescence of the spots where the hybridised sample stands is scanned, with the strength of the fluorescent signal indicating the number of transcripts present in each spot. This method is

most typically used to profile mRNA expression, but it can also be used to explore ncRNA expression or even find new ncRNAs (Huttenhofer 2006).

When it comes to methylation analysis there are two approaches that can be followed, which provide different levels of outcome specificity: the measurement of global DNA methylation, providing an unspecific indication on the methylation level over the entire genome; or methylation assessment on a nucleotide-by-nucleotide basis to retrieve information on the actual gene sequences that were methylated at the time of the sampling.

Global DNA methylation levels are most commonly assessed based on HPLC. Nucleotides are separated according to their size and are then further quantified both for cytosine and methylated cytosine by either UV spectrophotometry or mass spectrometry (e.g., LC-MS). Global methylation levels are depicted as the fraction of 5-methylC compared to total cytosine concentration. These techniques are only able to provide changes in global or total methylation levels and cannot provide single-base information or gene-level information (Kurdyukov and Bullock 2016).

There are several methods that can be used if the intent is to assess methylation at the base or gene level, all based on next-generation sequencing technology. Bisulfite sequencing is amongst the most used currently. It consists of the use of a bisulfite treatment of DNA before routine sequencing. Bisulfite treatment of DNA changes cytosine residues to uracil but has no effect on 5-methylcytosine residues. Only methylated cytosines are retained in bisulfite-treated DNA. Bisulfite treatment causes precise modifications in the DNA sequence that are dependent on the methylation status of individual cytosine residues, resulting in single-nucleotide resolution information about the methylation status of a DNA segment. The comparison of the treated sequence with the original sequence retrieves its methylation pattern (Urich et al. 2015). A variation of this method is reduced representation bisulfite sequencing (RRBS), which uses an enzymatic digestion with a methylation-insensitive restriction enzyme prior to the bisulfite treatment to enrich for CpG. RRBS is a more cost-effective method for analysing gene-specific methylation levels since only the enriched regions of the genome are sequenced and the costs of the technology are reduced (Meissner 2005). Methylated DNA

immunoprecipitation (MeDIP) is an often-used alternative technique. It uses monoclonal antibodies that specifically recognize 5-methylcytidines and/or 5-methylcytosines to enrich methylated DNA. The purified methylated DNA can then be used for locus-specific or genome-wide methylation studies (Kurdyukov and Bullock 2016). Methyl-CpG binding domain (MBD) protein-enriched genome sequencing (MBD-seq) is also an alternative using the methyl binding domain of methyl binding proteins MBD2 or MBD3L1 to capture methylated DNA fragments (Aberg et al. 2012).

While single-nucleotide resolution is critical to understand the DNA methylation responses to environmental stressors and their value as biomarkers of exposure in (eco)toxicological assessments, there are few studies aside the human health arena using them, most of them focusing on vertebrate species.

1.4. Structure and objectives of the Dissertation

This Dissertation is presented following a classical structure. After the Introduction chapter developed above, where a comprehensive revision of the meaning and application of epigenetic mechanisms to environmental assessment fields is made, a Methods chapter follows. Considering that a meta-analysis was carried out based on literature records of invertebrate genomes, this Methods chapter encloses a brief characterisation of the species listed for the analysis; this necessarily requires a less orthodox use of supporting literature, but it is justified by the need to appropriately contextualise the selection. Afterwards, a Results and Discussion chapter is presented, containing the main outcomes of the meta-analysis and their positioning considering previous evidence already reported in the literature. Finally, and immediately before the references listing, a chapter is presented with the final remarks and prospects within the field.

The meta-analysis carried out for this Dissertation was guided by the overarching hypothesis that invertebrates are able to respond to environmental challenges or to changing environmental conditions through their epigenome. In order to explore this hypothesis, several specific objectives were defined and tackled as follows:

(i) To collect publicly available reliable genomic data of various invertebrate species representing different environmental compartments and currently involved in research addressing ecological threats, evolutionary trends or eventually human health and wellbeing threats.

(ii) To quantify the distribution of the potential for gene body methylation based on the prevalence of CpG dinucleotides in the genome of each species.

(iii) Based on differential likelihood of gene body methylation, to elaborate on the differential ability of invertebrate species to respond or potentially adapt to challenging and/or highly fluctuating environments thought DNA methylation mediated phenotypic plasticity or mutation.

(iv) Based on a comparative approach of CpG prevalence in the genome of the analysed species, to conclude on the most favourable models to address gene body methylation within an ecotoxicological context towards the development of potential biomarkers of environmental exposure or effect.

2. MATERIAL AND METHODS

2.1. Species selection and justification

Considering the objectives of the present work, a set of invertebrate species was primarily selected to be part of the meta-analysis (see Table 3). This list is the result of the application of several criteria as follows, which were then refined and carefully reasoned to reach a final list for analysis.

Because the meta-analysis intended would require calculating and quantifying the probability of gene body methylation and the methylation of other genomic elements, access to the genomic data for each species would be a mandatory criterion. Therefore, a first selection criterion was the availability of a sequenced and annotated genome. In parallel, and considering the wideness of the study's focal field, the final set of species selected for the meta-analysis would need to cover different environmental compartments. In particular, the significance of gene body methylation as a potential ecotoxicological tool is a major focusof the study, so that an appraisal of the meaning of this endpoint for environmental monitoring can be made on the basis of the results. As an ecotoxicological assessment is made over different environmental compartments using different invertebrate species as indicators, the impact of the meta-analysis outcome would be severely compromised by a short selection covering a single environmental compartment.

The application of the two criteria above would result in a too long list of species which would be a priori unmanageable for the purposes of the study and considering its timeframe. So, in order to narrow down the list into a manageable size, several other criteria were reasoned. In this way, when a large array of species potentially representing a given environment compartment was available from the previous selection stage, further selection favoured (i) species that have epigenetic data available, as such data could assist the discussion of the results; (ii) species that are recognised as models, especially in ecotoxicology; (iii) species with economic and/or medical importance; or, finally, (iv) species that can be directly compared with a species already selected to represent another environmental compartment - for example, there are invertebrate groups that bear freshwater and saltwater counterparts (e.g. bivalves), thus providing the opportunity to

address whether the hosting environment is stronger than phylogenetic proximity in constraining methylation tendencies.

Table 3. List of species selected, according to the first two criteria as specified in the text, for analyses. a – removed for lack of appropriate files; b – removed to reduce list.

Species	Taxonomic Position	Environmental compartment	Ecological Position	Additional Information
[b]*Acanthoscurria geniculata*	Arachnida-Araneae	Soil	Consumer	
[b]*Aedes aegypti*	Insecta-Diptera	Air	Consumer	Disease vector
Aedes albopictus	Insecta-Diptera	Air	Consumer	Disease vector
[b]*Alatina alata*	Cubozoa-Carybdeida	Marine	Consumer	
Anopheles albimanus	Insecta-Diptera	Air	Consumer	Disease vector
Apis mellifera	Insecta-Hymenoptera	Air	Consumer	
[b]*Argiope bruennichi*	Arachnida-Araneae	Soil	Consumer	
Biomphalaria glabrata	Gastropoda-Basommatophora	Freshwater	Consumer	
Bombus terrestris	Insecta-Hymenoptera	Air	Consumer	
[b]*Bombyx mandarina*	Insecta-Lepidoptera	Air	Consumer	
Bombyx mori	Insecta-Lepidoptera	Soil	Consumer	
Caenorhabditis elegans	Secernentea-Rhabditida	Soil	Consumer	Model in Ecotox/epigenetics
Camponotus floridanus	Insecta-Hymenoptera	Soil	Consumer	
[b]*Caridina multidentata*	Malacostraca-Decapoda	Freshwater	Consumer	
[b]*Ceratophysella communis*	Entognatha-Poduromorpha	Soil	Detritivore	
[b]*Cherax destructor*	Malacostraca-Decapoda	Freshwater	Consumer	
[b]*Cherax quadricarinatus*	Malacostraca-Decapoda	Freshwater	Consumer	
[b]*Chionoecetes opilio*	Malacostraca-Decapoda	Marine	Consumer	
[a]*Chironomus riparius*	Insecta-Diptera	Freshwater	Consumer	
[b]*Chloridea virescens*	Insecta-Lepidoptera	Air	Consumer	
Ciona intestinalis	Ascidiacea-Enterogona	Marine	Consumer	
[a]*Corbicula flumínea*	Bivalvia-Venerida	Freshwater	Consumer	
[b]*Craspedacusta sowerbii*	Hydrozoa-Limnomedusae	Freshwater	Consumer	
Crassostrea gigas	Bivalvia-Ostreida	Marine	Consumer	Model in Epigenetics
Daphnia magna	Branchiopoda-Cladocera	Freshwater	Consumer	Model in Ecotox/epigenetics
Dreissena polymorpha	Bivalvia-Veneroida	Freshwater	Consumer	Invasive species
[a]*Dreissena rostriformis*	Bivalvia-Myida	Freshwater	Consumer	
Drosophila melanogaster	Insecta-Diptera	Air	Consumer	Model in many fields
[b]*Dysdera silvatica*	Arachnida-Araneae	Soil	Consumer	
[a]*Eisenia fetida*	Clitellata-Opisthopora	Soil	Detritivore	
[b]*Eriocheir sinensis*	Malacostraca-Decapoda	Freshwater	Consumer	
Folsomia candida	Entognatha-Poduramorpha	Soil	Detritivore	Model in Ecotoxicology
[b]*Galleria mellonella*	Insecta-Lepidoptera	Air	Consumer	
Harpegnathos saltator	Insecta-Hymenoptera	Soil	Consumer	
[b]*Helicoverpa armigera*	Insecta-Lepidoptera	Air	Consumer	
[b]*Helicoverpa zea*	Insecta-Lepidoptera	Air	Consumer	
[a]*Holacanthella duospinosa*	Entognatha-Poduromorpha	Soil	Detritivore	
Ixodes scapularis	Arachnida-Ixodida	Soil	Consumer	Disease vector
[a]*Limnoperna fortunei*	Bivalvia-Mytilida	Freshwater	Consumer	
[b]*Lipothrix lubbocki*	Entognatha-Symphypleona	Soil	Detritivore	
Litopenaeus vannamei	Malacostraca-Decapoda	Marine	Consumer	

[b]*Macrobrachium nipponense*	Malacostraca-Decapoda	Freshwater	Consumer	
[a]*Magallana hongkongensis*	Bivalvia-Ostreida	Marine	Consumer	
[a]*Margaritifera margaritifera*	Bivalvia-Unionida	Freshwater	Consumer	
[a]*Megalonaias nervosa*	Bivalvia-Unionida	Freshwater	Consumer	
[b]*Mesaphorura yosii*	Entognatha-Poduromorpha	Soil	Detritivore	
[b]*Mesobuthus martensii*	Arachnida-Scorpiones	Soil	Consumer	
Mytilus galloprovincialis	Bivalvia-Mytiloida	Marine	Consumer	Invasive species
Nasonia vitripennis	Insecta-Hymenoptera	Air	Consumer	
Nematostella vectensis	Anthozoa-Actiniaria	Marine	Consumer	
[b]*Nephila clavipes*	Arachnida-Araneae	Soil	Consumer	
Octopus bimaculoides	Cephalopoda-Octopoda	Marine	Consumer	
[b]*Oncopodura yosiiana*	Entognatha-Oncopoduridae	Soil	Detritivore	
Orchesella cincta	Entognatha-Orchesellidae	Soil	Detritivore	
[b]*Palaemon carinicauda*	Malacostraca-Decapoda	Marine	Consumer	
[b]*Pandalus platyceros*	Malacostraca-Decapoda	Marine	Consumer	
[b]*Papilio bianorc*	Insecta-Lepidoptera	Air	Consumer	
[b]*Papilio polytes*	Insecta-Lepidoptera	Air	Consumer	
[b]*Paralithodes platypus*	Malacostraca-Decapoda	Marine	Consumer	
[b]*Parasteatoda tepidariorum*	Arachnida-Araneae	Soil	Consumer	
[b]*Penaeus chinensis*	Malacostraca-Decapoda	Marine	Consumer	
[b]*Penaeus japonicus*	Malacostraca-Decapoda	Marine	Consumer	
[b]*Penaeus monodon*	Malacostraca-Decapoda	Marine	Consumer	
[a]*Physella acuta*	Gastropoda-Physidae	Freshwater	Consumer	
Polypedilum vanderplanki	Insecta-Diptera	Air	Consumer	
[b]*Portunus trituberculatus*	Malacostraca-Decapoda	Marine	Consumer	
[a]*Potamilus streckersoni*	Bivalvia-Unionida	Freshwater	Consumer	
[a]*Procambarus virginalis*	Malacostraca-Decapoda	Freshwater	Consumer	
[b]*Pseudachorutes palmiensis*	Entognatha-Poduromorpha	Soil	Detritivore	
[b]*Pseudobourletiella spinata*	Entognatha-Symphypleona	Soil	Detritivore	
[b]*Pygmarrhopalites habei*	Entognatha-Symphypleona	Soil	Detritivore	
[b]*Sinella curviseta*	Entognatha-Entomobryomorpha	Soil	Detritivore	
[a]*Sinonovacula constricta*	Bivalvia-Adapedonta	Freshwater	Consumer	
[b]*Sminthurides bifidus*	Entognatha-Symphypleona	Soil	Detritivore	
[b]*Stegodyphus mimosarum*	Arachnida-Araneae	Soil	Consumer	
Strongylocentrotus purpuratus	Echinoidea-Echinoida	Marine	Consumer	
Tetranychus urticae	Arachnida-Trombidiformes	Soil	Consumer	Disease vector
[b]*Thalassaphorura encarpata*	Entognatha-Poduromorpha	Soil	Detritivore	
Tigriopus californicus	Hexanauplia-Harpacticoida	Marine	Consumer	
[a]*Tigriopus japonicus*	Hexanauplia-Harpacticoida	Marine	Consumer	
[b]*Tomocerus minor*	Entognatha-Entomobryomorpha	Soil	Detritivore	
[b]*Tomocerus qinae*	Entognatha-Entomobryomorpha	Soil	Detritivore	
Tribolium castaneum	Insecta-Coleoptera	Soil	Consumer	Model in genetics
[b]*Tropilaelaps mercedesae*	Arachnida-Mesostigmata	Soil	Consumer	
[a]*Venustaconcha ellipsiformis*	Bivalvia-Unionida	Freshwater	Consumer	

The detailed reasoning for the inclusion of each species in the final short-list used for analysis is explored in detail in the sub-sections bellow. However, it is important to remark that there were species which were included in a first stage as they fulfilled all criteria and covered our rationale but were then excluded from the analysis for practical reasons. Namely, in the case of *Chironomus riparius, Corbicula flumínea, Dreissena rostriformis, Eisenia fetida, Holacanthella duospinosa, Limnoperna fortunei, Magallana hongkongensis, Margaritifera margaritifera, Megalonaias nervosa, Physella acuta, Potamilus streckersoni, Procambarus virginalis, Sinonovacula constricta, Tigriopus japonicus* and *Venustaconcha ellipsiformis*, exclusion was due to a deficient accessibility to the appropriate .gff files (see below for further details on the need for these files). In such a case, replacement by an appropriate alternative species was attempted whenever possible. Several other species were excluded from analyses in order to make the list more manageable, whenever a given group of organisms could be represented by many species. This was the case for example for *Ceratophysella communis, Lipothrix lubbocki, Mesaphorura yosii, Oncopodura yosiiana, Pseudachorutes palmiensis, Pseudobourletiella spinata, Pygmarrhopalites habei, Sinella curviseta, Sminthurides bifidus, Thalassaphorura encarpata, Tomocerus minor and Tomocerus qinae*, which were all removed leaving *Folsomia candida* and *Orchesella cincta* as the representatives of the springtails (these two species were distinguished also considering that epigenetic studies are available for them). The same scenario was in place regarding the representatives of arachnids, and those already epigenetically studied *Ixodes scapularis* and *Tetranychus urticae* were selected among the list of species fitting the requirements: *Acanthoscurria geniculata, Argiope bruennichi, Dysdera silvatica, Mesobuthus martensii, Nephila clavipes, Parasteatoda tepidariorum, Stegodyphus mimosarum* and *Tropilaelaps mercedesae.* Among the marine shrimps and crabs *Litopenaeus vannamei vannamei* and *Tigriopus californicus* were chosen over *Caridina multidentata, Cherax destructor, Cherax quadricarinatus, Chionoecetes opilio, Eriocheir sinensis, Macrobrachium nipponense, Palaemon carinicauda, Pandalus platyceros, Paralithodes platypus, Penaeus chinensis, Penaeus japonicus, Penaeus monodon* and *Portunus trituberculatus. Bombyx mori* was chosen over *Bombyx mandarina, Chloridea virescens, Galleria mellonella, Helicoverpa armigera, Helicoverpa zea, Papilio bianorc* and

Papilio polytes for its higher economic relevance and because it is a recognised validation species (see section 2.1.1. below). The jellyfish *Alatina alata* and *Craspedacusta sowerbii* were removed to make the list more manageable and because of their lack of economic, medical or epigenetic relevance as recognised in the literature. Lastly, *Anopheles albimanus* was chosen over *Aedes aegypti* as both fulfilled criteria but both are mosquitoes, thus redundancy in the analysis was avoided.

2.1.1. Validation species

In addition to meeting the criteria established for species selection (see above), four species (*Apis mellifera, Bombyx mori, Ciona intestinalis, Nematostella vectensis*) were forced inclusion in the analysis because they were addressed in a previous inspiring work (Sarda et al. 2012). Thus, these species also represent the means for the validation of our work through the comparison of data with previously published outcomes.

The western honeybee (*A. mellifera)*, also known as the European honeybee is the most common of the honeybees worldwide (Engel 1999; Lo et al. 2010). Western honeybees are an important model organism used in various scientific fields like social evolution, learning, memory and pesticide toxicity to non-target beneficial species (Denison and Raymond-Delpech 2008; Bloch 2010; Menzel 2012; Christen et al. 2016). This honeybee is also of high direct economic importance as it is the primary species maintained by beekeepers for both honey production and pollination (vanEngelsdorp and Meixner 2010). The western honeybee was the first insect to be discovered with a functional DNA methylation system (Wang et al. 2006), with the average CpG methylation for the whole genome being of about 1% (Harris et al. 2019). There are many studies focusing on gene methylation in *A. mellifera,* several of which regarding methylation differences constraining bee casts (Elango et al. 2009; Lyko et al. 2010; Wang et al. 2020). The intergenerational epigenetic inheritance of DNA methylation marks has also been demonstrated, indicating that there is no DNA methylation reprogramming in bees during embryogenesis and that DNA methylation marks are stably transferred between generations (Yagound et al. 2020). Several environmental pressures act on this species like exposure to pesticide cocktails, parasitic and infectious agents, habitat loss, beekeeping practices and climate change or decreased

abundance and diversity of floral resources (Goulson et al. 2015). Most studies are focused on abiotic pressures and the majority of these regard pesticides (Havard et al. 2019).

B. mori is a Bombycidae insect that is generally known as the domestic silk moth. The silkworm is the larval or caterpillar form of the silk moth, and it is very important economically because it is the main producer of silk (Rahmathulla 2012; Torres et al. 2021). The domestic silk moth is dependent on human intervention in order to reproduce, after the selective breeding over millennia; this species is very different from the wild silk moths (distinct *Bombyx* species), which are not as commercially viable, kept flying abilities and pigmentation (Torres et al. 2021). *B. mori* has, like most insects, a low methylation level of around 0.11% of genomic cytosines (Xiang et al. 2010) and there are some epigenetic studies on this species (Chen et al. 2020; Xu et al. 2020a; Gao et al. 2020). Most studies in the realm of environmental stress are related to oxidative stress, heat stress and address metal pollution (Yuan et al. 2016; Mandyam.D. and Muthangi 2020; Punyavathi and Manjunatha 2020; Parenti et al. 2020; Xu et al. 2020b).

C. intestinalis, also known as the vase tunicate, is an ascidian with a very soft tunic. It's a widely distributed species that has been employed as a model chordate in developmental biology and genetics (Satoh 2003; Satoh et al. 2003; Satou et al. 2005; Gallo and Tosti 2015). These sea squirts are known as invertebrate chordates as they may be the closest invertebrate relatives of vertebrates as suggested by molecular phylogenetic studies. It has a very small genome but has a gene corresponding to almost every family of genes in vertebrates (Dehal et al. 2002) and has a relatively large percentage of cytosine methylation compared to other invertebrates of around 30% (Suzuki et al. 2016). *C. intestinalis* is an invasive species that may form dense aggregations on any floating or submerged substrate, particularly in manmade constructions such as pilings, floats, and boat hulls, the larvae being transported by bilge or ballast water discharge (McDonald 2004; Micael et al. 2020). Studies on the ecotoxicological responses of the species are mostly focused on heat, oxidative stress and metal contamination (Fujikawa et al. 2010; Serafini et al. 2011; Franchi et al. 2012; Ferro et al. 2013; Gallo et al. 2016).

N. vectensis, commonly known as the starlet sea anemone is a species belonging to the Edwardsiidae family. Native to the east coast of the United States of America, its habitat is

the shallow and brackish water of coastal lagoons and salt marshes where it is usually buried in the mud with its tentacles exposed (Hand and Uhlinger 1994). Because of its ease of handling in the laboratory and high reproductive rates following gametogenesis induction, *N. vectensis* is commonly used as a model organism for the study of evolution, genomics, reproductive biology, developmental biology, and ecology (Darling et al. 2005; Genikhovich and Technau 2009; Watson et al. 2009). Genome analysis has been showing a remarkable degree of similarity in the gene sequence conservation and complexity between this sea anemone and vertebrates (Genikhovich and Technau 2009). Studies within the field of ecotoxicological responses address mostly oxidative stress and metal contamination (Tarrant et al. 2014; Elran et al. 2014; Friedman et al. 2018).

2.1.2. Non-validation terrestrial species

Eight species were considered for the meta-analysis that are intended to represent the soil compartment, namely *Caenorhabditis elegans, Camponotus floridanus, Harpegnathos saltator, Folsomia candida, Orchesella cincta, Ixodes scapularis, Tetranychus urticae* and *Tribolium castaneum.*

C. elegans is a translucent free-living nematode with a maximum length of around 1 mm that dwells in temperate soil habitats (Kenyon 1988). *C. elegans* has been used as a model organism in various research fields of molecular and developmental biology (Tejeda-Benitez and Olivero-Verbel 2016; Shen et al. 2018; Queirós et al. 2019; Goldstein and Nance 2020). It was the first multicellular organism whose entire genome was sequenced, and the only one whose connectome had been finished (The C. elegans Sequencing Consortium 1998; Cook et al. 2019). Several aspects facilitate its use in a laboratory environment like its transparency, the clear distinction of male and female, the low costs of production and possibility of long-term storage (Brenner 1974). *C. elegans* seems to have low to undetectable levels of CpG methylation (Suzuki et al. 2007). A single study has reported a low-level DNA methylation in *C. elegans* of 0.0019–0.0033% of cytosine methylated (Hu et al. 2015). Due to its prevalence as a model organism for so long, the type of studies that exist for its responses to environmental stress are varied, focusing on the effects of

pesticides, metals, nanoparticles, drugs, toxins and various other chemicals (Shen et al. 2009; Anbalagan et al. 2013; Moore et al. 2014; Starnes et al. 2015; Yu et al. 2015; Hu et al. 2015).

Due to its great size and prominent colour*, C. floridanus,* sometimes known as the Florida carpenter ant, is the most well-known ant species in Florida. Workers and queens are bicolored, having a head with a reddish-orange colour and a bright to dullish orange coloured mesosoma and legs and a deep black gaster. The usual length of a queen is 14–16 mm, and the highly polymorphic workers can reach similar lengths, with the largest majors reaching 11–13 mm. Minors and mediae are typically around 4–9 mm (Buckley 1867). The level of cytosine methylation in *C. floridanus* is around 0.14% to 0.16% (Bonasio et al. 2012). There appears to be no studies reporting its direct response to environmental stress.

H. saltator is also called by the name of the Indian jumping ant or Jerdon's jumping ant. It is an ant species found in India that has the ability to leap, being able to jump up to 2 cm high and 10 cm far. They are able to accomplish these leaps by synchronized abduction of the middle and hind pairs of legs, allowing them to escape and also catch flying prey (Urbani et al. 1994). They are encountered in small groups or as solitary individuals, unlike other ant species, with colonies comprising of only a few individuals; another remarkable trait is that the queen-worker distinction is quite limited, and some workers can mate and lay fertilized eggs just like the queen (Peeters et al. 2000). This renders the species interesting in the field of epigenetics for comparison with social ants with more defined casts that can possibly be constrained by differential methylation patterns. Such as *C. floridanus, H. saltator* has a low level of cytosine methylation of 0.11% to 0.12% (Bonasio et al. 2012). There only 2 studies that are in the realm of stress responses, one relating to social stress (Schneider et al. 2016) and another relating to longevity (Schneider et al. 2011).

F. candida is a springtail of the Isotomidae family found worldwide. It has an unpigmented, slender body up to 3 mm in length, it reproduces by parthenogenesis, and it is a model organism in ecotoxicological research. Its original distribution is unclear as it spread around the world in soil and in plant pots. Normally, *F. candida* occurs in sites with a high organic content, in farms, in leaf litter in forests and on-stream verges. By reproducing by

parthenogenesis, its population consists of only females (Fountain and Hopkin 2005). The nature of living in underground habitats allows them to tolerate higher levels of carbon dioxide, for example they are able to survive 10% of carbon dioxide for more than six weeks (Zinkler and Platthaus 1996). *F. candida* is simple to culture within the laboratory on granulated baker's yeast and it is extensively used in ecotoxicology to assess negative effects of insecticides and various other soil contaminants; in addition, *F. candida* is frequently used as a model to address cold tolerance, carbon dioxide tolerance and the outcomes of its grazing activities on soil fungi and mycorrhizae developing in roots, the decomposition of leaf litter, circadian rhythms, palatability and nutritious value of various plant residues and species of fungi, population dynamics and behavioural responses (Hopkin 1997; van Straalen et al. 1997; Fountain and Hopkin 2005; Maria et al. 2014; Szabó and Bakonyi 2017; Noordhoek et al. 2018; Szabó et al. 2019; Jensen et al. 2020; Dai et al. 2020). *F. candida* appears to have an absence of both total cytosine and locus-specific CpG methylation according to (Noordhoek et al. 2018).

O. cincta is a springtail species found in North America and Europe. It grows to a length of 4 millimetres on average, which is substantially longer than most springtails that rarely exceed 1 mm. They live among leaf litter and soils and have a high metabolic rate and a high fertility rate, being more mobile than many springtail species. The reproduction can be postponed until conditions are favourable and synchronised across the population as a strategy to allow this species to make the most of the available resources (Newbould et al. 1986). *O. cincta*, like all collembolans, seems to lack DNA methylation along with DNA methyltransferases (Provataris et al. 2018). The responses of this species to environmental stress are not remarkably known, but there are studies regarding temperature and metal stress (Bahrndorff et al. 2006; ROELOFS et al. 2009; Zizzari and Ellers 2011; Costa et al. 2012; Noer et al. 2020).

I. scapularis, commonly known as the deer tick, is a hard bodied tick found in the eastern and northern Midwest of the United States, as well as in south-eastern Canada (Dennis et al. 1998). It parasitizes the white-tailed deer, but it is also known to parasitize other species like mice, lizards and migratory birds (Mannelli et al. 1994; Levine et al. 1997; Ogden et al. 2008). Importantly, it is a vector of several diseases to humans like the Lyme disease,

babesiosiss and anaplasmosis (Brownstein et al. 2005; Gulia-Nuss et al. 2016). Although DNA methylation is yet to be confirmed on this species, the presence of DNA methyltransferases has been found in the species of the same genus *Ixodes ricinus,* thus it is speculated that the enzymes exist as well in *I. scapularis* (Kotsarenko et al. 2020). There is one study addressing *I. scapularis* survival under different environmental conditions (Ginsberg et al. 2017), but there are many focused on the interplay between environmental stress and its activity as a parasite and a disease vector (Busby et al. 2012; Villar et al. 2015; Bourret et al. 2016).

