# Genome-wide DNA methylation of the liver reveals delayed effects of early-life exposure to 17-α-ethinylestradiol in the self-fertilizing mangrove rivulus

Anne-Sophie Voisin, Victoria Suarez Ulloa, Peter Stockwell, Aniruddha Chatterjee & Frédéric Silvestre

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

Check for updates

# Genome-wide DNA methylation of the liver reveals delayed effects of early-life exposure to 17-α-ethinylestradiol in the self-fertilizing mangrove rivulus

Anne-Sophie Voisin[a], Victoria Suarez Ulloa[a], Peter Stockwell[b], Aniruddha Chatterjee [c,#], and Frédéric Silvestre [a,#]

[a]Laboratory of Evolutionary and Adaptive Physiology, Institute of Life, Earth and Environment, University of Namur, Namur, Belgium; [b]Department of Biochemistry, University of Otago, Dunedin, New Zealand; [c]Department of Pathology, Dunedin School of Medicine, University of Otago, Dunedin, New Zealand

**ABSTRACT**

Organisms exposed to endocrine disruptors in early life can show altered phenotype later in adulthood. Although the mechanisms underlying these long-term effects remain poorly understood, an increasing body of evidence points towards the potential role of epigenetic processes. In the present study, we exposed hatchlings of an isogenic lineage of the self-fertilizing fish mangrove rivulus for 28 days to 4 and 120 ng/L of 17-α-ethinylestradiol. After a recovery period of 140 days, reduced representation bisulphite sequencing (RRBS) was performed on the liver in order to assess the hepatic genome-wide methylation landscape. Across all treatment comparisons, a total of 146 differentially methylated fragments (DMFs) were reported, mostly for the group exposed to 4 ng/L, suggesting a non-monotonic effect of EE2 exposure. Gene ontology analysis revealed networks involved in lipid metabolism, cellular processes, connective tissue function, molecular transport and inflammation. The highest effect was reported for *nipped-B-like protein B* (NIPBL) promoter region after exposure to 4 ng/L EE2 (+ 21.9%), suggesting that NIPBL could be an important regulator for long-term effects of EE2. Our results also suggest a significant role of DNA methylation in intergenic regions and potentially in transposable elements. These results support the ability of early exposure to endocrine disruptors of inducing epigenetic alterations during adulthood, providing plausible mechanistic explanations for long-term phenotypic alteration. Additionally, this work demonstrates the usefulness of isogenic lineages of the self-fertilizing mangrove rivulus to better understand the biological significance of long-term alterations of DNA methylation by diminishing the confounding factor of genetic variability.

## Introduction

Early-life exposure to environmental stressors, encountered during the sensitive period of embryogenesis or in juveniles, can be critical in shaping the long-term control of tissue physiology and homoeostasis. Although this paradigm, referred to as the Developmental Origins of Health and Disease (DOHaD) [1], is now widely recognized, the molecular mechanisms by which early exposures influence the propensity of disease and phenotype later in life remain elusive. Identifying these mechanisms has become extremely important to understand the long-term effects of toxicants in human and wildlife and ensure the proper risk assessment of xenobiotic exposure [2].

Nowadays, endocrine disrupting chemicals (EDCs) – xenobiotics able to interfere with the proper functioning of the endocrine system – are ubiquitous in the environment and our everyday life. The timely release and tightly regulated concentrations of hormones in early life are crucial for the proper development of the organism, including the reproductive, nervous, and immune systems. Therefore, early-life exposure to EDCs can have dramatic consequences on homoeostasis and physiology. In fact, mounting epidemiological evidence link exposures to EDCs to the increased incidence of metabolic diseases, immune diseases, neurological disorders, cancer and alteration of fertility in humans [3–6]. Of the many EDCs

present in the aquatic environment, 17-α-ethinylestradiol (EE2) – a synthetic derivative of oestradiol used in oral contraceptives – is of particular concern due to its high potency and resistance to degradation [7]. In treated seawage, EE2 is often the major compound with oestrogenic activity and its impact on aquatic wildlife has been deeply investigated (reviewed in [8]). In aquatic species, field studies and laboratory experiments show that exposure to EE2 affects fecundity, fertility, reproductive behaviour and induces intersex, endangering natural populations [9].

Potential long-term and persistent effects of early-life EDC exposure involve the stable alteration of gene expression. This is possible through epigenetic regulation, involving histone modifications, non-coding RNAs, and DNA methylation. The latter refers to the transfer of a methyl group from a methyl-donor, S-adenosyl-methionine (SAM), to a cytosine in the context of CpG dinucleotides, forming a methylcytosine [10,11]. This process is catalysed by two families of DNA methyltransferases (DNMTs): whereas DNMT3 enzymes are responsible for de novo methylation during development and differentiation, the DNMT1 family is maintaining methylation through each cell division by copying hemimethylated DNA, making DNA methylation a stable and heritable modification. On the other hand, the addition of a hydroxyl group to methylcytosine by ten-eleven translocation (Tet) enzymes is one of the most common mechanisms of active demethylation [12,13]. In mammals, most DNA methylation occurs in the symmetric CpG context, while only a small amount of non-CpG methylation is observed [10,11]. DNA methylation regulates gene expression by altering chromatin state and accessibility to CpG sites. For instance, methylation occurring at promoter regions can prevent the binding of transcription factors, resulting in gene silencing. Yet, the relationship between methylation and gene expression is more complex, as methylation also occurs at intragenic sites, enhancers or suppressor elements [13,14]. DNA methylation is dynamically regulated throughout life: during gametogenesis and embryogenesis, two waves of demethylation/remethylation deeply reprogram the methylome to produce a totipotent zygote [14–16]. Then, DNA methylation plays an important role throughout development as it directs cellular differentiation after reprogramming. During early development and juvenile stages, DNA methylation is thus thought to be particularly sensitive to environmental factors [17].

An increasing number of studies highlight the potential roles of DNA methylation in mediating long-term effects of sub-toxic developmental exposure to xenobiotics and roles in the aetiology of diseases including cancer, obesity, cardiovascular disease, diabetes, hypertension, and neurodegenerative disorders [18–22]. Indeed, DNA methylation could act as a long-term memory of past exposures mediating stable changes in gene expression [23,24]. In ecotoxicology, the role of epigenomics is receiving more and more attention to explain the long-term delayed and potential transgenerational effects of xenobiotics [25–27]. In particular, accumulating evidence indicates that hormones and endocrine disrupting compounds can alter the epigenome [21,28,29]. For instance, the regulation of DNMT transcription by ESR1 (an oestrogen receptor that acts as a transcription factor) may represent one of the possible mechanisms by which hormones influence methylation [30]. Other potential mechanisms may involve the reduction of SAM availability [31], the alteration of histone activity as a result of membrane receptor oestrogenic signalling [32,33], the expression of miRNAs, or direct interactions between oestrogen receptors and enzymes involved in the methylation machinery, such as thymine-DNA glycosylase (TDG) [34].

So far, relatively few studies have examined the effects of endocrine disruptors on DNA methylation in aquatic species. Using bisulphite conversion and pyrosequencing, Strömqvist et al. [35] have shown decreased methylation levels of three CpG sites located in the 5' flanking region of the vitellogenin I gene, a known biomarker of oestrogenic exposure, in the liver of male and female zebrafish, following a 14-day exposure to 100 ng/L EE2. Using Methylated DNA Immunoprecipitation (MeDIP) and high-throughput sequencing, followed by validation using bisulphite sequencing PCR (BSP) and RT-PCR, Mirbahai et al. [36] investigated the whole-genome methylation and gene expression in

tumours in the liver of the Common dab (*Limanda limanda*), sampled in various sites in English rivers. Genes involved in pathways related to cancer, including apoptosis, wnt/β-catenin signalling and genomic and non-genomic oestrogen responses, were altered both in methylation and transcription. In that case the exact mixture of environmental contaminants causing the liver tumours was not identified but the molecular responses point towards oestrogenic disruption. On zebrafish early life stages, Falisse et al. [37] showed that exposure to the antibacterial agent and EDC triclosan modified the methylome after 7 days exposure, as well as the expression of related genes, using Reduced Representation Bisulphite Sequencing (RRBS).

The identification of environmentally-induced alterations in the epigenome ideally requires the use of individuals with no existing or very low genetic variability to rule out the confounding factors of the genotype on the observed phenotype [38]. Such a model species is provided by the mangrove rivulus, *Kryptolebias marmoratus*, one of the only known self-fertilizing vertebrates (the other one being its sister species, *K. hermaphroditus*). Although males and hermaphrodites are present in natural populations, most reproduction occurs by self-fertilization of hermaphrodites. Exclusive selfing of hermaphrodites during several generations results in highly homozygous and virtually 'clonal' lineages that can occur naturally or be bred in a laboratory setting [39,40]. A naturally clonal and sexually reproducing vertebrate species allows the study of epigenetic mechanisms ruling out the confounding factor of genetic variability, which is especially useful in ecotoxicology experiments in the sublethal range and/or dealing with delayed effects where non-dramatic molecular changes are expected. Little is known about the epigenetic mechanisms in *K. marmoratus* but it has been reported that the sex-ratio can be modulated by temperature-sensitive DNA methylation [41]. Recent studies described the DNA methylation reprogramming event during embryogenesis [42], as well as the time course expression of enzymes involved in epigenetics during its development [43,44]. Another study has shown that parasite

load can modify DNA methylation [45], while these authors also found a higher differentiation of the methylation landscape between different genotypes than between environments [46].

In previous studies, we successfully used the mangrove rivulus to assess the delayed effects of BMAA (beta-N-Methylamino-L-alanine) [47] and 17-α-ethinylestradiol (EE2) [48,49] in adults after an early-life stage exposure. Hatchlings were exposed to 4 and 120 ng/L EE2 for 28 days, and allowed to recover for 140 additional days in clean water. The 28 days exposure has been chosen to include a 4 weeks period of exposure during early life stages that would be long enough to potentially induce endocrine effects and to be environmentally relevant. At 168 dph, all the fish are sexually mature. This design represents a ratio 1:5 between the exposure and the recovery periods. We reported delayed and persistent effects of EE2 on growth, reproduction and steroid levels as well as liver, ovotestis and brain molecular phenotypes assessed by label-free quantitative proteomics. Effects of EE2 were tissue and dose dependent, with most effects occurring at the environmentally relevant concentration (4 ng/L) in brain and liver, and at the higher concentration in the gonads. Among the three studied organs, the liver showed the most sensitive response to EE2. In this organ, EE2 affected known oestrogen-responsive pathways such as lipid, fatty acid and steroid metabolism, apolipoproteins, innate immune system and inflammation. Interestingly, several proteins involved in SAM metabolism were affected in the liver, providing further indications of the potential effect of EE2 on DNA methylation. The liver serves an essential and conserved role for metabolic homoeostasis in all vertebrates, and reproduction in the case of oviparous species. Several studies have shown that the liver is an important target of endocrine disruption in mammals [50] and fish [36,51–53]. Using an RRBS approach, the present study aimed at characterizing the hepatic methylation landscape in adults and investigating the genome-wide changes in DNA methylation in the liver of 168 dph adults that were exposed to EE2 as hatchlings to test the hypothesis of the potential involvement of DNA methylation in the delayed effects of EDCs.

## Materials and methods

### Experimental fish, EE2 exposure and sample collection

Individuals used in this study were from the same experiment reported in Voisin et al. [48], which also described the generation of breeding stock population and collection of eggs. Upon hatching, mangrove rivulus were transferred to 300 mL glass jars filled with 100 mL of 25 ppt reconstituted salt water (Instant Ocean™ sea salt), and the corresponding treatment: vehicle control (0.000012% ethanol), 4 ng/L and 120 ng/L 17-α-ethinylestradiol (EE2) (Sigma-Aldrich E4876-1 G). Fish were raised at a temperature of 26°C and a 12:12 h photoperiod. Detailed preparation of EE2 solutions and the measurement of actual exposure concentrations by ELISA are reported in Voisin et al. [48]. Average measured concentration of the nominal 120 ng/L exposure solution was 132.3 ± 14.7 ng/L. Since 4 ng/L is situated below the detection limit (20 ng/L), we refer to the measure of the stock solution, which was 10.15 ± 0.15 µg/L for the nominal 10 µg/L stock solution. Throughout the paper, we will refer to the nominal concentrations of 4 and 120 ng/L. Fish were exposed for a period of 28 days with a 100% water renewal 3 times a week. At 28 dph, fish were transferred to individual 1.2 L plastic containers filled with 400 mL of 25 ppt clean salt water and reared until 168 dph. Rivulus were fed 1 mL of concentrated newly hatched brine shrimp (Artemia nauplii) until 56 dph, 2 mL starting at 56 dph and 4 mL starting at 91 dph. At 168 dph, fish were sacrificed by a rapid transfer to 4°C 25 ppt salt water and sectioning the spinal cord. For each experimental condition, livers from five individuals were dissected out at 168 dph, snap-frozen in liquid nitrogen and stored at −80°C until DNA extraction. All rivulus husbandry and experimental procedures were performed in accordance with the Belgian animal protection standards and were approved by the University of Namur Local Research Ethics Committee (UN KE14/230). The agreement number of the laboratory for fish experiments is the LA1900048.

### DNA extraction, RRBS library preparation and post-processing

Genomic DNA from individual livers was extracted using the NucleoSpin Tissue XS kit (Macherey-Nagel, Germany) following the manufacturer's protocol. DNA concentration and quality were assessed using the NanoDrop 8000 spectrophotometer (Thermo Scientific, Waltham, MA) and 1% agarose gel electrophoresis. A total of 15 RRBS libraries (5 for each experimental condition) were prepared following established protocols [54,55]. In brief, genomic DNA was digested with MSPI (New England Biolabs, Ipswich, MA) followed by end repair, addition of 3 A overhangs and addition of methylated adapters (Illumina, San Diego, CA) to the digested fragments. Following adapter ligation, size selection was performed by cutting 40–220 bp DNA fragments (pre-ligation size) from a 3% (w/v) NuSieve GTG agarose gel (Lonza, Basel, Switzerland). Subsequently, libraries were bisulphite converted using the EZ DNA methylation kit (Zymo Research, Irvine, CA) with an extended incubation time of 18–20 h. Bisulphite converted libraries were amplified by PCR and sequenced (100-bp single ended reads) by New Zealand Genomics Limited (University of Otago, New Zealand) on an Illumina HiSeq2000 sequencer. Sequences were obtained in FASTQ format.

The quality of sequenced reads was performed using the FastQC application (Babraham Institute, Cambridge, UK). Adaptor sequences were removed from the reads using the *cleanadaptors* program of the DMAP package [56]. Single-ended bisulphite reads were aligned against the *Kryptolebias marmoratus* assembly (GCF_001649575.1_ASM164957v1) using the Bismark software v0.17.0 (default mode: alignments to complementary strands were ignored – i.e., not performed) [57]. Proximal genes and genomic positions of the fragments were identified using the *identgeneloc* programme from the DMAP package [58]. Fragments were linked to the nearest protein-coding gene and the genomic position was attributed as follows: internal (within the gene body), promoter (0–5 kb upstream of the transcription start site), upstream (5–10 kb

upstream of the TSS) and intergenic (>10 kb). The annotation of intergenic fragments or fragments located upstream of genes (5–10 kb) and in intergenic regions (>10 kb) is provided as an indication and should be considered with caution as methylation of these regions may not have any influence on the corresponding gene expression. Fragments with a methylation level ≥ 80% are qualified as highly methylated, while those with a methylation level ≤ 20% are qualified as lowly methylated.

### Analysis of differential DNA methylation

Base resolution differences in DNA methylation following developmental exposure to EE2 were assessed based on MSPI fragments (40–220 bp range) as a unit of analysis. To filter the fragments suitable for differential methylation analysis, we selected the fragments with ≥ 2 CpG sites, ≥10 reads per fragment (CpG10) and present in at least three of five individuals for each experimental condition. These fragments are referred to as 'analysed fragments'. The 102,018 analysed fragments form our reduced representative genome (RR genome) and corresponds to a total of 14,679 unique RefSeq identifiers, as more than one fragment often maps to a same gene. Differential methylation analysis was performed as previously described using the diffmeth program from DMAP [56,58] between each experimental group, resulting in three group comparisons: Ctl versus 4 ng/L EE2 exposed individuals, Ctl versus 120 ng/L and 4 versus 120 ng/L. Determination of significantly differentially methylated fragments (DMFs) was based on ≥ 10% differential methylation between the groups and P-value <0.01, and are qualified as hyper- or hypomethylated fragments.

### Network analysis

RefSeq identifiers were retrieved using BLAST for all analysed DMFs. To establish a reference gene list for the network analysis, a list including all DMFs was filtered to include only gene body and promoter genomic locations and contained 68,762 fragments mapped to rivulus RefSeq and human RefSeq identifiers. As more than one fragment often mapped to a same gene, this list corresponded to 10,440 unique human RefSeq identifiers. The list, containing P-values and methylation difference for each of the three treatment comparisons was uploaded in the Ingenuity Pathway Analysis software (IPA, QIAGEN, www.qiagen.com/ingenuity). To gain an overall view of the pathways and functions associated with promoter and gene body DMFs, as well as the connections between differentially abundant proteins, we performed a network analysis for each group comparison.

## Results

### Characterization of the hepatic DNA methylation landscape in *K. marmoratus*

Using RRBS, we generated 15 *K. marmoratus* liver methylomes with an average of 29.5 million sequenced reads (between 19.7 and 40 million reads) (Table 1, Supplementary Table 1). Unique alignment efficiency of these libraries ranged from 40.3% to 46.8% (except the library 23–4 with 28.5%, which was discarded in further analyses), with an average of 44.4%. This corresponds to a total of 102,018 fragments suited for the analysis after MSPI digestion of genomic DNA and 40–220 bp size-selection (Figure 1a). The total size of the RR genome was 9.37 Mb, which accounted for 1.4% of the published full genome (GCF_001649575.1_ASM164957v1) [59]. Within this RR genome, we counted 701,862 CpG dinucleotides, representing 6.1% of the 11.5 million CpG in the reference genome. Consequently, the

**Table 1.** Characteristics of the reduced representative genome and comparisons with the genome.

| Genome | |
|---|---|
| Size in scaffolds (Mb) | 680 |
| GC content | 37% |
| Number of CpGs (millions) | 11.5 |
| **Reduced genome** | |
| Size (Mb) | 9.37 |
| Number of reads (min-max; millions) | 19.7–40 |
| Uniquely mapped reads (min-max; millions) | 8.9–17.8 |
| Mapping efficiency (min-max) | 40.3–47.1% |
| Number of analysed fragments | 102 018 |
| % of analysed genome | 1.4 |
| Number of CpGs | 701,862 |
| % of genomic CpGs | 6.1 |
| CpG enrichment factor | 4.3 |
| % CpG methylation (min-max) | 63.5–66.4% |
| % non-CpG methylation | 1.7–2.4% |
| % of hypomethylated fragments | 20 |
| % hypermethylated fragments | 67 |
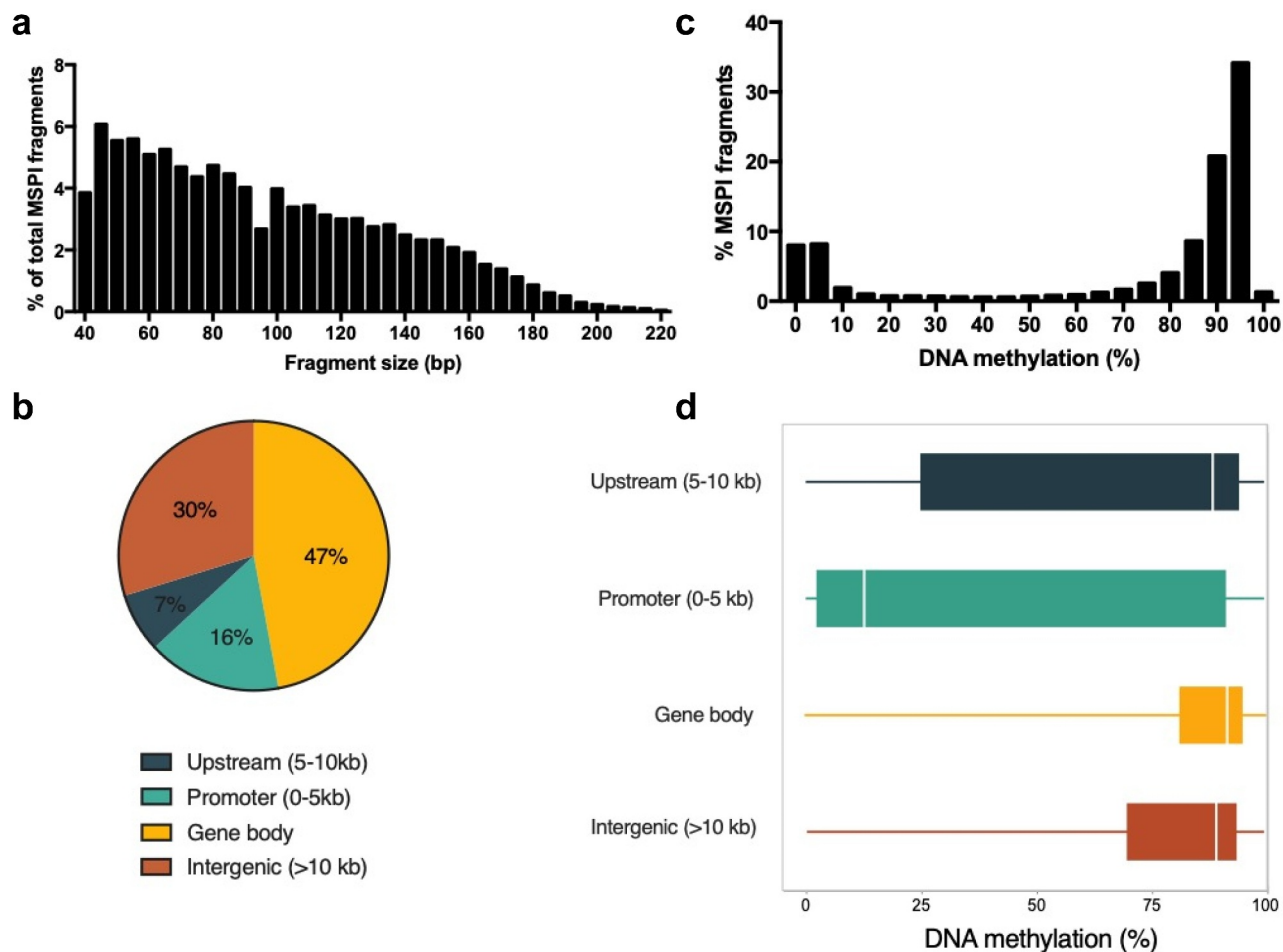
**Figure 1.** Hepatic DNA methylation landscape in *K. marmoratus* . (a) Frequency distribution of MSPI fragment size. Mean fragment size was 95 bp and median 88 bp; (b) Overall distribution of analysed fragments in genomic elements; (c:)) Frequency distribution of DNA methylation % for MSPI fragments. Distribution of MSPI fragments follows a bimodal distribution, which is expected given the binary character of DNA methylation in cells; (d) Boxplot of DNA methylation (%) of fragments according to their location in genomic elements (including medians and quantile 25–75%).

enrichment factor in CpGs of the RR genome was 4.3 fold. Of all the analysed fragments, the majority was located within gene bodies (47%), followed by intergenic regions (30%), promoter (16%) and upstream regions (7%) (Figure 1b). The distribution of CpGs followed exactly the same pattern as the fragments (data not shown). We consider this dataset of about 6% of all genomic CpGs, albeit not uniformly distributed across the genome, as a representative sample that permits the investigation of changes in DNA methylation in liver due to exposure to toxicants. Global CpG methylation levels in liver tissue ranged from 63.5% to 66.4% in all samples. On the contrary, methylation level of non-CpG cytosines remained low between 1.7 and 2.4%. It should be noted that non-CpG methylation calculated here included the true genomic

non-CpG methylation and the CpGs that failed to convert during bisulphite treatment.

For the following description, we took only the control group into consideration, in order to avoid any possible effect of the EE2 treatment. The average DNA methylation level of most MSPI fragments in mangrove rivulus were above 80% (67% of all fragments – considered as highly methylated) or below 20% of methylation (20% of all fragments – considered as lowly methylated) (Figure 1c). The level of DNA methylation was the lowest in fragments located in promoter regions (mean = 42.3%; median = 12.8%) compared to upstream (mean = 66.2%; median = 88.4%), intergenic (mean = 72.6%; median = 89.1%) and gene body (mean = 78.0%; median = 91.5%). Figure 1d depicts the distribution and spread (25–75

quantiles) of DNA methylation level for fragments within each region. We can observe that the methylation range was the highest for fragments located in the promoter region, showing that even if the majority of these fragments were lowly methylated (52%), a considerable part was highly methylated (38%). Despite the fact that fragments in promoter regions represented only 16% of the total analysed fragments, they accounted for 42% of all lowly methylated fragments (Figure 2a), compared to fragments located in gene bodies (27%), intergenic (23%) and upstream (8%) regions. In contrast, only 9% of all highly methylated fragments were located in the promoter regions, compared to 53% in gene bodies, 31% in intergenic regions and 7% in upstream regions (Figure 2b). Supplementary Figure 1 depicts the coefficient of variability of the methylation level within each genomic region, for lowly methylated fragments on the one hand, and for highly methylated fragments on the other hand. Highly methylated fragments presented a low level of variability (< 3%), in opposition to lowly methylated fragments for which the coefficient of variability was higher than 33%. This pattern was the same for all the fragments, independently of the genomic region.

## Changes in DNA methylation following developmental exposure to EE2

Global DNA methylation levels of rivulus liver averaged $65.1 \pm 1.0\%$ in Ctl samples, $66.0 \pm 0.2\%$ in 4 ng/L and $65.5 \pm 0.6\%$ in 120 ng/L exposed individuals and were not significantly different (Anova on arcsin squareroot transformed values). Compared to the Ctl group, 58 DMFs were reported for the low EE2 concentration of 4 ng/L, 33 for the higher concentration of 120 ng/L, while 62 DMFs were significant between 120 and 4 ng/L groups. The majority of DMFs had a higher methylation level at 4 ng/L compared to Ctl and 120 ng/L groups, and at 120 ng/L compared to Ctl (Table 2, Figure 3a). Figure 3b depicts the DMF location within the RR genome and showed a similar pattern as the general distribution of the analysed fragments (Figure 1b) with the DMFs located in the order gene body > intergenic > promoter > upstream regions. None of the significant DMF was common to all three group comparisons (Figure 3c). However, 2 DMFs were significant at both EE2 treatments compared to Ctl: one located in an intergenic region, in proxi-

Table 2. Number of analysed fragments, analysed CpGs and number of DMFs in each group comparison.

|  | 4 ng/L vs Ctl | 120 ng/L vs Ctl | 120 vs 4 ng/L |
|---|---|---|---|
| Fragments with P < 0.01 | 1,674 (1.6%) | 787 (0.75%) | 1,471 (1.4%) |
| Fragments with ≥ 10% | 383 (0.36%) | 311 (0.30%) | 471 (0.46%) |
| Differentially methylated fragments (DMFs; P < 0.01 and methylation difference ≥ 10%)** | **58** (0.06%) (46↑, 12↓) | **33** (0.03%) (20↑, 13↓) | **62** (0.06%) (13↑,49↓) |
| Gene body DMFs | 26 (22↑, 4↓) | 12 (9↑, 3↓) | 31 (6↑, 25↓) |
| Promoter (>5kb) DMFs | 6 (4↑, 2↓) | 4 (3↑, 1↓) | 11(3↑, 8↓) |
| Upstream (5–10 kb) DMFs | 3↑ | 4 (2↑, 2↓) | 3↓ |
| Intergenic (>10 kb) DMFs | 23 (17↑, 6↓) | 13 (6↑, 7↓) | 17 (4↑, 13↓) |
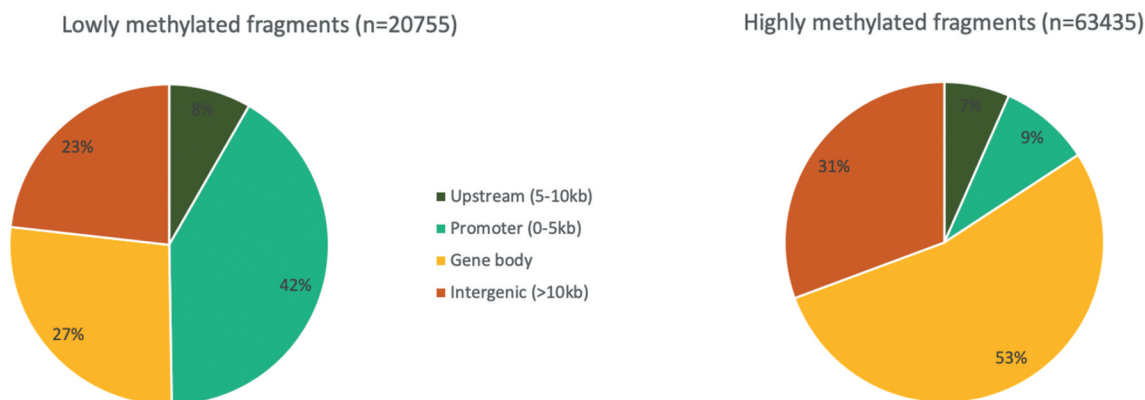