T. urticae is a plant-feeding mite that is generally considered to be a pest. It is the most well-known member of the Tetranychidae family, which is native to Eurasia but now has a global distribution range. It is extremely small, with adult females measuring about 0.4 mm long (Horn 1994; Raworth et al. 2001), and polyphagous, i.e., it can feed on hundreds of different plants, including most vegetable and food crops (over 1100 plant species that belong to more than 140 different plant families). This fact reveals the economic importance of this pest. On its own, a single *T. urticae* individual wouldn't do much damage but they attack in the numbers of hundreds or thousands, which significantly damages the plant's capacity to survive (Attia et al. 2013; Rincón et al. 2019). *T. urticae* has a reported level of global methylation of around 25 to 30%, with a cytosine methylation level of 15% to 20% (Yang et al. 2018b). There are a couple studies relating to UV stress and predation stress, but the majority of the studies investigate on how stress improves plant resistance to the pest (Aucejo-Romero et al. 2004; Ximénez-Embún et al. 2017; Murata and Osakabe 2017; Yoshioka et al. 2018; Li and Zhang 2019).

T. castaneum, or the red flour beetle, is a beetle of Tenebrionidae family that grows up to around 3-4 mm long and bears a uniform rust, brown or black colour (Newell 1936). Having an Indo-Australian origin, the species has now a worldwide distribution, following human commercial trading routes, as a pest attacking stored grain and other food products including flour, cereals, pasta, biscuits, beans, and nuts (FAO 2001). Along with *Tribolium confusum, T. castaneum* is the most common secondary pest of all plant commodities in store throughout the world. *T. castaneum* is also a model organism for ethological and food safety research, as well in development and functional genomics (Richards et al. 2008;

Grünwald et al. 2013). The existence of a DNA methylation system in *T. castaneum* hasn't been agreed on, and while some studies report the existence of methylation (Feliciello et al. 2013; Song et al. 2017), others report its absence (Zemach et al. 2010; Schulz et al. 2018). The studies that have been made regarding environmental stress responses address mostly temperature as a stressor and oxidative stress as recorded endpoint (Scharf et al. 2016; Pajaro-Castro et al. 2019; Jiang et al. 2019).

2.1.3. Non- validation aerial species

Six species were selected for the meta-analysis representing aerial invertebrates. These are exclusively insects and the group joined species with flying adult stages of considerable length: *Aedes albopictus*, *Anopheles albimanus, Bombus terrestris, Drosophila melanogaster, Nasonia vitripennis* and *Polypedilum vanderplanki.*

The Asian tiger mosquito (*A. albopictus* or *Stegomyia albopicta*) belongs to the Culicidae mosquito family and is also known as the forest mosquito. This mosquito, which has white stripes on its legs and body (Skuse 1894) and is endemic to Southeast Asia's tropical and subtropical zones, has spread to many nations via international trade and human travel routes, conquering every continent except Antarctica in the last 30–40 years (Benedict et al. 2007; Bonizzoni et al. 2013; Kraemer et al. 2015). It has become an epidemiologically important vector in many communities for different pathogens, including those promoting yellow fever, dengue fever, and Chikungunya fever, as well as several filarial nematodes such as *Dirofilaria immitis* and the Zika virus (Cancrini et al. 2003; Bonilauri et al. 2008; Rezza 2012; Kamgang et al. 2019; McKenzie et al. 2019). Levels of global DNA methylation in this species were early reported to be around 0.03% (Adams et al. 1979), but more recently a wide range was found for global methylation in this species of 0.021% to 0.674% methylated cytosine relative to guanine (Oppold et al. 2015). Environmental stress studies with this species mostly fall within the category of temperature stress studies (da Gloria da Costa Carvalho and Freitas 1988; Sivan et al. 2017; Reinhold et al. 2018).

A. albimanus is another Culicidae mosquito. It is native of coastal areas up to 500 meters altitude in Central and South America, the Caribbean, and Mexico, but has been dispersing

widely due to its generalist characteristics (Pinault and Hunter 2012; Gómez et al. 2014; Cázares-Raga et al. 2014). *A. albimanus* is an important species for the humankind as it is the vector of *Plasmodium vivax*, which causes malaria (Chareonviriyaphap et al. 1997; Martínez-Barnetche et al. 2012). With global methylation levels between 0.85% and 1.73%, *A. albimanus* meets the generally expected methylation levels of invertebrates (Claudio-Piedras et al. 2020). Most studies that have been carried out with this mosquito species relate to its control rather than address directly negative effects of environmental stressors.

B. terrestris, often known as the buff-tailed bumblebee or the large earth bumblebee, is one of Europe's most common bumblebee species and one of the most important pollinator species, thus it may be found in a wide range of non-native habitats (Dafni et al. 2010) being considered an invasive species in many temperate areas (Torretta et al. 2006; Inoue et al. 2008). Nests are usually found underground forming a comb-like structure where eggs are contained (Duchateau and Velthuis 1988). *B. terrestris* is of economic importance since it is bred commercially for use in pollination of greenhouse crops; the global trade of bumblebee colonies likely exceeds 1 million nests per year (Dafni et al. 2010; Nation 2010). In the wild, *B. terrestris* exhibits greater nest growth in suburban areas rather than in farmland because of the plant diversity available in suburban gardens (Goulson et al. 2002). There is a population decline in many bumblebee species due to agriculture but *B. terrestris* is still widespread, which is likely because it can forage through very long distances thus rendering changes in biodiversity and environment less important for their success (Goulson et al. 2002). Records of methylation of 0.5 to 0.6% of all CpGs are reported for the *B. terrestris* genome (Sadd et al. 2015). The effects of a wide range of environmental stressors like insecticides, fungicides, parasites, nutritional aspects and temperature have been studied in this species (Owen et al. 2013; Bishop et al. 2016; Roger et al. 2017; Wang et al. 2019a; Botías et al. 2021).

D. melanogaster belongs to the family Drosophilidae, and it is often referred to as the fruit fly. It is used as a model organism widely and in various fields like genetics, physiology, microbial pathogenesis, and life-history evolution, owing to the fact that it has a rapid life cycle, simple genetics (with only four pairs of chromosomes), large number of offspring per

generation and cheap maintenance in the laboratory (Roberts 2006). It can be found worldwide, and it is often seen as a common pest concerning food safety. In the genome of fruit flies, around 75% of known human illness genes have an identifiable match, and 50% of fly protein sequences have mammalian homologs (Reiter 2001). Environmental factors can influence several developmental aspects in *D. melanogaster*. Fruit flies raised in hypoxia, for example, have shorter thorax lengths, while hyperoxia causes smaller flying muscles (Harrison et al. 2018). *D. melanogaster* has a very low level of methylation, with 0.001% of cytosines being generally methylated (Deshmukh et al. 2018). Due to its role as a model species, there are many studies on the effects of different environmental stressors in *D. melanogaster*, addressing pressing factors ranging from temperature to nanotoxicity and desiccation, and recording on different endpoints (Ong et al. 2015; Thorat et al. 2016; Garcia and Teets 2019; Jayapalan et al. 2020; Oliveira et al. 2021).

N. vitripennis is one of the four known species of the genus *Nasonia.* It is a small parasitoid wasp that affects the larvae of parasitic carrion flies such as blowflies and flesh flies. It is known as the most widely and best studied of the parasitoid wasps and has a cosmopolitan distribution (Whiting 1967). Due to its parasitic behaviour, the wasp has been of interest to study the development of biopesticides and biological systems for controlling unwanted insects (Zhang et al. 2005). *N. vitripennis* has been also used for comparison with other hymenopterans, most often *A. mellifera,* regarding the study of biosynthetic pathways of sex pheromones, determination of sex in development and many proteins and gene products (Abdel-latief et al. 2008). Its genome has been sequenced and released since 2010, and it is known that *N. vitripennis* has an average global methylation level of 0.18% over the whole genome and a level 0.63% among CpGs (Beeler et al. 2014). Only studies relating to heat stress are available regarding the impacts of environmental conditions in this species (Wang et al. 2012; Chirault et al. 2015; Hidalgo et al. 2019).

P. vanderplanki, often known as the sleeping chironomid, is a dipteran that lives in African semi-arid regions and belongs to the Chironomidae family. The larval state of *P. vanderplanki* usually develops under dry conditions since nests are generally placed in the mud, at the bottom of temporary pools that frequently dry out. During the larval stage, water content can reach levels as low as 3% rendering the organism a high tolerance to

extreme environmental conditions like very high temperatures, gamma-rays, and vacuum (Hinton 1960); this is indeed one of few metazoans that can withstand nearly complete desiccation or anhydrobiosis. *P. vanderplanki* is also one of the most cold-tolerant insect species, being for example able to survive liquid helium (−270 °C) for up to 5 min (Hinton 1960). There are apparently no studies regarding methylation levels in this species, and studies considering environmental stressors regard more specifically the species tolerance to heat, radiation and desiccation stress (Watanabe et al. 2006a, b; Gusev et al. 2011; Cornette and Kikawada 2011; Mazin et al. 2018).

2.1.4. Non-validation freshwater species

Two molluscs and a crustacean were assembled to represent freshwater ecosystems in the meta-analysis, namely the snail *Biomphalaria glabrata*, the bivalve *Dreissena polymorpha* and the cladoceran *Daphnia magna*.

The snail *B. glabrata* is a Planorbidae freshwater pulmonated gastropod mollusc that lives in small streams, ponds, and marshes. It is a Neotropical species native to South America and the Caribbean (Pointier et al. 2005). When removed from their watery environment or when the ecosystem dries out, these snails may survive in aestivation for a few months (Majoros et al. 2008). It is also the intermediate host for the trematode *Schistosoma mansoni*, infecting about 83.31 million people worldwide, and this is why is the species is economically important (Crompton 1999; Mitta et al. 2017) *B. glabrata* is an important model organism for studying the interactions between parasites and their hosts (Lockyer et al. 2008). About 2% of the cytosines are methylated in the snail's genome (Fneich et al. 2013). Stress response studies with this species include the appraisal of the effects of heat and metal contaminants, but also relate environmental conditions to the infection by *S. mansoni* due to its importance in human health (Ittiprasert et al. 2009; Knight et al. 2016; da Silva Cantinha et al. 2017; Niederwanger et al. 2017; Allan and Blouin 2020).

D. polymorpha, commonly known as the zebra mussel, is a freshwater mussel that can grow to a maximum length of around 5.1 cm, filter-feeding from the water column (Reeders and Bij de Vaate 1990). They can attach to most substrates and are prone to build colonies on

native unionid clams; this feature, together with their superior filtering abilities have been driving the elimination of other freshwater bivalves (Schloesser and Nalepa 1994; Karatayev et al. 1997). For this reason, this native species of Russia and Ukraine has become a feral invader worldwide, for example in North America, Great Britain, Ireland, Italy, Spain, and Sweden (Pollux et al. 2010; Ożgo et al. 2020). Besides their ecological effects as an invasive species, the zebra mussel is also a strong macrofouler that can promote severe economic damage in the water-dependent industry (Connelly et al. 2007). No information seems to exist regarding DNA methylation in *D. polymorpha*, but there is a considerable body of literature reporting on the responses of *D. polymorpha* to environmental stressors such as heating, metals, microplastics, and various other xenobiotics (Riva et al. 2012; Pedriali et al. 2013; Magni et al. 2017; Potet et al. 2018; Brand et al. 2019; Weber et al. 2020).

D. magna is a small planktonic crustacean (Order Cladocera) that inhabits freshwater ecosystems and is distributed throughout the Northern Hemisphere and South Africa (de GELAS and de MEESTER 2005). It is a model organism in several fields of research like in ecological and evolutionary studies, and particularly in ecotoxicology (Altshuler et al. 2011; Tkaczyk et al. 2021). *D. magna* has plastic phenotypic responses to diverse environmental challenges. For example, when in the presence of kairomones, *D. magna* develops protective structures like an elongated spine and a large body size (de Meester 1993). Its status as a model organism derives from research-facilitating features such as body transparency that allows the observation of its inner anatomy, parthenogenetic reproduction allowing the easy use of clonal organisms isolating genetic variation, fast generation time and easy, cheap and simple maintenance. In regulatory ecotoxicology, *D. magna* is recommended for several standard test methods (e.g., OECD 2004, 2008). This species has a global CpG methylation of 0.7–0.9% (Hearn et al. 2019), and its molecular responses to a wide range of environmental stressors have been conspicuously studied, using different endpoints at different levels of biological organization (Fan et al. 2015, 2020; Jeremias et al. 2018b; Yuxuan et al. 2019; Shahid et al. 2019; Hearn et al. 2019; Ates et al. 2020; Samanta et al. 2020; Ellis et al. 2020).

2.1.5. Non-validation marine species

Three molluscs, two crustaceans and one echinoderm were assembled to represent marine ecosystems in the meta-analysis, namely *Crassostrea gigas, Litopenaeus vannamei, Mytilus galloprovincialis, Octopus bimaculoides, Strongylocentrotus purpuratus* and *Tigriopus californicus.*

The Pacific oyster, *C. gigas* is native to the Pacific coast of Asia, more specifically Japan, and was introduced in North America, Australia, Europe, and New Zealand, mostly for commercial purposes. Aside from deliberate introductions, the Pacific oyster has spread by accident, through larvae either in ballast water or on ship hulls and is now considered an invasive species. They usually have a pale white or off-white shell and measure from 80 to 400 mm long as adults. *C. gigas* is primarily an estuarine species, but it also inhabits the intertidal and subtidal zones. When the ideal environment is unavailable, they have been observed to settle in muddy or sandy locations; in shallow waters they like to stick to hard or rocky surfaces, including other shells. The Pacific oyster is a temperature-tolerant species, capable of withstanding temperatures from −1.8 to 35 °C (Leffler and Greer 1991). It is now the world's most frequently farmed and commercially important oyster, thanks to its ease of cultivation, wide environmental tolerance, and ability to spread from one location to another. *C. gigas* has a genome where 15% of the CpG dinucleotides are methylated (Wang et al. 2021). Most studies about stress responses to environmental stressors regard heat and salinity variations (Zhao et al. 2012; Yang et al. 2016, 2017; Li et al. 2020; Elia et al. 2020; Wang et al. 2021).

L. vannamei, commonly known as the whiteleg shrimp or the Pacific white shrimp, can be found in the eastern Pacific Ocean where water temperature remains above 20 °C throughout the year. *L. vannamei* grows to a maximum length of 230 mm and lives at depths up to 72 m. *L. vannamei* is an economically important species as it is commonly caught or farmed for food. *L. vannamei* production reached 53% of the total production of farmed crustaceans in 2016 globally (Holthuis et al. 1998; FAO 2018). There are apparently no DNA methylation studies about *L. vannamei*, and the species responses to environmental stress have been studied mostly regarding the effects of temperature and

ammonia stress (Duan et al. 2017, 2021; Wang et al. 2019b; Fan et al. 2019; Liu et al. 2020; Bautista-Covarrubias et al. 2020; Zhuo et al. 2021).

The Mediterranean mussel, *M. galloprovincialis* is a Mytilidae bivalve growing to up to 140 mm shell length by filter-feeding from the water column. As the name may suggest, this species is native from Europe, particularly from Mediterranean coastal areas, but has spread to many parts of the world, where it can be considered an invasive species (Bownes and McQuaid 2006). Mussels are generally considered good bioindicators and biosentinels of ecosystem quality and their immune system is often inspected for such evaluation (Malagoli et al. 2008). The immune system of this mussel is influenced by changing environmental parameters such as water acidification, temperature and salinity variation, and consequently this species is often used as a model for predicting the impact of climate change on the health status of both wild and farmed organisms (Matozzo et al. 2012). The global methylation level reported for this species is of around 70% according to Ardura et al. (2018). Studies related to environmental challenges are in the realm of thermal and metal stress (Li et al. 2019b; Collins et al. 2020; Feidantsis et al. 2020; Braga et al. 2020; Giuliani et al. 2020; Santovito et al. 2021).

O. bimaculoides, commonly known as the California two-spot octopus, inhabits the Pacific Ocean in subtidal zones up to a depth of 20 m and tolerates a wide temperature range 15-26 °C. It is identifiable by the circular blue eyespots on each side of its head, which are false eye spots under each real eye. It can reach a mantle size of 17.5 cm with arms reaching up to 58 cm length. These octopuses live one to two years, and the end of their life is signalled by egg-laying in the female and senescence in both males and female (Roper et al. 2010; Villegas et al. 2014). Cephalopods in general are important for their potential use as models in the fields of developmental biology, evolution, neuroscience (Yekutieli et al. 2005; di Cristo 2013; Maldonado et al. 2019; Hanke and Kelber 2020). They are regarded as the most intelligent of the invertebrates, having well developed senses and large brains (Tricarico et al. 2014). For this uniqueness within the invertebrates' landscape, it was important to include a representative of this group in the analysis. There are no methylation studies in *O. bimaculoides* and also there are no studies about the effects of environmental stress on this species.

S. purpuratus, commonly known as the purple sea urchin, lives in lower intertidal and nearshore sub-tidal communities along the eastern edge of the Pacific Ocean. It normally grows to a diameter of about 10 cm (4 inches) and may live as long as 70 years (Workman 1999; Ebert and Southon 2003). The genome of *S. purpuratus* is comparable to that of vertebrates regarding its non-redundancy level but without the complexity of vertebrate genomes (Sodergren et al. 2006). Humans and *S. purpuratus* share about 7,700 genes, many involved in sensing the environment, which is somewhat surprising considering that the sea urchin lacks a head-like structure (Burke et al. 2006; Materna et al. 2006). The epigenome, gene expression, and phenotypic responses of the purple sea urchin are all affected by rising carbon dioxide levels. Carbon dioxide concentration also reduces the size of its larvae, which indicates a negative impact in the organisms fitness (Kelly et al. 2013; Doney et al. 2020). In *S. purpuratus*, around 22 % of CpGs are methylated (Strader et al. 2020). There are some studies reporting the effects of metal contamination, thermal stress, radiation stress and ocean acidification in this species (Hammond and Hofmann 2010; Adams et al. 2012; Evans and Watson-Wynn 2014; Klein et al. 2019; Garrett et al. 2020).

T. californicus is an intertidal copepod distributed along the Pacific coast of North America in splash pools where they can achieve high population densities from 800 to 20,000 individuals per litter (Dethier 1980; Powlik 1998). Due to the isolation of the splash pools from the ocean, salinity and temperature fluctuate significantly therein over the course of hours or days. *T. californicus* has the ability to thrive under these fluctuating environmental conditions (Dethier 1980). Populations of *T. californicus* show a striking pattern of genetic differentiation often exceeding 20% total sequence divergence in mitochondrial DNA between populations separated by as little as 500 m, suggesting that inter-population gene flow must be relatively rare for this copepod (Burton et al. 2007; Willett and Ladner 2009). In response to salinity variation, *T. californicus* changes the number of amino acids within its cells to maintain water balance, especially regarding proline, and this is thought to be a common mechanism of osmoregulation across crustaceans (Burton and Feldman 1982). No studies were found regarding methylation patterns in this species, while studies investigating the effects of thermal stress and salinity stress exist (DeBiasse et al. 2018; Li et al. 2019a; Tangwancharoen et al. 2020; Lee et al. 2021).

2.2. Data Processing

In this section, the flow of the data analysis towards achieving the specific objectives of this Dissertation is described. Occasionally, examples are given for one species while the same procedure was naturally applied to all species included in the meta-analysis.

2.2.1. Preparation of files sourcing bioinformatics analysis

The immediate goal of this data processing was to calculate the observed-to-expected CpG ratio (CpG O/E), the use of this metric is further explained in the results section. But to explain briefly CpG O/E is a good proxy when there is a lack of methylation data (Sarda et al. 2012). First, a data file containing 2 columns was required, one with the gene name and another one with the corresponding gene sequence. In order to create these data files, two files per species were first retrieved from the National Center for Biotechnology Information (NCBI) database: the FASTA file (.fna) containing gene sequences, and the General Feature Format (GFF) file (.gff) containing the corresponding gene names. The exception was for the files regarding *Dreissena polymorpha*, which were obtained from a different source (see table 4 below). In some cases, more than one genome assembly was available and contained the required files, thus the selection of the most suitable one was necessary. This was done based on the following criteria: I) Assemblies that had a higher level were chosen over others, e.g., chromosome level assemblies were chosen over those that were presented at the scaffold level; II) when assemblies presented the same level, several other relevant aspects were considered at the same time (usually in a case-by-case basis), including the number of scaffolds, genome coverage and number of citations. In this regard, it was considered that a high coverage indicates a high-quality assembly, and a high number of scaffolds indicates a fragmented assembly and of lower quality (Gurevich et al. 2013; Wajid and Serpedin 2016; Whibley et al. 2021). On the other hand, the number of citations was considered a less important criterion since some papers are more recent, therefore having less time to be cited. The final list of which assemblies were chosen and for what reason is provided in Table 4.

Table 4. List of the data used for each species and reason for selection of specific files.

Species	Link to genome assembly used in analysis	Reason for selection over other sources
Aedes albopictus	https://www.ncbi.nlm.nih.gov/assembly/GCA_001444175.2/	More reliable sequencing technology (Illumina) that has less errors at a base level; better coverage; more citations.
Anopheles albimanus	https://www.ncbi.nlm.nih.gov/assembly/GCA_013758885.1	NA
Apis mellifera	https://www.ncbi.nlm.nih.gov/assembly/GCF_003254395.2/	Less scaffold and higher coverage
Biomphalaria glabrata	https://www.ncbi.nlm.nih.gov/assembly/GCF_000457365.1/	Higher assembly level
Bombus terrestris	https://www.ncbi.nlm.nih.gov/assembly/GCF_000214255.1/	NA
Bombyx mori	https://www.ncbi.nlm.nih.gov/assembly/GCF_014905235.1/	Higher assembly level
Caenorhabditis elegans	https://www.ncbi.nlm.nih.gov/assembly/GCF_000002985.6/	NA
Camponotus floridanus	https://www.ncbi.nlm.nih.gov/assembly/GCF_000147175.1/	Higher coverage, more citations
Ciona intestinalis	https://www.ncbi.nlm.nih.gov/assembly/GCF_000224145.3/	NA
Crassostrea gigas	https://www.ncbi.nlm.nih.gov/assembly/GCF_902806645.1/	NA
Daphnia magna	https://www.ncbi.nlm.nih.gov/assembly/GCF_003990815.1/	Higher assembly level
Dreissena polymorpha	https://www.biorxiv.org/node/810487.external-links.html	NA
Drosophila melanogaster	https://www.ncbi.nlm.nih.gov/assembly/GCF_000001215.4/	NA
Folsomia candida	https://www.ncbi.nlm.nih.gov/assembly/GCF_002217175.1/	NA
Harpegnathos saltator	https://www.ncbi.nlm.nih.gov/assembly/GCF_003227715.1/	Less scaffolds
Ixodes scapularis	https://www.ncbi.nlm.nih.gov/assembly/GCF_000208615.1/	Higher assembly level
Litopenaeus vannamei	https://www.ncbi.nlm.nih.gov/assembly/GCF_003789085.1/	NA
Mytilus galloprovincialis	https://www.ncbi.nlm.nih.gov/assembly/GCA_900618805.1/	Less scaffold and higher coverage
Nasonia vitripennis	https://www.ncbi.nlm.nih.gov/assembly/GCF_009193385.2/	Less scaffold and higher coverage
Nematostella vectensis	https://www.ncbi.nlm.nih.gov/assembly/GCF_000209225.1/	NA
Octopus bimaculoides	https://www.ncbi.nlm.nih.gov/assembly/GCF_001194135.1/	NA
Orchesella cincta	https://www.ncbi.nlm.nih.gov/assembly/GCA_001718145.1/	NA
Polypedilum vanderplanki	https://www.ncbi.nlm.nih.gov/assembly/GCA_018290095.1/	NA
Strongylocentrotus purpuratus	https://www.ncbi.nlm.nih.gov/assembly/GCF_000002235.5/	NA
Tetranychus urticae	https://www.ncbi.nlm.nih.gov/assembly/GCF_000239435.1/	NA
Tigriopus californicus	https://www.ncbi.nlm.nih.gov/assembly/GCA_007210705.1/	NA
Tribolium castaneum	https://www.ncbi.nlm.nih.gov/assembly/GCF_000002335.3/	NA

NA – Not Applicable, i.e., only one source available for the species.

Following the download of the appropriate files for each species, the polishing of the data from the .gff files was operated by removing all data but those referring to the gene bodies. This step was manually achieved by opening the .gff files in MS®Excel and removing all rows except the ones with the “mRNA” tag. The polished files were saved as tab delimited files for downstream compilation of information on gene sequences contained in the FASTA files and the corresponding gene names contained in the .CSV (previously .GFF) files. The

bedtools toolset (version 2.30.0), one of the most used toolsets for genome arithmetic (Quinlan and Hall 2010), was used for the purpose following the prior installation of the full virtualizer VirtualBox (version 6.1.22) in Windows (this allowed Linux emulation to run the bedtools toolset). The code run in bedtools was as follows: $ bedtools getfasta -fi <FASTA> -bed <GFF> output.fasta, where <FASTA> is the name of the .fna file, <GFF> is the name of the .gff file and output.fasta is the output file were "output" should be replaced by the desired name for this file.

This command on Bedtools made the intersection of the genome FASTA file, that contained the gene sequences, with the GFF file, that contained the corresponding gene names, thus creating a single file with the information previously contained in two different files. The output files were further processed/edited for compliance with the requirements of later analysis where a new structural organization of the files was needed: all the names in one file and sequences in another, with full line correspondence. For the purpose, the "output" file was opened in notepad++ (version 7.9.2) and the ">" symbol present before every line being replaced with an empty space. Then two files were created from this "output", one named "speciesname_genesname" containing the names of the genes and a second entitled "speciesname_genesequence" containing the gene sequences. To create either of these files from the original it was necessary to either delete the odd or the even lines from the original text. This was achieved in two steps. First by finding with the code thread "([^\n]*\n)[^\n]*\n" and replacing with the code thread "$1", which made all the even numbered lines disappear; the odd lines were left, collecting the gene names only in a file. Second, in order to obtain the gene sequences, the previous step was repeated over the original "output" but adding a blank line as the first line in the file beforehand, which relocated gene sequences to the odd number lines; a file with gene sequences only was generated.

Further data analysis was run in R (https://www.r-project.org/) and R studio (https://www.rstudio.com/), versions 4.0.3 and 1.4.1717, respectively.

2.2.2. Bioinformatic analysis

Both the "speciesname_genesequence" and "speciesname_genesname" files were imported into R by selecting the "From text (readr)" option, deselecting "First row as Names" and choosing "character" as class. Then, in order to make data within the files uniform, the characters in the sequences were all turned into uppercase (Code Box 1).

Code Box 1. Lower-to-uppercase transformation

```
speciesname_genesequence$X1=toupper(speciesname_genesequence$X1)
```

Afterwards the "stringr" package was installed (Wickham 2019) and upload from the R library. The number of C, G and CGs per gene was then calculated, then divided by the total number of base pairs of each gene sequence (Code Box 2), using the file with the gene sequences.

Code Box 2. Counting and normalizing C, G and CG

```
speciesname_genesequence$C<-str_count(speciesname_genesequence$X1,pattern =
"C")/nchar(speciesname_genesequence$X1)

speciesname_genesequence$G<-str_count(speciesname_genesequence$X1,pattern =
"G")/nchar(speciesname_genesequence$X1)

speciesname_genesequence$CG<-str_count(speciesname_genesequence$X1,pattern =
"CG")/nchar(speciesname_genesequence$X1)
```

The Expected-to-observed CpG ratio (CpG O/E) was then calculated (Equation 1) for each gene sequence (Code Box 3). Afterwards, the file where genes were identified was integrated using the appropriate code (Code Box 4).

Equation 1: $\text{CpG O/E} = \frac{P_{CpG}}{P_{C}*P_{G}}$, where P_{CpG} is the frequency of CG dinucleotides, P_{C} is the frequency of C nucleotides and P_{G} is the frequency of G nucleotides.