Figure 2. Distribution of lowly methylated (<20% methylation) (a) and highly methylated (> 80% methylation) (b) fragments among the genomic regions (gene body, promoter, upstream, intergenic) of the liver reduced representative genome of the mangrove rivulus.
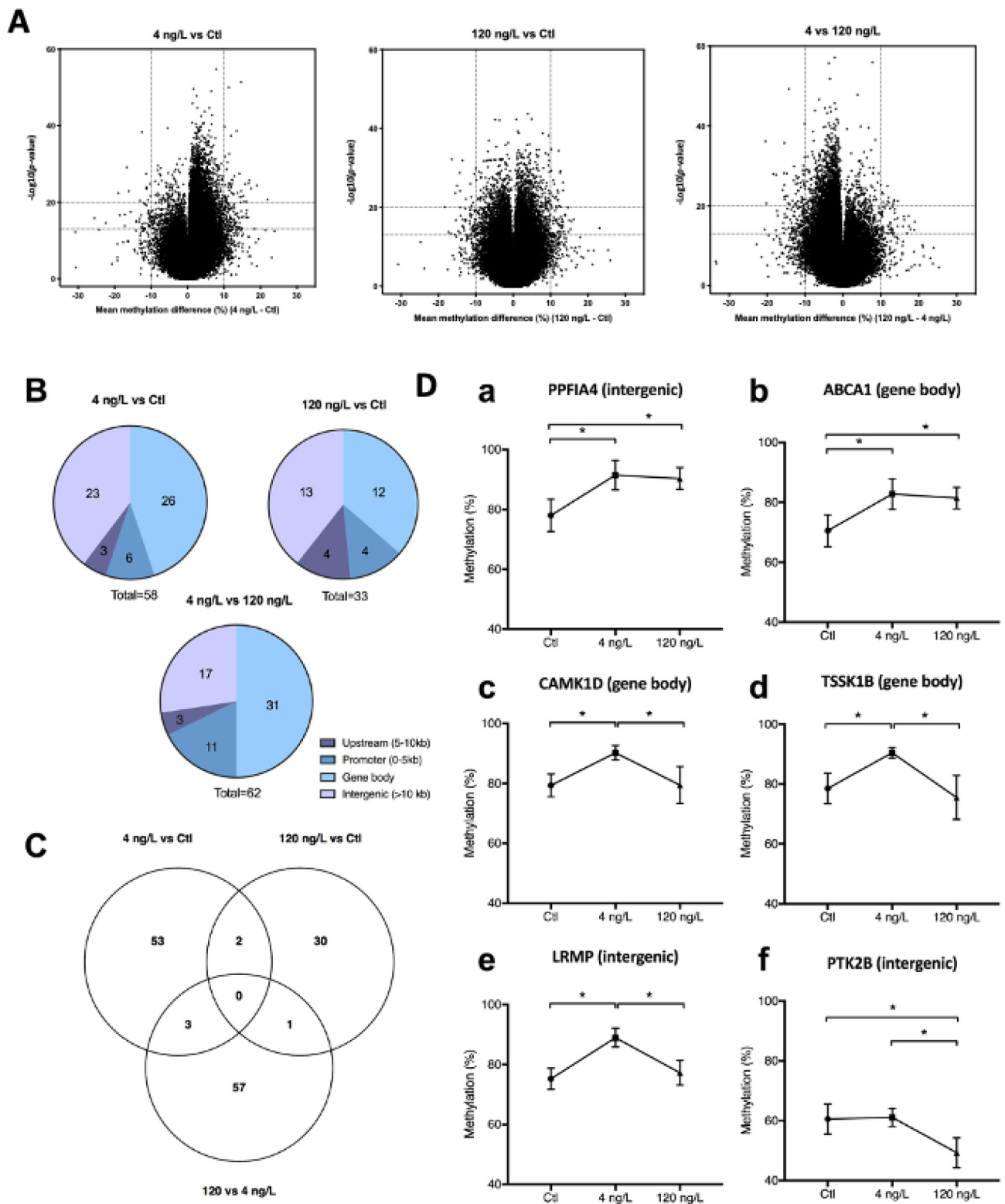
**Figure 3.** (a) Volcano plot representations of significance and differential methylation in analysed fragments. Dotted lines represent thresholds of significance: −10LogP20, corresponding to P0.01 and mean methylation difference 10% (a) 4 ng/L vs control; B:120 ng/L vs control; C: 4 ng/L vs 120 ng/L. (b) Overall distribution of DMFs in genomic elements; (c) Venn Diagram showing specific and common DMFs between group comparisons; (d) Methylation levels (%) of common fragments between 4 vs Ctl and 120 vs Ctl comparisons (Da,Db), between 4 vs Ctl and 4 vs 120 comparisons (c–e) and between 120 vs Ctl and 4 vs 120 ng/L comparisons (f). For clearer representation of methylation differences, y axes start at 40%.

mity of the *liprin-alpha-4* (PPFIA4) gene, and one located in the gene body of the *ATP-binding cassette sub-family A member 1* gene (ABCA1) (Figure 3d). In addition, there were three common DMFs to 4 ng/L vs Ctl and 4 ng/L vs 120 ng/L comparisons, of which two were found within gene bodies: *calcium/calmodulin-dependent protein kinase type 1D* (CAMK1D), *testis-specific serine/threonine-protein kinase* (TSSK1B) and one in the intergenic region in proximity of the *lymphoid-restricted membrane protein-like* gene (LRMP). Finally, one fragment was significantly differentially methylated in both 120 vs Ctl and 120 vs 4 ng/L, located in the intergenic region in proximity to the PTK2B (*protein-tyrosine kinase 2*) gene. Tables 3–5 report the DMFs for the three comparisons: 4 ng/L vs Ctl, 120 ng/L vs Ctl, and 4 ng/L vs 120 ng/L, respectively. Among all group comparisons, the maximum effect size (difference in mean methylation) observed for hypermethylation was 21.9% (promoter region of *nipped-B-like protein B*, NIPBL) and −19.0% for hypomethylation (intergenic region, *eukaryotic translation initiation factor 4 gamma 3* EIF4G3). Both DMFs were significant between 4 ng/L and Ctl groups. It is also important to report the fact that very few DMFs were classified as lowly methylated fragments. When comparing exposed groups to Ctl, only NIPBL was lowly methylated in Ctl at 19.6% and increased to 41.5% after exposure to 4 ng/L EE2.

## Networks associated with gene body and promoter DMFs

A network analysis was performed with the IPA software using genes corresponding to DMFs located in gene body and promoter regions. Only networks with the highest number of associated DMFs are reported (Table 6). The highest network scores were found in the 4 ng/L vs Ctl comparison: the first network involved 16 molecules, with functions related to organismal abnormalities, cell morphology and nervous system function and is represented in Figure 4. The second network was composed of 15 molecules and was related to gene expression, drug metabolism and lipid metabolism. In 120 ng/L exposed individuals compared to Ctl, only one network with more than nine molecules was established in IPA, related to lipid metabolism,

molecular transport and small molecule biochemistry. Finally, four networks were created when comparing 120 ng/L to 4 ng/L exposed individuals, endorsing the differences in the response at these two concentrations. The first network was related to antigen presenting, cell-to-cell signalling and the inflammatory response. The 2nd and 3rd networks were mostly associated with cellular processes (e.g., cellular organization and morphology, cell-to-cell signalling) and the 4th network was linked to cellular death and survival and connective tissue function.

## DNA methylation of genes involved in growth, steroidogenesis and reproduction

As mentioned above, a total of 14,679 unique RefSeq identifiers were attributed among the 102,018 analysed fragments. It is therefore possible to search for genes of interest regardless of their significance. Based on the effects observed on growth, steroid hormone levels and reproduction reported in Voisin et al. [48] following EE2 exposure, we extracted DNA methylation levels of selected genes or biomarkers involved in the response to EE2 (oestrogen receptor), reproduction (vitellogenin), steroidogenesis (aromatase) and the somatotropic axis (insulin-like growth factor I and II, Igf1 receptor, Igf binding protein and Igf2 mRNA binding proteins) (Table 7). This list is non-exhaustive and contains 24 fragments located in gene bodies [21] or promoter regions [3]. The methylation difference for all the fragments is very low (below 5%). However, a fragment of 153 bp located in vitellogenin-1-like gene body showed a hypomethylation of 9.4% after exposure to EE2 at 120 ng/L compared to control.

## Discussion

### Characterization of the mangrove rivulus RR genome

Cytosine methylation is a predominant epigenetic regulation of gene expression. It occurs mainly on CpG dinucleotides and is classically associated with gene silencing when it occurs at CpG rich promoter regions [60]. Whole genome bisulphite sequencing (WGBS) provides global genome coverage of DNA methylation at single-base

**Table 3.** Differentially methylated fragments (DMFs) between 4 ng/L EE2 and control individuals. Methylation difference represents the mean methylation at 4 ng/L minus the mean methylation in controls. Fragments discussed in the text are marked with *.

| GeneID | Len | CpGs | Description | Human gene symbol | Ctl (% Meth.) | 4 ng/L (% Meth.) | Meth. Difference (%) | P-value | Test | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Gene body** | | | | | | | | | | |
| XP_017282219.1 | 126 | 4 | Protein kinase C-binding protein NELL1 | NELL1 | 62.6 | 79.4 | 16.8 | 3.20E-03 | F(1.8) = 17.24 | |
| XP_017293079.1 | 186 | 11 | Sphingosine 1-phosphate receptor 1 | S1PR1 | 59.4 | 74.9 | 15.5 | 9.64E-03 | F(1.6) = 13.98 | |
| XP_017282327.1 | 144 | 8 | Drebrin | DBN1 | 72.2 | 86.9 | 14.7 | 5.12E-03 | F(1.8) = 14.56 | |
| XP_017297110.1 | 139 | 7 | Protein phosphatase 1 regulatory subunit 12 C-like | PPP1R12C | 51.8 | 66.0 | 14.3 | 1.79E-03 | F(1.7) = 23.81 | |
| XP_017263166.1 | 57 | 2 | Astrotactin-2 | ASTN2 | 23.2 | 36.7 | 13.5 | 1.38E-04 | F(1.8) = 46.24 | |
| XP_017269942.1 | 112 | 6 | ATP-binding cassette sub-family A member 1-like | ABCA1 | 70.5 | 83.0 | 12.4 | 5.99E-03 | F(1.8) = 13.73 | * |
| XP_017273856.1 | 72 | 2 | Rhomboid-related protein 1 | RHBDL1 | 71.5 | 83.5 | 12.0 | 5.64E-03 | F(1.5) = 21.51 | |
| XP_017287079.1 | 95 | 2 | AT-rich interactive domain-containing protein 2 | ARID2 | 53.3 | 65.3 | 12.0 | 3.68E-03 | F(1.4) = 37.07 | |
| XP_017290801.1 | 47 | 2 | Nuclear receptor corepressor 2 | NCOR2 | 53.9 | 65.4 | 11.5 | 3.41E-03 | F(1.8) = 16.86 | |
| XP_017259611.1 | 127 | 8 | Protein CASC3 | CASC3 | 79.0 | 90.5 | 11.5 | 4.87E-03 | F(1.6) = 18.84 | |
| XP_017284950.1 | 143 | 6 | Zinc finger protein 385 C | ZNF385C | 81.1 | 92.5 | 11.4 | 5.43E-04 | F(1.8) = 30.77 | |
| XP_017290159.1 | 154 | 5 | Testis-specific serine/threonine-protein kinase 1-like | TSSK1B | 79.0 | 90.4 | 11.4 | 3.12E-03 | F(1.7) = 19.44 | * |
| XP_017272906.1 | 101 | 5 | Calcipressin-3 | RCAN3 | 71.5 | 82.8 | 11.3 | 9.95E-06 | F(1.8) = 95.83 | |
| XP_017269734.1 | 152 | 12 | Neogenin | NEO1 | 79.8 | 90.8 | 11.1 | 6.99E-03 | F(1.8) = 12.95 | |
| XP_017263758.1 | 72 | 6 | Disks large homolog 3 | DLG3 | 71.4 | 82.2 | 10.8 | 1.39E-03 | F(1.8) = 22.84 | |
| XP_017270265.1 | 98 | 7 | Round spermatid basic protein 1-like protein | RSBN1 | 65.6 | 76.3 | 10.7 | 1.37E-03 | F(1.8) = 22.96 | |
| XP_017280982.1 | 184 | 13 | Whirlin | WHRN | 74.4 | 85.1 | 10.7 | 8.98E-03 | F(1.4) = 22.56 | |
| XP_017294972.1 | 126 | 5 | Friend leukaemia integration 1 transcription factor-like | FLI1 | 63.3 | 73.8 | 10.5 | 8.18E-03 | F(1.8) = 12.19 | |
| XP_017274567.1 | 173 | 6 | Matrix metalloproteinase-15-like | MMP15 | 82.8 | 93.1 | 10.3 | 3.53E-03 | F(1.6) = 21.56 | |
| XP_017269638.1 | 153 | 8 | Late secretory pathway protein AVL9 homolog | AVL9 | 61.6 | 71.9 | 10.3 | 9.28E-03 | F(1.6) = 14.22 | |
| XP_017261650.1 | 148 | 13 | Calcium/calmodulin-dependent protein kinase type 1D-like | CAMK1D | 79.5 | 89.8 | 10.3 | 6.41E-04 | F(1.8) = 29.22 | * |
| XP_017288959.1 | 106 | 4 | Neurotensin receptor type 1 | NTSR1 | 80.6 | 90.7 | 10.2 | 1.07E-04 | F(1.8) = 49.78 | |
| XP_017273249.1 | 136 | 9 | Lethal [3]malignant brain tumour-like protein 1 | L3MBTL1 | 87.8 | 77.6 | −10.2 | 3.30E-03 | F(1.6) = 22.17 | |
| XP_017275235.1 | 161 | 14 | PH and SEC7 domain-containing protein 2 | PSD2 | 37.6 | 26.7 | −10.9 | 8.71E-03 | F(1.8) = 11.89 | |
| XP_017288462.1 | 144 | 5 | Adhesion G-protein coupled receptor D2 | ADGRL2 | 92.5 | 81.3 | −11.2 | 4.30E-03 | F(1.7) = 17.22 | |
| XP_017293232.1 | 169 | 8 | Amyloid beta A4 protein | APP | 38.5 | 27.2 | −11.3 | 4.05E-03 | F(1.8) = 15.85 | |
| **Promoter** | | | | | 0.0 | 0.0 | | | | |
| XP_017263260.1 | 53 | 3 | Nipped-B-like protein B | NIPBL | 19.6 | 41.6 | 21.9 | 8.51E-03 | F(1.8) = 12.00 | * |
| XP_017266946.1 | 172 | 10 | G1/S-specific cyclin-D2-like | CCND2 | 53.5 | 72.5 | 19.0 | 3.66E-03 | F(1.6) = 21.24 | * |
| XP_017267503.1 | 176 | 13 | Translation initiation factor IF-3, mitochondrial | MTIF3 | 73.3 | 85.2 | 11.9 | 4.16E-03 | F(1.7) = 17.43 | |
| XP_017259959.1 | 165 | 9 | Protein SCAI | SCAI | 83.6 | 94.2 | 10.6 | 9.07E-03 | F(1.6) = 14.36 | |
| XP_017283065.1 | 126 | 4 | Tyrosine-protein kinase STYK1-like | STYK1 | 57.7 | 46.4 | −11.3 | 4.51E-03 | F(1.8) = 15.25 | |
| XP_017283302.1 | 138 | 4 | Natterin-3-like | | 83.0 | 66.3 | −16.7 | 1.22E-03 | F(1.6) = 32.87 | |
| **Upstream** | | | | | 0.0 | 0.0 | | | | |
| XP_017279598.1 | 153 | 5 | Homeobox protein CDX-1 | CDX1 | 51.5 | 68.9 | 17.4 | 4.36E-03 | F(1.7) = 17.12 | |
| XP_017280347.1 | 97 | 7 | Peroxisome proliferator-activated receptor alpha | PPARA | 52.5 | 64.7 | 12.1 | 2.58E-03 | F(1.8) = 18.58 | |
| XP_017295153.1 | 154 | 5 | Cortactin-binding protein 2-like | CTTNBP2 | 83.3 | 94.7 | 11.5 | 3.35E-03 | F(1.6) = 22.04 | |
| **Intergenic** | | | | | 0.0 | 0.0 | | | | |
| XP_017271697.1 | 116 | 3 | Sorting nexin-29 | SNX29 | 66.2 | 83.1 | 16.9 | 2.67E-03 | F(1.6) = 24.16 | |
| XP_017286782.1 | 132 | 4 | Dedicator of cytokinesis protein 8 | DOCK8 | 57.3 | 73.4 | 16.2 | 9.74E-03 | F(1.8) = 11.38 | |
| XP_017275594.1 | 95 | 4 | Ephrin-A1-like | EFNA1 | 54.0 | 69.8 | 15.8 | 3.34E-03 | F(1.8) = 16.98 | |
| XP_017270106.1 | 162 | 7 | Armadillo repeat protein deleted in velo-cardio-facial syndrome | ARVCF | 73.4 | 88.7 | 15.3 | 6.87E-03 | F(1.6) = 16.26 | |
| XP_017275116.1 | 105 | 4 | Teashirt homolog 3 | TSHZ3 | 35.0 | 49.8 | 14.8 | 7.20E-03 | F(1.7) = 14.03 | |

(*Continued*)

**Table 3.** (Continued).

| GeneID | Len | CpGs | Description | Human gene symbol | Ctl (% Meth.) | 4 ng/L (% Meth.) | Meth. Difference (%) | P-value | Test |
|---|---|---|---|---|---|---|---|---|---|
| XP_017275238.1 | 182 | 8 | TGF-beta receptor type-2-like | TGFBR2 | 46.1 | 60.7 | 14.7 | 7.26E-06 | F (1.6) = 205.01 |
| XP_017272930.1 | 98 | 2 | Uncharacterized protein C4orf45 homolog | | 60.7 | 75.0 | 14.4 | 7.28E-03 | F(1.7) = 13.97 |
| XP_017268339.1 | 172 | 6 | Zinc finger protein 511 | ZNF511 | 68.3 | 80.9 | 12.6 | 6.16E-03 | F(1.8) = 13.59 |
| XP_017272056.1 | 139 | 11 | Protein Wiz-like | WIZ | 81.7 | 94.0 | 12.4 | 2.63E-03 | F(1.7) = 20.72 |
| XP_017271447.1 | 135 | 4 | Liprin-alpha-4 | PPFIA4 | 78.6 | 90.8 | 12.2 | 3.28E-03 | F(1.8) = 17.10 |
| XP_017285408.1 | 52 | 4 | Lymphoid-restricted membrane protein-like | LRMP | 75.7 | 87.7 | 11.9 | 1.83E-04 | F(1.8) = 42.61 |
| XP_017272912.1 | 179 | 9 | Protocadherin Fat 4 | FAT4 | 84.6 | 95.7 | 11.1 | 3.98E-03 | F(1.5) = 25.36 |
| XP_017263052.1 | 157 | 6 | Protein C9orf69 homolog | TMEM250 | 70.9 | 81.9 | 10.9 | 8.97E-04 | F(1.8) = 26.31 |
| XP_017287805.1 | 122 | 4 | Collagen alpha-1(XI) chain | COL11A1 | 67.4 | 78.1 | 10.8 | 8.22E-03 | F(1.8) = 12.17 |
| XP_017279615.1 | 103 | 4 | Transmembrane protein 126A-like | TMEM126A | 60.1 | 70.5 | 10.4 | 7.63E-03 | F(1.8) = 12.52 |
| XP_017286707.1 | 100 | 3 | Homeobox protein SIX3 | SIX3 | 34.8 | 45.1 | 10.3 | 6.88E-03 | F(1.7) = 14.30 |
| XP_017272585.1 | 118 | 13 | Transmembrane protein 132 C | TMEM132C | 43.4 | 53.5 | 10.2 | 7.24E-03 | F(1.8) = 12.78 |
| XP_017293370.1 | 155 | 10 | Lethal(2) giant larvae protein homolog 2 | LLGL1 | 90.4 | 80.3 | −10.2 | 6.85E-03 | F(1.8) = 13.06 |
| XP_017288632.1 | 176 | 7 | Large neutral amino acids transporter small subunit 2-like | SLC7A8 | 89.8 | 78.4 | −11.4 | 1.80E-03 | F(1.7) = 23.80 |
| XP_017268028.1 | 127 | 4 | Protein FAM222B | FAM222B | 41.6 | 29.1 | −12.5 | 1.47E-04 | F(1.8) = 45.41 |
| XP_017269623.1 | 166 | 17 | Cadherin-20-like | CDH20 | 42.7 | 29.6 | −13.0 | 9.02E-03 | F(1.7) = 12.79 |
| XP_017270388.1 | 147 | 8 | Protein Wnt-5b | WNT5B | 56.0 | 39.9 | −16.1 | 5.14E-03 | F(1.7) = 16.05 |
| XP_017264233.1 | 167 | 3 | Eukaryotic translation initiation factor 4 gamma 3 | EIF4G3 | 61.4 | 42.4 | −19.0 | 5.78E-03 | F(1.8) = 13.92 |

**Table 4.** Differentially methylated fragments (DMFs) between 120 ng/L EE2 and control individuals. Methylation difference represents the mean methylation at 120 ng/L minus the mean methylation in controls. Fragments discussed in the text are marked with *.

| GeneID | Len | CpGs | Description | Human gene symbol | Ctl (% Meth.) | 120 ng/L (% Meth.) | Meth. Difference (%) | P-value | Test | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Gene body** | | | | | | | | | | |
| XP_017286106.1 | 136 | 4 | Mitogen-activated protein kinase kinase kinase kinase 3 0 | MAP3K3 | 74.2 | 87.5 | 13.3 | 2.34E-03 | F(1.8) = 19.20 | |
| XP_017285971.1 | 74 | 2 | Nuclear factor 1 X-type-like | NFIX | 41.8 | 55.2 | 13.3 | 3.76E-03 | F(1.8) = 16.29 | |
| XP_017259709.1 | 123 | 6 | LMBR1 domain-containing protein 2-B-like | LMBRD2 | 79.7 | 92.0 | 12.3 | 9.18E-03 | F(1.7) = 12.70 | |
| XP_017291042.1 | 172 | 15 | Neuronal PAS domain-containing protein 2-like | NPAS2 | 72.9 | 85.0 | 12.1 | 3.67E-04 | F(1.8) = 34.66 | |
| XP_017286691.1 | 65 | 2 | A disintegrin and metalloproteinase with thrombospondin motifs 6 | ADAMTS6 | 79.0 | 91.0 | 12.0 | 6.47E-03 | F(1.8) = 13.34 | |
| XP_017287636.1 | 158 | 7 | Leucine-rich PPR motif-containing protein, mitochondrial | LRPPRC | 81.8 | 93.7 | 11.9 | 6.48E-03 | F(1.6) = 16.68 | |
| XP_017259612.1 | 145 | 6 | Potassium Voltage-gated channel subfamily H member 5-like | KCNH5 | 79.2 | 90.7 | 11.5 | 5.41E-03 | F(1.7) = 15.75 | |
| XP_017284196.1 | 159 | 8 | AT-rich interactive domain-containing protein 3B-like | ARID3B | 66.9 | 77.7 | 10.8 | 2.05E-03 | F(1.8) = 20.08 | |
| XP_017269942.1 | 112 | 6 | ATP-binding cassette sub-family A member 1-like | ABCA1 | 70.5 | 81.1 | 10.6 | 5.20E-03 | F(1.8) = 14.48 | * |
| XP_017282238.1 | 72 | 2 | Neuron navigator 2 | NAV2 | 86.7 | 76.3 | −10.4 | 1.43E-03 | F(1.7) = 25.80 | |
| XP_017275214.1 | 174 | 12 | Teneurin-1-like | TENM1 | 86.5 | 75.3 | −11.2 | 5.48E-03 | F(1.7) = 15.66 | * |
| XP_017280416.1 | 108 | 5 | Ephrin-A3-like | EFNA3 | 73.9 | 62.4 | −11.4 | 2.53E-03 | F(1.8) = 18.70 | |
| **Promoter** | | | | | | | | | | |
| XP_017296617.1 | 106 | 5 | LIM domain Transcription factor LMO4.1-like | LMO4 | 72.2 | 84.2 | 12.0 | 4.29E-03 | F(1.8) = 15.53 | |
| XP_017280337.1 | 48 | 2 | Protein NDRG1-like | NDRG1 | 29.2 | 41.1 | 11.9 | 3.72E-03 | F(1.7) = 18.19 | * |
| XP_017273723.1 | 137 | 3 | Solute carrier family 22 member 6-like | SLC22A6 | 85.0 | 96.5 | 11.5 | 9.45E-03 | F(1.6) = 14.10 | |
| XP_017278883.1 | 158 | 9 | Collectrin | CLTRN | 73.2 | 59.0 | −14.2 | 3.86E-03 | F(1.5) = 25.71 | |
| **Upstream** | | | | | | | | | | |

**Table 4.** (Continued).

| GeneID | Len | CpGs | Description | Human gene symbol | Ctl (% Meth.) | 120 ng/L (% Meth.) | Meth. Difference (%) | P-value | Test | |
|---|---|---|---|---|---|---|---|---|---|---|
| XP_017276084.1 | 156 | 4 | Protein FAM102B | FAM102B | 84.5 | 97.0 | 12.6 | 1.33E-03 | F(1.6) = 31.80 | |
| XP_017266436.1 | 65 | 2 | Supervillin-like | SVIL | 79.8 | 89.9 | 10.1 | 4.29E-03 | F(1.7) = 17.23 | |
| XP_017277864.1 | 46 | 2 | Receptor-interacting serine/ threonine-protein kinase 1 | RIPK1 | 80.5 | 67.0 | −13.6 | 6.50E-04 | F(1.8) = 29.11 | |
| XP_017261155.1 | 151 | 7 | Golgin subfamily A member 6-like protein 2 | GOLGA6L2 | 73.8 | 58.8 | −15.0 | 2.85E-03 | F(1.7) = 20.11 | |
| **Intergenic** | | | | | | | | | | |
| XP_017279686.1 | 122 | 5 | Ras-GEF domain-containing family member 1 C | RASGEF1B | 72.8 | 85.5 | 12.7 | 9.78E-03 | F(1.7) = 12.36 | |
| XP_017277156.1 | 163 | 11 | Coiled-coil domain-containing protein 85A | CCDC85A | 81.3 | 93.9 | 12.7 | 1.20E-03 | F(1.5) = 43.51 | |
| XP_017271447.1 | 135 | 4 | Liprin-alpha-4 | PPFIA4 | 78.6 | 89.7 | 11.2 | 5.86E-03 | F(1.7) = 15.25 | |
| XP_017288708.1 | 170 | 5 | Protein turtle homolog B-like | IGSF9B | 78.0 | 88.3 | 10.2 | 8.82E-03 | F(1.7) = 12.91 | |
| XP_017297201.1 | 96 | 3 | Williams-Beuren syndrome chromosomal region 27 protein-like | | 69.3 | 79.5 | 10.2 | 1.17E-03 | F(1.8) = 24.14 | |
| XP_017263439.1 | 143 | 4 | Rho GTPase-activating protein 24 | ARHGAP24 | 79.6 | 89.8 | 10.2 | 3.37E-04 | F(1.7) = 42.13 | |
| XP_017282790.1 | 124 | 10 | Protein-tyrosine kinase 2-beta | PTK2B | 60.3 | 49.6 | −10.7 | 7.58E-03 | F(1.8) = 12.56 | |
| XP_017292488.1 | 165 | 11 | Fibrinogen alpha chain | FGA | 56.9 | 45.8 | −11.1 | 8.99E-03 | F(1.6) = 14.42 | * |
| XP_017281773.1 | 112 | 2 | Alpha-crystallin A chain-like | CRYAA | 75.2 | 62.5 | −12.7 | 6.19E-03 | F(1.7) = 14.92 | |
| XP_017270131.1 | 151 | 10 | Semaphorin-4 C | SEMA4C | 62.6 | 48.3 | −14.3 | 1.33E-03 | F(1.6) = 31.80 | |
| XP_017283010.1 | 154 | 6 | General transcription factor II–I repeat domain-containing protein 2-like | GTF2IRD2 | 53.2 | 37.7 | −15.5 | 4.35E-03 | F(1.8) = 15.45 | |
| XP_017276780.1 | 127 | 4 | THAP domain-containing protein 10 | THAP10 | 93.8 | 77.4 | −16.4 | 6.03E-04 | F(1.6) = 42.98 | |
| XP_017292195.1 | 140 | 6 | Cullin-5 | CUL5 | 88.4 | 70.1 | −18.3 | 9.85E-03 | F(1.6) = 13.84 | |

resolution, and is consequently referred as the 'gold standard' method. Despite its benefits, including a high coverage of CpGs (>90% in human) and an unbiased representation, it remains expensive and difficult to use with large number of samples, mostly because the detection of changes in methylation between samples demands high sequencing depth [61]. In the present study, we opted for reduced representation bisulphite sequencing (RRBS), an approach that also resolves single-base DNA methylation, combining the use of restriction enzymes and bisulphite sequencing. Unlike WGBS, RRBS can only sequence rich-CpG regions, and has limited coverage of the genome [62]. However, the fact that it enriches the analysed fragments with a high density of CpGs makes this technique cost-effective and allows to work on a higher number of samples with a high sequencing depth [63]. Here we reported the DNA methylation pattern of mangrove rivulus liver using RRBS. This species shows a very peculiar reproduction strategy as it is the only vertebrate which alternates between cross-fertilization and selfing, the latter being the most common. It results in several lineages of

isogenic individuals, and epigenetic mechanisms have been proposed to explain the occurrence of phenotypic variability despite a low level of genetic diversity [42]. We reported a RR genome representing 1.4% of the whole genome, which was comparable to what was obtained in mouse (1.4%) or in zebrafish (2.2%) brain [54,64]. The percentage of CpG covered in the RR genome reached 6.1% of the total number of genomic CpGs, reaching a 4.3-fold enrichment in CpG sites. This CpG coverage is similar to what was found in mouse and zebrafish (7.0% and 5.3%, respectively [54],), but higher than what other studies reported for diverse fish species such as stickleback [1% [65]], Atlantic salmon [2.75% [66]], rainbow trout [<1% [67]] and guppies [1.5–2% [68]]. In mangrove rivulus brain, Berbel-Filho et al. [46] also reported a lower CpG coverage at 1.2%. This might be astonishing considering that different tissues of the same species usually show the same methylation profile [69] and methodological aspects can not be rejected to explain this discrepancy between those studies. Most noticeable is the fact that we obtained significantly more reads in the present study (explaining that

**Table 5.** Differentially methylated fragments (DMFs) between 120 ng/L EE2 and 4 ng/L EE2 exposed individuals. Methylation difference represents the mean methylation at 120 ng/L minus the mean methylation at 4 ng/L. Fragments discussed in the text are marked with *.