Code Box 3. CpG O/E calculation

```
speciesname_genesequence$CpGOE<-
as.numeric(speciesname_genesequence$C)*as.numeric(speciesname_genesequence$G)

speciesname_genesequence$CpGOE<-
as.numeric(speciesname_genesequence$CG)/as.numeric(speciesname_genesequence$CpGOE)
```

Code Box 4. Integration of gene names

```
speciesname_genesequence$name<-speciesname_genesname;remove(speciesname_genesname)
```

2.2.3. Data analysis

CpG O/E data organized in a data frame (see above) were first inspected using a boxplot approach (Code Box 5). Boxplots were created to generally analyse the data distribution and to identify potential outliers in particular. CpG O/E values that were found to be too high when compared to most of the data for each species (this classification was qualitative and only applied when there was a clear discrepancy comparing to the general dataset). The identified outliers were removed to avoid data skewness (Code box 6, where "x" is the number of the line where the outlier stands, easily retrieved from the data frame).

Code Box 5. CpG O/E Boxplot building

```
boxplot(speciesname_genesequence$CpGOE)
```

Code Box 6. Outliers' removal

```
speciesname_genesequence<-as.data.frame(speciesname_genesequence[-c(x, ),])
```

For some species, missing values were found for the CpG O/E record, represented by "NaN". "NaN" is provided when a given gene sequence bears no Cs or Gs, which mathematically results in a division by zero for the CpG O/E calculation. Because text values

would prevent further numerical analysis, all "NaN" values were replaced by zero (Code Box 7).

Code Box 7. Correction of text CpG O/E values.

```
speciesname_genesequence$CpGOE[is.na(speciesname_genesequence $CpGOE)] <- 0
```

Following on the analysis above, visualization of CpG O/E distribution was improved through the creation of density plots using the ratio values although a histogram type of plot was selected; given that histograms artificially break continuous datasets such as the CpG O/E dataset into nominal classes, a density curve was added to each plot for consistency (Code Box 8).

Code Box 8. Histogram-type density plot and density line addition

```
hist(speciesname_genesequence$CpGOE, freq = F, xlab="CpG Observed/Expected", main = "Density plot of species name", breaks = seq(from=0, to=3.5, by=0.05))

lines(density(speciesname_genesequence$CpGOE), col="blue", lwd=2)
```

Prior to a systematic assessment of the type of distribution shown for CpG O/E by each species (see below), the corresponding CpG O/E records were exported from R studio in .txt files (Code Box 9).

Code Box 9. Export of gene sequences added CpG O/E

```
speciesname_genesequencefinal<-speciesname_genesequence[,c("CpGOE")]

write.table(speciesname_genesequencefinal,"speciesname.txt",col.names=F,row.names=F,sep="\t",quote=F)
```

The CpGO/E distributions were tested using the NOCOM program (Ott 1992; e.g., https://www.jurgott.org/linkage/nocom.htm). This was necessary because CpG O/E data can either reflect a unimodal or a bimodal distribution, and any estimations and conclusions made on the basis of such datasets are naturally dependent on the assumed distribution. This software estimates parameters like means, variance, proportions of components of a normal distribution or a mixture of normal distributions for independent observations, following the fitting of the data to both unimodal and bimodal distribution models. Under a null hypothesis reflecting the equality of two means reflecting a bimodal distribution, the G2 statistic applies (equation 2) approximately following a $\chi 2$ distribution with 2 degrees of freedom, to test whether a bimodal distribution provides a better fit to the data than a unimodal distribution (Hasselblad 1966; Ott 1979; Thode et al. 1988).

Equation 2: $G^2 = 2\,[\ln(L1) - \ln(L0)]$, with L_i being the maximum likelihood obtained under the ith hypothesis.

The first step in NOCOM involved opening the .txt file containing the CpG O/E values for a given species. After this, testing of the unimodal model occurred, followed by bimodal model testing. For the unimodal model, and for all the species, the value 1 for the exponential option was selected (indicating to the program that data transformation should not be performed), as well as the option of keeping this exponential constant (option 0) during NOCOM mathematical estimations. Also, the number of distribution components (or modes) was selected as 1, i.e., reflecting a unimodal distribution, which ultimately allowed the estimation of the parameters of interest for the unimodal distribution adjusted to the input data (see Results and Discussion). Then, for testing the bimodal model, NOCOM options were set as follows: the exponential option was selected as 1 (indicating no data transformation), as well as the option of keeping this exponential constant (option 0 in NOCOM). Then, the number of components (or modes) was selected as 2 (i.e., bimodal distribution), and a fixed ratio of standard deviation was selected for the two components by choosing the option 2 in NOCOM. Afterwards, a visual estimation of the means, common standard deviation and proportions (weights) of the two distributions was made by analysing the previously created histograms in R. It is important to notice that when the data presented a distinct unimodal pattern, very similar values of means and

proportions were attributed to the two components. This data was used to further feed NOCOM. Accordingly, the mean of the first component was inserted, while also selecting the option to estimate its value in NOCOM (option 1). The shared value of standard deviation of the two components was then inserted in NOCOM, as well as the weight (probability) of the first component (this value was also selected to be estimated by NOCOM by choosing option 1). After this, the values of the second component were inserted, starting by the mean (also selected to be estimated by NOCOM - option 1). Next, the ratio of standard deviation of the second component to that of the first was inserted in NOCOM as 1, as a common value of standard deviation for the two distributions had already been calculated by visual inspection and inserted in the software, as described above. Then, the weight (probability) of the second component was inserted and the option of estimating the value in NOCOM selected. The final prompt question required the indication for the need or not to estimate the previously inserted value of common standard deviation, and option 1 (estimation) was select. This procedure allowed us to obtain the values of the parameters estimated by NOCOM for the unimodal and bimodal scenarios, while also providing the values for EXPO LN(L) in each scenario. This EXPO LN(L) refers to the logarithmic likelihood obtained under each hypothesis. The G^2 statistic is then applied, and the test-value is compared to a threshold calculated as shown in equation 3. If the G^2 is smaller than threshold then there is a non-significant presence of two components with different means, meaning that distribution is unimodal but if G^2 is bigger than the threshold then there is a significant presence of two components with different means which means a bimodal distribution.

Equation 3: $6.08 + 4.51/\sqrt{n}$, where n is the number of observations.

Further steps were taken in order to get a more complete picture of the results in R. First, skewness of each distribution was calculated using the package “moments” (Lukasz Komsta 2015) as detailed in Code box 10. This informed more specifically on how to categorize the distributions in some particular cases (see results). Second, the analysis was extended to the calculation of the percent of CG dinucleotides within the genome of each individual species, using the code detailed in Code Box 11.

Code Box 10. Calculation of skewness

```
skewness (speciesname_genesequence$CpGOE)
```

Code Box 11. Calculation of average level of CG frequency

```
mean(speciesname_genesequence[,"CG"])
```

3. RESULTS AND DISCUSSION

The present work addressed the likelihood of gene body methylation of invertebrate genomes, which confers to different species the phenotypic plasticity allowing responses to environmental challenges or to changing environmental conditions. It addresses this topic thought the analyses of the distribution of the potential for gene body methylation based on the prevalence of CpG dinucleotides. Additionally, a comparative approach on CpG prevalence in the genome was employed to conclude on which are the more favourable model species to address gene body methylation within an ecotoxicological context and develop potential biomarkers of environmental exposure or effect.

As the databases for methylomes are not comprehensive so far, except maybe concerning human methylomes, the appraisal of trends in DNA methylation within closely related organisms and between unrelated groups, e.g., to address the role of the environment as a constraint of epigenetically-based adaptation, must be inferred from genomic information. One way of approaching this inference is based on the fact that methylated cytosines are hyper-mutable. This is because methylated cytosines turn into thymine residues through deamination at a higher rate when compared with unmethylated cytosines (Coulondre et al. 1978). Namely, the mutation rate of methylated cytosines into thymine is 10- to 50-fold higher than that for other point mutations (Fryxell and Moon 2005). This specific type of mutation is not easily corrected by the DNA repair machinery and, as a result, it is widely assumed that regions of DNA that once in the past were consistently methylated are currently depleted of CpG dinucleotides, the privileged sites for DNA methylation over evolutionary time (Schorderet and Gartler 1992). A practical way of addressing the evolutionary trends in methylation among different species, or simply to assess whether the genome of a given species is more or less prone to be methylated, is through the use of the observed-to-expected ratio of CpG dinucleotides in their genome (CpG O/E). This ratio relates the frequency of the CpG dinucleotides with the frequency of cytosines and guanines available for pairing. Low CpG O/E values (far below 1.0) hence indicate that gene sequences are likely to be methylated, whereas high CpG O/E values (approaching 1.0 or higher) indicate that sequences should be sparsely methylated. CpG O/E has been demonstrated to be consistently informative regarding the presence of DNA

methylation in the genome of several organisms (Shimizu et al. 1997; Elango et al. 2009; Yi and Goodisman 2009; Glastad et al. 2011; Sarda et al. 2012; Aliaga et al. 2019). As very well pictured by Aliaga et al. (2019), low CpG O/E is not a condition for methylation but a consequence of it; also, the whole rationale applies to stable methylation signatures that are inherited through several generations, thus that are part of the germline methylome and that can characterise a given species in its current evolutionary momentum.

3.1. Likelihood of gene body methylation of invertebrate genomes

As a general outcome of the analyses of the genomes of 27 species representing several environmental compartments, viz. the soil, air, freshwater and marine environments, it is clear that there is not a common pattern that can be observed in all invertebrates as detailed below.

The interpretation of CpG O/E at the genome level (note that the ratio is calculated for every gene body sequence, in the present work) demands the assumption of an appropriate distribution that feasibly describes the dataset. From the 27 species analysed herein, a unimodal distribution was found to be that feasibly describing the CpG O/E dataset for 16 species (Table 5): *A. albopictus, A. albimanus, B. mori, C. elegans, C. floridanus, D. magna, D. melanogaster, F. candida, H. saltator, I. scapularis, N. vectensis, O. cincta, P. vanderplanki, T. urticae, T. californicus, T. castaneum*. The CpG O/E dataset was better described by a bimodal distribution for the remaining 11 species (Table 5): *A. mellifera, B. glabrata, B. terrestris, C. intestinalis, C. gigas, D. polymorpha, L. vannamei, M. galloprovincialis, N. vitripennis, O. bimaculoides, S. purpuratus*.

It is important to reiterate how the values that are represented in table 5 were obtained and how they constrain the interpretation of the results. As mentioned in the material and methods section, both the unimodal and bimodal models were tested in NOCOM, where the "LN(L) – Unimodal" and "LN(L) – Bimodal" were obtained for each respective test. These values represent the logarithmic maximum likelihood for their respective hypothesis and are used to calculate the G^2 value presented in the table. The actual number of gene sequences found in each species was used to calculate the threshold. Then this threshold

value was compared to the G^2 value to inform on whether there is a statistically significant fit of the distribution to the density data or not. A G^2 value higher than the threshold means that there is a significant fit to the bimodal distribution with two components. On the other hand, if the G^2 value is smaller than the threshold then there is a non-significant fit to the bimodal distribution, hence a unimodal distribution better describes the dataset. This distinction is an important step because although some distributions can be clearly identified through simple visual inspection of gene density histograms concerning CpG O/E others are less clear. The analysis for the presence of one or two components in the distribution can corroborate what can be deduced visually and, in some cases, it can clarify what is difficult to discern visually, preventing less accurate interpretation of the data (Hasselblad 1966; Ott 1979; Thode et al. 1988).

The interpretation of presumably methylated and presumably non-methylated gene bodies based on CpG O/E distributions is not yet fully defined, particularly for those cases where the distributions offer room for differential interpretation. Therefore, and to avoid personal bias as much as possible, the present study generally followed the criteria established by Aliaga et al. (2019) as follows. Bimodal distributions denote a mosaic pattern of DNA methylation, as well as unimodal distributions strongly negatively skewed (skewness lower than -0.04). Unimodal, non-negatively skewed distributions can represent one of three scenarios. Distributions with a CpG O/E mean for the single mode lower than 0.69 represent genomes prone to gene body methylation; if the mean for the single mode is equal or higher than 0.69, low gene body methylation or ultra-low gene body methylation is expected, with the distinction being given by standard deviation - distributions with low gene body methylation have a standard deviation equal or larger than 0.12, while distribution characterized by ultra-low gene body methylation have a standard deviation lower than 0.12. Based on the criteria followed, different groups of species can be assembled that bear distinct patterns regarding the likelihood of gene body methylation. This grouping ruled the organisation of text and figures below and, to assist the interpretation and provide an overview of the results, the corresponding parameters are summarised in Table 7 for each species included in the meta-analysis.

Table 5. Summary statistics for the fitting of unimodal or bimodal distributions to the CpG O/E density data retrieved from the genomes of the selected invertebrate species. The values of "LN(L) – Unimodal" are the logarithmic maximum likelihood obtained under the Unimodal hypothesis and the "LN(L) – Bimodal" values are the logarithmic maximum likelihood obtained under the Bimodal hypothesis. Observations is the actual number of sequences used in the calculation of the threshold for comparison with the G^2 value.

Species	LN(L) – Unimodal	LN(L) – Bimodal	Observations	G^2	Threshold	
Aedes albopictus	**23046.13**4904	**23046.12**6766	17141	0.016	6.11444	
Anopheles albimanus	**47489.92**2987	**47489.90**1033	23947	0.044	6.10914	
Apis mellifera	**7807.55**9327	**10645.88**3355	23458	5676.648	6.10944	*
Biomphalaria glabrata	**35417.73**0893	**37787.89**6518	36662	4740.331	6.10355	*
Bombus terrestris	**9407.42**9333	**9878.46**2308	22092	942.066	6.11034	*
Bombyx mori	**30942.38**6091	**30942.46**6413	27303	0.161	6.10729	
Caenorhabditis elegans	**28357.58**8793	**28357.58**7418	28338	0.003	6.10679	
Camponotus floridanus	**6238.45**8203	**6238.45**1162	14862	0.014	6.11699	
Ciona intestinalis	**17778.34**7161	**18579.69**3517	21086	1602.693	6.11105	*
Crassostrea gigas	**62914.20**0284	**72378.03**6735	63341	18927.673	6.09791	*
Daphnia magna	**23131.31**3042	**23131.30**5224	23570	0.016	6.10937	
Dreissena polymorpha	**80804.55**2578	**85767.38**8528	188502	9925.672	6.09038	*
Drosophila melanogaster	**53166.82**2409	**53166.78**9275	30701	0.066	6.10573	
Folsomia candida	**53669.47**5938	**53669.47**5281	37114	0.001	6.10341	
Harpegnathos saltator	**21201.98**6979	**21202.01**8407	26783	0.063	6.10755	
Ixodes scapularis	**4862.40**2152	**4862.39**7943	20479	0.008	6.11151	
Litopenaeus vannamei	**26278.51**4596	**27028.05**3109	33260	1499.077	6.10472	*
Mytilus galloprovincialis	**67709.07**3835	**71102.90**7376	77409	6787.667	6.09620	*
Nasonia vitripennis	**28798.91**2421	**30280.45**7315	34173	2963.090	6.10439	*
Nematostella vectensis	**25576.48**3700	**25576.47**6914	34307	0.014	6.10434	
Octopus bimaculoides	**19220.15**7456	**23467.34**1359	23992	8494.368	6.10911	*
Orchesella cincta	**32452.25**3162	**32452.25**8622	20247	0.011	6.11169	
Polypedilum vanderplanki	**18165.15**2004	**18165.13**7641	19127	0.029	6.11261	
Strongylocentrotus purpuratus	**44822.93**7544	**46885.71**0144	38427	4125.545	6.10300	*
Tetranychus urticae	**24121.20**0134	**24121.22**3185	15670	0.046	6.11602	
Tigriopus californicus	**24037.50**3330	**24037.49**6345	15576	0.014	6.11613	
Tribolium castaneum	**24391.01**5745	**24391.01**0580	22588	0.010	6.11000	

* Statistically significant fit of the distribution to the density data, so a bimodal distribution should be assumed

3.1.1. Validation Species

Four species within our dataset were previously analysed by others using the same methodology and the same software for distribution testing, namely *A. mellifera, B. mori, C. intestinalis and N. vectensis* (Sarda et al. 2012). This provides the opportunity of validating the methodology and general conclusions in the present study through direct comparison of the results. A direct comparison of the density plots obtained herein (Figure 6) and in (Sarda et al. 2012), highlights that there are no remarkable differences between the two studies in general. In (Sarda et al. 2012), sequences density in *A. mellifera* highlighted two groups of genes, one centred around the value of 0.7 CpG O/E and the other one around 1.5; for *B. mori*, only one group of genes was identified with CpG O/E values of around 1.1; for *C. intestinalis*, two groups of genes were identified, one centred around the CpG O/E value of 0.6 and the other one around 1; *N. vectensis* seems to have 2 groups of genes one around the value of 0.5 and the other one around 0.9. An exception to the consistency between the results of the present study and those by (Sarda et al. 2012) occurred for *N. vectensis*. A unimodal distribution was found herein to more appropriately describe the data (Table 5), thus only one group of genes should be recognised in our dataset (Figure 6d). The same inconsistency was found when comparing with the predictive approach for this species by (Nanty et al. 2011). It is also noteworthy that there are some aspects where full consistency was not reached. The peaks, i.e., the density values in the present work are higher. For example, in (Sarda et al. 2012) the highest peak found for the density in *A. mellifera* is around 1 while in the present work is slightly below 1.5 CpG O/E; similar inconsistencies can be verified for the remaining 3 species. These differences between the two studies are likely due to the use of different sourcing data, but the essential trends that can be interpreted and consequently conclusions bear no remarkable differences, which validates the methodology and conclusions presented herein.

The bee *A. mellifera* has a previous report of an average whole genome CpG methylation level of about 1% (Harris et al. 2019). This actual quantification evidence along with previous works using predictive approaches (Nanty et al. 2011; Glastad et al. 2011; Provataris et al. 2018) corroborate the results of the present study (Figure 6a). The distribution fitted for *A. mellifera* reveals that a narrow group of genes bear a CpG O/E level slightly below 1 (mean for mode 1 of 0.747), but where low methylation levels are still expected (mode mean below 0.69), but

most of the genes (around 62% of the genes) were found to have a CpG O/E value around 2 (mean for mode 2 of 1.520), reinforcing the tendency for low methylation levels (Table 7). Despite this uneven distribution within the mosaic pattern, both groups suggest low likelihood of methylation of the gene bodies, which is consistent with the previously reported low overall level of methylation quantified in the genomes of this species. In the case of the tunicate *C. intestinalis* the distribution fitted herein (Figure 6c) is also consistent with previous ones that have been produced (Nanty et al. 2011), and the species has indeed a previously reported high level of 30% cytosine methylation (Suzuki et al. 2016). Similarly, previous predictive studies on the domestic silk moth, *B. mori*, are consistent with the results obtained herein (Nanty et al. 2011; Glastad et al. 2011; Provataris et al. 2018) and *B. mori* has a reported low methylation level with 0.11% of the genomic cytosines being regularly methylated (Xiang et al. 2010) that is also consistent with the data here presented.

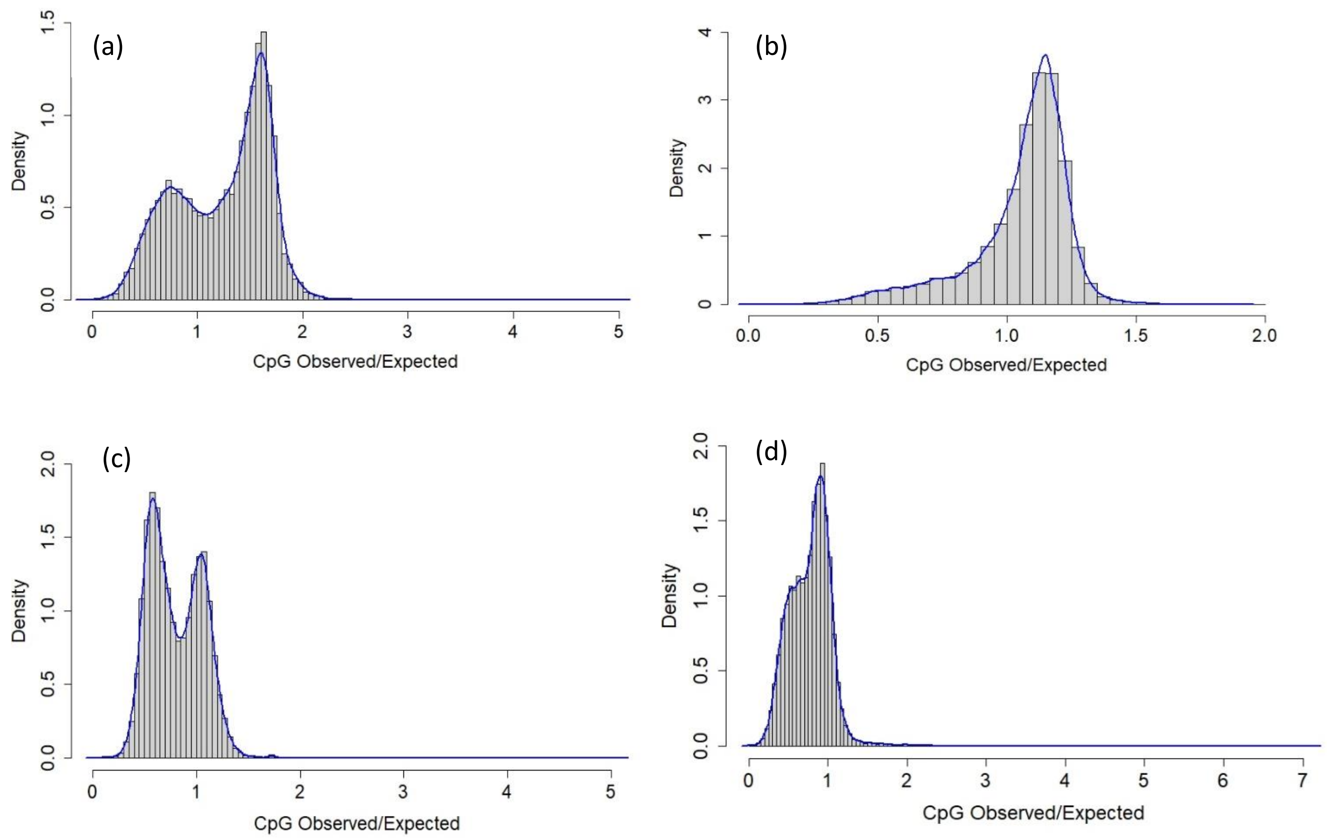

Figure 6. Density plot compiling the CpG O/E density as observed in the genome of *Apis mellifera* (a), *Bombyx mori* (b), *Ciona intestinalis* (c) and *Nematostella vectensis* (d), with the corresponding unimodal or bimodal distributions best describing the datasets represented by a blue solid line and the bimodal distribution. The descriptive statistics for the fitted distribution is given in the grey-shadowed panel of each graph, namely the component-associated means and proportions, as well as the common standard deviation of the components.

3.1.2. Non-Validation Species

The first group, comprising 14 of the 27 species analysed, were found to have genomes fitting unimodal CpG O/E distributions with means for the single mode that are equal or larger than 0.69: *A. albopictus, C. elegans, C. floridanus, D. magna* and *F. candida* as depicted in Figure 7; *I. scapularis, O. cincta, P. vanderplanki, T. californicus, T. castaneum* as depicted in Figure 8. Thus, these species are expected to have low gene body methylation throughout the genome. These results are consistent with previous studies as discussed below.

The Asian tiger mosquito *A. albopictus* has been reported to have a low level of global DNA methylation of around 0.03% according to (Adams et al. 1979), with a more recent report by (Oppold et al. 2015) reporting a wider range of methylation from 0.021% to 0.674% of

methylated cytosine relative to guanine. The nematode *C. elegans* seems to have low to undetectable levels of CpG methylation (Suzuki et al. 2007), a single study being available that reported more specifically a low-level (0.0019 to 0.0033%) of methylated cytosines (Hu et al. 2015). Both the ant *C. floridanus* and the freshwater cladoceran *D. magna* have low methylation levels of 0.14%-0.16% (Bonasio et al. 2012) and 0.7–0.9% (Hearn et al. 2019), respectively. Both *F. candida* and *O. cincta* are collembolans whose genomes apparently lack DNA methylation (Noordhoek et al. 2018; Provataris et al. 2018). All of these reports are consistent with the data achieved through the analysis made in the present dissertation with the exception of *C. elegans*. Herein, the worm is pictured as bearing low gene body methylation (Figure 7b), rather than ultra-low to undetectable methylation levels as argued in the literature (Hu et al. 2015). The CpG O/E distributions presented herein are also consistent with previous similar analyses for *F. candida* and *O. cincta* (Provataris et al. 2018).

There are studies reporting the occurrence of DNA methylation in the red flour beetle *T. castaneum* (Feliciello et al. 2013; Song et al. 2017; Provataris et al. 2018) but others rather report on its absence (Zemach et al. 2010; Schulz et al. 2018). The results obtained herein suggests that, if found, DNA methylation in *T. castaneum* is expected to occur sparingly (Figure 8e). The deer tick *I. scapularis*, the sleeping chironomid *P. vanderplanki* and the intertidal copepod *T. californicus* have not yet been inspected for DNA methylation levels, but the CpG O/E distribution for the *I. scapularis* genome was already addressed and it is in line with the results of the present work (Provataris et al. 2018). The outcomes of the analyses over these species suggest that low gene body methylation should be expected in future actual quantification attempts.

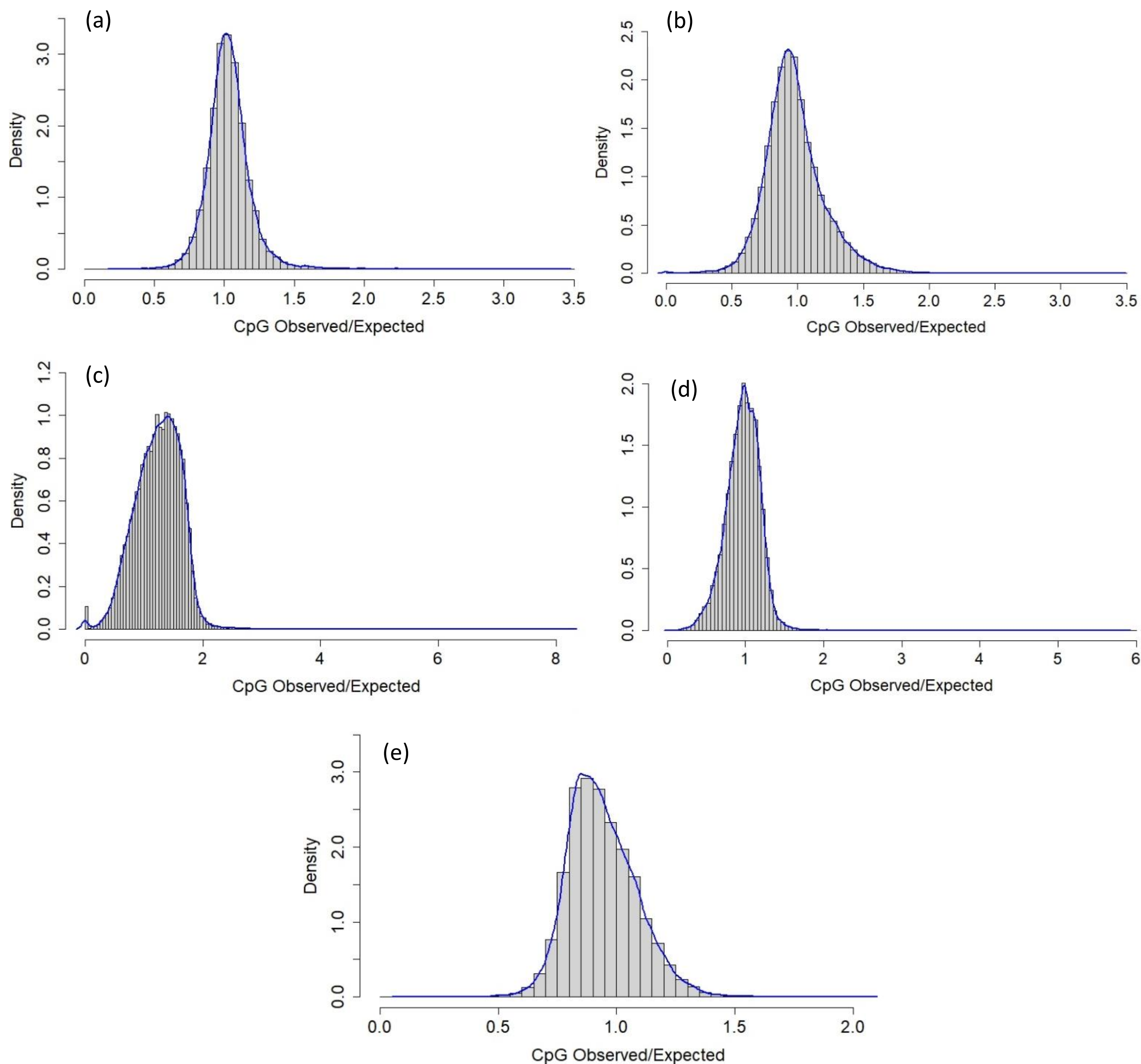

Figure 7. Density plot compiling the CpG O/E density as observed in the genome of *Aedes albopictus* (a), *Caenorhabditis elegans* (b), *Camponotus floridanus* (c), *Daphnia magna* (d) and *Folsomia candida* (e), with the corresponding unimodal or bimodal distributions best describing the datasets represented by a blue solid line and the bimodal distribution. The descriptive statistics for the fitted distribution is given in the grey-shadowed panel of each graph, namely the component-associated means and proportions, as well as the common standard deviation of the components.