| GeneID | Len | CpGs | Description | Human gene symbol | 4 ng/L Meth (%) | 120 ng/L Meth (%) | Diffmeth (%) | Pr | Test | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Gene body** | | | | | | | | | | |
| XP_017288867.1 | 154 | 13 | Scavenger receptor class F member 1 | SCARF1 | 21.9 | 35.2 | 13.3 | 8.77E-03 | $F_{(1,8)}$ = 11.86 | |
| XP_017272657.1 | 145 | 13 | Dihydropyrimidinase-related protein 2-like | DPYSL2 | 22.9 | 34.9 | 12.0 | 6.20E-03 | $F_{(1,8)}$ = 13.55 | |
| XP_017280995.1 | 45 | 2 | Glutamate receptor ionotropic, NMDA 1-like | GRIN2D | 58.2 | 70.0 | 11.8 | 7.75E-03 | $F_{(1,8)}$ = 12.45 | |
| XP_017271462.1 | 61 | 3 | Selenocysteine-specific elongation factor | EEFSEC | 29.4 | 41.2 | 11.8 | 3.18E-03 | $F_{(1,8)}$ = 17.28 | |
| XP_017286963.1 | 148 | 8 | G patch domain and ankyrin repeat-containing protein 1 | ANKRD1 | 84.4 | 95.8 | 11.4 | 9.38E-03 | $F_{(1,6)}$ = 14.15 | |
| XP_017289196.1 | 101 | 7 | Pre-B-cell leukaemia transcription factor 1 | PBX1 | 40.2 | 50.8 | 10.6 | 8.05E-03 | $F_{(1,8)}$ = 12.27 | |
| XP_017290510.1 | 51 | 4 | CREB-regulated transcription coactivator 3 | CRTC3 | 45.2 | 35.1 | -10.0 | 3.68E-03 | $F_{(1,8)}$ = 16.40 | |
| XP_017285590.1 | 138 | 4 | Fibroblast growth factor 14-like | FGF14 | 85.6 | 75.5 | -10.1 | 6.02E-04 | $F_{(1,8)}$ = 34.77 | |
| XP_017261320.1 | 140 | 9 | Tight junction protein ZO-1-like | TJP1 | 95.8 | 85.6 | -10.2 | 1.65E-03 | $F_{(1,7)}$ = 57.07 | |
| XP_017295247.1 | 155 | 11 | Protein AF-17-like | MLLT6 | 92.7 | 82.4 | -10.3 | 9.40E-03 | $F_{(1,5)}$ = 16.77 | |
| XP_017288897.1 | 108 | 5 | Methyltransferase-like protein 16 | METTL16 | 69.2 | 58.8 | -10.4 | 6.98E-03 | $F_{(1,8)}$ = 12.96 | |
| XP_017283711.1 | 156 | 15 | Large neutral amino acids transporter small subunit 3 | SLC43A1 | 92.5 | 81.9 | -10.6 | 1.29E-03 | $F_{(1,7)}$ = 26.73 | |
| XP_017261650.1 | 148 | 13 | Calcium/calmodulin-dependent protein kinase type 1D-like | CAMK1D | 89.8 | 79.1 | -10.7 | 6.59E-03 | $F_{(1,8)}$ = 13.25 | * |
| XP_017261484.1 | 88 | 4 | Zinc finger protein AEBP2 | AEBP2 | 95.2 | 84.2 | -11.1 | 3.41E-03 | $F_{(1,8)}$ = 16.86 | |
| XP_017290127.1 | 142 | 7 | Phospholysine phosphohistidine inorganic pyrophosphate phosphatase | LHPP | 95.1 | 83.9 | -11.2 | 2.65E-03 | $F_{(1,5)}$ = 30.59 | |
| XP_017286486.1 | 60 | 4 | Rho guanine nucleotide exchange factor 17 | ARHGEF17 | 41.4 | 29.6 | -11.8 | 4.21E-03 | $F_{(1,8)}$ = 15.64 | |
| XP_017269775.1 | 137 | 5 | C-Jun-amino-terminal kinase-interacting protein 1 | MAPK8IP1 | 83.3 | 71.5 | -11.9 | 7.69E-03 | $F_{(1,7)}$ = 13.66 | |
| XP_017283446.1 | 101 | 3 | Multiple C2 and transmembrane domain-containing protein 1 | MCTP1 | 87.0 | 75.0 | -12.0 | 2.07E-03 | $F_{(1,8)}$ = 20.03 | |
| XP_017291190.1 | 148 | 9 | Nuclear factor 1 X-type-like | NFIX | 89.8 | 77.5 | -12.3 | 8.63E-03 | $F_{(1,6)}$ = 14.69 | |
| XP_017296764.1 | 142 | 6 | Serine-rich coiled-coil domain-containing protein 2-like | CCSER2 | 91.8 | 79.4 | -12.4 | 1.78E-03 | $F_{(1,6)}$ = 28.40 | |
| XP_017284506.1 | 137 | 5 | Neutral amino acid transporter B(0) | SLC1A5 | 19.2 | 6.8 | -12.4 | 5.05E-03 | $F_{(1,8)}$ = 14.64 | |
| XP_017272705.1 | 90 | 3 | G-protein coupled receptor 98 | GPR98 | 90.8 | 78.3 | -12.5 | 2.84E-03 | $F_{(1,6)}$ = 23.57 | |
| XP_017263193.1 | 131 | 3 | Regulator of G-protein signalling 3 | RGS3 | 88.3 | 75.8 | -12.6 | 8.92E-03 | $F_{(1,6)}$ = 14.48 | |
| XP_017295850.1 | 52 | 6 | Cathepsin E-like | CTSE | 75.2 | 62.2 | -13.0 | 2.30E-03 | $F_{(1,8)}$ = 19.32 | |
| XP_017279844.1 | 102 | 7 | Rab11 family-interacting protein 2 | RAB11FIP2 | 91.7 | 77.4 | -14.3 | 1.18E-05 | $F_{(1,8)}$ = 91.50 | |
| XP_017289883.1 | 140 | 5 | Partitioning defective 3 homolog B | PARD3B | 80.4 | 65.9 | -14.5 | 8.65E-03 | $F_{(1,8)}$ = 14.68 | |
| XP_017290159.1 | 154 | 5 | Testis-specific serine/threonine-protein kinase 1-like | TSSK1B | 90.4 | 75.9 | -14.6 | 5.69E-03 | $F_{(1,7)}$ = 15.43 | * |
| XP_017279357.1 | 128 | 3 | Ephrin type-A receptor 6 | EPHA6 | 77.8 | 62.7 | -15.1 | 5.71E-03 | $F_{(1,8)}$ = 13.98 | |
| XP_017262478.1 | 135 | 4 | Unconventional myosin-IXb | MYO9B | 66.6 | 51.3 | -15.3 | 6.94E-03 | $F_{(1,8)}$ = 12.99 | |
| XP_017265799.1 | 136 | 5 | Serine/threonine-protein phosphatase 2A regulatory subunit B'' subunit beta-like | PPP2R2B | 38.1 | 22.6 | -15.4 | 5.97E-03 | $F_{(1,7)}$ = 15.14 | |
| **Promoter** | | | | | | | | | | |
| XP_017268748.1 | 131 | 7 | Zinc finger protein ZFPM2 | ZFPM2 | 68.2 | 52.7 | -15.5 | 1.51E-03 | $F_{(1,6)}$ = 22.22 | |
| XP_017285362.1 | 44 | 2 | Death-associated protein 1 | DAP | 80.3 | 62.0 | -18.4 | 1.59E-03 | $F_{(1,6)}$ = 29.70 | * |
| XP_017274091.1 | 81 | 3 | Inactive rhomboid protein 1-like | RHBDF1 | 64.2 | 78.7 | 14.5 | 7.55E-03 | $F_{(1,6)}$ = 15.59 | |
| XP_017294893.1 | 50 | 2 | Ribosomal RNA processing protein 36 homolog | RRP36 | 50.7 | 62.8 | 12.1 | 8.34E-03 | $F_{(1,8)}$ = 12.10 | |
| XP_017292323.1 | 84 | 6 | E3 ubiquitin/ISG15 ligase TRIM25-like | ZFHX3 | 19.0 | 29.0 | 10.0 | 3.02E-03 | $F_{(1,8)}$ = 17.59 | |
| XP_017296523.1 | 125 | 6 | Synaptotagmin-1-like | SYTL1 | 86.0 | 75.7 | -10.3 | 3.01E-03 | $F_{(1,8)}$ = 17.62 | |
| XP_017290352.1 | 136 | 9 | Synaptotagmin-14 | SYT14 | 81.5 | 70.7 | -10.8 | 7.25E-03 | $F_{(1,8)}$ = 12.78 | |
| XP_017272703.1 | 99 | 3 | Transcriptional activator MN1 | MN1 | 82.5 | 71.6 | -10.9 | 8.16E-03 | $F_{(1,8)}$ = 12.20 | |
| XP_017294308.1 | 158 | 4 | Uncharacterized protein LOC108249428 | TCR alpha | 89.1 | 77.3 | -11.9 | 4.12E-03 | $F_{(1,5)}$ = 24.94 | |
| XP_017280009.1 | 157 | 8 | Collagen alpha-2(VI) chain-like | COL6A2 | 85.8 | 70.4 | -15.4 | 2.67E-04 | $F_{(1,7)}$ = 45.45 | |
| XP_017261271.1 | 125 | 4 | Repulsive guidance molecule A-like | RGMA | 88.9 | 73.0 | -15.9 | 4.05E-03 | $F_{(1,7)}$ = 17.62 | |
| XP_017285485.1 | 168 | 12 | Synaptotagmin-12 | SYT12 | 75.0 | 57.4 | -17.6 | 1.68E-03 | $F_{(1,7)}$ = 24.38 | |
| XP_017294454.1 | 168 | 15 | Zinc finger and BTB domain-containing protein 4-like | ZBTB4 | 65.3 | 45.2 | -20.2 | 8.65E-03 | $F_{(1,6)}$ = 14.68 | |

(Continued)

Table 5. (Continued).

| GeneID | Len | CpGs | Description | Human gene symbol | 4 ng/L Meth (%) | 120 ng/L Meth (%) | Diffmeth (%) | Pr | Test |
|---|---|---|---|---|---|---|---|---|---|
| **Upstream** | | | | | | | | | |
| XP_017284506.1 | 126 | 8 | Neutral amino acid transporter B(0) | SLC1A5 | 54.0 | 43.9 | −10.1 | 8.77E-05 | $F_{(1.8)} = 52.63$ |
| XP_017265262.1 | 179 | 9 | Nuclear receptor subfamily 2 group C member 2 | NR2C2 | 93.6 | 83.1 | −10.5 | 1.63E-03 | $F_{(1.6)} = 29.41$ |
| XP_017279091.1 | 165 | 10 | T-lymphoma invasion and metastasis-inducing protein 2-like | TIAM2 | 74.7 | 60.7 | −14.1 | 8.95E-03 | $F_{(1.8)} = 11.77$ |
| **Intergenic** | | | | | | | | | |
| XP_017264766.1 | 68 | 3 | Leucine-rich repeat transmembrane protein FLRT2-like | FLRT2 | 29.3 | 42.8 | 13.5 | 7.29E-03 | $F_{(1.7)} = 13.96$ |
| XP_017295097.1 | 105 | 3 | Uncharacterized protein LOC108249937 | | 71.5 | 85.1 | 13.5 | 7.43E-03 | $F_{(1.6)} = 15.70$ |
| XP_017280372.1 | 78 | 3 | Sugar phosphate exchanger 3 | SLC37A3 | 36.3 | 46.9 | 10.6 | 5.03E-03 | $F_{(1.8)} = 14.66$ |
| XP_017282790.1 | 85 | 4 | Protein-tyrosine kinase 2-beta | PTK2B | 53.1 | 63.4 | 10.3 | 4.62E-03 | $F_{(1.8)} = 15.12$ |
| XP_017263868.1 | 112 | 4 | Serine protease 23-like | PRSS23 | 89.1 | 79.0 | −10.1 | 1.30E-03 | $F_{(1.8)} = 23.38$ |
| XP_017268118.1 | 115 | 6 | RNA-binding protein Musashi homolog 2 | MSI2 | 72.8 | 62.6 | −10.2 | 2.68E-03 | $F_{(1.8)} = 18.33$ |
| XP_017285408.1 | 52 | 4 | Lymphoid-restricted membrane protein-like | LRMP | 87.7 | 77.2 | −10.4 | 9.80E-04 | $F_{(1.8)} = 25.58$ |
| XP_017274772.1 | 103 | 5 | Hepatic leukaemia factor | HLF | 78.1 | 67.6 | −10.5 | 2.36E-03 | $F_{(1.7)} = 21.55$ |
| XP_017295868.1 | 162 | 8 | Protein diaphanous homolog 3 | DIAPH3 | 85.8 | 74.9 | −10.8 | 8.12E-03 | $F_{(1.5)} = 18.03$ |
| XP_017261032.1 | 108 | 3 | NACHT, LRR and PYD domains-containing protein 12-like | NLRP12 | 94.8 | 83.9 | −10.9 | 5.96E-03 | $F_{(1.6)} = 17.28$ |
| XP_017282790.1 | 124 | 10 | Protein-tyrosine kinase 2-beta | PTK2B | 61.3 | 49.6 | −11.7 | 1.92E-03 | $F_{(1.8)} = 20.53$ |
| XP_017293105.1 | 121 | 4 | Uncharacterized protein LOC108248683 | | 81.2 | 69.3 | −11.9 | 2.76E-03 | $F_{(1.8)} = 18.15$ |
| XP_017284405.1 | 137 | 8 | Aryl hydrocarbon receptor nuclear translocator-like protein 1 | ARNTL | 95.6 | 82.8 | −12.9 | 1.03E-03 | $F_{(1.7)} = 28.95$ |
| XP_017277282.1 | 134 | 7 | Stromal membrane-associated protein 2-like | SMAP2 | 68.9 | 55.6 | −13.3 | 7.38E-03 | $F_{(1.8)} = 12.69$ |
| XP_017296945.1 | 92 | 3 | DNA repair protein RAD51 homolog 3 | RAD51 C | 61.1 | 46.3 | −14.8 | 1.54E-03 | $F_{(1.8)} = 22.08$ |
| XP_017264564.1 | 145 | 11 | GDNF-inducible zinc finger protein 1-like | GZF1 | 80.1 | 64.7 | −15.4 | 3.02E-03 | $F_{(1.8)} = 17.58$ |
| XP_017260964.1 | 158 | 6 | Myotubularin-related protein 10 | MTMR10 | 66.8 | 46.4 | −20.4 | 2.43E-04 | $F_{(1.7)} = 46.87$ |

**Table 6.** Top molecular networks and functions associated with our dataset. Only differentially methylated fragments situated in promoter or gene body and their corresponding annotated genes were taken into account.

| | Top Diseases and Functions | Score | # | Molecules in Network |
|---|---|---|---|---|
| 4 ng/L vs CTL | Organismal Injury and Abnormalities, Cell Morphology, Nervous System Development and Function | 40 | 16 | 26s Proteasome, ↑ABCA1, ↓ADGRL2, Akt, ↓APP, Calmodulin, ↑CAMK1D, ↑CASC3, ↑CCND2, ↑DBN1, ↑DLG3, ERK, ERK1/2, oestrogen receptor, ↑FLI1, GPR61, GPR78, HISTONE, Hsp90, IgG, Jnk, LDL, ↑MMP15, ↑NCOR2, ↑NIPBL, Nr1h, P38 MAPK, PI3K (complex), ↑RHBDL1, RNA polymerase II, ↑S1PR1, stigmasterol, ↓STYK1, Tgf beta, ↑TSSK1B |
| | Gene Expression, Drug Metabolism, Lipid Metabolism | 37 | 15 | AGT, ↑ARID2, ↑ASTN2, ↑AVL9, CREB1, EGFR, ESR1, FOXP3, GPR12, GPR85, HNRNPK, IFNG, IGSF10, ↓L3MBTL1, MAPK1, MATN2, ↑MTIF3, ↑NELL1, ↑NEO1, ↑NTSR1, ↑PPP1R12C, ↓PSD2, Pvr, ↑RCAN3, RELA, ↑RSBN1, S1PR, SAMD11, ↑SCAI, SMAD3, TNFRSF1A, TP53, UBL3, ↑WHRN, ↑ZNF385C |
| 120 ng/L vs CTL | Lipid Metabolism, Molecular Transport, Small Molecule Biochemistry | 23 | 9 | ↑ABCA1, Act1, ↑ARID3B, benzoic acid, CHL1, cholesterol sulphate, Cxcl15, ERK, FPR2, GPNMB, GPR119, hydrocortisone, Jnk, LMBRD2, LXR ligand-LXR-Retinoic acid-RXRŒ±, ↑MAP3K3, mir-33, MTX3, ↓NAV2, ↑NDRG1, ↑NPAS2, palmitoleic acid, PAPPA, SCARB1, SFXN5, SLC1A4, ↑SLC22A6, SPARCL1, STEAP4, stigmasterol, ↓TENM1, TGFB1, TNF, tretinoin, UTS2R |
| 120 ng/L vs 4 ng/L | Antigen Presentation, Cell-To-Cell Signalling and Interaction, Inflammatory Response | 29 | 13 | ACOX1, ↓ADGRV1, ↓AEBP2, APP, Aspartyl Protease, Camk, ↓CAMK1D, CAMK1G, ↓CTSE, DOCK3, GPR50, ↑GRIN2D, HSP90AA1, ↓LHPP, ↓MCTP1, ↓MLLT6, NMDA Receptor, NTSR2, PPARG, ↓PPP2R2B, QPCTL, QRFPR, RASL11B, ↑RRP36, ↑SCARF1, SET, SLC5A7, SNAP25, TAS1R3, TMOD4, TSSK2, ↓TSSK1B, ↑ZFHX3, ZKSCAN8, ZNF445 |
| | Cellular Assembly and Organization, Cellular Movement, Cell-To-Cell Signalling and Interaction | 21 | 10 | 3 beta HSD, Akt, ↑ANKRD1, ↓ARHGEF17, ARHGEF18, ↓DAP, ↑DPYSL2, EFNB, ↓EPHA6, ERK, gelatinase, GJC1, GJD2, Gsk3, indole, Insulin, KLF11, MAP3K13, ↓MAPK8IP1, miR-130a-3p (and other miRNAs w/seed AGUGCAA), MYZAP, NFkB (complex), NTN4, PDGF (family), PDK2, ↓RGS3, SETD6, SH3RF1, TACSTD2, TDO2, ↓TJP1, Vegf, VMP1, ↓ZBTB4, ↓ZFPM2 |
| | Cell Morphology, Cellular Compromise, Cellular Function and Maintenance | 21 | 10 | ANO3, APP, ARHGEF17, BCL2L11, CCSER2, CLK4, DIXDC1, DOCK3, ↑EEFSEC, EFNB3, ENTPD7, FBXL7, GRN, HIPK2, MAPK14, MASTL, METTL16, MN1, NBEAL1, NTN4, NXF1, NXPE3, PARD3B, PARD6B, PDK2, PDK3, QRICH1, RAC1, RGMA, ↑RHBDF1, RTN4R, SAMD4B, SYT12, SYTL1, ZNF445 |
| | Cell Death and Survival, Cellular Compromise, Connective Tissue Development and Function | 18 | 9 | AR, ARFGEF1, COL5A2, ↓COL6A2, COL9A1, Collagen type VI, ↓CRTC3, CTNNB1, DHH, DQX1, EHD3, ↓FGF14, FOS, Glucocorticoid-GCR, HSPA5, KSR1, LYPD1, MYO5B, ↓MYO9B, MYZAP, ↓NFIX, Nuclear factor 1, ↑PBX1, PKNOX2, PLAU, ↓RAB11FIP2, RBM33, SCGB1D4, ↓SLC1A5, SLC25A46, ↓SLC43A1, SPDEF, TDO2, TP53, WNT10B |

fragments aligned to a higher proportion of the genome) but also a lower alignment efficiency (44.4% in average vs 62.9%) as many fragments could not be mapped. This relatively low level of alignment efficiency was also observed in zebrafish brain following the exact same technical workflow [54]. RRBS reads are relatively short (40–220 bp with majority being < 100 bp) and its plausible that this contributes to large proportion of unmapped reads. Another possible explanation is the proportion of RRBS reads that are derived from repeat regions in the genome as the RRBS reads from repeat region aligns with less efficiency [70,71].

The distribution of the analysed fragments across the RR genome showed 16% located in promoter region. The analysed fragments in this region are also low in zebrafish liver at 12%, and even lower in Atlantic salmon (4%) [66]. On the contrary, this percentage is higher in mammal species such as mouse liver (52%) [69]. This proportion of analysed fragments located in promoter region can explain the observed difference of the mean global methylation level reported in fish species and in mammals, the latter having a lower value (28% in mouse vs 65% in the mangrove rivulus, 70% in zebrafish, 75% in Atlantic salmon). The promoter regions being less methylated on average than other regions in both fish and mammals, a higher proportion of fragments from this region decreases the average global methylation as observed in mouse RR genome. These differences between mammals and fish can also be reflected in the proportion of highly and lowly methylated fragments. If the majority of fragments are either highly or lowly methylated
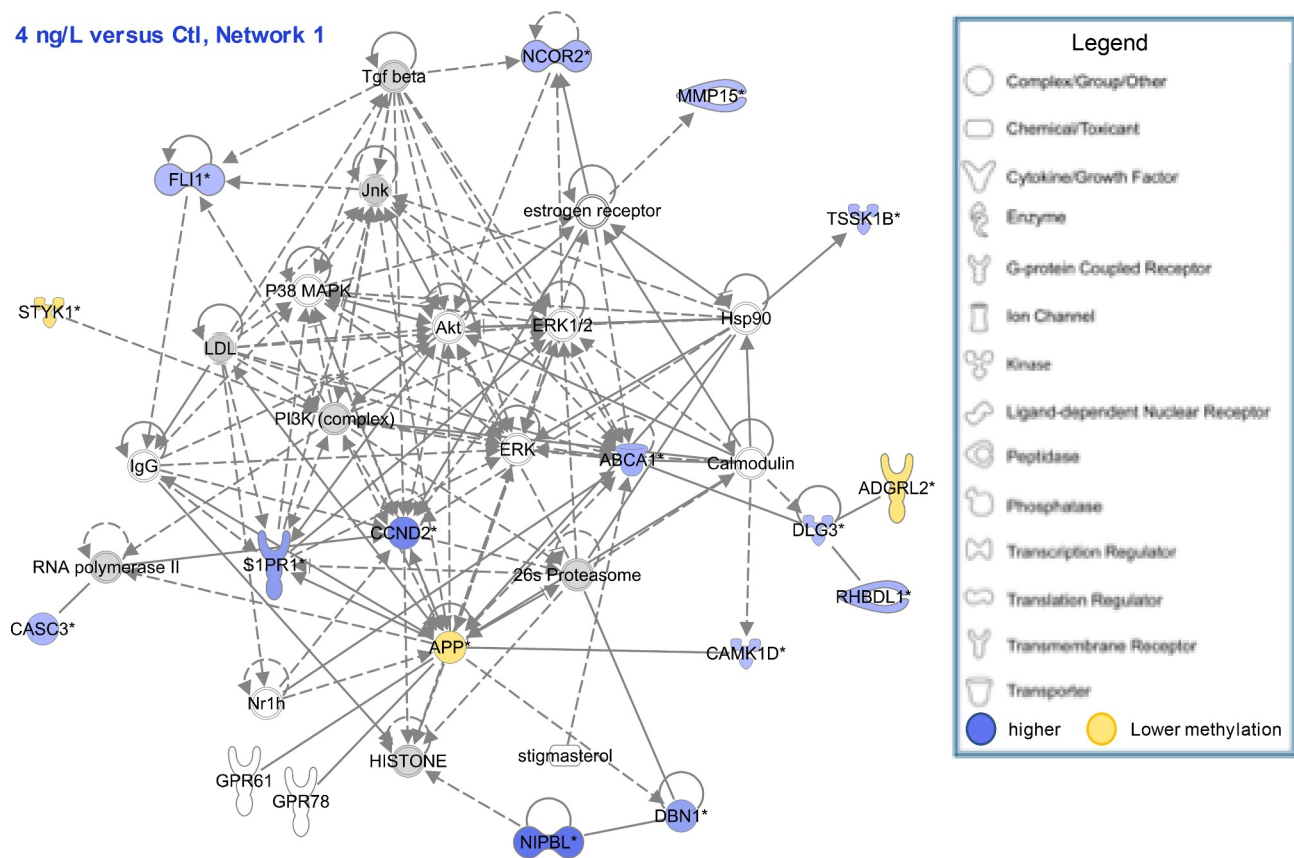
**Figure 4.** Network of genes corresponding to significant DMFs (located in promoter region or in gene body) at 4 ng/L compared to control. Blue: hypermethylation at 4 ng/L compared to control, Yellow: hypomethylation at 4 ng/L compared to control.

in all reported vertebrates, a larger proportion of fragments present a methylation level higher than 80% in fish than in mammals (67% of fragments in our study) [69].

Regarding the mangrove rivulus, even if the majority of promoter fragments were lowly methylated, a significant proportion were highly methylated (38%), which indicates an elevated range of the methylation level in promoter regions, and consequently a regulatory activity of gene expression. In contrast, other genomic regions presented a narrower range in methylation levels, with the majority of fragments being highly methylated. When compared to zebrafish, an interesting difference appeared regarding the proportion of analysed fragments located in gene bodies. While it reached 47% in mangrove rivulus, zebrafish has only 25% of CpGs in introns and exons. Conversely, the proportion of analysed intergenic fragments is higher in zebrafish (63%) than in mangrove rivulus (37% including intergenic and upstream regions). A possible explanation could

reside in the higher proportion of transposable elements (TEs) in zebrafish genome. There is increasing evidence that TEs are important targets of DNA methylation to silence their expression and consequently show a high level of CpG methylation. The zebrafish genome is rich in TEs as 52% of its genomic sequence is composed of the different classes of TEs [72,73], compared to 27.3% in the mangrove rivulus [59]. This difference could explain the relatively lower proportion of analysed fragments located in intergenic regions in the mangrove rivulus compared to zebrafish. When considering other fish species, the part of TEs in genomes is very variable and values as low as 5% have been reported in *Tetraodon nigroviridis* [73]. This variability of TE amount is explained by the large range of genome size among fish, and Chalopin et al. [72] reported a positive correlation between TE content and genome size in teleosts species, which also applies to the mangrove rivulus. Contrary to what was advanced by Zhang et al. [69], it is therefore unlikely that the pattern of high

Table 7. DNA methylation levels of fragments related to genes involved in growth, steroidogenesis or reproduction.

| ID | Len | CpGs | Description | DMF | % Methylation | | | Methylation difference (%) | | | P-value | Test |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Ctl | 4 ng/L | 120 ng/L | 4 – Ctl | 120 – Ctl | 120–4 | | |
| XP_017294595.1 | 85 | 9 | Oestrogen receptor beta-like | Gene body | 94.71 | 96.28 | 95.30 | 1.57 | 0.59 | −0.98 | 0.07493 | F(2.12) = 3.24 |
| XP_017267390.1 | 54 | 8 | Oestrogen receptor beta-like | Promoter | 97.06 | 97.82 | 96.54 | 0.76 | −0.52 | −1.28 | 0.13056 | F(2.12) = 2.42 |
| XP_017267390.1 | 72 | 6 | Oestrogen receptor beta-like | Gene body | 94.50 | 95.72 | 95.28 | 1.22 | 0.78 | −0.44 | 0.14477 | F(2.12) = 2.28 |
| XP_017267390.1 | 101 | 6 | Oestrogen receptor beta-like | Gene body | 29.99 | 33.30 | 29.14 | 3.31 | −0.85 | −4.16 | 0.39458 | F(2.12) = 1.01 |
| XP_017260454.1 | 104 | 4 | Oestrogen-related receptor gamma-like | Gene body | 85.19 | 89.50 | 86.40 | 4.31 | 1.21 | −3.10 | 0.0425 | F(2.12) = 4.16 |
| XP_017291458.1 | 96 | 4 | Oestrogen-related receptor gamma-like | Gene body | 81.14 | 85.59 | 81.55 | 4.45 | 0.41 | −4.04 | 0.0446 | F(2.11) = 4.18 |
| XP_017262455.1 | 153 | 5 | Vitellogenin-1-like | Gene body | 91.30 | 89.95 | 81.87 | −1.35 | −9.43 | −8.08 | 0.0050 | F(2.10) = 9.45 |
| XP_017262455.1 | 89 | 6 | Vitellogenin-1-like | Gene body | 90.45 | 93.45 | 89.99 | 3.00 | −0.46 | −3.46 | 0.0105 | F(2.12) = 6.82 |
| XP_017262455.1 | 75 | 6 | Vitellogenin-1-like | Gene body | 92.58 | 92.50 | 92.81 | −0.08 | 0.23 | 0.31 | 0.6645 | F(2.12) = 0.42 |
| XP_017295559.1 | 72 | 6 | Aromatase | Gene body | 87.84 | 87.62 | 84.35 | −0.22 | −3.49 | −3.27 | 0.0093 | F(2.12) = 7.08 |
| XP_017295559.1 | 137 | 12 | Aromatase | Gene body | 92.69 | 93.60 | 94.71 | 0.91 | 2.02 | 1.11 | 0.1607 | F(2.12) = 2.14 |
| XP_017267185.1 | 131 | 5 | Insulin-like growth factor I | Gene body | 4.51 | 3.68 | 2.51 | −0.83 | −2.00 | −1.17 | 0.2681 | F(2.11) = 1.49 |
| XP_017267185.1 | 150 | 9 | insulin-like growth factor I | Promoter | 91.94 | 93.76 | 94.54 | 1.82 | 2.60 | 0.78 | 0.3066 | F(2.12) = 1.31 |
| XP_017259346.1 | 44 | 3 | Insulin-like growth factor II | Gene body | 95.35 | 96.63 | 93.88 | 1.28 | −1.47 | −2.75 | 0.1261 | F(2.12) = 2.47 |
| XP_017274518.1 | 64 | 4 | Insulin-like growth factor II | Gene body | 5.53 | 4.72 | 5.07 | −0.81 | −0.46 | 0.35 | 0.7446 | F(2.12) = 0.30 |
| XP_017274518.1 | 44 | 3 | Insulin-like growth factor II | Gene body | 4.86 | 5.91 | 4.55 | 1.05 | −0.31 | −1.36 | 0.1565 | F(2.12) = 2.17 |
| XP_017284163.1 | 86 | 7 | Insulin-like growth factor 1 receptor | Gene body | 92.18 | 94.49 | 91.72 | 2.31 | −0.46 | −2.77 | 0.0390 | F(2.12) = 4.30 |
| XP_017284163.1 | 153 | 13 | Insulin-like growth factor 1 receptor | Gene body | 94.08 | 94.82 | 91.76 | 0.74 | −2.32 | −3.06 | 0.0701 | F(2.12) = 3.34 |
| XP_017261561.1 | 164 | 13 | Insulin-like growth factor 1 receptor | Gene body | 91.60 | 95.83 | 92.94 | 4.23 | 1.34 | −2.89 | 0.0823 | F(2.10) = 3.24 |
| XP_017284163.1 | 138 | 10 | Insulin-like growth factor 1 receptor | Gene body | 88.79 | 92.11 | 90.51 | 3.32 | 1.72 | −1.60 | 0.1541 | F(2.12) = 2.19 |
| XP_017279935.1 | 55 | 7 | Insulin-like growth factor 2 mRNA-binding protein 2 | Gene body | 95.81 | 97.37 | 95.80 | 1.56 | −0.01 | −1.57 | 0.0029 | F(2.12) = 9.85 |
| XP_017279935.1 | 98 | 10 | insulin-like growth factor 2 mRNA-binding protein 2 | Gene body | 90.32 | 93.25 | 88.80 | 2.93 | −1.52 | −4.45 | 0.0129 | F(2.12) = 6.40 |
| XP_017279586.1 | 97 | 6 | Insulin-like growth factor 2 mRNA-binding protein 3 | Gene body | 91.79 | 93.45 | 90.36 | 1.66 | −1.43 | −3.09 | 0.0669 | F(2.12) = 3.42 |
| XP_017292570.1 | 75 | 10 | Insulin-like growth factor-binding protein 1 | Promoter | 1.26 | 1.66 | 1.19 | 0.40 | −0.07 | −0.47 | 0.1667 | F(2.12) = 2.09 |

global methylation observed in fish RR genomes is the result of an enrichment in TEs. Nevertheless, the role of DNA methylation in TEs should be further investigated, including a precise look at the different types of TEs. For example, the rolling-circle transposons, known as Helitrons, are particularly abundant in the mangrove rivulus genome (0.65%) compared to other fish species including its phylogenetic relatives [Japanese medaka (0.03%) and Turquoise killifish (0.06%)].