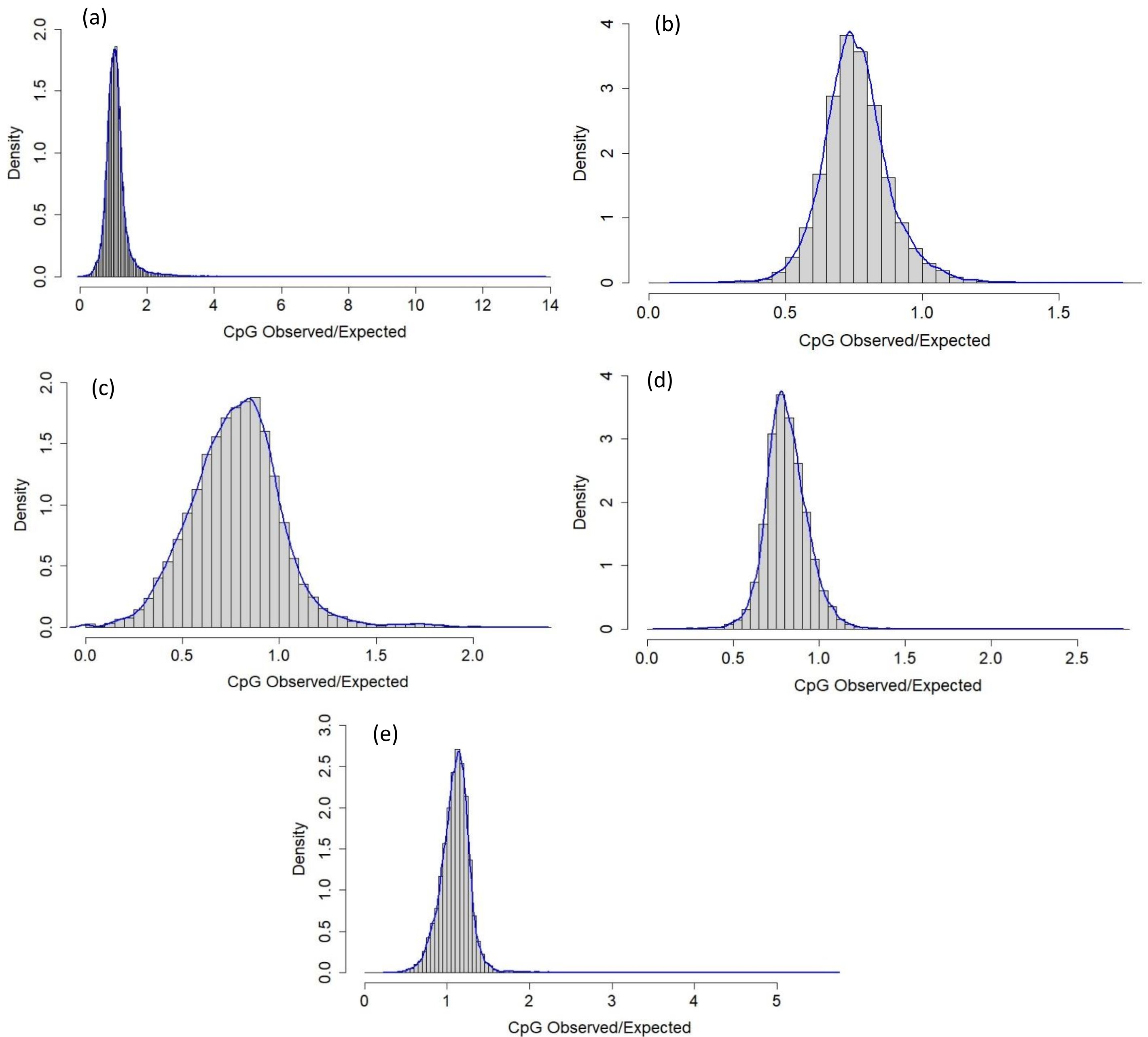

Figure 8. Density plot compiling the CpG O/E density as observed in the genome of *Ixodes scapularis* (a), *Orchesella cincta* (b), *Polypedilum vanderplanki* (c), *Tigriopus californicus* (d) and *Tribolium castaneum* (e), with the corresponding unimodal or bimodal distributions best describing the datasets represented by a blue solid line and the bimodal distribution. The descriptive statistics for the fitted distribution is given in the grey-shadowed panel of each graph, namely the component-associated means and proportions, as well as the common standard deviation of the components.

The mite *T. urticae* stands alone as the single representative of a genome clearly meeting a unimodal distribution of CpG O/E with expectedly high DNA methylation in gene bodies: its CpG O/E distribution has a mean position for the single mode below 0.69 (0.671; Table 7). This interpretation is consistent with the data reported in previous studies showing global

methylation levels for this species of around 25 to 30%, with a cytosine methylation level of 15% to 20% (Yang et al. 2018b).

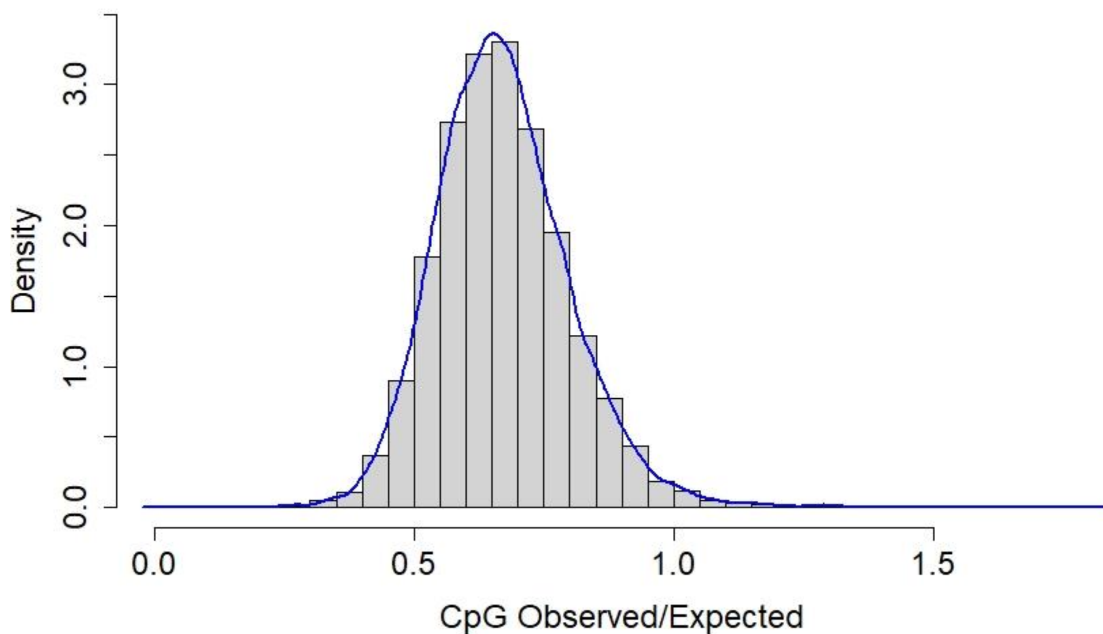

Figure 9. Density plot compiling the CpG O/E density as observed in the genome of *Tetranychus urticae* and the unimodal distribution significantly describing the dataset. The descriptive statistics for the fitted distribution is given in the right-hand, grey-shadowed panel of the figure, namely the component-associated mean as well as the standard deviation of the component.

Lastly, another group of species should be recognised by showing a mosaic pattern for CpG O/E distributions. This means that these species present density distributions with two components, representative of two groups of genes where one is more methylated than the other (bimodal distribution) or have a single component highly left-skewed (Aliaga et al. 2019). This group includes *A. albimanus, B. glabrata, B. terrestris, C. gigas, D. melanogaster, D. polymorpha, H. saltator, L. vannamei, M. galloprovincialis, N. vitripennis, O. bimaculoides, S. purpuratus* (Figures 10 and 11)

This mosaic pattern was found in both species where the overall DNA methylation level is expectedly high and low. The Pacific oyster *C. gigas,* the coastal mussel *M. galloprovincialis* and the purple sea urchin *S. purpuratus* have high reported levels of methylation of 15% of the CpG dinucleotides (Wang et al. 2021), 70% global methylation level (Ardura et al. 2018) and CpGs methylation of 22 % (Strader et al. 2020), respectively. These literature records are

consistent with the outcomes of the CpG O/E analysis presented herein as the density distributions of these species tend all to peak in the left-hand side, suggesting that at least part of the genome tend to bear high levels of DNA methylation (Figures 10a, b, c).

The freshwater snail *B. glabrata* has 2% of the cytosines in its genome regularly methylated (Fneich et al. 2013), the jumping ant *H. saltator* was reported to have a low level of cytosine methylation of 0.11% to 0.12% (Bonasio et al. 2012) and the mosquito *A. albimanus* showed a global methylation level between 0.85% and 1.73% (Claudio-Piedras et al. 2020). The CpG O/E distribution for *B. glabrata* is bimodal, revealing a group of genes that tends to be methylated and another group of a similar proportion that are likely to be less methylated than the first (Figure 10d). Distinctly, the single mode in CpG O/E distributions fit to the genome of both *H. saltator* and *A. albimanus* spread through the right-side of the benchmark CpG O/E value of 0.69 (Figure 10e, f; Table 7), which is coherent with the lower records for measured DNA methylation in these two species compared to *B. glabrata*. The distributions of *B. glabrata* and *H. saltator* found herein are also consistent with previous works using similar predictive approaches (Fneich et al. 2013; Provataris et al. 2018) while there are no studies available in this regard for the mosquito. The fruit fly *D. melanogaster* is generally considered to have no DNA methylation and a consistent extremely low level of methylated cytosines (0.001%) was already reported (Deshmukh et al. 2018). The CpG O/E distribution presented herein (Figure 11a) is also consistent with previous similar analyses for *D. melanogaster* (Glastad et al. 2011; Provataris et al. 2018). There is a particular situation regarding the classification of the distributions for both *A. albimanus* and *D. melanogaster* that is worth exploring further. When analysed visually, in both cases distributions appear to be normal and to have no skewness (Figures 10f and 11a), but calculated skewness values denote that the distributions are rather left-skewed, which lead to the inclusion of the species in this group comprising species that show mosaic CpG O/E distributions (Table 7). This classification was made according to Aliaga et al. (2019) as clarified above but is noteworthy that other literature sources, use other measures and calculations of skewness and are more tolerant regarding the said skewness (Groeneveld 1991; Hosking 1992; Doane and Seward 2011). In this way, despite these two species being initially classified has having mosaic distributions for CpG O/E, they are likely to reflect low methylation following integrated interpretation.

The bumble bee *B. terrestris* and the wasp *N. vitripennis* show, according to previous data, low levels of methylation. *B. terrestris* was reported to have 0.5 to 0.6% of all CpGs methylated (Sadd et al. 2015) and *N. vitripennis* was found to have an average global methylation level of 0.18% over the whole genome and methylation of 0.63% of the CpGs (Beeler et al. 2014). For the first species, these literature records are consistent with the predictive distributions found in the present study given the position of both modes in the CpG O/E scale, with means above 0.69 suggesting low methylation levels for both groups of genes (Figure 11a; Table 7). In the case of *N. vitripennis*, both components of the distribution mosaic pattern have CpG O/E means of nearly or above 1 (Figure 11b; Table 7); this means that there are two recognisable groups of genes with differential likelihood of methylation, but both show CpG O/E records suggesting a tendency for low methylation levels. This is generally in line with previous works that recognise mosaic CpG O/E distributions, the difference being a slight moving of distributions therein to the left (Glastad et al. 2011; Park et al. 2011; Provataris et al. 2018).

Regarding the freshwater bivalve *D. polymorpha*, the shrimp *L. vannamei* and the cephalopod *O. bimaculoides*, no previous studies exist actually quantifying DNA methylation that could be used for comparison to the predictive outcomes of the present study. However, a previous study already addressed the CpG O/E distribution for *L. vannamei* (Provataris et al. 2018) and reached similar results. Herein the genome of these species have a typical mosaic distribution as to the CpG O/E ratio. For the shrimp and the octopus, one group of genes is highly susceptible to methylation (CpG O/E mean below 0.69) and the other shows a pattern consistent with a low susceptibility to methylation (Figure 11d, e), whereas in the octopus this latter group is proportionally much smaller (Table 7). Differently, *D. polymorpha* has both groups of genes with a mode means above 0.69 (Figure 11c), denoting that although at different scales, the genome of this tends generally for accommodating low DNA methylation (Table 7). Something that needs to be pointed out in regard to *D. polymorpha* is that the data presented in Table 7, particularly the values of the means and proportions, do not 100% line up with what is shown in the figure (11d). This can be seen as the values of the second mean that according to the image should be around the value of 1.1 are instead presented has 2.1. This does not change our conclusions has those are mainly derived from the analyses of the image, but it does present with something that needs to be further investigated to identify

what is skewing the values making the values obtained from NOCOM different from those presented in the image.

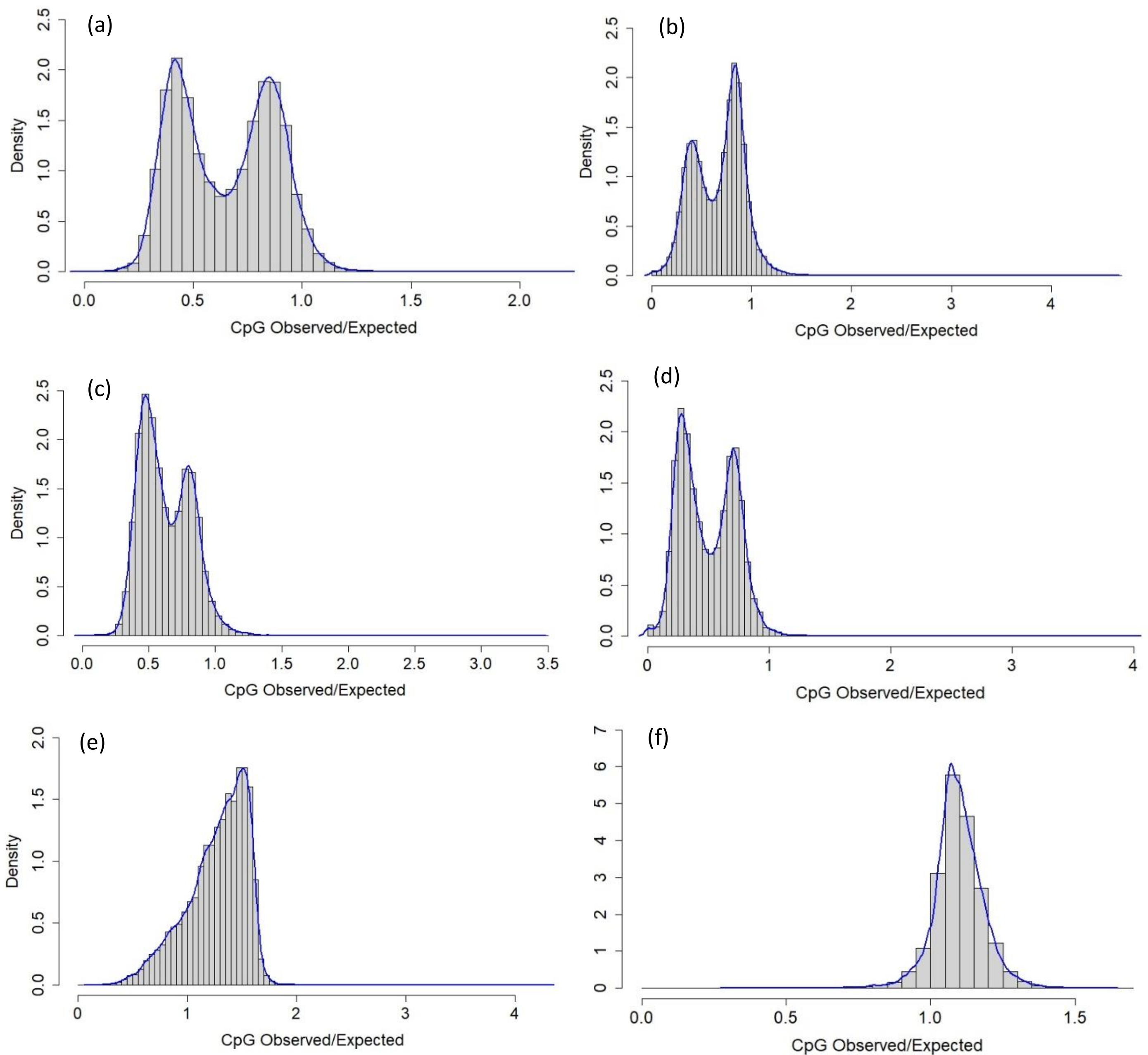

Figure 10. Density plot compiling the CpG O/E mosaic density as observed in the genome of *Crassostrea gigas (a), Mytilus galloprovincialis (b), Strongylocentrotus purpuratus (c), Biomphalaria glabrata (d) Harpegnathos saltator (e) and Anopheles albimanus* (f) and with the corresponding unimodal but highly left-skewed (skewness for *A. albimanus* of -0.0490733 and for *H. saltator* of -0.7419896) or bimodal distributions best describing the datasets represented by a blue solid line and the bimodal distribution. The descriptive statistics for the fitted distribution is given in the grey-shadowed panel of each graph, namely the component-associated means and proportions, as well as the common standard deviation of the components.

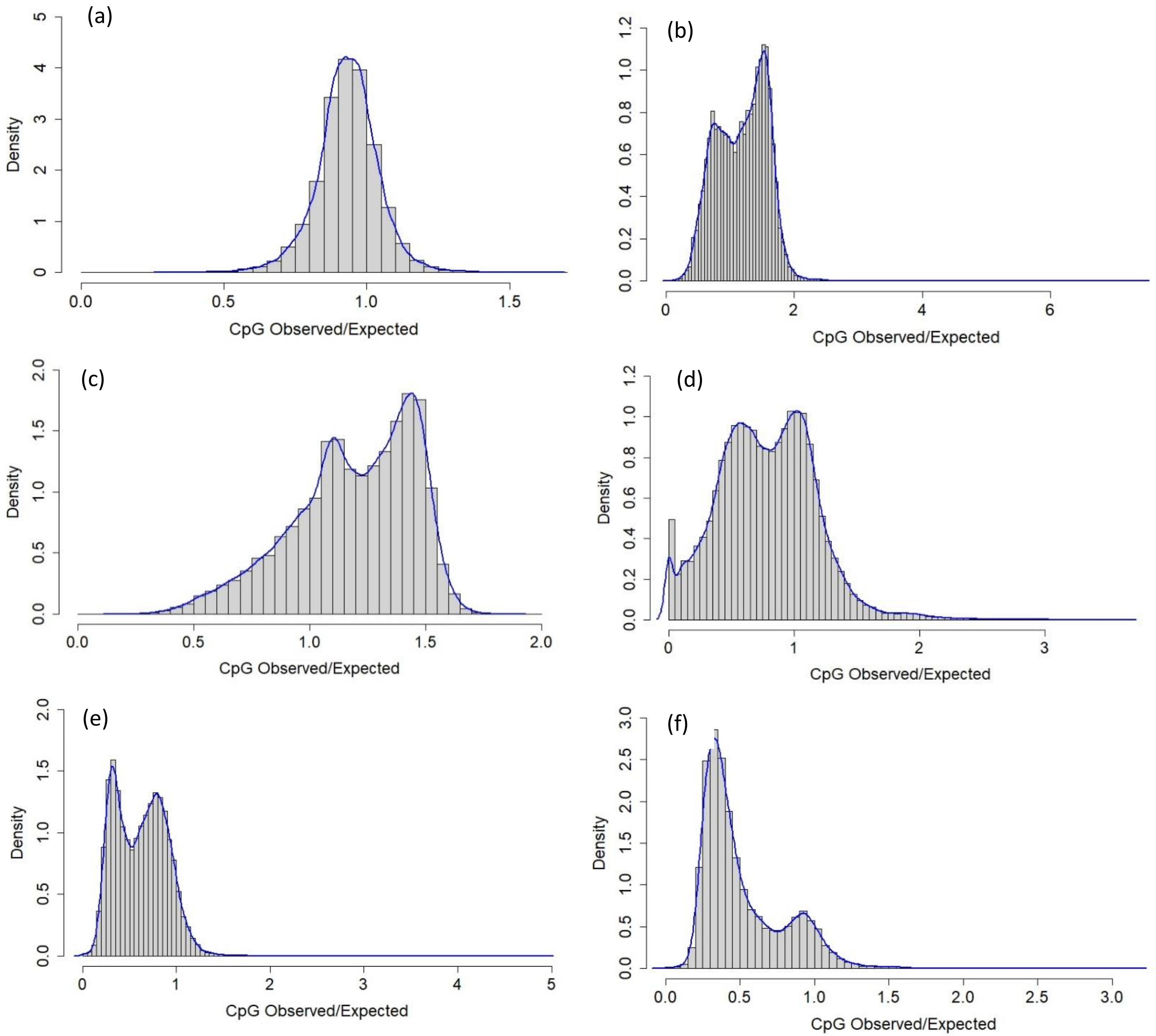

Figure 11. Density plot compiling the CpG O/E density as observed in the genome of *D. melanogaster(a), Bombus terrestris (b), Nasonia vitripennis (c), Dreissena polymorpha (d), Litopenaeus vannamei (e)* and *Octopus bimaculoides (f),* with the corresponding unimodal but highly left-skewed (skewness for *D. melanogaster* of -0.1234172) or bimodal distributions best describing the datasets represented by a blue solid line and the bimodal distribution. The descriptive statistics for the fitted distribution is given in the grey-shadowed panel of each graph, namely the component-associated means and proportions, as well as the common standard deviation of the components.

3.2. Differences in gene body methylation in closely related species

On a more specific view of the outcome, even species more closely related may present genomes which are differentially susceptible to gene body methylation.

Interestingly, when appraising gene body methylation across environmental compartments, aquatic species can be clearly highlighted as generally bearing a higher gene methylation likelihood, with 63.6% of the aquatic species showing higher density of genes around CpG O/E records below 0.69, compared to only 11% of the terrestrial species bearing a genome susceptible to methylation (Table 7). Within the aquatic compartment, freshwater species are underrepresented in the dataset since only 3 species representatives could be included in the meta-analysis. Although general extrapolations regarding freshwater species would be abusive, it is noticeable from the three representative species that only *B. glabrata* has part of the genome highly susceptible to methylation, with both *D. polymorpha* and *D. magna* bearing likely lowly methylated genomes. Considering that methylated CpG dinucleotides tend to mutate easily (Coulondre et al. 1978), hypermethylated genomes in the past can currently be hypomethylated. Based on this feature and in the observation that the genomes of terrestrial species tend to be less methylated currently than the current genomes of aquatic species, it is reasonable to postulate an evolutionary path behind this pattern. From a macroevolutionary point of view, the living world tended to follow the exit from water to land. In this course, methylation may have played an important role in species adaptation to all the new niches available on land, enhancing phenotypic plasticity ranges through the increase of transcriptional opportunities (Roberts and Gavery 2012; Dimond and Roberts 2016; Gatzmann et al. 2018). The past susceptibility to methylation allowing the transition to land may have driven increased methylation-driven mutation rates in CpG dinucleotides, thereby eroding the genomes of preferential sites for methylation and leading to the lower overall susceptibility to methylation observed in current terrestrial invertebrate genomes.

In terms of specific comparisons between related species from different compartments a few can be made based on the selected genomes in the present work. First, one can compare three bivalves that were analysed, namely the marine species *C. gigas* (Figure 10a), the freshwater species *D. polymorpha* (Figure 11c) and the marine species *M. galloprovincialis* (Figure 10b) can be compared. All three species, as previously described, have a bimodal distribution with on two groups of genes one being predicted to be more susceptible to

methylation than the other. *C. gigas* and *M. galloprovincialis* are very similar in terms of terms of their distributions and their genomes bear at least one group of genes with CpG O/E means below 0.69. *D. polymorpha*, by contrast, shows both modes of the CpG O/E distribution with a mean above 0.69, denoting that the genome is able to accommodate little methylation (Table 7). This seems to indicate that the genome of the freshwater species can become methylated at a much littler scale than those of the two marine ones. According to the hypothesis we presented above in terms of evolutionary path, this suggests that *D. polymorpha* may have needed the scope for phenotypic plasticity given by DNA methylation opportunities while its ancestors colonized freshwater from the seas, and then undergone CpG erosion following mutation towards the current genome less prone to methylation than the compared marine bivalves. This is consistent with what is known about the evolutionary course of *D. polymorpha*, whose ancestors were marine species (Gelembiuk et al. 2006; Strong et al. 2008; Graf 2013).

Several other comparisons are meaningful among closely related species. For example, here is the one between the two crustaceans *D. magna* and *T. californicus. D. magna* is a freshwater cladoceran while *T. californicus* is an intertidal copepod. Both of these species have a unimodal distribution with a mean close to 1 indicating low methylation. Both species in terms of their distributions are fairly similar, with *T. californicus* having a distribution slighter more centred to the left of the graft with a mean of 0.81 while *D. magna* has a mean of 0.96 (Table 7). The fact of that *T. californicus* it is intertidal is interesting due to the fact that they have to deal with fluctuating conditions in the intertidal zone, and so leading to possibility of them having to resort to methylation to be able to quickly adapt leading to an erosion of CpG levels and leading to the current small level of methylation. Another crustacean that is present here is *L. vannamei* that unlike *D. magna* and *T. californicus* has a bimodal distribution being much more shifted to the left and so having a higher predicted methylation than both *D. magna* and *T. californicus.* Other species closely related that share similar patterns regarding the distribution of CpG O/E, thus likelihood of methylation, can be identified: like the ants *C. floridanus* and *H. saltator* that both have a unimodal distributions of means around 1.2 indicating low methylation with *H. saltator* having its distribution skewed to the left (Table 7). In a similar fashion both the bees of *A. mellifera* and *B. terrestris* similar bimodal distributions, with components of similar values, indicating low methylation with

areas of high methylation. Close to the bees is the wasp *N. vitripennis* that unlike the previous present a distribution with high values of components means indicating a smaller level of methylation than the bees but present areas of higher methylation (Table 7). Finally, we can also compare the two mosquitos, *A. albimanus* and *A. albopictus* which are both predicted to have low methylation due to their respective unimodal distributions with means around the value of 1 (Table 7). The sleeping chironomid *P. vanderplanki* being somewhat related can also be compared with these two mosquitos and when we look at its distribution we can see that it is wider and is more to the left than the previous two indicating a higher level of methylation than the two mosquitos.

Table 7. Overview of the parameters applying to CpG O/E distributions of the studied species, methylation levels reported in the literature and conclusions on methylation likelihood based on CpG O/E analysis run herein.

Species	1º Mean	2º Mean	1º Proportion	2º Proportion	Standard Deviation	Skewness	Reported methylation level	Conclusions
Terrestrial Species								
Bombyx mori	1.053	-	1	-	0.195	0.5689	0.11% [b]	low gene body methylation
Caenorhabditis elegans	0.977	-	1	-	0.223	0.7105661	0.0019%-0.0033% [b]	low gene body methylation
Camponotus floridanus	1.239	-	1	-	0.399	1.149869	0.14%-0.16% [a]	low gene body methylation
Harpegnathos saltator	1.271	-	1	-	0.275	-0.7419896	0.11%-0.12% [b]	low gene body methylation
Folsomia candida	0.940	-	1	-	0.143	0.4702213	No methylation	low gene body methylation
Orchesella cincta	0.758	-	1	-	0.122	0.4042925	No methylation	low gene body methylation
Ixodes scapularis	1.098	-	1	-	0.478	7.935341	No reports	low gene body methylation
Tetranychus urticae	0.671	-	1	-	0.130	0.7361029	25-30% [a], 15%-20% [b]	high gene body methylation
Tribolium castaneum	1.101	-	1	-	0.206	4.160719	Conflicting Reports	low gene body methylation
Air Species								
Aedes albopictus	1.031	-	1	-	0.158	1.659051	0.03% [a]	low gene body methylation
Apis mellifera	0.747	1.520	0.376	0.624	0.221	-0.353119	1% c	low gene body methylation
Anopheles albimanus	1.10	-	1	-	0.084	-0.0490733	0.85%-1.73% [a]	low gene body methylation
Bombus terrestris	0.825	1.450	0.426	0.574	0.248	0.4549392	0.5%-0.6% [c]	low gene body methylation
Drosophila melanogaster	0.933	-	1	-	0.107	-0.1234172	0.001% [b]	low gene body methylation
Nasonia vitripennis	0.857	1.299	0.249	0.751	0.178	-0.6272069	0.18% [a], 0.63% [c]	low gene body methylation
Polypedilum vanderplanki	0.775	-	1	-	0.235	0.4749372	No Reports	low gene body methylation
Freshwater Species								
Biomphalaria glabrata	0.318	0.710	0.547	0.453	0.123	0.646204	2% [b]	high gene body methylation
Dreissena polymorpha	0.767	2.117	0.987	0.013	0.363	0.5650045	No Reports	low gene body methylation
Daphnia magna	0.966	-	1	-	0.227	1.231337	0.7–0.9% [a]	low gene body methylation
Marine Species								
Ciona intestinalis	0.619	1.049	0.570	0.430	0.151	0.8218976	30% [b]	high gene body methylation
Crassostrea gigas	0.448	0.846	0.496	0.504	0.105	0.0769514	15% [c]	high gene body methylation
Litopenaeus vannamei	0.411	0.841	0.519	0.481	0.172	0.5378573	No Reports	high gene body methylation
Mytilus galloprovincialis	0.418	0.841	0.409	0.591	0.144	0.1172125	70% [a]	high gene body methylation
Nematostella vectensis	0.777	-	1	-	0.288	3.052101	No Reports	low gene body methylation
Octopus bimaculoides	0.389	0.942	0.771	0.229	0.142	1.305207	No Reports	high gene body methylation
Strongylocentrotus purpuratus	0.502	0.606	0.819	0.394	0.108	0.5299487	22 % [c]	high gene body methylation
Tigriopus californicus	0.813	-	1	-	0.130	1.346726	No Reports	low gene body methylation

a- Global Methylation, b- Cytosine methylation, c- CpG Methylation

3.3. Species appropriateness for the development of epigenetic biomarkers supporting environmental assessment

The analysis also highlighted that the use of DNA methylation as a biomarker of exposure and/or effects can be constrained by the level of CpG enrichment of the genomes of the species. In order to appraise on what species would be better to apply in environmental assessments studies using the epigenome, the potential for methylation, i.e., the level of CG within the genome should primarily be considered. This is because the more CG dinucleotides there are, the higher the odds that the corresponding sequences become methylated as a response to a given environmental stressor, thus the best chances of finding a responsive epigenome that can support the development of epigenetic biomarkers. Although is important to point out that a high level of CG doesn't guarantee a high level of methylation, but only provides more locations here it can occur.

Considering the CG frequency for the studied species (table 8), some general recommendations can be provided regarding the focus towards development of epigenetic biomarkers within each compartment. Regarding the soil compartment, *I. scapularis* and *H. saltator* should be the species most focused; regarding aerial species, *A. albimanus, A. albopictus* and *N. vitripennis* should be considered; species such as *N. vectensis* and *T. californicus* would likely be the most responsive to represent marine ecosystems; and *D. magna* or *D. polymorpha* are freshwater species that seem appropriate for the development of epigenetic biomarkers.

While these general recommendations are logical if one looks only at the CG frequency, the feasible application of DNA methylation as a biomarker within fields relating to environmental assessment, e.g., ecotoxicology, requires further reasoning. In this context, it is worth to highlight for example the case of the deer tick *I. scapularis* and the jumping ant *H. saltator*. The CG content of their genomes is high, and this would suggest that both could be good indicator species where the development of responsive DNA methylation biomarkers would be successful. However, they lack ecotoxicology relevance as there is very little information on the responses of these species to environmental stressors at different levels of biological organization. While a limited number studies can be found in

this regard with *I. scapularis* and *H. saltator* (Schneider et al. 2011, 2016; Busby et al. 2012; Villar et al. 2015; Bourret et al. 2016; Ginsberg et al. 2017), much more is known for other model species representing the same environmental compartment, such as the nematode *C. elegans* (Sochová et al. 2006; Queirós et al. 2019) or springtails, e.g., *O. cincta* (Faddeeva-Vakhrusheva et al. 2017), that for a long time have been used as models in ecotoxicology hence accumulating a large and relevant body of evidence on responses to environmental stressors, namely at the phenotypic level and especially at the individual and population levels. *C.elegans* shows a notable body of evidence on responses to environmental contaminants but was reported to have low to no DNA methylation (Sochová et al. 2006; Suzuki et al. 2007; Hu et al. 2015). This is consistent with the CG frequency and the CpG O/E ratio analysis. But it is still reasonable to hold this species as a potentially good candidate for the development of epigenetic biomarkers to apply in environmental assessment studies due to the large number of work previously done.

The same critical reasoning applies regarding aerial species. *A. albimanus* has little to no studies in the realm of environmental stressors, while *A. albopictus* and *N. vitripennis* are more documented in this context (da Gloria da Costa Carvalho and Freitas 1988; Wang et al. 2012; Chirault et al. 2015; Sivan et al. 2017; Reinhold et al. 2018; Hidalgo et al. 2019). *D. melanogaster* is even more documented, again, due to its role as a model organism that has been used for many years. As to marine species, *T. californicus* has been investigated regarding the effects of thermal stress and salinity stress (DeBiasse et al. 2018; Li et al. 2019a; Tangwancharoen et al. 2020; Lee et al. 2021) but *N. vectensis* is far more studied since it is a recognised model organism for the study of evolution, genomics, reproductive biology, developmental biology and ecology (Darling et al. 2005; Genikhovich and Technau 2009; Watson et al. 2009), thus it should be better suited for the development of feasible epigenetic biomarkers of exposure and effect. Another option for the marine compartment is *M. galloprovincialis* due to the fact that mussels are generally good bioindicators mostly because of their biosentinel role, which is appropriate for ecosystem quality assessment. *M. galloprovincialis* has been used as such (Malagoli et al. 2008; Kandŭca et al. 2011; Solé et al. 2020), thus there is more background knowledge available for the interpretation of the responses of potential DNA methylation biomarkers. Finally, for the freshwater

compartment, the GC frequency points towards *D. magna* or *D. polymorpha* as the most suitable species to develop biomarkers based on DNA methylation. *D. magna* is perhaps the best of these two species given that it has a slightly higher CG frequency but especially because it is a model organism in a several fields of research like in ecological and evolutionary studies, and particularly in ecotoxicology (Altshuler et al. 2011; Tkaczyk et al. 2021), being already used in several standard test methods (e.g., OECD 2004, 2008) and having already its molecular responses (including differential DNA methylation) to a wide range of environmental stressors studied (Fan et al. 2015, 2020; Jeremias et al. 2018b; Yuxuan et al. 2019; Shahid et al. 2019; Hearn et al. 2019; Ates et al. 2020; Samanta et al. 2020; Ellis et al. 2020).

Table 8. Average level of CG frequency by species organized by compartment.

Species	CG Frequency Average
Terrestrial Species	
Bombyx mori	0.03841567
Caenorhabditis elegans	0.03353303
Camponotus floridanus	0.04151784
Harpegnathos saltator	0.05763937
Folsomia candida	0.03447081
Orchesella cincta	0.02856794
Ixodes scapularis	0.05672893
Tetranychus urticae	0.01859712
Tribolium castaneum	0.0378923
Air Species	
Aedes albopictus	0.05302826
Apis mellifera	0.03196508
Anopheles albimanus	0.06727366
Bombus terrestris	0.03672484
Drosophila melanogaster	0.04777528
Nasonia vitripennis	0.05130614
Polypedilum vanderplanki	0.01842584
Freshwater Species	
Biomphalaria glabrata	0.01579216
Dreissena polymorpha	0.03513598
Daphnia magna	0.03963171
Marine Species	
Ciona intestinalis	0.026249
Crassostrea gigas	0.01924947
Litopenaeus vannamei	0.02575081
Mytilus galloprovincialis	0.01881304
Nematostella vectensis	0.03079966
Octopus bimaculoides	0.01446947
Strongylocentrotus purpuratus	0.02323292
Tigriopus californicus	0.04219962

4. CONCLUSIONS AND FUTURE PERSPECTIVES

Regarding the specific objects that were defined for this work, it is noteworthy that all were satisfactorily tackled. Publicly available reliable genomic data were successfully collected on 27 invertebrate species representing different environmental compartments and that are currently involved in research addressing ecological threats, evolutionary trends or eventually human health and wellbeing threats. These genomic datasets allowed the quantification of the distributions of CpG O/E and consequently the appraisal on the potential for gene body methylation, along with the prediction of the corresponding methylation pattern for all species. This analysis on the distribution of CpG O/E across the genomes confirmed previous evidence in the literature on actual quantification of DNA methylation levels in general. However, for several species, this is the first study where DNA methylation patterns are addressed even if following a predictive approach. Also based on likelihood of gene body methylation, it was elaborated on the potential of different invertebrate species to respond or potentially adapt to challenging and/or highly fluctuating environments thought DNA methylation mediated phenotypic plasticity or ultimately mutation. Overall, this analyses showed that there is not a common tendency applying to all invertebrates as to genome susceptibility to DNA methylation. However, some consistent associations were made suggesting that phylogenetically closely related species belonging to the same environmental compartment evolved their epigenome similarly. It was also speculated that adaptation to new ecological niches may have involved DNA methylation followed by mutation, ultimately resulting in genomes currently depleted of DNA methylation susceptibility. Finally, based on a comparative analysis of CG frequency in the genomes, reasoned along with knowledge availability on phenotypic responses to environmental stress, the present study advanced on the identification of species that can more efficiently be used to address gene body methylation within an ecotoxicological context, namely for the development and implementation of epigenetic biomarkers of environmental exposure or effect.

Regarding future perspectives, the most general indication that can be given following the present study regards the need to enlarge the database on the epigenome of invertebrate species. Next-generation sequencing techniques such as whole genome bisulfite

sequencing among others are available and increasingly accessible with moderate budgets that can provide the characterization of actual DNA methylation patterns at a nucleotide resolution level. This is critical to understand the evolutionary path of DNA methylation in different species, as well as to confirm estimations on the level of methylation of invertebrate genomes, for example to appraise on the sensitivity and responsiveness of the epigenome to environmental stressors and consequently the adequacy of each model species to provide epigenetic biomarkers (based on DNA methylation) of exposure and/or effect to ecotoxicological toolboxes integrated in Environmental Risk Assessment frameworks.

Considering the specific study presented herein, the immediate step to improve the conclusions reached would be the identification of highly methylated genes and poorly methylated genes. This involves the systematic retrieval of the sequences belonging to each of these two groups and their integration with Gene Ontology annotations for selected species (those with globally hypomethylated genomes are not relevant for such an extended analysis). The outcome would inform on which types of genes are more likely to be methylated, thus possible on the genes that can be more responsive to environmental stressors. Moreover, it would allow to appraise on the functional role of genes belonging to either group, namely on whether they serve more housekeeping functions, like those involved in transcription and translation, or whether are they are most likely involved processes like in cellular signalling, reproductive processes and stress responses. This would largely increase the current knowledge on DNA methylation patterns in invertebrates, confirming or not the current assumptions that can be found in the literature. Finally, and under a meta-analysis perspective, the identification of genes that are likely to be methylated and the comparison of species in this context accounting to their phylogenetic relationships can inform on the evolutionary roles of DNA methylation.

REFERENCES

Abdel-latief M, Garbe LA, Koch M, Ruther J (2008) An epoxide hydrolase involved in the biosynthesis of an insect sex attractant and its use to localize the production site. Proceedings of the National Academy of Sciences 105:8914–8919. https://doi.org/10.1073/pnas.0801559105

Aberg KA, McClay JL, Nerella S, et al (2012) MBD-seq as a cost-effective approach for methylome-wide association studies: demonstration in 1500 case–control samples. Epigenomics 4:605–621. https://doi.org/10.2217/epi.12.59

Adams NL, Campanale JP, Foltz KR (2012) Proteomic Responses of Sea Urchin Embryos to Stressful Ultraviolet Radiation. Integrative and Comparative Biology 52:665–680. https://doi.org/10.1093/icb/ics058

Adams RLP, McKay EL, Craig LM, Burdon RH (1979) Methylation of mosquito DNA. Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis 563:72–81. https://doi.org/10.1016/0005-2787(79)90008-X

Akcha F, Barranger A, Bachère E (2021) Genotoxic and epigenetic effects of diuron in the Pacific oyster: in vitro evidence of interaction between DNA damage and DNA methylation. Environmental Science and Pollution Research 28:8266–8280. https://doi.org/10.1007/s11356-020-11021-6

Akhila JS, Shyamjith, Deepa, Alwar MC (2007) Acute toxicity studies and determination of median lethal dose. Current Science 93

Alan Buis (2019) The Atmosphere: Getting a Handle on Carbon Dioxide – Climate Change: Vital Signs of the Planet. In: NASA’s Jet Propulsion Laboratory. https://climate.nasa.gov/news/2915/the-atmosphere-getting-a-handle-on-carbon-dioxide/. Accessed 16 Dec 2021

Aliaga B, Bulla I, Mouahid G, et al (2019) Universality of the DNA methylation codes in Eucaryotes. Scientific Reports 9:173. https://doi.org/10.1038/s41598-018-37407-8

Allan ERO, Blouin MS (2020) Heat shock increases hydrogen peroxide release from circulating hemocytes of the snail Biomphalaria glabrata. Fish & Shellfish Immunology 105:203–208. https://doi.org/10.1016/j.fsi.2020.07.029

Altshuler I, Demiri B, Xu S, et al (2011) An Integrated Multi-Disciplinary Approach for Studying Multiple Stressors in Freshwater Ecosystems: Daphnia as a Model Organism. Integrative and Comparative Biology 51:623–633. https://doi.org/10.1093/icb/icr103

Anbalagan C, Lafayette I, Antoniou-Kourounioti M, et al (2013) Use of transgenic GFP reporter strains of the nematode Caenorhabditis elegans to investigate the patterns of stress responses induced by pesticides and by organic extracts from agricultural soils. Ecotoxicology 22:72–85. https://doi.org/10.1007/s10646-012-1004-2

Aravin AA, Sachidanandam R, Bourc’his D, et al (2008) A piRNA Pathway Primed by Individual Transposons Is Linked to De Novo DNA Methylation in Mice. Molecular Cell 31:785–799. https://doi.org/10.1016/j.molcel.2008.09.003

Ardura A, Clusa L, Zaiko A, et al (2018) Stress related epigenetic changes may explain opportunistic success in biological invasions in Antipode mussels. Scientific Reports 8:10793. https://doi.org/10.1038/s41598-018-29181-4

Aronson RB, Thatje S, Mcclintock JB, Hughes KA (2011) Anthropogenic impacts on marine ecosystems in Antarctica. Annals of the New York Academy of Sciences 1223:82–107

Asselman J, de Coninck DI, Beert E, et al (2017) Bisulfite Sequencing with Daphnia Highlights a Role for Epigenetics in Regulating Stress Response to Microcystis through Preferential Differential Methylation of Serine and Threonine Amino Acids. Environmental Science & Technology 51:924–931. https://doi.org/10.1021/acs.est.6b03870

Asselman J, de Coninck DIM, Pfrender ME, de Schamphelaere KAC (2016) Gene body methylation patterns in daphnia are associated with gene family size. Genome Biology and Evolution 8:1185–1196. https://doi.org/10.1093/gbe/evw069

Asselman J, de Coninck DIM, Vandegehuchte MB, et al (2015) Global cytosine methylation in Daphnia magna depends on genotype, environment, and their interaction. Environmental Toxicology and Chemistry 34:1056–1061. https://doi.org/10.1002/etc.2887

Ates M, Danabas D, Ertit Tastan B, et al (2020) Assessment of Oxidative Stress on Artemia salina and Daphnia magna After Exposure to Zn and ZnO Nanoparticles. Bulletin of Environmental Contamination and Toxicology 104:206–214. https://doi.org/10.1007/s00128-019-02751-6

Athanasio CG, Sommer U, Viant MR, et al (2018) Use of 5-azacytidine in a proof-of-concept study to evaluate the impact of pre-natal and post-natal exposures, as well as within generation persistent DNA methylation changes in Daphnia. Ecotoxicology 27:556–568. https://doi.org/10.1007/s10646-018-1927-3

Attia S, Grissa KL, Lognay G, et al (2013) A review of the major biological approaches to control the worldwide pest Tetranychus urticae (Acari: Tetranychidae) with special reference to natural pesticides. Journal of Pest Science 86:361–386. https://doi.org/10.1007/s10340-013-0503-0

Aucejo-Romero S, Gómez-Cadenas A, Jacas-Miret J-A (2004) Effects of NaCl-stressed citrus plants on life-history parameters of Tetranychus urticae (Acari: Tetranychidae). Experimental and Applied Acarology 33:55–67. https://doi.org/10.1023/B:APPA.0000030026.77800.0c

Bachère E, Barranger A, Bruno R, et al (2017) Parental diuron-exposure alters offspring transcriptome and fitness in Pacific oyster Crassostrea gigas. Ecotoxicology and Environmental Safety 142:51–58. https://doi.org/10.1016/j.ecoenv.2017.03.030

Bahrndorff S, Holmstrup M, Petersen H, Loeschcke V (2006) Geographic variation for climatic stress resistance traits in the springtail Orchesella cincta. Journal of Insect Physiology 52:951–959. https://doi.org/10.1016/j.jinsphys.2006.06.005

Bal N, Kumar A, Nugegoda D (2017) Assessing multigenerational effects of prednisolone to the freshwater snail, Physa acuta (Gastropoda: Physidae). Journal of Hazardous Materials 339:281–291. https://doi.org/10.1016/j.jhazmat.2017.06.024

Bautista-Covarrubias JC, Zamora-Ibarra PA, Apreza-Burgos E, et al (2020) Immune response and oxidative stress of shrimp Litopenaeus vannamei at different moon phases. Fish & Shellfish Immunology 106:591–595. https://doi.org/10.1016/j.fsi.2020.08.040

Beeler SM, Wong GT, Zheng JM, et al (2014) Whole-Genome DNA Methylation Profile of the Jewel Wasp (Nasonia vitripennis). G3 Genes|Genomes|Genetics 4:383–388. https://doi.org/10.1534/g3.113.008953

Benedict MQ, Levine RS, Hawley WA, Lounibos LP (2007) Spread of The Tiger: Global Risk of Invasion by The Mosquito Aedes albopictus. Vector-Borne and Zoonotic Diseases 7:76–85. https://doi.org/10.1089/vbz.2006.0562

Bicho RC, Faustino AMR, Rêma A, et al (2021) Confirmatory assays for transient changes of omics in soil invertebrates – Copper materials in a multigenerational exposure. Journal of Hazardous Materials 402:123500. https://doi.org/10.1016/j.jhazmat.2020.123500

Bicho RC, Roelofs D, Mariën J, et al (2020a) Epigenetic effects of (nano)materials in environmental species – Cu case study in Enchytraeus crypticus. Environment International 136:105447. https://doi.org/10.1016/j.envint.2019.105447

Bicho RC, Scott-Fordsmand JJ, Amorim MJB (2020b) Multigenerational Exposure to WCCo Nanomaterials—Epigenetics in the Soil Invertebrate Enchytraeus crypticus. Nanomaterials 10:836. https://doi.org/10.3390/nano10050836

Bingsohn L, Knorr E, Vilcinskas A (2016) The model beetle Tribolium castaneum can be used as an early warning system for transgenerational epigenetic side effects caused by pharmaceuticals. Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology 185–186:57–64. https://doi.org/10.1016/j.cbpc.2016.03.002

Bird AP, Taggart MH (1980) Variable patterns of total DNA and rDNA methylation in animals. Nucleic Acids Research 8:1485–1497. https://doi.org/10.1093/nar/8.7.1485

Bishop J, Jones HE, Lukac M, Potts SG (2016) Insect pollination reduces yield loss following heat stress in faba bean (Vicia faba L.). Agriculture, Ecosystems & Environment 220:89–96. https://doi.org/10.1016/j.agee.2015.12.007

Bloch G (2010) The Social Clock of the Honeybee. Journal of Biological Rhythms 25:307–317. https://doi.org/10.1177/0748730410380149

Bonasio R, Li Q, Lian J, et al (2012) Genome-wide and Caste-Specific DNA Methylomes of the Ants Camponotus floridanus and Harpegnathos saltator. Current Biology 22:1755–1764. https://doi.org/10.1016/j.cub.2012.07.042

Bonilauri P, Bellini R, Calzolari M, et al (2008) Chikungunya Virus in Aedes albopictus, Italy. Emerging Infectious Diseases 14:852–854. https://doi.org/10.3201/eid1405.071144

Bonizzoni M, Gasperi G, Chen X, James AA (2013) The invasive mosquito species Aedes albopictus: current knowledge and future perspectives. Trends in Parasitology 29:460–468. https://doi.org/10.1016/j.pt.2013.07.003

Borsani O, Zhu J, Verslues PE, et al (2005) Endogenous siRNAs Derived from a Pair of Natural cis-Antisense Transcripts Regulate Salt Tolerance in Arabidopsis. Cell 123:1279–1291. https://doi.org/10.1016/j.cell.2005.11.035

Bošković A, Rando OJ (2018) Transgenerational Epigenetic Inheritance. Annual Review of Genetics 52:21–41. https://doi.org/10.1146/annurev-genet-120417-031404

Botías C, Jones JC, Pamminger T, et al (2021) Multiple stressors interact to impair the performance of bumblebee Bombus terrestris colonies. Journal of Animal Ecology 90:415–431. https://doi.org/10.1111/1365-2656.13375

Bourret TJ, Lawrence KA, Shaw JA, et al (2016) The nucleotide excision repair pathway protects Borrelia burgdorferi from nitrosative stress in Ixodes scapularis ticks. Frontiers in Microbiology 7:. https://doi.org/10.3389/fmicb.2016.01397

Bownes SJ, McQuaid CD (2006) Will the invasive mussel Mytilus galloprovincialis Lamarck replace the indigenous Perna perna L. on the south coast of South Africa? Journal of Experimental Marine Biology and Ecology 338:140–151. https://doi.org/10.1016/j.jembe.2006.07.006

Braga AC, Pereira V, Marçal R, et al (2020) DNA damage and oxidative stress responses of mussels Mytilus galloprovincialis to paralytic shellfish toxins under warming and acidification conditions – Elucidation on the organ-specificity. Aquatic Toxicology 228:105619. https://doi.org/10.1016/j.aquatox.2020.105619

Brand SJ, Erasmus JH, Labuschagne M, et al (2019) Bioaccumulation and metal-associated biomarker responses in a freshwater mussel, Dreissena polymorpha, following short-term platinum exposure. Environmental Pollution 246:69–78. https://doi.org/10.1016/j.envpol.2018.11.061

Brenner S (1974) THE GENETICS OF CAENORHABDITIS ELEGANS. Genetics 77:71–94. https://doi.org/10.1093/genetics/77.1.71

Bridges EM, Oldeman LR (1999) Global Assessment of Human-Induced Soil Degradation. Arid Soil Research and Rehabilitation 13:319–325. https://doi.org/10.1080/089030699263212

Brownstein JS, Holford TR, Fish D (2005) Effect of Climate Change on Lyme Disease Risk in North America. EcoHealth 2:38–46. https://doi.org/10.1007/s10393-004-0139-x

Buckley SB (1867) Descriptions of new species of North American Formicidae (continued from page 172.). Proc Entomol Soc Phila 6:

Burke RD, Angerer LM, Elphick MR, et al (2006) A genomic view of the sea urchin nervous system. Developmental Biology 300:434–460. https://doi.org/10.1016/j.ydbio.2006.08.007

Burkhart BA, Kennett SB, Archer TK (2007) Osmotic stress-dependent repression is mediated by histone H3 phosphorylation and chromatin structure. Journal of Biological Chemistry 282:4400–4407. https://doi.org/10.1074/jbc.M609041200

Burton RS, Byrne RJ, Rawson PD (2007) Three divergent mitochondrial genomes from California populations of the copepod Tigriopus californicus. Gene 403:53–59. https://doi.org/10.1016/j.gene.2007.07.026

Burton RS, Feldman MW (1982) Changes in free amino acid concentrations during osmotic response in the intertidal copepod Tigriopus californicus. Comparative Biochemistry and Physiology Part A: Physiology 73:441–445. https://doi.org/10.1016/0300-9629(82)90182-7

Busby AT, Ayllón N, Kocan KM, et al (2012) Expression of heat shock proteins and subolesin affects stress responses, Anaplasma phagocytophilum infection and questing behaviour in the tick, Ixodes scapularis. Medical and Veterinary Entomology 26:92–102. https://doi.org/10.1111/j.1365-2915.2011.00973.x

Bushati N, Cohen SM (2007) microRNA Functions. Annual Review of Cell and Developmental Biology 23:175–205. https://doi.org/10.1146/annurev.cellbio.23.090506.123406

Camacho J, Truong L, Kurt Z, et al (2018) The Memory of Environmental Chemical Exposure in C. elegans Is Dependent on the Jumonji Demethylases jmjd-2 and jmjd-3/utx-1. Cell Reports 23:2392–2404. https://doi.org/10.1016/j.celrep.2018.04.078

Cancrini G, Frangipane di Regalbono A, Ricci I, et al (2003) Aedes albopictus is a natural vector of Dirofilaria immitis in Italy. Veterinary Parasitology 118:195–202. https://doi.org/10.1016/j.vetpar.2003.10.011

Carpenter SR, Stanley EH, vander Zanden MJ (2011a) State of the World's Freshwater Ecosystems: Physical, Chemical, and Biological Changes. Annual Review of Environment and Resources 36:75–99. https://doi.org/10.1146/annurev-environ-021810-094524

Carpenter SR, Stanley EH, vander Zanden MJ (2011b) State of the World's Freshwater Ecosystems: Physical, Chemical, and Biological Changes. Annual Review of Environment and Resources 36:75–99. https://doi.org/10.1146/annurev-environ-021810-094524

Cázares-Raga FE, Chávez-Munguía B, González-Calixto C, et al (2014) Morphological and proteomic characterization of midgut of the malaria vector Anopheles albimanus at early time after a blood feeding. Journal of Proteomics 111:100–112. https://doi.org/10.1016/j.jprot.2014.07.037

Cedar H, Bergman Y (2012) Programming of DNA Methylation Patterns. Annual Review of Biochemistry 81:97–117. https://doi.org/10.1146/annurev-biochem-052610-091920

Cesar H, Burke L, Pet-soede L (2003) The Economics of Worldwide Coral Reef Degradation. Cesar Environmental Economics Consulting, Arnhem, and WWF-Netherlands 14:

Chareonviriyaphap T, Roberts DR, Andre RG, et al (1997) Pesticide avoidance behavior in Anopheles albimanus, a malaria vector in the Americas. Journal of the American Mosquito Control Association 13:

Chatterjee N, Choi S, Kwon OK, et al (2019) Multi-generational impacts of organic contaminated stream water on Daphnia magna: A combined proteomics, epigenetics and ecotoxicity approach. Environmental Pollution 249:217–224. https://doi.org/10.1016/j.envpol.2019.03.028

Chatterjee N, Walker GC (2017) Mechanisms of DNA damage, repair, and mutagenesis. Environmental and Molecular Mutagenesis 58:235–263