From all studied fish species, only zebrafish showed a higher level of Helitrons at 1.5% [73]. It is known that TEs in general, and the Helitron subfamily in particular, are mediator of host genes regulation and contribute to genome evolution and adaptation [74]. It has been suggested that the mobilization of TEs and the changes in epigenetic pattern could be important in the process of organism's rapid adaptation to environmental changes [75]. The exact role of these elements

and their methylation level in the mangrove rivulus is a promising perspective to investigate how this species adapt and evolve using a mixed-mating reproduction system and consequently a low genetic diversity within lineages.

## Effects of EE2 on the liver methylome

Growing evidence suggests that EDCs can modify DNA methylation and consequently gene expression pattern. Recently, experiments on zebrafish embryos clearly established a link between exposure to bisphenol A, DNA methylation and behavioural impairments [76]. Data also showed that EE2 can modify the methylation status of target genes [35]. In a previous study, we showed that the liver proteome of adult mangrove rivulus was impaired by exposure to EE2 during early life stages (ELS) [49]. It has been hypothesized that exposure to EE2 during ELS can impair DNA methylation and induce long-term consequences later in life as observed on the proteome and on the phenotype [48]. To test this hypothesis, we aimed at detecting potential long-term changes in the hepatic DNA methylation of 168 dph adults following a 28-day early-life exposure of mangrove rivulus hatchlings. We obtained a total of 146 DMFs among all group comparisons, from which a majority were observed between the two EE2 treatments and between 4 ng/L exposed individuals and Ctl. The larger observed impacts of the lowest EE2 concentration on DNA methylation is in accordance with a non-monotonic dose–response relationship often reported for exposure to endocrine disruptors. Hormones are effective at very low concentration, having a high affinity for their receptors, and so are the endocrine disrupting chemicals [77]. The present results are reinforced by a similar response detected in brain and liver protein expression profiles of the mangrove rivulus exposed in the same conditions to EE2 [49]. To our knowledge, only one other study investigated the long-term effects of early-life exposure to chemicals on fish methylome [78]. This study investigated the effects of two compounds, mono(2-ethylhexyl) phthalate (MEHP, 30 μM) and 5-azacytidine (5AC, 10 μM, an inhibitor of DNMT1) in zebrafish. They found that most methylation changes occurred at conserved non-genic regions with cis-regulatory functions (i.e., enhancers, silencers, transcription factor binding sites).

Among the differentially methylated fragments (DMFs) found in adult mangrove rivulus liver following early-life exposure to EE2, none was significantly affected in all three group comparisons (Ctl vs 4 ng/L; Ctl vs 120 ng/L; 4 vs 120 ng/L). Six DMFs were significant in two out of three comparisons, of which three were located in gene bodies. A DMF corresponding to *ATP-binding cassette sub-family A member 1* (ABCA1) was hypermethylated at both 4 ng/L and 120 ng/L compared to Ctl. ABCA1 is a known oestrogen-responsive gene [79]. It was included in networks in both comparisons and appeared to play a central role in the response to EE2, showing interactions with the oestrogen receptor and several other molecules (Figure 4). It is a major regulator of cholesterol and phospholipid homoeostasis by regulating efflux of cholesterol and phospholipids outside of the cell, that are then taken up by apolipoproteins A1 and E to form high density lipoprotein (HDL) molecules. Namely, oestradiol exposure of smooth muscle cells has been shown to enhance cholesterol efflux to APOA1 and HDL, which was associated to ABCA1 overexpression via ESR2 and liver X receptor (LXR) activation [80]. We suggest that the alteration of methylation level of ABCA1 gene might be related to the altered abundances of liver apolipoproteins and other downstream proteins involved in lipid metabolism previously reported [49]. The fact that both concentrations of EE2 impacted this gene methylation level indicates that this mechanism might be a general mode of action of EE2, and could induce effects at environmental concentrations as low as 4 ng/L.

Two DMFs were significantly hypermethylated at 4 ng/L in comparison to both Ctl and 120 ng/L treatments: the *calcium/calmodulin-dependent protein kinase type 1D* (CAMK1D), and the *testis-specific serine/threonine-protein kinase* (TSSK1B). Both DMFs were included in the first network of 4 ng/L vs Ctl and 120 vs 4 ng/L comparisons. CAMK1D is a protein kinase that participates in the calcium-regulated calmodulin-dependent kinase cascade, regulating calcium-mediated granulocyte function and respiratory burst and

activating the transcription factor CREB1. According to Figure 4, CAMK1D interacts not only with calmodulin, but also with the amyloid precursor APP, which can be induced by oestrogens through extracellular-regulated kinase 1 and 2 (ERK1/2) phosphorylation [81]. Finally, TSSK1B is essential for spermatid development and spermatogenesis. While TSSK1B is likely not expressed in the liver, it can be possible that an epigenetic alteration of this gene could be found in the ovotestis following EE2 exposure, with potential consequences on reproduction. Although no study to date has linked EE2 and TSSK1B, a study in rats has shown that bisphenol A exposure decreases the expression of TSSK1B mRNA [82].

We must nevertheless be careful when interpreting the changes of methylation in these 3 DMFs. First, the observed differences in methylation, if significant, ranged between 10.3 and 14.6%. Further studies should determine whether these differences have a biological impact on gene expression. Moreover, these three DMFs occurred in gene bodies, as did the majority of DMFs in this study (and MSPI fragments). While methylation of DNA at promoters is well known to suppress transcription, the role of DNA methylation in gene bodies is more difficult to interpret [83]. Gene body methylation is conserved among plants and animals, with a general preference for exons [84]. Potential roles include the stimulation or inhibition of transcription elongation [85] and the regulation of splicing [86]. In mammals, most genes possess alternative transcription start sites that can be located within the gene bodies, complicating the link between methylation and expression [86]. Nevertheless, it is now recognized that gene body methylation is positively correlated with transcriptional activity in most species [12] and we should not disregard the possible implication of these observed changes in methylation after exposure to EE2.

Beside significant changes of DNA methylation within gene bodies, we also reported several DMFs belonging to intergenic regions. Interpreting alterations in DNA methylation in these regions is challenging as the gene annotation may not directly apply. Intergenic regions are typically highly methylated in all vertebrates and bear a conserved role of maintaining genome stability, by preventing the transposition of repetitive TEs and silencing cryptic promoters and cryptic splice sites [87]. Moreover, TEs possess strong promoters that may be able to transcribe neighbouring genes (transcriptional interference). Finally, a compact chromatin state in these regions can prevent recombination. Most (>90%) methylcytosines occur in transposons, and low methylation of these regions could reduce genomic stability [88]. As above mentioned, the diversity and composition of TEs in the mangrove rivulus could contribute to the genetic recombination and evolutionary adaptation in this self-fertilizing species [59]. To gain further insights into the role of intergenic methylation, future research should investigate the overlapping of DMFs with potential TEs.

In addition, we pointed out the very low proportion of DMFs classified as lowly methylated fragment (< 20% methylation). The *nipped-B-like protein B* (NIPBL) is the exception as it showed 19.6% methylation in Ctl but increased to 41.5% in fish exposed to 4 ng/L EE2. It is remarkable that this fragment is also the one showing the highest effect size of our study (+ 21.9%) and is located in the promoter region. Few more lowly methylated fragments were significant when comparing the 2 treatments, but all others were either highly methylated or showed an intermediary methylation level. Two explanations can be advanced. First, the lowly methylated fragments were about 1/3 less numerous than the highly methylated ones in the RR genome as reported in Figure 2. This is related to the proportionally low occurrence of fragments located in the promoter region, which have the lowest methylation level. Second, we also showed that lowly methylated fragments have a 10-fold higher coefficient of variability (33% vs 3% in highly methylated fragments). It results that the statistical power to detect changes of lowly methylated fragments is lower and that we can consequently report only DMFs with a high effect size, such as NIPBL. To overstep this limitation, one should include a higher number of replicates in future RRBS analyses. It is even more important for this kind of studies for which effects are searched long after the end of the exposure, which decrease the probability to observe high

effect sizes. Nevertheless, our study is the first to report the effect of an endocrine disruptor exposure on the methylation level of NIPBL promoter region. This effect, which was significant while the exposure had stopped 140 days earlier, could have serious long-term consequences on adult fish. NIPBL is associated to cohesin in order to mediate chromatin cohesion, important for mitosis. This complex plays an important role in histone acetylation, chromosome architecture and therefore in regulation of gene expression [89,90]. In human, a mutation of NIPBL is associated to the 'Cornelia de Lange' syndrome, a developmental disorder [91]. In zebrafish, the zygotic genome activation is facilitated by cohesin and was a key for the embryogenesis [92]. Intriguingly, the cohesin complex has been proposed to modulate oestrogen-dependent gene transcription and a link has been established with breast cancer [93]. Our results suggest that NIPBL hypermethylation could be a mode of action of EE2 exposure and could potentially modify gene regulation. Further studies should investigate the long-term links between the change in NIPBL methylation level and the overall changes in gene expression in adults, long after the exposure to EE2.

Additionally, we reported the methylation level of fragments corresponding to oestrogen receptors, the aromatase, vitellogenin and several genes involved in the somatotropic axis, which is the main regulator of growth in vertebrates. Most differences in mean methylation of these fragments between experimental conditions were comprised between 0 and 5% and were not significant. Consequently, we cannot link the previously observed effects of EE2 exposure on growth, steroid hormone levels and reproduction [48] with the DNA methylation of these genes. However, one fragment located in gene body of vitellogenin 1 was 9.4% less methylated at 120 ng/L compared to control, while there was no significant effect at 4 ng/L. Vitellogenin is the precursor of egg yolk proteins, produced in the liver of sexually mature female fish in response to endogenous oestrogens. It is well known that it can also be induced by exogenous oestrogens such as EE2 [94]. Stromqvist et al. [35] reported a significant hypomethylation of three CpGs sites located between two potential predicted oestrogen responsive elements (EREs)

upstream of the first vitellogenin exon, in the liver of adult zebrafish after exposure to 100 ng/L EE2 during 14 days. In mangrove rivulus, further studies should investigate whether the methylation changes of the vitellogenin 1 fragment could influence vitellogenin expression and further the reproduction of hermaphrodites. The different effect between the two tested concentrations can be associated with the reported impact on reproduction by Voisin et al. [48] who showed fewer laid eggs by hermaphrodites after exposure to 4 ng/L, but not to 120 ng/L. The implication of some compensatory mechanisms at the highest tested concentration involving DNA methylation of vitellogenin 1 CpGs might be hypothesized.

As previously mentioned, DMFs reported as being significant in the present study mostly showed different methylation level between 10 and 20%. These effect sizes could be considered as low, mostly if we compare with experiments comparing normal and pathological tissue such as tumour [95] which can be over 20%. Berbel-Filho et al. [46] also reported higher effect sizes on the rivulus reared during 10 months in differentially enriched environments (up to 48.9% of hypermethylation). If it is necessary to question the biological relevance of changes in DNA methylation below 20%, it is not surprising to have low effect sizes using an experimental design where fish were reared in the same environment for most of their life after an initial exposure during 28 days. This is a similar approach as in studies searching for the role of pregnancy and early-life exposures on later-in-life health outcomes in human. These studies show small-magnitude epigenetic effect sizes below 5%, while it can still persist and be replicated across populations and time and be correlated with gene expression [96]. The individuals used in the present study being issued from the same isogenic lineage, the confounding factor of genetic variability is reduced, which may help to reveal more subtle difference in DNA methylation level, and stress the usefulness of this species in epigenetic studies as model to investigate the developmental origin of health and diseases.

Little correspondence was found between the proteome previously published [49] and the epigenome (Supplementary Figure 2). Although the

number of identified genes was about 10 x greater in the methylome analysis, only about half of the total identified proteins in the liver was found in the methylome analysis. Similarly, we could retrieve the methylation level of fragments belonging to approximately half (52%) of the differentially abundant proteins. None of these differentially abundant proteins could be related to a change in DNA methylation level of the promotor region. Conversely, only five of the significantly DMFs were also identified, but not significant, in the proteomic analysis in liver: *G1/S-specific cyclin-D2-like* (CCND2, promoter region, significant between 4 ng/L and Ctl), *teneurin-1-like* (TENM1, gene body, significant between 120 ng/L and Ctl), *protein NDRG1-like* (NDRG1, promoter, significant between 120 ng/L and Ctl), *fibrinogen alpha chain* (FGA, intergenic, significant between 120 ng/L and Ctl), and *death-associated protein 1* (DAP, gene body, significant between 4 and 120 ng/L). Comparison between DNA methylation and protein expression pattern is rare but is fully justified. The proteome is the main functional product of gene expression and is closely related to the molecular phenotype [97]. As such, the proteomic level can generate new hypothesis on the modes of action of xenobiotics and on the adaptive responses of organisms [98]. Epigenetics being at the interface between the genotype and the environment, a more systematic link with the proteome would improve the understanding of the adverse outcome pathways (AOPs) and the outcomes in terms of adaptation to the environment. However, it is challenging to correlate these two levels, because of technical variability, and because differentially abundant proteins are not necessarily regulated by differential methylation, but rather the end product of upstream regulators, themselves differentially methylated. Also, protein abundance results not only from transcription, which can be influenced by DNA methylation, but also from mRNA stability, alternative splicing, translation, protein turnover and post-translational modifications. Nevertheless, comparison of proteome and methylation data can be eased by the analysis of gene ontology. Some biological processes were identified in both datasets, such as lipid metabolism, inflammation, connective tissue development and molecular transport,

which makes their long-term impairment plausible mechanisms of EE2 toxicity.

## Conclusions

In conclusion, our study provides further evidence on the capacity of EDCs to alter the methylome and shows that these changes can be apparent several months after the exposure, supporting the hypothesis of possible long-term modulation of gene expression through epigenetics. It also confirms the non-monotonic response to EDCs as most significant effects were observed at the environmental concentration of 4 ng/L. The DMFs belonging to promoter and gene bodies revealed networks involved in several biological functions potentially regulated by oestrogens, including lipid metabolism, cellular processes (death and survival), connective tissue function, molecular transport and inflammation, many of which were also identified in a previous proteomic analysis. Importantly, we reported that methylation of *nipped-B-like protein B* (NIPBL) might be an important mode of action of EE2 and, as involved in chromosome organization and gene expression, could have long-term impact on gene regulation and on organism's general functions. To gain a more comprehensive view of the significance of DMFs in promoter, gene body, upstream and intergenic regions, future studies should investigate the overlapping of the DMFs with potential enhancers, insulators, transcription start sites, TEs, oestrogen-responsive elements and other transcription factor binding sites that may interact with oestrogen receptors. Among them, TEs, and more particularly helitrons, are potentially important targets of environmental stressors that could be involved in the response of organisms to environmental changes. Despite its limitations, RRBS has proved to be an efficient method to investigate effects of EE2 on DNA methylation in the mangrove rivulus. This self-fertilizing fish should be further considered as a model species to investigate the interplay between environmental stressors, epigenetics and adaptation. The possibility to naturally discard the confounding factor of genetic variability makes this species a top model to better understand the roles of epigenetics in the long-term response to environmental xenobiotics

and in the developmental origin of health and diseases.