Chaudhry M (2013) Expression Pattern of Small Nucleolar RNA Host Genes and Long Non-Coding RNA in X-rays-Treated Lymphoblastoid Cells. International Journal of Molecular Sciences 14:9099–9110. https://doi.org/10.3390/ijms14059099

Chen P, Xiao W-F, Pan M-H, et al (2020) Comparative genome-wide DNA methylation analysis reveals epigenomic differences in response to heat-humidity stress in Bombyx mori. International Journal of Biological Macromolecules 164:3771–3779. https://doi.org/10.1016/j.ijbiomac.2020.08.251

Chirault M, Lucas C, Goubault M, et al (2015) A Combined Approach to Heat Stress Effect on Male Fertility in Nasonia vitripennis: From the Physiological Consequences on Spermatogenesis to the Reproductive Adjustment of Females Mated with Stressed Males. PLOS ONE 10:e0120656. https://doi.org/10.1371/journal.pone.0120656

Christen V, Mittner F, Fent K (2016) Molecular Effects of Neonicotinoids in Honey Bees (Apis mellifera). Environmental Science & Technology 50:4071–4081. https://doi.org/10.1021/acs.est.6b00678

Christophorou MA, Castelo-Branco G, Halley-Stott RP, et al (2014) Citrullination regulates pluripotency and histone H1 binding to chromatin. Nature 507:104–108. https://doi.org/10.1038/nature12942

Claudio-Piedras F, Recio-Tótoro B, Condé R, et al (2020) DNA Methylation in Anopheles albimanus Modulates the Midgut Immune Response Against Plasmodium berghei. Frontiers in Immunology 10:. https://doi.org/10.3389/fimmu.2019.03025

Collins CL, Burnett NP, Ramsey MJ, et al (2020) Physiological responses to heat stress in an invasive mussel Mytilus galloprovincialis depend on tidal habitat. Marine Environmental Research 154:104849. https://doi.org/10.1016/j.marenvres.2019.104849

Colovic MB, Krstic DZ, Lazarevic-Pasti TD, et al (2013) Acetylcholinesterase Inhibitors: Pharmacology and Toxicology. Current Neuropharmacology 11:315–335. https://doi.org/10.2174/1570159X11311030006

Connelly NA, O’Neill CR, Knuth BA, Brown TL (2007) Economic Impacts of Zebra Mussels on Drinking Water Treatment and Electric Power Generation Facilities. Environmental Management 40:105–112. https://doi.org/10.1007/s00267-006-0296-5

Cook SJ, Jarrell TA, Brittin CA, et al (2019) Whole-animal connectomes of both Caenorhabditis elegans sexes. Nature 571:63–71. https://doi.org/10.1038/s41586-019-1352-7

Cornette R, Kikawada T (2011) The induction of anhydrobiosis in the sleeping chironomid: Current status of our knowledge. IUBMB Life 63:419–429. https://doi.org/10.1002/iub.463

Costa D, Mariën J, Janssens TKS, et al (2012) Influence of adaptive evolution of cadmium tolerance on neutral and functional genetic variation in Orchesella cincta. Ecotoxicology 21:2078–2087. https://doi.org/10.1007/s10646-012-0961-9

Coulondre C, Miller JH, Farabaugh PJ, Gilbert W (1978) Molecular basis of base substitution hotspots in Escherichia coli. Nature 274:775–780. https://doi.org/10.1038/274775a0

Cribiu P, Chaumot A, Geffard O, et al (2018) Natural variability and modulation by environmental stressors of global genomic cytosine methylation levels in a freshwater crustacean, Gammarus fossarum. Aquatic Toxicology 205:11–18. https://doi.org/10.1016/j.aquatox.2018.09.015

Crompton DWT (1999) How Much Human Helminthiasis Is There in the World? The Journal of Parasitology 85:397. https://doi.org/10.2307/3285768

Cuthbert GL, Daujat S, Snowden AW, et al (2004) Histone Deimination Antagonizes Arginine Methylation. Cell 118:545–553. https://doi.org/10.1016/j.cell.2004.08.020

da Gloria da Costa Carvalho M, Freitas MS (1988) Effect of continuous heat stress on cell growth and protein synthesis in Aedes albopictus. Journal of Cellular Physiology 137:455–461. https://doi.org/10.1002/jcp.1041370309

da Silva Cantinha R, Borrely SI, Oguiura N, et al (2017) HSP70 expression in Biomphalaria glabrata snails exposed to cadmium. Ecotoxicology and Environmental Safety 140:18–23. https://doi.org/10.1016/j.ecoenv.2017.02.026

Dafni A, Kevan P, Gross CL, Goka K (2010) Bombus terrestris, pollinator, invasive and pest: An assessment of problems associated with its widespread introductions for commercial purposes. Applied Entomology and Zoology 45:101–113. https://doi.org/10.1303/aez.2010.101

Dai W, Holmstrup M, Slotsbo S, et al (2020) Compartmentation and effects of lead (Pb) in the collembolan, Folsomia candida. Environmental Science and Pollution Research 27:43638–43645. https://doi.org/10.1007/s11356-020-10300-6

Darling JA, Reitzel AR, Burton PM, et al (2005) Rising starlet: the starlet sea anemone,Nematostella vectensis. BioEssays 27:211–221. https://doi.org/10.1002/bies.20181

de GELAS K, de MEESTER L (2005) Phylogeography of Daphnia magna in Europe. Molecular Ecology 14:753–764. https://doi.org/10.1111/j.1365-294X.2004.02434.x

de Meester L (1993) Genotype, Fish-Mediated Chemical, and Phototactic Behavior in Daphnia Magna. Ecology 74:1467–1474. https://doi.org/10.2307/1940075

de Mendoza A, Hatleberg WL, Pang K, et al (2019) Convergent evolution of a vertebrate-like methylome in a marine sponge. Nature Ecology and Evolution 3:1464–1473. https://doi.org/10.1038/s41559-019-0983-2

de Vos JM, Joppa LN, Gittleman JL, et al (2015) Estimating the normal background rate of species extinction. Conservation Biology 29:452–462. https://doi.org/10.1111/cobi.12380

Dean W (2016) Pathways of DNA Demethylation. In: Advances in Experimental Medicine and Biology. Springer New York LLC, pp 247–274

DeBiasse MB, Kawji Y, Kelly MW (2018) Phenotypic and transcriptomic responses to salinity stress across genetically and geographically divergent Tigriopus californicus populations. Molecular Ecology 27:1621–1632. https://doi.org/10.1111/mec.14547

Dehal P, Satou Y, Campbell RK, et al (2002) The Draft Genome of Ciona intestinalis: Insights into Chordate and Vertebrate Origins. Science 298:2157–2167. https://doi.org/10.1126/science.1080049

Denison R, Raymond-Delpech V (2008) Insights into the molecular basis of social behaviour from studies on the honeybee, Apis mellifera. Invertebrate Neuroscience 8:1–9. https://doi.org/10.1007/s10158-008-0066-6

Dennis DT, Nekomoto TS, Victor JC, et al (1998) Forum: Reported Distribution of Ixodes scapularis and Ixodes pacificus (Acari: Ixodidae) in the United States. Journal of Medical Entomology 35:629–638. https://doi.org/10.1093/jmedent/35.5.629

Deshmukh S, Ponnaluri VC, Dai N, et al (2018) Levels of DNA cytosine methylation in the Drosophila genome. PeerJ 6:e5119. https://doi.org/10.7717/peerj.5119

Dethier MN (1980) Tidepools as refuges: Predation and the limits of the harpacticoid copepod Tigriopus californicus (Baker). Journal of Experimental Marine Biology and Ecology 42:99–111. https://doi.org/10.1016/0022-0981(80)90169-0

di Cristo C (2013) Nervous control of reproduction in Octopus vulgaris: a new model. Invertebrate Neuroscience 13:27–34. https://doi.org/10.1007/s10158-013-0149-x

Dietz DM, Nestler EJ (2012) From Father to Offspring: Paternal Transmission of Depressive-Like Behaviors. Neuropsychopharmacology 37:311–312. https://doi.org/10.1038/npp.2011.167

Dimond JL, Roberts SB (2016) Germline DNA methylation in reef corals: patterns and potential roles in response to environmental change. Molecular Ecology 25:1895–1904. https://doi.org/10.1111/mec.13414

Ding G, Wang L, Zhang J, et al (2015) Toxicity and DNA methylation changes induced by perfluorooctane sulfonate (PFOS) in sea urchin Glyptocidaris crenularis. Chemosphere 128:225–230. https://doi.org/10.1016/j.chemosphere.2015.01.045

Dirzo R, Young HS, Galetti M, et al (2014) Defaunation in the Anthropocene. Science 345:401–406. https://doi.org/10.1126/science.1251817

Doane DP, Seward LE (2011) Measuring skewness: A forgotten statistic? Journal of Statistics Education 19:. https://doi.org/10.1080/10691898.2011.11889611

Doney SC, Busch DS, Cooley SR, Kroeker KJ (2020) The Impacts of Ocean Acidification on Marine Ecosystems and Reliant Human Communities. Annual Review of Environment and Resources 45:83–112. https://doi.org/10.1146/annurev-environ-012320-083019

Duan Y, Dong H, Wang Y, et al (2017) Intestine oxidative stress and immune response to sulfide stress in Pacific white shrimp Litopenaeus vannamei. Fish & Shellfish Immunology 63:201–207. https://doi.org/10.1016/j.fsi.2017.02.013

Duan Y, Xiong D, Wang Y, et al (2021) Toxic effects of ammonia and thermal stress on the intestinal microbiota and transcriptomic and metabolomic responses of Litopenaeus vannamei. Science of The Total Environment 754:141867. https://doi.org/10.1016/j.scitotenv.2020.141867

Dubois O (2011) The State Of The Worls's Land And Water Resources For Food And Agriculture Managing systems: Managing systems at risk from Routledge

Duchateau MJ, Velthuis HHW (1988) Development and Reproductive Strategies in Bombus Terrestris Colonies. Behaviour 107:186–207. https://doi.org/10.1163/156853988X00340

Dudgeon D, Arthington AH, Gessner MO, et al (2006a) Freshwater biodiversity: importance, threats, status and conservation challenges. Biological Reviews 81:163. https://doi.org/10.1017/S1464793105006950

Dudgeon D, Arthington AH, Gessner MO, et al (2006b) Freshwater biodiversity: importance, threats, status and conservation challenges. Biological Reviews 81:163. https://doi.org/10.1017/S1464793105006950

DURANCE I, ORMEROD SJ (2009) Trends in water quality and discharge confound long-term warming effects on river macroinvertebrates. Freshwater Biology 54:388–405. https://doi.org/10.1111/j.1365-2427.2008.02112.x

Ebert TA, Southon JR (2003) Red sea urchins (Strongylocentrotus franciscanus) can live over 100 years: Confirmation with A-bomb 14carbon. Fishery Bulletin 101:

Ehrenfeld JG (2010) Ecosystem Consequences of Biological Invasions. Annual Review of Ecology, Evolution, and Systematics 41:59–80. https://doi.org/10.1146/annurev-ecolsys-102209-144650

Elango N, Hunt BG, Goodisman MAD, Yi S v. (2009) DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, Apis mellifera. Proceedings of the National Academy of Sciences of the United States of America 106:11206–11211. https://doi.org/10.1073/pnas.0900301106

Elia AC, Burioli E, Magara G, et al (2020) Oxidative stress ecology on Pacific oyster Crassostrea gigas from lagoon and offshore Italian sites. Science of The Total Environment 739:139886. https://doi.org/10.1016/j.scitotenv.2020.139886

Ellis L-JayneA, Kissane S, Lynch I (2020) Maternal Responses and Adaptive Changes to Environmental Stress via Chronic Nanomaterial Exposure: Differences in Inter and

Transgenerational Interclonal Broods of Daphnia magna. International Journal of Molecular Sciences 22:15. https://doi.org/10.3390/ijms22010015

Elran R, Raam M, Kraus R, et al (2014) Early and late response of Nematostella vectensis transcriptome to heavy metals. Molecular Ecology 23:4722–4736. https://doi.org/10.1111/mec.12891

Ernst C, Morton CC (2013) Identification and function of long non-coding RNA. Frontiers in Cellular Neuroscience 7:. https://doi.org/10.3389/fncel.2013.00168

Evans TG, Watson-Wynn P (2014) Effects of Seawater Acidification on Gene Expression: Resolving Broader-Scale Trends in Sea Urchins. The Biological Bulletin 226:237–254. https://doi.org/10.1086/BBLv226n3p237

Faddeeva-Vakhrusheva A, Kraaijeveld K, Derks MFL, et al (2017) Coping with living in the soil: The genome of the parthenogenetic springtail Folsomia candida. BMC Genomics 18:. https://doi.org/10.1186/s12864-017-3852-x

Falkenberg LJ, Bellerby RGJ, Connell SD, et al (2020) Ocean acidification and human health. International Journal of Environmental Research and Public Health 17:1–20

Fan L, Wang L, Wang Z (2019) Proteomic characterization of the hepatopancreas in the Pacific white shrimp Litopenaeus vannamei under cold stress: Revealing the organism homeostasis mechanism. Fish & Shellfish Immunology 92:438–449. https://doi.org/10.1016/j.fsi.2019.06.037

Fan W, Ren J, Li X, et al (2015) Bioaccumulation and oxidative stress in *Daphnia magna* exposed to arsenite and arsenate. Environmental Toxicology and Chemistry 34:2629–2635. https://doi.org/10.1002/etc.3119

Fan W, Zhang Y, Liu S, et al (2020) Alleviation of copper toxicity in Daphnia magna by hydrogen nanobubble water. Journal of Hazardous Materials 389:122155. https://doi.org/10.1016/j.jhazmat.2020.122155

FAO (2001) Insect damage: damage on post-harvest. INPhO-Post-harvest Compendium.

FAO (2018) The state of world fisheries and aquaculture 2018 : meeting the sustainable development goals. Food and Agriculture Organization of the United Nations

Feidantsis K, Giantsis IA, Vratsistas A, et al (2020) Correlation between intermediary metabolism, Hsp gene expression, and oxidative stress-related proteins in long-term thermal-stressed Mytilus galloprovincialis. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology 319:R264–R281. https://doi.org/10.1152/ajpregu.00066.2020

Feliciello I, Parazajder J, Akrap I, Ugarković Đ (2013) First evidence of DNA methylation in insect Tribolium castaneum . Epigenetics 8:534–541. https://doi.org/10.4161/epi.24507

Fenoglio S, Bo T, Cucco M, et al (2010) Effects of global climate change on freshwater biota: A review with special emphasis on the Italian situation. Italian Journal of Zoology 77:374–383. https://doi.org/10.1080/11250000903176497

Ferro D, Franchi N, Mangano V, et al (2013) Characterization and metal-induced gene transcription of two new copper zinc superoxide dismutases in the solitary ascidian Ciona intestinalis. Aquatic Toxicology 140–141:369–379. https://doi.org/10.1016/j.aquatox.2013.06.020

Field CB, Barros VR, Dokken DJ, et al (2014) Climate Change 2014 Impacts, Adaptation, and Vulnerability. Cambridge University Press, Cambridge

Fneich S, Dheilly N, Adema C, et al (2013) 5-methyl-cytosine and 5-hydroxy-methyl-cytosine in the genome of Biomphalaria glabrata, a snail intermediate host of Schistosoma mansoni. Parasites & Vectors 6:167. https://doi.org/10.1186/1756-3305-6-167

Food and Agriculture Organization of the United Nations (2020) The State of World Fisheries and Aquaculture 2020. FAO

Fountain MT, Hopkin SP (2005) Folsomia candida (Collembola): A “standard” soil arthropod. Annual Review of Entomology 50:201–222

Franchi N, Ferro D, Ballarin L, Santovito G (2012) Transcription of genes involved in glutathione biosynthesis in the solitary tunicate Ciona intestinalis exposed to metals. Aquatic Toxicology 114–115:14–22. https://doi.org/10.1016/j.aquatox.2012.02.007

Friedman LE, Gilmore TD, Finnerty JR (2018) Intraspecific variation in oxidative stress tolerance in a model cnidarian: Differences in peroxide sensitivity between and within populations of Nematostella vectensis. PLOS ONE 13:e0188265. https://doi.org/10.1371/journal.pone.0188265

Fryxell KJ, Moon W-J (2005) CpG Mutation Rates in the Human Genome Are Highly Dependent on Local GC Content. Molecular Biology and Evolution 22:650–658. https://doi.org/10.1093/molbev/msi043

Fujikawa T, Munakata T, Kondo S, et al (2010) Stress response in the ascidian Ciona intestinalis: transcriptional profiling of genes for the heat shock protein 70 chaperone system under heat stress and endoplasmic reticulum stress. Cell Stress and Chaperones 15:193–204. https://doi.org/10.1007/s12192-009-0133-x

Gallo A, Boni R, Buttino I, Tosti E (2016) Spermiotoxicity of nickel nanoparticles in the marine invertebrate Ciona intestinalis (ascidians). Nanotoxicology 10:1096–1104. https://doi.org/10.1080/17435390.2016.1177743

Gallo A, Tosti E (2015) The Ascidian Ciona Intestinalis as Model Organism for Ecotoxicological Bioassays. Journal of Marine Science: Research & Development 05: https://doi.org/10.4172/2155-9910.1000e138

Gao R, Li C-L, Tong X-L, et al (2020) Identification, expression, and artificial selection of silkworm epigenetic modification enzymes. BMC Genomics 21:740. https://doi.org/10.1186/s12864-020-07155-z

Garcia MJ, Teets NM (2019) Cold stress results in sustained locomotor and behavioral deficits in Drosophila melanogaster. Journal of Experimental Zoology Part A: Ecological and Integrative Physiology 331:192–200. https://doi.org/10.1002/jez.2253

Garrett AD, Brennan RS, Steinhart AL, et al (2020) Unique Genomic and Phenotypic Responses to Extreme and Variable pH Conditions in Purple Urchin Larvae. Integrative and Comparative Biology 60:318–331. https://doi.org/10.1093/icb/icaa072

Gatzmann F, Falckenhayn C, Gutekunst J, et al (2018) The methylome of the marbled crayfish links gene body methylation to stable expression of poorly accessible genes. Epigenetics & Chromatin 11:57. https://doi.org/10.1186/s13072-018-0229-6

Gaw S, Thomas K v., Hutchinson TH (2014) Sources, impacts and trends of pharmaceuticals in the marine and coastal environment. Philosophical Transactions of the Royal Society B: Biological Sciences 369:20130572

Gelembiuk GW, May GE, Lee CE (2006) Phylogeography and systematics of zebra mussels and related species. Molecular Ecology 15:. https://doi.org/10.1111/j.1365-294X.2006.02816.x

Genikhovich G, Technau U (2009) The Starlet Sea Anemone Nematostella vectensis: An Anthozoan Model Organism for Studies in Comparative Genomics and Functional Evolutionary Developmental Biology. Cold Spring Harbor Protocols 2009:pdb.emo129-pdb.emo129. https://doi.org/10.1101/pdb.emo129

Gibney ER, Nolan CM (2010) Epigenetics and gene expression. Heredity 105:4–13. https://doi.org/10.1038/hdy.2010.54

Gilmour PS, Rahman I, Donaldson K, MacNee W (2003) Histone acetylation regulates epithelial IL-8 release mediated by oxidative stress from environmental particles. American Journal of Physiology - Lung Cellular and Molecular Physiology 284:L533–L540. https://doi.org/10.1152/ajplung.00277.2002

Ginsberg HS, Albert M, Acevedo L, et al (2017) Environmental Factors Affecting Survival of Immature Ixodes scapularis and Implications for Geographical Distribution of Lyme Disease: The Climate/Behavior Hypothesis. PLOS ONE 12:e0168723. https://doi.org/10.1371/journal.pone.0168723

Giuliani ME, Filippini G, Nardi A (2020) Season specific influence of projected ocean changes on the response to cadmium of stress-related genes in Mytilus galloprovincialis. Marine Environmental Research 162:105091. https://doi.org/10.1016/j.marenvres.2020.105091

Glastad KM, Hunt BG, Yi S v., Goodisman MAD (2011) DNA methylation in insects: on the brink of the epigenomic era. Insect Molecular Biology 20:553–565. https://doi.org/10.1111/j.1365-2583.2011.01092.x

Goldberg L, Lagomasino D, Thomas N, Fatoyinbo T (2020) Global declines in human-driven mangrove loss. Global Change Biology 26:5844–5855. https://doi.org/10.1111/gcb.15275

Goldstein B, Nance J (2020) Caenorhabditis elegans Gastrulation: A Model for Understanding How Cells Polarize, Change Shape, and Journey Toward the Center of an Embryo. Genetics 214:265–277. https://doi.org/10.1534/genetics.119.300240

Gómez GF, Márquez EJ, Gutiérrez LA, et al (2014) Geometric morphometric analysis of Colombian Anopheles albimanus (Diptera: Culicidae) reveals significant effect of environmental factors on wing traits and presence of a metapopulation. Acta Tropica 135:75–85. https://doi.org/10.1016/j.actatropica.2014.03.020

Goulson D, Hughes W, Derwent L, Stout J (2002) Colony growth of the bumblebee, Bombus terrestris, in improved and conventional agricultural and suburban habitats. Oecologia 130:267–273. https://doi.org/10.1007/s004420100803

Goulson D, Nicholls E, Botías C, Rotheray EL (2015) Bee declines driven by combined stress from parasites, pesticides, and lack of flowers. Science 347:. https://doi.org/10.1126/science.1255957

Graf DL (2013) Patterns of freshwater bivalve global diversity and the state of phylogenetic studies on the Unionoida, Sphaeriidae, and Cyrenidae. In: American Malacological Bulletin

Greenberg MVC, Bourc'his D (2019) The diverse roles of DNA methylation in mammalian development and disease. Nature Reviews Molecular Cell Biology 20:590–607. https://doi.org/10.1038/s41580-019-0159-6

Groeneveld RA (1991) An Influence Function Approach to Describing the Skewness of a Distribution. The American Statistician 45:. https://doi.org/10.2307/2684367

Grunstein M (1997) Histone acetylation in chromatin structure and transcription. Nature 389:349–352. https://doi.org/10.1038/38664

Grünwald S, Adam I v., Gurmai A-M, et al (2013) The Red Flour Beetle Tribolium castaneum as a Model to Monitor Food Safety and Functionality. In: Advances in biochemical engineering/biotechnology. pp 111–122

Guan D-L, Ding R-R, Hu X-Y, et al (2019) Cadmium-induced genome-wide DNA methylation changes in growth and oxidative metabolism in Drosophila melanogaster. BMC Genomics 20:356. https://doi.org/10.1186/s12864-019-5688-z

Gulia-Nuss M, Nuss AB, Meyer JM, et al (2016) Genomic insights into the Ixodes scapularis tick vector of Lyme disease. Nature Communications 7:10507. https://doi.org/10.1038/ncomms10507

Gurevich A, Saveliev V, Vyahhi N, Tesler G (2013) QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086

Gusev O, Cornette R, Kikawada T, Okuda T (2011) Expression of heat shock protein-coding genes associated with anhydrobiosis in an African chironomid Polypedilum vanderplanki. Cell Stress and Chaperones 16:81–90. https://doi.org/10.1007/s12192-010-0223-9

Hammond LM, Hofmann GE (2010) Thermal tolerance of Strongylocentrotus purpuratus early life history stages: mortality, stress-induced gene expression and biogeographic patterns. Marine Biology 157:2677–2687. https://doi.org/10.1007/s00227-010-1528-z

Hand C, Uhlinger KR (1994) The Unique, Widely Distributed, Estuarine Sea Anemone, Nematostella vectensis Stephenson: A Review, New Facts, and Questions. Estuaries 17:501. https://doi.org/10.2307/1352679

Hanke FD, Kelber A (2020) The Eye of the Common Octopus (Octopus vulgaris). Frontiers in Physiology 10:. https://doi.org/10.3389/fphys.2019.01637

Harris KD, Lloyd JPB, Domb K, et al (2019) DNA methylation is maintained with high fidelity in the honey bee germline and exhibits global non-functional fluctuations during somatic development. Epigenetics & Chromatin 12:62. https://doi.org/10.1186/s13072-019-0307-4

Harrison JF, Waters JS, Biddulph TA, et al (2018) Developmental plasticity and stability in the tracheal networks supplying Drosophila flight muscle in response to rearing oxygen level. Journal of Insect Physiology 106:189–198. https://doi.org/10.1016/j.jinsphys.2017.09.006

Hasselblad V (1966) Estimation of Parameters for a Mixture of Normal Distributions. Technometrics 8:431–444. https://doi.org/10.1080/00401706.1966.10490375

Havard T, Laurent M, Chauzat M-P (2019) Impact of Stressors on Honey Bees (Apis mellifera; Hymenoptera: Apidae): Some Guidance for Research Emerge from a Meta-Analysis. Diversity 12:7. https://doi.org/10.3390/d12010007

Hearn J, Pearson M, Blaxter M, et al (2019) Genome-wide methylation is modified by caloric restriction in Daphnia magna. BMC Genomics 20:197. https://doi.org/10.1186/s12864-019-5578-4

Hidalgo K, Beaugeard E, Renault D, et al (2019) Physiological and biochemical responses to thermal stress vary among genotypes in the parasitic wasp Nasonia vitripennis. Journal of Insect Physiology 117:103909. https://doi.org/10.1016/j.jinsphys.2019.103909

Hinton HE (1960) A Fly Larva that tolerates Dehydration and Temperatures of −270° to +102° C. Nature 188:336–337. https://doi.org/10.1038/188336a0

Holthuis LB, Farfante IP, Kensley B, Farfante IP (1998) Penaeoid and Sergestoid Shrimps and Prawns of the World. Keys and Diagnoses for the Families and Genera. Journal of Crustacean Biology 18:621. https://doi.org/10.2307/1549427

Hopkin SP (1997) Biology of the springtails (Insecta: Collembola)

Horn DJ (1994) Destructive and Useful Insects, 5th edition. American Entomologist 40:255–255. https://doi.org/10.1093/ae/40.4.255a

Hosking JRM (1992) Moments or l moments? An example comparing two measures of distributional shape. American Statistician 46:. https://doi.org/10.1080/00031305.1992.10475880

Hu C-W, Chen J-L, Hsu Y-W, et al (2015) Trace analysis of methylated and hydroxymethylated cytosines in DNA by isotope-dilution LC–MS/MS: first evidence of DNA methylation in Caenorhabditis elegans. Biochemical Journal 465:39–47. https://doi.org/10.1042/BJ20140844

Huttenhofer A (2006) Experimental approaches to identify non-coding RNAs. Nucleic Acids Research 34:635–646. https://doi.org/10.1093/nar/gkj469

Hyun K, Jeon J, Park K, Kim J (2017) Writing, erasing and reading histone lysine methylations. Experimental & Molecular Medicine 49:e324–e324. https://doi.org/10.1038/emm.2017.11

Im J, Chatterjee N, Choi J (2019) Genetic, epigenetic, and developmental toxicity of Chironomus riparius raised in metal-contaminated field sediments: A multi-generational study with arsenic as a second challenge. Science of the Total Environment 672:. https://doi.org/10.1016/j.scitotenv.2019.04.013

Inoue MN, Yokoyama J, Washitani I (2008) Displacement of Japanese native bumblebees by the recently introduced Bombus terrestris (L.) (Hymenoptera: Apidae). Journal of Insect Conservation 12:135–146. https://doi.org/10.1007/s10841-007-9071-z

Ittiprasert W, Nene R, Miller A, et al (2009) Schistosoma mansoni infection of juvenile Biomphalaria glabrata induces a differential stress response between resistant and susceptible snails. Experimental Parasitology 123:203–211. https://doi.org/10.1016/j.exppara.2009.07.015

Jaenisch R, Bird A (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nature Genetics 33:245–254. https://doi.org/10.1038/ng1089

Jayapalan JJ, Subramanian P, Kani A, et al (2020) Hesperidin modulates the rhythmic proteomic profiling in Drosophila melanogaster under oxidative stress. Archives of Insect Biochemistry and Physiology 105:. https://doi.org/10.1002/arch.21738

Jensen TG, Holmstrup M, Madsen RB, et al (2020) Effects of α-pinene on life history traits and stress tolerance in the springtail Folsomia candida. Comparative Biochemistry and Physiology Part - C: Toxicology and Pharmacology 229:108681. https://doi.org/10.1016/j.cbpc.2019.108681

Jeremias G, Barbosa J, Marques SM, et al (2018a) Synthesizing the role of epigenetics in the response and adaptation of species to climate change in freshwater ecosystems. Molecular Ecology 27:2790–2806. https://doi.org/10.1111/mec.14727

Jeremias G, Barbosa J, Marques SM, et al (2018b) Transgenerational Inheritance of DNA Hypomethylation in Daphnia magna in Response to Salinity Stress. Environmental Science & Technology 52:10114–10123. https://doi.org/10.1021/acs.est.8b03225

Jeremias G, Gonçalves FJM, Pereira JL, Asselman J (2020) Prospects for incorporation of epigenetic biomarkers in human health and environmental risk assessment of chemicals. Biological Reviews 95:822–846. https://doi.org/10.1111/brv.12589

Jiang H, Zhang N, Chen M, et al (2019) Transcriptional and post-translational activation of AMPKα by oxidative, heat, and cold stresses in the red flour beetle, Tribolium castaneum. Cell Stress and Chaperones 24:1079–1089. https://doi.org/10.1007/s12192-019-01030-x

Joyner CC, Frew S (1991) Plastic pollution in the marine environment. Ocean Development and International Law 22:33–69. https://doi.org/10.1080/00908329109545949

Kamgang B, Vazeille M, Yougang AP, et al (2019) Potential of Aedes albopictus and Aedes aegypti (Diptera: Culicidae) to transmit yellow fever virus in urban areas in Central Africa. Emerging Microbes & Infections 8:1636–1641. https://doi.org/10.1080/22221751.2019.1688097

Kandŭca T, Medakovíc D, Hamer B (2011) Mytilus galloprovincialis as a bioindicator of environmental conditions: The case of the eastern coast of the Adriatic Sea. Isotopes in Environmental and Health Studies 47:. https://doi.org/10.1080/10256016.2011.548866

Karatayev AY, Burlakova LE, Padilla DK (1997) The effects of dreissena polymorpha (Pallas) invasion on aquatic communities in eastern Europe. Journal of Shellfish Research 16:187–203

Kelly MW, Padilla-Gamiño JL, Hofmann GE (2013) Natural variation and the capacity to adapt to ocean acidification in the keystone sea urchin Strongylocentrotus purpuratus. Global Change Biology 19:2536–2546. https://doi.org/10.1111/gcb.12251

Kenyon C (1988) The Nematode Caenorhabditis elegans. Science 240:1448–1453. https://doi.org/10.1126/science.3287621

Kille P, Andre J, Anderson C, et al (2013) DNA sequence variation and methylation in an arsenic tolerant earthworm population. Soil Biology and Biochemistry 57:524–532. https://doi.org/10.1016/j.soilbio.2012.10.014

Klattenhoff C, Xi H, Li C, et al (2009) The Drosophila HP1 Homolog Rhino Is Required for Transposon Silencing and piRNA Production by Dual-Strand Clusters. Cell 138:1137–1149. https://doi.org/10.1016/j.cell.2009.07.014

Klein RD, Nogueira LS, Domingos-Moreira FXV, et al (2019) Effects of sublethal Cd, Zn, and mixture exposures on antioxidant defense and oxidative stress parameters in early life stages of the purple sea urchin Strongylocentrotus purpuratus. Aquatic Toxicology 217:105338. https://doi.org/10.1016/j.aquatox.2019.105338

Knight M, Ittiprasert W, Arican-Goktas HD, Bridger JM (2016) Epigenetic modulation, stress and plasticity in susceptibility of the snail host, Biomphalaria glabrata, to Schistosoma mansoni infection. International Journal for Parasitology 46:389–394. https://doi.org/10.1016/j.ijpara.2016.03.003

Kohli RM, Zhang Y (2013) TET enzymes, TDG and the dynamics of DNA demethylation. Nature 502:472–479. https://doi.org/10.1038/nature12750

Kong APS, Xiao K, Choi KC, et al (2012) Associations between microRNA (miR-21, 126, 155 and 221), albuminuria and heavy metals in Hong Kong Chinese adolescents. Clinica Chimica Acta 413:1053–1057. https://doi.org/10.1016/j.cca.2012.02.014

Kotsarenko K, Vechtova P, Hammerova Z, et al (2020) Newly identified DNA methyltransferases of Ixodes ricinus ticks. Ticks and Tick-borne Diseases 11:101348. https://doi.org/10.1016/j.ttbdis.2019.101348

Kraemer MU, Sinka ME, Duda KA, et al (2015) The global distribution of the arbovirus vectors Aedes aegypti and Ae. albopictus. eLife 4:. https://doi.org/10.7554/eLife.08347

Kumar SV, Wigge PA (2010) H2A.Z-Containing Nucleosomes Mediate the Thermosensory Response in Arabidopsis. Cell 140:136–147. https://doi.org/10.1016/j.cell.2009.11.006

Kurdyukov S, Bullock M (2016) DNA Methylation Analysis: Choosing the Right Method. Biology 5:3. https://doi.org/10.3390/biology5010003

Kurimoto K, Yabuta Y, Ohinata Y, et al (2008) Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice. Genes & Development 22:1617–1635. https://doi.org/10.1101/gad.1649908

Lang-Mladek C, Popova O, Kiok K, et al (2010) Transgenerational Inheritance and Resetting of Stress-Induced Loss of Epigenetic Gene Silencing in Arabidopsis. Molecular Plant 3:594–602. https://doi.org/10.1093/mp/ssq014

Lee J, Phillips MC, Lobo M, Willett CS (2021) Tolerance Patterns and Transcriptomic Response to Extreme and Fluctuating Salinities across Populations of the Intertidal Copepod Tigriopus californicus. Physiological and Biochemical Zoology 94:50–69. https://doi.org/10.1086/712031

Lee S-W, Chatterjee N, Im J-E, et al (2018) Integrated approach of eco-epigenetics and eco-metabolomics on the stress response of bisphenol-A exposure in the aquatic midge Chironomus riparius. Ecotoxicology and Environmental Safety 163:111–116. https://doi.org/10.1016/j.ecoenv.2018.06.084

Leffler M, Greer J (1991) Prepared for the Oyster Ecology Workshop

Lennartsson A, Ekwall K (2009) Histone modification patterns and epigenetic codes. Biochimica et Biophysica Acta (BBA) - General Subjects 1790:863–868. https://doi.org/10.1016/j.bbagen.2008.12.006

Levine JF, Apperson CS, Howard P, et al (1997) Lizards as Hosts for Immature Ixodes scapularis (Acari: Ixodidae) in North Carolina. Journal of Medical Entomology 34:594–598. https://doi.org/10.1093/jmedent/34.6.594

Li A, Li L, Wang W, Zhang G (2020) Acetylome Analysis Reveals Population Differentiation of the Pacific Oyster Crassostrea gigas in Response to Heat Stress. Marine Biotechnology 22:233–245. https://doi.org/10.1007/s10126-020-09947-6

Li G-Y, Zhang Z-Q (2019) Development, lifespan and reproduction of spider mites exposed to predator-induced stress across generations. Biogerontology 20:871–882. https://doi.org/10.1007/s10522-019-09835-0

Li N, Arief N, Edmands S (2019a) Effects of oxidative stress on sex-specific gene expression in the copepod Tigriopus californicus revealed by single individual RNA-seq. Comparative Biochemistry and Physiology Part D: Genomics and Proteomics 31:100608. https://doi.org/10.1016/j.cbd.2019.100608

Li Y-F, Xu J-K, Chen Y-W, et al (2019b) Characterization of Gut Microbiome in the Mussel Mytilus galloprovincialis in Response to Thermal Stress. Frontiers in Physiology 10:. https://doi.org/10.3389/fphys.2019.01086

Lindeman LC, Thaulow J, Song Y, et al (2019) Epigenetic, transcriptional and phenotypic responses in two generations of Daphnia magna exposed to the DNA methylation inhibitor 5-azacytidine. Environmental Epigenetics 5:. https://doi.org/10.1093/eep/dvz016

Liu F, Li S, Yu Y, et al (2020) Effects of ammonia stress on the hemocytes of the Pacific white shrimp Litopenaeus vannamei. Chemosphere 239:124759. https://doi.org/10.1016/j.chemosphere.2019.124759

Lockyer AE, Spinks J, Kane RA, et al (2008) Biomphalaria glabrata transcriptome: cDNA microarray profiling identifies resistant- and susceptible-specific gene expression in haemocytes from snail strains exposed to Schistosoma mansoni. BMC Genomics 9:634. https://doi.org/10.1186/1471-2164-9-634

Louis YD, Bhagooli R, Kenkel CD, et al (2017) Gene expression biomarkers of heat stress in scleractinian corals: Promises and limitations. Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology 191:63–77. https://doi.org/10.1016/j.cbpc.2016.08.007

Lukasz Komsta FN (2015) Moments, cumulants, skewness, kurtosis and related tests

Lyko F, Foret S, Kucharski R, et al (2010) The Honey Bee Epigenomes: Differential Methylation of Brain DNA in Queens and Workers. PLoS Biology 8:e1000506. https://doi.org/10.1371/journal.pbio.1000506

Mack GS (2007) MicroRNA gets down to business. Nature Biotechnology 25:631–638. https://doi.org/10.1038/nbt0607-631

Madin EMP, Dill LM, Ridlon AD, et al (2016a) Human activities change marine ecosystems by altering predation risk. Global Change Biology 22:44–60. https://doi.org/10.1111/gcb.13083

Madin EMP, Dill LM, Ridlon AD, et al (2016b) Human activities change marine ecosystems by altering predation risk. Global Change Biology 22:44–60. https://doi.org/10.1111/gcb.13083

Magni S, Parolini M, della Torre C, et al (2017) Multi-biomarker investigation to assess toxicity induced by two antidepressants on Dreissena polymorpha. Science of The Total Environment 578:452–459. https://doi.org/10.1016/j.scitotenv.2016.10.208

Majoros G, Fehér Z, Deli T, Földvári G (2008) Establishment of Biomphalaria tenagophila Snails in Europe. Emerging Infectious Diseases 14:1812–1814. https://doi.org/10.3201/eid1411.080479

Malagoli D, Casarini L, Fiori F, Ottaviani E (2008) Cytotoxic activity by the mussel Mytilus galloprovincialis and the Venus clam Chamelea gallina in the Adriatic sea in 2007. Invertebrate Survival Journal 5:50–53

Maldonado E, Rangel-Huerta E, González-Gómez R, et al (2019) Octopus insularis as a new marine model for EvoDevo. Biology Open 8:. https://doi.org/10.1242/bio.046086

Mandyam.D. V, Muthangi S (2020) Survival of silk worm, Bombyx mori in azaserine induced oxidative stress. Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology 228:108637. https://doi.org/10.1016/j.cbpc.2019.108637

Mannelli A, Kitron U, Jones CJ, Slajchert TL (1994) Influence of Season and Habitat on Ixodes scapularis Infestation on White-Footed Mice in Northwestern Illinois. The Journal of Parasitology 80:1038. https://doi.org/10.2307/3283457

Maria VL, Ribeiro MJ, Amorim MJB (2014) Oxidative stress biomarkers and metallothionein in Folsomia candida - responses to Cu and Cd. Environmental Research 133:164–169. https://doi.org/10.1016/j.envres.2014.05.027

Martínez-Barnetche J, Gómez-Barreto RE, Ovilla-Muñoz M, et al (2012) Transcriptome of the adult female malaria mosquito vector Anopheles albimanus. BMC Genomics 13:207. https://doi.org/10.1186/1471-2164-13-207

Materna SC, Berney K, Cameron RA (2006) The S. purpuratus genome: A comparative perspective. Developmental Biology 300:485–495. https://doi.org/10.1016/j.ydbio.2006.09.033

Matozzo V, Chinellato A, Munari M, et al (2012) First Evidence of Immunomodulation in Bivalves under Seawater Acidification and Increased Temperature. PLoS ONE 7:e33820. https://doi.org/10.1371/journal.pone.0033820

May RM (1988) How Many Species are There on Earth?

Mazin P v., Shagimardanova E, Kozlova O, et al (2018) Cooption of heat shock regulatory system for anhydrobiosis in the sleeping chironomid Polypedilum vanderplanki. Proceedings of the National Academy of Sciences 115:E2477–E2486. https://doi.org/10.1073/pnas.1719493115

McDonald J (2004) The invasive pest species Ciona intestinalis (Linnaeus, 1767) reported in a harbour in southern Western Australia. Marine Pollution Bulletin 49:868–870. https://doi.org/10.1016/j.marpolbul.2004.08.024

McGinty RK, Tan S (2015) Nucleosome Structure and Function. Chemical Reviews 115:2255–2273. https://doi.org/10.1021/cr500373h

McKenzie BA, Wilson AE, Zohdy S (2019) Aedes albopictus is a competent vector of Zika virus: A meta-analysis. PLOS ONE 14:e0216794. https://doi.org/10.1371/journal.pone.0216794

Mearns AJ, Reish DJ, Oshida PS, et al (2012) Effects of Pollution on Marine Organisms. Water Environment Research 84:1737–1823. https://doi.org/10.2175/106143012x13407275695751

Meissner A (2005) Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic Acids Research 33:5868–5877. https://doi.org/10.1093/nar/gki901

Menzel R (2012) The honeybee as a model for understanding the basis of cognition. Nature Reviews Neuroscience 13:758–768. https://doi.org/10.1038/nrn3357

Micael J, Rodrigues P, Halldórsson HP, Gíslason S (2020) Distribution and abundance of the invasive tunicate Ciona intestinalis (Linnaeus, 1767) in Icelandic harbours. Regional Studies in Marine Science 34:101039. https://doi.org/10.1016/j.rsma.2020.101039

Millennium Ecosystem Assessment (2005) Ecosystems and human well-being : synthesis. Island Press 137

Minshull TC, Cole J, Dockrell DH, et al (2016) Analysis of histone post translational modifications in primary monocyte derived macrophages using reverse phase×reverse phase chromatography in conjunction with porous graphitic carbon stationary phase. Journal of Chromatography A 1453:43–53. https://doi.org/10.1016/j.chroma.2016.05.025

Mitta G, Gourbal B, Grunau C, et al (2017) The Compatibility Between Biomphalaria glabrata Snails and Schistosoma mansoni. In: Advances in Parasitology. pp 111–145

Moore C, Lein P, Puschner B (2014) Microcystins Alter Chemotactic Behavior in Caenorhabditis elegans by Selectively Targeting the AWA Sensory Neuron. Toxins 6:1813–1836. https://doi.org/10.3390/toxins6061813

Mulholland CB, Nishiyama A, Ryan J, et al (2020) Recent evolution of a TET-controlled and DPPA3/STELLA-driven pathway of passive DNA demethylation in mammals. Nature Communications 11:5972. https://doi.org/10.1038/s41467-020-19603-1

Müller R, Charaf S, Scherer C, et al (2016) Phenotypic and epigenetic effects of vinclozolin in the gastropod Physella acuta. Journal of Molluscan Studies 82:320–327. https://doi.org/10.1093/mollus/eyv069

Murata Y, Osakabe M (2017) Photo-enzymatic repair of UVB-induced DNA damage in the two-spotted spider mite Tetranychus urticae. Experimental and Applied Acarology 71:15–34. https://doi.org/10.1007/s10493-016-0100-z

Nambudiri VE, Widlund HR (2013) Small Interfering RNA. Journal of Investigative Dermatology 133:1–4. https://doi.org/10.1038/jid.2013.411

Nanty L, Carbajosa G, Heap GA, et al (2011) Comparative methylomics reveals gene-body H3K36me3 in Drosophila predicts DNA methylation and CpG landscapes in other invertebrates. Genome Research 21:. https://doi.org/10.1101/gr.121640.111

Naser HA (2013) Assessment and management of heavy metal pollution in the marine environment of the Arabian Gulf: A review. Marine Pollution Bulletin 72:6–13

Nation JL (2010) Bumblebees: Behaviour, Ecology, and Conservation. Florida Entomologist 93:477–477. https://doi.org/10.1653/024.093.0332

Newbold LK, Robinson A, Rasnaca I, et al (2019) Genetic, epigenetic and microbiome characterisation of an earthworm species (Octolasion lacteum) along a radiation exposure gradient at Chernobyl. Environmental Pollution 255:113238. https://doi.org/10.1016/j.envpol.2019.113238

Newbould PJ, Macfadyen A, Ford ED (1986) Advances in Ecological Research, Volume 14. The Journal of Applied Ecology 23:1066. https://doi.org/10.2307/2403962

Newell G (1936) The Flour Beetles of the Genus Tribolium

Newman MC (2019) Fundamentals of Ecotoxicology. CRC Press, Fifth edition. | Boca Raton : CRC Press, 2020.

Nica D, Popescu C, Draghici G, et al (2017a) Effect of cadmium on cytosine hydroxymethylation in gastropod hepatopancreas. Environmental Science and Pollution Research 24:15187–15195. https://doi.org/10.1007/s11356-017-9104-4

Nica D v., Popescu C, Draghici GA, et al (2017b) High-level dietary cadmium exposure is associated with global DNA hypermethylation in the gastropod hepatopancreas. PLOS ONE 12:e0184221. https://doi.org/10.1371/journal.pone.0184221

Niederwanger M, Dvorak M, Schnegg R, et al (2017) Challenging the Metallothionein (MT) Gene of Biomphalaria glabrata: Unexpected Response Patterns Due to Cadmium Exposure and Temperature Stress. International Journal of Molecular Sciences 18:1747. https://doi.org/10.3390/ijms18081747

Noer NK, Pagter M, Bahrndorff S, et al (2020) Impacts of thermal fluctuations on heat tolerance and its metabolomic basis in Arabidopsis thaliana, Drosophila melanogaster, and Orchesella cincta. PLOS ONE 15:e0237201. https://doi.org/10.1371/journal.pone.0237201

Noordhoek JW, Koning JT, Mariën J, et al (2018) Exploring DNA methylation patterns in copper exposed Folsomia candida and Enchytraeus crypticus. Pedobiologia 66:52–57. https://doi.org/10.1016/j.pedobi.2017.10.001

Norouzitallab P, Baruah K, Vandegehuchte M, et al (2014) Environmental heat stress induces epigenetic transgenerational inheritance of robustness in parthenogenetic Artemia model. The FASEB Journal 28:3552–3563. https://doi.org/10.1096/fj.14-252049

Novo M, Verdú I, Trigo D, Martínez-Guitarte JL (2018) Endocrine disruptors in soil: Effects of bisphenol A on gene expression of the earthworm Eisenia fetida. Ecotoxicology and Environmental Safety 150:159–167. https://doi.org/10.1016/j.ecoenv.2017.12.030

OECD (2004) Test No. 202: Daphnia sp. Acute Immobilisation Test. OECD Guideline for the Testing og Chemicals, Section 2. https://doi.org/10.1787/9789264069947-en

OECD (2008) Test No. 211: Daphnia magna Reproduction Test. OECD Publishing

Ogden NH, Lindsay LR, Hanincová K, et al (2008) Role of Migratory Birds in Introduction and Range Expansion of Ixodes scapularis Ticks and of Borrelia burgdorferi and Anaplasma phagocytophilum in Canada. Applied and Environmental Microbiology 74:1780–1790. https://doi.org/10.1128/AEM.01982-07

Oliveira DS, Rosa MT, Vieira C, Loreto ELS (2021) Oxidative and radiation stress induces transposable element transcription in Drosophila melanogaster. Journal of Evolutionary Biology 34:628–638. https://doi.org/10.1111/jeb.13762

Ong C, Yung L-YL, Cai Y, et al (2015) Drosophila melanogaster as a model organism to study nanotoxicity. Nanotoxicology 9:396–403. https://doi.org/10.3109/17435390.2014.940405

Oppold A, Kreß A, vanden Bussche J, et al (2015) Epigenetic alterations and decreasing insecticide sensitivity of the Asian tiger mosquito Aedes albopictus. Ecotoxicology and Environmental Safety 122:45–53. https://doi.org/10.1016/j.ecoenv.2015.06.036

Ott J (1979) Detection of rare major genes in lipid levels. Human Genetics 51:79–91. https://doi.org/10.1007/BF00278296

Otto SP (2018) Adaptation, speciation and extinction in the Anthropocene. Proceedings of the Royal Society B: Biological Sciences 285:20182047. https://doi.org/10.1098/rspb.2018.2047

Owen EL, Bale JS, Hayward SAL (2013) Can Winter-Active Bumblebees Survive the Cold? Assessing the Cold Tolerance of Bombus terrestris audax and the Effects of Pollen Feeding. PLoS ONE 8:e80061. https://doi.org/10.1371/journal.pone.0080061

Ożgo M, Urbańska M, Hoos P, et al (2020) Invasive zebra mussel (Dreissena polymorpha) threatens an exceptionally large population of the depressed river mussel (Pseudanodonta complanata) in a postglacial lake. Ecology and Evolution 10:4918–4927. https://doi.org/10.1002/ece3.6243

Pajaro-Castro N, Caballero-Gallardo K, Olivero-Verbel J (2019) Toxicity and expression of oxidative stress genes in Tribolium castaneum induced by toluene, xylene, and thinner. Journal of Toxicology and Environmental Health, Part A 82:28–36. https://doi.org/10.1080/15287394.2018.1546245

Parenti CC, Binelli A, Caccia S, et al (2020) Ingestion and effects of polystyrene nanoparticles in the silkworm Bombyx mori. Chemosphere 257:127203. https://doi.org/10.1016/j.chemosphere.2020.127203

Park J, Peng Z, Zeng J, et al (2011) Comparative Analyses of DNA Methylation and Sequence Evolution Using Nasonia Genomes. Molecular Biology and Evolution 28:3345–3354. https://doi.org/10.1093/molbev/msr168

Park PJ (2009) ChIP–seq: advantages and challenges of a maturing technology. Nature Reviews Genetics 10:669–680. https://doi.org/10.1038/nrg2641

Pecl GT, Araújo MB, Bell JD, et al (2017) Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. Science 355:. https://doi.org/10.1126/science.aai9214

Pedriali A, Riva C, Parolini M, et al (2013) A redox proteomic investigation of oxidative stress caused by benzoylecgonine in the freshwater bivalve Dreissena polymorpha. Drug Testing and Analysis 5:646–656. https://doi.org/10.1002/dta.1409

Peeters C, Liebig J, Hölldobler B (2000) Sexual reproduction by both queens and workers in the ponerine ant Harpegnathos saltator. Insectes Sociaux 47:325–332. https://doi.org/10.1007/PL00001724

Peng L, Fu D, Qi H, et al (2020) Micro- and nano-plastics in marine environment: Source, distribution and threats — A review. Science of the Total Environment 698:134254

Pereira HM, Leadley PW, Proença V, et al (2010) Scenarios for Global Biodiversity in the 21st Century. Science 330:1496–1501. https://doi.org/10.1126/science.1196624

Pestana JLT, Novais SC, Norouzitallab P, et al (2016) Non-lethal heat shock increases tolerance to metal exposure in brine shrimp. Environmental Research 151:663–670. https://doi.org/10.1016/j.envres.2016.08.037

Pinault LL, Hunter FF (2012) Characterization of larval habitats of Anopheles albimanus, Anopheles pseudopunctipennis, Anopheles punctimacula, and Anopheles oswaldoi s.l. populations in lowland and highland Ecuador. Journal of Vector Ecology 37:124–136. https://doi.org/10.1111/j.1948-7134.2012.00209.x

Pointier JP, David P, Jarne P (2005) Biological invasions: the case of planorbid snails. Journal of Helminthology 79:249–256. https://doi.org/10.1079/JOH2005292

Pollux, Bart JA, Gerard Van der Velde, Abraham Bij de Vaate (2010) A perspective on global spread of Dreissena polymorpha: a review on possibilities and limitations. Backhuys Publishers 45–58

Potet M, Giambérini L, Pain-Devin S, et al (2018) Differential tolerance to nickel between Dreissena polymorpha and Dreissena rostriformis bugensis populations. Scientific Reports 8:700. https://doi.org/10.1038/s41598-018-19228-x

Powlik JJ (1998) Seasonal Abundance and Population Flux of Tigriopus Californicus (Copepoda: Harpacticoida) in Barkley Sound, British Columbia. Journal of the Marine Biological Association of the United Kingdom 78:467–481. https://doi.org/10.1017/S0025315400041564

Provataris P, Meusemann K, Niehuis O, et al (2018) Signatures of DNA Methylation across Insects Suggest Reduced DNA Methylation Levels in Holometabola. Genome Biology and Evolution 10:1185–1197. https://doi.org/10.1093/gbe/evy066

Punyavathi, Manjunatha HB (2020) Comprehensive analysis of differentially expressed proteins in the male and female Bombyx mori larval instars exposed to thermal stress. Archives of Insect Biochemistry and Physiology 105:. https://doi.org/10.1002/arch.21719

Queirós L, Pereira JL, Gonçalves FJM, et al (2019) Caenorhabditis elegans as a tool for environmental risk assessment: emerging and promising applications for a "nobelized worm." Critical Reviews in Toxicology 49:411–429. https://doi.org/10.1080/10408444.2019.1626801

Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842. https://doi.org/10.1093/bioinformatics/btq033

Quinn JJ, Chang HY (2016) Unique features of long non-coding RNA biogenesis and function. Nature Reviews Genetics 17:47–62. https://doi.org/10.1038/nrg.2015.10

Rager JE, Bailey KA, Smeester L, et al (2014) Prenatal arsenic exposure and the epigenome: Altered microRNAs associated with innate and adaptive immune signaling in newborn cord blood. Environmental and Molecular Mutagenesis 55:196–208. https://doi.org/10.1002/em.21842

Rahmathulla VK (2012) Management of Climatic Factors for Successful Silkworm (Bombyx mori L.) Crop and Higher Silk Production: A Review. Psyche: A Journal of Entomology 2012:1–12. https://doi.org/10.1155/2012/121234

Raworth DA, Gillespie DR, Roy M, Thistlewood HMA (2001) Tetranychus urticae Koch, twospotted spider mite (Acari: Tetranychidae). In: Biological Control Programmes in Canada, 1981-2000. CABI, Wallingford, pp 259–265

Reeders HH, Bij de Vaate A (1990) Zebra mussels (Dreissena polymorpha): a new perspective for water quality management. Hydrobiologia 200–201:437–450. https://doi.org/10.1007/BF02530361

Reid AJ, Carlson AK, Creed IF, et al (2019) Emerging threats and persistent conservation challenges for freshwater biodiversity. Biological Reviews 94:849–873. https://doi.org/10.1111/brv.12480

Reinhold JM, Lazzari CR, Lahondère C (2018) Effects of the environmental temperature on Aedes aegypti and Aedes albopictus mosquitoes: A review. Insects 9:158

Reish DJ, Oshida PS, Mearns AJ, et al (2002) EFFECTS OF POLLUTION ON MARINE ORGANISMS Review by

Reiter LT (2001) A Systematic Analysis of Human Disease-Associated Gene Sequences In Drosophila melanogaster. Genome Research 11:1114–1125. https://doi.org/10.1101/gr.169101

Rezza G (2012) Aedes albopictus and the reemergence of Dengue. BMC Public Health 12:72. https://doi.org/10.1186/1471-2458-12-72

Richards S, Gibbs RA, Weinstock GM, et al (2008) The genome of the model beetle and pest Tribolium castaneum. Nature 452:949–955. https://doi.org/10.1038/nature06784

Rincón RA, Rodríguez D, Coy-Barrera E (2019) Botanicals Against Tetranychus urticae Koch Under Laboratory Conditions: A Survey of Alternatives for Controlling Pest Mites. Plants 8:272. https://doi.org/10.3390/plants8080272

Riva C, Cristoni S, Binelli A (2012) Effects of triclosan in the freshwater mussel Dreissena polymorpha: A proteomic investigation. Aquatic Toxicology 118–119:62–71. https://doi.org/10.1016/j.aquatox.2012.03.013

Roberts DB (2006) Drosophila melanogaster: the model organism. Entomologia Experimentalis et Applicata 121:93–103. https://doi.org/10.1111/j.1570-8703.2006.00474.x

Roberts SB, Gavery MR (2012) Is There a Relationship between DNA Methylation and Phenotypic Plasticity in Invertebrates? Frontiers in Physiology 2:. https://doi.org/10.3389/fphys.2011.00116

ROELOFS D, JANSSENS TKS, TIMMERMANS MJTN, et al (2009) Adaptive differences in gene expression associated with heavy metal tolerance in the soil arthropod Orchesella cincta. Molecular Ecology 18:3227–3239. https://doi.org/10.1111/j.1365-294X.2009.04261.x

Roger N, Michez D, Wattiez R, et al (2017) Diet effects on bumblebee health. Journal of Insect Physiology 96:128–133. https://doi.org/10.1016/j.jinsphys.2016.11.002

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

Roper CFE, de Angelis N, Roper I, et al (2010) Cephalopods of the World - An annotated and illustrated catalogue of cephalopod species known to date

Rossetto D, Avvakumov N, Côté J (2012) Histone phosphorylation. Epigenetics 7:1098–1108. https://doi.org/10.4161/epi.21975

Rudgalvyte M, Peltonen J, Lakso M, Wong G (2017) Chronic MeHg exposure modifies the histone H3K4me3 epigenetic landscape in Caenorhabditis elegans. Comparative Biochemistry and Physiology Part - C: Toxicology and Pharmacology 191:109–116. https://doi.org/10.1016/j.cbpc.2016.10.001

Sadd BM, Barribeau SM, Bloch G, et al (2015) The genomes of two key bumblebee species with primitive eusocial organization. Genome Biology 16:76. https://doi.org/10.1186/s13059-015-0623-3

Samanta P, Im H, Shim T, et al (2020) Linking multiple biomarker responses in Daphnia magna under thermal stress. Environmental Pollution 263:114432. https://doi.org/10.1016/j.envpol.2020.114432

Sanchez-Arguello P, Aparicio N, Guevara MA, et al (2016) Effects on Reproduction, Genotoxicity and DNA Methylation Pattern after Chronic Exposure of the Freshwater Snail Physa acuta (Gastropoda, Pulmonata) to Vinclozolin. Austin J Environ Toxicol 2:

Santos JI, Vidal T, Gonçalves FJM, et al (2021) Challenges to water quality assessment in Europe – Is there scope for improvement of the current Water Framework Directive bioassessment scheme in rivers? Ecological Indicators 121:107030. https://doi.org/10.1016/j.ecolind.2020.107030

Santovito G, Trentin E, Gobbi I, et al (2021) Non-enzymatic antioxidant responses of Mytilus galloprovincialis: Insights into the physiological role against metal-induced oxidative stress. Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology 240:108909. https://doi.org/10.1016/j.cbpc.2020.108909

Santoyo MM, Flores CR, Torres AL, et al (2011) Global DNA methylation in earthworms: A candidate biomarker of epigenetic risks related to the presence of metals/metalloids in terrestrial environments. Environmental Pollution 159:2387–2392. https://doi.org/10.1016/j.envpol.2011.06.041

Sarda S, Zeng J, Hunt BG, Yi S v. (2012) The Evolution of Invertebrate Gene Body Methylation. Molecular Biology and Evolution 29:1907–1916. https://doi.org/10.1093/molbev/mss062

Sardina JL, Collombet S, Tian T v., et al (2018) Transcription Factors Drive Tet2-Mediated Enhancer Demethylation to Reprogram Cell Fate. Cell Stem Cell 23:727-741.e9. https://doi.org/10.1016/j.stem.2018.08.016

Satoh N (2003) The ascidian tadpole larva: comparative molecular development and genomics. Nature Reviews Genetics 4:285–295. https://doi.org/10.1038/nrg1042

Satoh N, Satou Y, Davidson B, Levine M (2003) Ciona intestinalis: an emerging model for whole-genome analyses. Trends in Genetics 19:376–381. https://doi.org/10.1016/S0168-9525(03)00144-6

Satou Y, Kawashima T, Shoguchi E, et al (2005) An Integrated Database of the Ascidian, Ciona intestinalis: Towards Functional Genomics. Zoological Science 22:837–843. https://doi.org/10.2108/zsj.22.837

Scharf I, Wexler Y, MacMillan HA, et al (2016) The negative effect of starvation and the positive effect of mild thermal stress on thermal tolerance of the red flour beetle, Tribolium castaneum. The Science of Nature 103:20. https://doi.org/10.1007/s00114-016-1344-5

Schirmer K, Fischer BB, Madureira DJ, Pillai S (2010) Transcriptomics in ecotoxicology. Analytical and Bioanalytical Chemistry 397:917–923. https://doi.org/10.1007/s00216-010-3662-3

Schloesser DW, Nalepa TF (1994) Dramatic Decline of Unionid Bivalves in Offshore Waters of Western Lake Erie After Infestation by the Zebra Mussel, Dreissena polymorpha. Canadian Journal of Fisheries and Aquatic Sciences 51:2234–2242. https://doi.org/10.1139/f94-226

Schneider SA, Scharffetter C, Wagner AE, et al (2016) Social stress increases the susceptibility to infection in the ant Harpegnathos saltator. Scientific Reports 6:25800. https://doi.org/10.1038/srep25800

Schneider SA, Schrader C, Wagner AE, et al (2011) Stress Resistance and Longevity Are Not Directly Linked to Levels of Enzymatic Antioxidants in the Ponerine Ant Harpegnathos saltator. PLoS ONE 6:e14601. https://doi.org/10.1371/journal.pone.0014601

Schorderet DF, Gartler SM (1992) Analysis of CpG suppression in methylated and nonmethylated species. Proceedings of the National Academy of Sciences 89:957–961. https://doi.org/10.1073/pnas.89.3.957

Schulz NKE, Wagner CI, Ebeling J, et al (2018) Dnmt1 has an essential function despite the absence of CpG DNA methylation in the red flour beetle Tribolium castaneum. Scientific Reports 8:16462. https://doi.org/10.1038/s41598-018-34701-3

Serafini L, Hann JB, Kültz D, Tomanek L (2011) The proteomic response of sea squirts (genus Ciona) to acute heat stress: A global perspective on the thermal stability of proteins.

Comparative Biochemistry and Physiology Part D: Genomics and Proteomics 6:322–334. https://doi.org/10.1016/j.cbd.2011.07.002

Shahid N, Liess M, Knillmann S (2019) Environmental Stress Increases Synergistic Effects of Pesticide Mixtures on Daphnia magna. Environmental Science & Technology 53:12586–12593. https://doi.org/10.1021/acs.est.9b04293

Shea D, Thorsen W (2012) Ecological Risk Assessment. In: Progress in Molecular Biology and Translational Science. Elsevier B.V., pp 323–348

Sheldon CC, Finnegan EJ, Rouse DT, et al (2000) The control of flowering by vernalization. Current Opinion in Plant Biology 3:418–422. https://doi.org/10.1016/S1369-5266(00)00106-0

Shen L, Xiao J, Ye H, Wang D (2009) Toxicity evaluation in nematode Caenorhabditis elegans after chronic metal exposure. Environmental Toxicology and Pharmacology 28:125–132. https://doi.org/10.1016/j.etap.2009.03.009

Shen P, Yue Y, Park Y (2018) A living model for obesity and aging research: Caenorhabditis elegans. Critical Reviews in Food Science and Nutrition 58:741–754. https://doi.org/10.1080/10408398.2016.1220914

Shimizu TS, Takahashi K, Tomita M (1997) CpG distribution patterns in methylated and non-methylated species. In: Gene. pp 103–107

Sivan A, Shriram AN, Muruganandam N, Thamizhmani R (2017) Expression of heat shock proteins (HSPs) in Aedes aegypti (L) and Aedes albopictus (Skuse) (Diptera: Culicidae) larvae in response to thermal stress. Acta Tropica 167:121–127. https://doi.org/10.1016/j.actatropica.2016.12.017

Skuse FAA (1894) The banded mosquito of Bengal. Indian Museum Notice 3:

Smallwood SA, Kelsey G (2012) De novo DNA methylation: a germ cell perspective. Trends in Genetics 28:33–42. https://doi.org/10.1016/j.tig.2011.09.004

Snell TW, Brogdon SE, Morgan MB (2003) Gene Expression Profiling in Ecotoxicology. Ecotoxicology 12:475–483. https://doi.org/10.1023/B:ECTX.0000003033.09923.a8

Sochová I, Hofman J, Holoubek I (2006) Using nematodes in soil ecotoxicology. Environment International 32:. https://doi.org/10.1016/j.envint.2005.08.031

Sodergren E, Weinstock GM, Davidson EH, et al (2006) The Genome of the Sea Urchin Strongylocentrotus purpuratus. Science 314:941–952. https://doi.org/10.1126/science.1133609

Solé M, Freitas R, Viñas L, Rivera-Ingraham GA (2020) Biomarker considerations in monitoring petrogenic pollution using the mussel Mytilus galloprovincialis. Environmental Science and Pollution Research 27:. https://doi.org/10.1007/s11356-020-09427-3

Song X, Huang F, Liu J, et al (2017) Genome-wide DNA methylomes from discrete developmental stages reveal the predominance of non-CpG methylation in Tribolium castaneum. DNA Research 24:445–457. https://doi.org/10.1093/dnares/dsx016

Šrut M (2021) Ecotoxicological epigenetics in invertebrates: Emerging tool for the evaluation of present and past pollution burden. Chemosphere 282:131026. https://doi.org/10.1016/j.chemosphere.2021.131026

Šrut M, Drechsel V, Höckner M (2017) Low levels of Cd induce persisting epigenetic modifications and acclimation mechanisms in the earthworm Lumbricus terrestris. PLOS ONE 12:e0176047. https://doi.org/10.1371/journal.pone.0176047

Starnes DL, Unrine JM, Starnes CP, et al (2015) Impact of sulfidation on the bioavailability and toxicity of silver nanoparticles to Caenorhabditis elegans. Environmental Pollution 196:239–246. https://doi.org/10.1016/j.envpol.2014.10.009

Strader ME, Kozal LC, Leach TS, et al (2020) Examining the Role of DNA Methylation in Transcriptomic Plasticity of Early Stage Sea Urchins: Developmental and Maternal Effects in a Kelp Forest Herbivore. Frontiers in Marine Science 7:. https://doi.org/10.3389/fmars.2020.00205

Strimbu K, Tavel JA (2010) What are biomarkers? Current Opinion in HIV and AIDS 5:463–466. https://doi.org/10.1097/COH.0b013e32833ed177

Strong EE, Gargominy O, Ponder WF, Bouchet P (2008) Global diversity of gastropods (Gastropoda; Mollusca) in freshwater. Hydrobiologia 595

Suarez-Ulloa V, Gonzalez-Romero R, Eirin-Lopez JM (2015) Environmental epigenetics: A promising venue for developing next-generation pollution biomonitoring tools in marine invertebrates. Marine Pollution Bulletin 98:5–13. https://doi.org/10.1016/j.marpolbul.2015.06.020

Sussarellu R, Lebreton M, Rouxel J, et al (2018) Copper induces expression and methylation changes of early development genes in Crassostrea gigas embryos. Aquatic Toxicology 196:70–78. https://doi.org/10.1016/j.aquatox.2018.01.001

Suzuki MM, Bird A (2008) DNA methylation landscapes: Provocative insights from epigenomics. Nature Reviews Genetics 9:465–476

Suzuki MM, Kerr ARW, de Sousa D, Bird A (2007) CpG methylation is targeted to transcription units in an invertebrate genome. Genome Research 17:625–631. https://doi.org/10.1101/gr.6163007

Suzuki MM, Mori T, Satoh N (2016) The Ciona intestinalis cleavage clock is independent of DNA methylation. Genomics 108:168–176. https://doi.org/10.1016/j.ygeno.2016.10.001

Szabó B, Bakonyi G (2017) Multigenerational and transgenerational side-effects of an insecticide on eggs of Folsomia Candida (Collembola). Polish Journal of Ecology 65:110–121. https://doi.org/10.3161/15052249PJE2017.65.1.010

Szabó B, Lang Z, Bakonyi G, et al (2019) Transgenerational and multigenerational stress gene responses to the insecticide etofenprox in Folsomia candida (Collembola). Ecotoxicology and Environmental Safety 175:181–191. https://doi.org/10.1016/j.ecoenv.2019.03.052

Tangwancharoen S, Semmens BX, Burton RS (2020) Allele-Specific Expression and Evolution of Gene Regulation Underlying Acute Heat Stress Response and Local Adaptation in the Copepod Tigriopus californicus. Journal of Heredity 111:539–547. https://doi.org/10.1093/jhered/esaa044

Tarrant AM, Reitzel AM, Kwok CK, Jenny MJ (2014) Activation of the cnidarian oxidative stress response by ultraviolet light, polycyclic aromatic hydrocarbons and crude oil. Journal of Experimental Biology 217:. https://doi.org/10.1242/jeb.093690

Tejeda-Benitez L, Olivero-Verbel J (2016) Caenorhabditis elegans, a Biological Model for Research in Toxicology. In: Reviews of Environmental Contamination and Toxicology. pp 1–35

The C. elegans Sequencing Consortium (1998) Genome Sequence of the Nematode C. elegans: A Platform for Investigating Biology. Science 282:2012–2018. https://doi.org/10.1126/science.282.5396.2012

Thode HC, Finch SJ, Mendell NR (1988) Simulated Percentage Points for the Null Distribution of the Likelihood Ratio Test for a Mixture of Two Normals. Biometrics 44:1195. https://doi.org/10.2307/2531747

Thorat L, Oulkar DP, Banerjee K, Nath BB (2016) Desiccation stress induces developmental heterochrony in Drosophila melanogaster. Journal of Biosciences 41:331–339. https://doi.org/10.1007/s12038-016-9628-7

Tkaczyk A, Bownik A, Dudka J, et al (2021) Daphnia magna model in the toxicity assessment of pharmaceuticals: A review. Science of The Total Environment 763:143038. https://doi.org/10.1016/j.scitotenv.2020.143038

Torres AR, Beatriz Caballero, Adileidys Ruiz (2021) Description of the main aspects influencing Bombyx mori L.(Lepidoptera: Bombycidae) rearing. Cuban Journal of Agricultural Science 55:

Torretta JP, Medan D, Abrahamovich AH (2006) First record of the invasive bumblebee Bombus terrestris (L.) (Hymenoptera, Apidae) in Argentina. Transactions of the American Entomological Society 132:285–289. https://doi.org/https://doi.org/10.3157/0002-8320(2006)132[285:FROTIB]2.0.CO;2

Tricarico E, Amodio P, Ponte G, Fiorito G (2014) Cognition and Recognition in the Cephalopod Mollusc Octopus vulgaris: Coordinating Interaction with Environment and Conspecifics. In: Biocommunication of Animals. Springer Netherlands, Dordrecht, pp 337–349

Trijau M, Asselman J, Armant O, et al (2018) Transgenerational DNA Methylation Changes in Daphnia magna Exposed to Chronic γ Irradiation. Environmental Science & Technology 52:4331–4339. https://doi.org/10.1021/acs.est.7b05695

Tsankova NM, Berton O, Renthal W, et al (2006) Sustained hippocampal chromatin regulation in a mouse model of depression and antidepressant action. Nature Neuroscience 9:519–525. https://doi.org/10.1038/nn1659

Tweedie S, Charlton J, Clark V, Bird A (1997) Methylation of genomes and genes at the invertebrate-vertebrate boundary. Molecular and Cellular Biology 17:1469–1475. https://doi.org/10.1128/mcb.17.3.1469

Urbani CB, Boyan GS, Blarer A, et al (1994) A novel mechanism for jumping in the indian antHarpegnathos saltator (Jerdon) (Formicidae, Ponerinae). Experientia 50:63–71. https://doi.org/10.1007/BF01992052

Urich MA, Nery JR, Lister R, et al (2015) MethylC-seq library preparation for base-resolution whole-genome bisulfite sequencing. Nature Protocols 10:475–483. https://doi.org/10.1038/nprot.2014.114

van Straalen, Nico M, Hans Løkke (1997) Ecological Risk Assessment of Contaminants in Soil. Springer US, Boston, MA

Vandegehuchte MB, Janssen CR (2011) Epigenetics and its implications for ecotoxicology. Ecotoxicology 20:607–624. https://doi.org/10.1007/s10646-011-0634-0

Vandegehuchte MB, Kyndt T, Vanholme B, et al (2009a) Occurrence of DNA methylation in Daphnia magna and influence of multigeneration Cd exposure. Environment International 35:700–706. https://doi.org/10.1016/j.envint.2009.01.002

Vandegehuchte MB, Lemière F, Janssen CR (2009b) Quantitative DNA-methylation in Daphnia magna and effects of multigeneration Zn exposure. Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology 150:343–348. https://doi.org/10.1016/j.cbpc.2009.05.014

Vandegehuchte MB, Lemière F, Vanhaecke L, et al (2010) Direct and transgenerational impact on Daphnia magna of chemicals with a known effect on DNA methylation. Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology 151:278–285. https://doi.org/10.1016/j.cbpc.2009.11.007

vanEngelsdorp D, Meixner MD (2010) A historical review of managed honey bee populations in Europe and the United States and the factors that may affect them. Journal of Invertebrate Pathology 103:S80–S95. https://doi.org/10.1016/j.jip.2009.06.011

Villar M, Ayllón N, Kocan KM, et al (2015) Identification and characterization of anaplasma phagocytophilum proteins involved in infection of the tick vector, ixodes scapularis. PLoS ONE 10:e0137237. https://doi.org/10.1371/journal.pone.0137237

Villegas EJA, Ceballos-Vázquez BP, Markaida U, et al (2014) Diet of Octopus bimaculatus Verril, 1883 (Cephalopoda: Octopodidae) in Bahía De Los Ángeles, Gulf of California. Journal of Shellfish Research 33:305–314. https://doi.org/10.2983/035.033.0129

Wajid B, Serpedin E (2016) Do it yourself guide to genome assembly. Briefings in Functional Genomics 15:1–9. https://doi.org/10.1093/bfgp/elu042

Walker CH (1992) Biochemical responses as indicators of toxic effects of chemicals in ecosystems. Toxicology Letters 64–65:527–533. https://doi.org/10.1016/0378-4274(92)90228-C

Wamucho A, Heffley A, Tsyusko O v. (2020) Epigenetic effects induced by silver nanoparticles in Caenorhabditis elegans after multigenerational exposure. Science of the Total Environment 725:138523. https://doi.org/10.1016/j.scitotenv.2020.138523

Wang H, Li K, Zhu J-Y, et al (2012) CLONING AND EXPRESSION PATTERN OF HEAT SHOCK PROTEIN GENES FROM THE ENDOPARASITOID WASP, Pteromalus puparum IN RESPONSE TO ENVIRONMENTAL STRESSES. Archives of Insect Biochemistry and Physiology 79:247–263. https://doi.org/10.1002/arch.21013

Wang H, Liu Z, Wang Y, et al (2020) Genome-Wide Differential DNA Methylation in Reproductive, Morphological, and Visual System Differences Between Queen Bee and Worker Bee (Apis mellifera). Frontiers in Genetics 11:. https://doi.org/10.3389/fgene.2020.00770

Wang L, Meeus I, Rombouts C, et al (2019a) Metabolomics-based biomarker discovery for bee health monitoring: A proof of concept study concerning nutritional stress in Bombus terrestris. Scientific Reports 9:11423. https://doi.org/10.1038/s41598-019-47896-w

Wang X, Li A, Wang W, et al (2021) DNA methylation mediates differentiation in thermal responses of Pacific oyster (Crassostrea gigas) derived from different tidal levels. Heredity 126:10–22. https://doi.org/10.1038/s41437-020-0351-7

Wang Y, Jorda M, Jones PL, et al (2006) Functional CpG Methylation System in a Social Insect. Science 314:645–647. https://doi.org/10.1126/science.1135213

Wang Z, Qu Y, Zhuo X, et al (2019b) Investigating the physiological responses of Pacific white shrimp Litopenaeus vannamei to acute cold-stress. PeerJ 7:e7381. https://doi.org/10.7717/peerj.7381

Watanabe M, Sakashita T, Fujita A, et al (2006a) Biological effects of anhydrobiosis in an African chironomid, Polypedilum vanderplanki on radiation tolerance. International Journal of Radiation Biology 82:587–592. https://doi.org/10.1080/09553000600876652

Watanabe M, Sakashita T, Fujita A, et al (2006b) Estimation of radiation tolerance to high LET heavy ions in an anhydrobiotic insect, Polypedilum vanderplanki. International Journal of Radiation Biology 82:835–842. https://doi.org/10.1080/09553000600979100

Watson GM, Mire P, Kinler KM (2009) Mechanosensitivity in the model sea anemone Nematostella vectensis. Marine Biology 156:2129–2137. https://doi.org/10.1007/s00227-009-1243-9

Watt F, Molloy PL (1988) Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. Genes & Development 2:1136–1143. https://doi.org/10.1101/gad.2.9.1136

Weber A, Jeckel N, Wagner M (2020) Combined effects of polystyrene microplastics and thermal stress on the freshwater mussel Dreissena polymorpha. Science of The Total Environment 718:137253. https://doi.org/10.1016/j.scitotenv.2020.137253

Weissmannová HD, Pavlovský J (2017) Indices of soil contamination by heavy metals – methodology of calculation for pollution assessment (minireview). Environmental Monitoring and Assessment 189:616. https://doi.org/10.1007/s10661-017-6340-5

Whibley A, Kelley JL, Narum SR (2021) The changing face of genome assemblies: Guidance on achieving high-quality reference genomes. Molecular Ecology Resources 21:641–652. https://doi.org/10.1111/1755-0998.13312

Whiting AR (1967) The Biology of the Parasitic Wasp Mormoniella vitripennis [=Nasonia brevicornis] (Walker). The Quarterly Review of Biology 42:333–406. https://doi.org/10.1086/405402

Wickham H (2019) Package stringr: Simple, Consistent Wrappers for Common String Operations. Cran

Willett CS, Ladner JT (2009) Investigations of fine-scale phylogeography in Tigriopus californicus reveal historical patterns of population divergence. BMC Evolutionary Biology 9:139. https://doi.org/10.1186/1471-2148-9-139

Wilson MC, Chen X-Y, Corlett RT, et al (2016) Habitat fragmentation and biodiversity conservation: key findings and future challenges. Landscape Ecology 31:219–227. https://doi.org/10.1007/s10980-015-0312-3

Wołejko E, Jabłońska-Trypuć A, Wydro U, et al (2020) Soil biological activity as an indicator of soil pollution with pesticides – A review. Applied Soil Ecology 147:103356. https://doi.org/10.1016/j.apsoil.2019.09.006

Workman G (1999) A Review of the Biology and Fisheries for Purple Sea Urchin (Strongylocentrotus purpuratus, Stimpson, 1857) and Discussion of the Assessment Needs for a Proposed Fishery. Ottawa: Fisheries an oceans Canada

Wu X, Zhang Y (2017) TET-mediated active DNA demethylation: mechanism, function and beyond. Nature Reviews Genetics 18:517–534. https://doi.org/10.1038/nrg.2017.33

Xiang H, Zhu J, Chen Q, et al (2010) Single base–resolution methylome of the silkworm reveals a sparse epigenomic map. Nature Biotechnology 28:516–520. https://doi.org/10.1038/nbt.1626

Ximénez-Embún MG, Castañera P, Ortego F (2017) Drought stress in tomato increases the performance of adapted and non-adapted strains of Tetranychus urticae. Journal of Insect Physiology 96:73–81. https://doi.org/10.1016/j.jinsphys.2016.10.015

Xu G, Yi Y, Lyu H, et al (2020a) DNA methylation suppresses chitin degradation and promotes the wing development by inhibiting Bmara-mediated chitinase expression in the silkworm, Bombyx mori. Epigenetics & Chromatin 13:34. https://doi.org/10.1186/s13072-020-00356-6

Xu Y, Wang W, Ma L, et al (2020b) Acute toxicity of Zinc Oxide nanoparticles to silkworm (Bombyx mori L.). Chemosphere 259:127481. https://doi.org/10.1016/j.chemosphere.2020.127481

Yagound B, Remnant EJ, Buchmann G, Oldroyd BP (2020) Intergenerational transfer of DNA methylation marks in the honey bee. Proceedings of the National Academy of Sciences 117:32519–32527. https://doi.org/10.1073/pnas.2017094117

Yang C, Gao Q, Liu C, et al (2017) The transcriptional response of the Pacific oyster Crassostrea gigas against acute heat stress. Fish & Shellfish Immunology 68:132–143. https://doi.org/10.1016/j.fsi.2017.07.016

Yang C-Y, Sierp MT, Abbott CA, et al (2016) Responses to thermal and salinity stress in wild and farmed Pacific oysters Crassostrea gigas. Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology 201:22–29. https://doi.org/10.1016/j.cbpa.2016.06.024

Yang J, Chatterjee N, Kim Y, et al (2018a) Histone methylation-associated transgenerational inheritance of reproductive defects in Caenorhabditis elegans exposed to crude oil under various exposure scenarios. Chemosphere 200:358–365. https://doi.org/10.1016/j.chemosphere.2018.02.080

Yang S-X, Guo C, Zhao X-T, et al (2018b) Divergent methylation pattern in adult stage between two forms of Tetranychus urticae (Acari: Tetranychidae). Insect Science 25:667–678. https://doi.org/10.1111/1744-7917.12444

Yekutieli Y, Sagiv-Zohar R, Aharonov R, et al (2005) Dynamic Model of the Octopus Arm. I. Biomechanics of the Octopus Reaching Movement. Journal of Neurophysiology 94:1443–1458. https://doi.org/10.1152/jn.00684.2004

Yi S v., Goodisman MAD (2009) Computational approaches for understanding the evolution of DNA methylation in animals. Epigenetics 4:551–556

Yoshioka Y, Gotoh T, Suzuki T (2018) UV-B susceptibility and photoreactivation in embryonic development of the two-spotted spider mite, Tetranychus urticae. Experimental and Applied Acarology 75:155–166. https://doi.org/10.1007/s10493-018-0263-x

Yu Z, Yin D, Deng H (2015) The combinational effects between sulfonamides and metals on nematode Caenorhabditis elegans. Ecotoxicology and Environmental Safety 111:66–71. https://doi.org/10.1016/j.ecoenv.2014.09.026

Yuan H, Qin F, Guo W, et al (2016) Oxidative stress and spermatogenesis suppression in the testis of cadmium-treated Bombyx mori larvae. Environmental Science and Pollution Research 23:5763–5770. https://doi.org/10.1007/s11356-015-5818-3

Yuxuan Z, Peiyong G, Yanmei W, et al (2019) Evaluation of the subtle effects and oxidative stress response of chloramphenicol, thiamphenicol, and florfenicol in Daphnia magna. Environmental Toxicology and Chemistry 38:575–584. https://doi.org/10.1002/etc.4344

Zemach A, McDaniel IE, Silva P, Zilberman D (2010) Genome-Wide Evolutionary Analysis of Eukaryotic DNA Methylation. Science 328:916–919. https://doi.org/10.1126/science.1186366

Zhang X, Wang W, Zhu W, et al (2019) Mechanisms and Functions of Long Non-Coding RNAs at Multiple Regulatory Levels. International Journal of Molecular Sciences 20:5573. https://doi.org/10.3390/ijms20225573

Zhang Z, Ye G, Cai J, Hu C (2005) Comparative venom toxicity between Pteromalus puparum and Nasonia vitripennis (Hymenoptera: Pteromalidae) toward the hemocytes of their natural hosts, non-target insects and cultured insect cells. Toxicon 46:337–349. https://doi.org/10.1016/j.toxicon.2005.05.005

Zhao X, Yu H, Kong L, Li Q (2012) Transcriptomic Responses to Salinity Stress in the Pacific Oyster Crassostrea gigas. PLoS ONE 7:e46244. https://doi.org/10.1371/journal.pone.0046244

Zhuo X, Qin Y, He P, et al (2021) Transcriptomic analysis of Litopenaeus vannamei hepatopancreas under cold stress in cold-tolerant and cold-sensitive cultivars. Gene 764:145090. https://doi.org/10.1016/j.gene.2020.145090

Zilberman D, Cao X, Jacobsen SE (2003) ARGONAUTE4 Control of Locus-Specific siRNA Accumulation and DNA and Histone Methylation. Science 299:716–719. https://doi.org/10.1126/science.1079695

Zinkler D, Platthaus J (1996) Tolerance of soil-dwelling Collembola to high carbon dioxide concentrations. European Journal of Entomology 93:

Zizzari ZV, Ellers J (2011) Effects of exposure to short-term heat stress on male reproductive fitness in a soil arthropod. Journal of Insect Physiology 57:421–426. https://doi.org/10.1016/j.jinsphys.2011.01.002