## Data accessibility statement

All the data generated by the RRBS workflow will be submitted to Dryad repository and will be publicly accessible. The BAM files after Bismark alignment and the excel sheet containing the values for all analysed fragments have already been submitted and are available at this link: https://datadryad.org/stash/share/k7Lq6mHbxBMF5BHdk8cOp1EFGSGVK3tfJoBx4tPYnzc.

The DOI is: https://doi.org/10.5061/dryad.v41ns1rsv

The Fastq files will be submitted when a decision will have been given for the acceptance of the manuscript.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Author contributions

Anne-Sophie Voisin: investigation, manuscript writing and editing, experimental design, data acquisition and analysis

Victoria Suarez Ulloa: data analysis, manuscript editing

Peter Stockwell: bioinformatic data analyses

Aniruddha Chatterjee: supervision, methodology, RRBS analyses, manuscript editing

Frédéric Silvestre: investigation, supervision, funding acquisition, manuscript writing and editing, experimental design, submission and revision

The authors report no conflict of interest

## ORCID

Aniruddha Chatterjee 🔾 http://orcid.org/0000-0001-7276-2248

Frédéric Silvestre 🔾 http://orcid.org/0000-0001-5889-7728

## References

[1] Barker DJ. The developmental origins of adult disease. J Am Coll Nutr. 2004;23:588S–595S.

[2] Brockmeier EK, Hodges G, Hutchinson TH, et al. The role of omics in the application of adverse outcome pathways for chemical risk assessment. Toxicol Sci. 2017;158:252–262.

[3] Braun JM. Early-life exposure to EDCs: role in child-hood obesity and neurodevelopment. Nat Rev Endocrinol. 2017;13:161.

[4] Heindel JJ, Newbold R, Schug TT. Endocrine disruptors and obesity. Nat Rev Endocrinol. 2015;11:653–661.

[5] Kajta M, Wójtowicz AK. Impact of endocrine-disrupting chemicals on neural development and the onset of neurological disorders. Pharmacol Rep. 2013;65:1632–1639.

[6] Kuo CH, Yang SN, Kuo PL, et al. Immunomodulatory effects of environmental endocrine disrupting chemicals. Kaohsiung J Med Sci. 2012;28:S37–42.

[7] Clouzot L, Marrot B, Doumenq P, et al. 17α-Ethinylestradiol: an endocrine disrupter of great concern. Analytical methods and removal processes applied to water purification. A review. Environ Prog. 2008;27:383–396.

[8] Matthiessen P, Wheeler JR, Weltje L. A review of the evidence for endocrine disrupting effects of current-use chemicals on wildlife populations. Crit Rev Toxicol. 2018;48:195–216.

[9] Aris AZ, Shamsuddin AS, Praveena SM. Occurrence of 17α-ethynylestradiol (EE2) in the environment and effect on exposed biota: a review. Environ Int. 2014;69:104–119.

[10] Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet. 2010;11:204–220.

[11] Lister R, Ecker JR. Finding the fifth base: genome-wide sequencing of cytosine methylation. Genome Res. 2009;19:959–966.

[12] De Mendoza A, Lister R, Bogdanovic O. Evolution of DNA methylome diversity in eukaryotes. J Mol Biol. 2020;432:1687–1705.

[13] Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology. 2013;38:23–38.

[14] Edwards JR, Yarychkivska O, Boulard M, et al. DNA methylation and DNA methyltransferases. Epigenetics Chromatin. 2017;10. DOI:10.1186/s13072-017-0130-8

[15] Guo H, Zhu P, Yan L, et al. The DNA methylation landscape of human early embryos. Nature. 2014;511:606–610.

[16] Smith ZD, Chan MM, Humm KC, et al. DNA methylation dynamics of the human preimplantation embryo. Nature. 2014;511:611–615.

[17] Dorts J, Falisse E, Schoofs E, et al. DNA methyltransferases and stress-related genes expression in zebrafish larvae after exposure to heat and copper during reprogramming of DNA methylation. Sci Rep. 2016;6:34254.

[18] Eid A, Zawia N. Consequences of lead exposure, and it's emerging role as an epigenetic modifier in the aging brain. NeuroToxicology. 2016;56:254–261.

[19] Gore AC, Walker DM, Zama AM, et al. Early life exposure to endocrine-disrupting chemicals causes life-long molecular reprogramming of the hypothalamus and premature reproductive aging. Mol Endocrinol. 2011;25:2157–2168.

[20] Goyal D, Limesand SW, Goyal R. Epigenetic responses and the developmental origins of health and disease. J Endocrinol. 2019;242:T105–T119.

[21] Stel J, Legler J. The role of epigenetics in the latent effects of early life exposure to obesogenic endocrine disrupting chemicals. Endocrinology. 2015;156:3466–3472.

[22] Wadhwa PD, Buss C, Entringer S, et al. Developmental origins of health and disease: brief history of the approach and current focus on epigenetic mechanisms. Semin Reprod Med. 2009;27:358–368.

[23] Barouki R, Melén E, Herceg Z, et al. Epigenetics as a mechanism linking developmental exposures to long-term toxicity. Environ Int. 2018;114:77–86.

[24] Mirbahai L, Chipman JK. Epigenetic memory of environmental organisms: a reflection of lifetime stressor exposures. Mutat Res Genet Toxicol Environ Mutagen. 2014;764-765:10–17.

[25] Brander SM, Biales AD, Connon RE. The role of epigenomics in aquatic toxicology. Environ Toxicol Chem. 2017;36:2565–2573.

[26] Vandegehuchte MB, Janssen CR. Epigenetics and its implications for ecotoxicology. Ecotoxicology. 2011;20:607–624.

[27] Vandegehuchte MB, Janssen CR. Epigenetics in an ecotoxicological context. Mutat Res Genet Toxicol Environ Mutagen. 2014;764-765:36–45.

[28] Walker CL. Minireview: epigenomic plasticity and vulnerability to EDC exposures. Mol Endocrinol. 2016;30:848–855.

[29] Zhang X, Ho SM. Epigenetics meets endocrinology. J Mol Endocrinol. 2011;46:R11–32.

[30] Shi JF, Li XJ, Si XX, et al. ERα positively regulated DNMT1 expression by binding to the gene promoter region in human breast cancer MCF-7 cells. Biochem Biophys Res Commun. 2012;427:47–53.

[31] Lee DH, Jacobs DR, Porta M. Hypothesis: a unifying mechanism for nutrition and chemicals as lifelong modulators of DNA hypomethylation. Environ Health Perspect. 2009;117:1799–1802.

[32] Anderson AM, Carter KW, Anderson D, et al. Coexpression of nuclear receptors and histone methylation modifying genes in the testis: implications for endocrine disruptor modes of action. PLoS One. 2012;7:e34158.

[33] Casati L, Sendra R, Sibilia V, et al. Endocrine disrupters: the new players able to affect the epigenome. Front Cell Dev Biol. 2015;3:37.

[34] Liu Y, Duong W, Krawczyk C, et al. Oestrogen receptor β regulates epigenetic patterns at specific genomic loci through interaction with thymine DNA glycosylase. Epigenetics Chromatin. 2016;9:7.

[35] Strömqvist M, Tooke N, Brunström B. DNA methylation levels in the 5' flanking region of the vitellogenin I gene in liver and brain of adult zebrafish (Danio rerio)–sex and tissue differences and effects of 17alpha-ethinylestradiol exposure. Aquat Toxicol. 2010;98:275–281.

[36] Mirbahai L, Yin G, Bignell JP, et al. DNA methylation in liver tumorigenesis in fish from the environment. Epigenetics. 2011;6:1319–1333.

[37] Falisse E, Ducos B, Stockwell PA, et al. DNA methylation and gene expression alterations in zebrafish early-life stages exposed to the antibacterial agent triclosan. Environ Pollut. 2018;243:1867–1877.

[38] Heard E, Martienssen RA. Transgenerational epigenetic inheritance: myths and mechanisms. Cell. 2014;157:95–109.

[39] Avise JC, Tatarenkov A. Population genetics and evolution of the mangrove rivulus Kryptolebias marmoratus, the world's only self-fertilizing hermaphroditic vertebrate. J Fish Biol. 2015;87:519–538.

[40] Mackiewicz M, Tatarenkov A, Perry A, et al. Microsatellite documentation of male-mediated outcrossing between inbred laboratory strains of the self-fertilizing mangrove killifish (Kryptolebias Marmoratus). J Hered. 2006;97:508–513.

[41] Ellison A, Rodríguez López CM, Moran P, et al. Epigenetic regulation of sex ratios may explain natural variation in self-fertilization rates. Proc Biol Sci. 2015;282. DOI:10.1098/rspb.2015.1900

[42] Fellous A, Voisin A-S, Labed-veydert T, et al. DNA methylation in adults and during development of the self- fertilizing mangrove rivulus, Kryptolebias marmoratus. Ecol Evol. 2018;8:6016–6033.

[43] Fellous A, Earley RL, Silvestre F. Identification and expression of mangrove rivulus (Kryptolebias marmoratus) histone deacetylase (HDAC) and lysine acetyltransferase (KAT) genes. Gene. 2019;691:56–69.

[44] Fellous A, Earley RL, Silvestre F. The Kdm/Kmt gene families in the self-fertilizing mangrove rivulus fish, Kryptolebias marmoratus, suggest involvement of histone methylation machinery in development and reproduction. Gene. 2019;687:173–187.

[45] Berbel-Filho WM, De Leaniz CG, Moran P, et al. Local parasite pressures and host genotype modulate epigenetic diversity in a mixed - mating fish. Ecol Evol. 2019;9:8736–8748.

[46] Berbel-Filho WM, Rodríguez-Barreto D, Berry N, et al. Contrasting DNA methylation responses of inbred fish

lines to different rearing environments. Epigenetics. 2019;14:939–948.

[47] Carion A, Markey A, Hétru J, et al. Behavior and gene expression in the brain of adult self-fertilizing mangrove rivulus fish (Kryptolebias marmoratus) after early life exposure to the neurotoxin β-N-methylamino-L-alanine (BMAA). Neurotoxicology. 2020;79:110–121.

[48] Voisin A-S, Fellous A, Earley RL, et al. Delayed impacts of developmental exposure to 17-α-ethinylestradiol in the self-fertilizing fish Kryptolebias marmoratus. Aquatic Toxicol. 2016;180:247–257.

[49] Voisin A-S, Kültz D, Silvestre F. Early-life exposure to the endocrine disruptor 17-α-ethinylestradiol induces delayed effects in adult brain, liver and ovotestis proteomes of a self-fertilizing fish. J Proteomics. 2018;1–13.

[50] Foulds CE, Treviño LS, York B, et al. Endocrine-disrupting chemicals and fatty liver disease. Nat Rev Endocrinol. 2017;13:445–457.

[51] De Wit M, Keil D, Remmerie N, et al. Molecular targets of TBBPA in zebrafish analysed through integration of genomic and proteomic approaches. Chemosphere. 2008;74:96–105.

[52] Feswick A, Munkittrick KR, Martyniuk CJ. Estrogen-responsive gene networks in the teleost liver: what are the key molecular indicators. Environ Toxicol Pharmacol. 2017;56:366–374.

[53] Humble JL, Hands E, Saaristo M, et al. Characterisation of genes transcriptionally upregulated in the liver of sand goby (Pomatoschistus minutus) by 17α-ethinyloestradiol: identification of distinct vitellogenin and zona radiata protein transcripts. Chemosphere. 2013;90:2722–2729.

[54] Chatterjee A, Ozaki Y, Stockwell PA, et al. Mapping the zebrafish brain methylome using reduced representation bisulfite sequencing. Epigenetics. 2013;8:979–989.

[55] Chatterjee A, Stockwell PA, Horsfield JA, et al. Base-resolution DNA methylation landscape of zebrafish brain and liver. Genom Data. 2014;2:342–344.

[56] Chatterjee A, Stockwell PA, Rodger EJ, et al. Comparison of alignment software for genome-wide bisulphite sequence data. Nucleic Acids Res. 2012;40:e79.

[57] Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics. 2011;27:1571–1572.

[58] Stockwell PA, Chatterjee A, Rodger EJ, et al. DMAP: differential methylation analysis package for RRBS and WGBS data. Bioinformatics. 2014;30:1814–1822.

[59] Rhee JS, Choi BS, Kim J, et al. Diversity, distribution, and significance of transposable elements in the genome of the only selfing hermaphroditic vertebrate Kryptolebias marmoratus. Sci Rep. 2017;7:40121.

[60] Bock C, Beerman I, Lien WH, et al. DNA methylation dynamics during in vivo differentiation of blood and skin stem cells. Mol Cell. 2012;47:633–647.

[61] Wreczycka K, Gosdschan A, Yusuf D, et al. Strategies for analyzing bisulfite sequencing data. J Biotechnol. 2017;261:105–115.

[62] Meer A, Gnirke A, Bell GW, et al. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic Acids Res. 2005;33:5868–5877.

[63] Chatterjee A, Rodger EJ, Stockwell PA, et al. Technical considerations for reduced representation bisulfite sequencing with multiplexed libraries. J Biomed Biotechnol. 2012;2012:741542.

[64] Smith ZD, Gu H, Bock C, et al. High-throughput bisulfite sequencing in mammalian genomes. Methods. 2009;48:226–232.

[65] Metzger DCH, Schulte PM. Persistent and plastic effects of temperature on DNA methylation across the genome of threespine stickleback (Gasterosteus aculeatus). Proc Biol Sci. 2017;284.

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

[67] Baerwald MR, Meek MH, Stephens MR, et al. Migration-related phenotypic divergence is associated with epigenetic modifications in rainbow trout. Mol Ecol. 2016;25:1785–1800.

[68] Hu J, Pérez-Jvostov F, Blondel L, et al. Genome-wide DNA methylation signatures of infection status in Trinidadian guppies (Poecilia reticulata). Mol Ecol. 2018;27:3087–3102.

[69] Zhang C, Hoshida Y, Sadler KC. Comparative epigenomic profiling of the DNA methylome in mouse and zebrafish uncovers high interspecies divergence. Front Genet. 2016;7:110.

[70] Chatterjee A, Rodger EJ, Stockwell PA, et al. Generating multiple base-resolution DNA methylomes using reduced representation bisulfite sequencing. In: Seymour G, Cullinan M, Heng N, editors. Oral biology. New York, N.Y: Humana Press; 2017. p. 279–298.

[71] Chatterjee A, Rodger EJ, Morison IM, et al. Tools and strategies for analysis of genome-wide and gene-specific DNA methylation patterns. In: Seymour G, Cullinan M, Heng N, editors. Oral biology. New York, N.Y: Humana Press; 2017. p. 249–277.

[72] Chalopin D, Naville M, Plard F, et al. Comparative analysis of transposable elements highlights mobilome diversity and evolution in vertebrates. Genome Biol Evol. 2015;7:567–580.

[73] Shao F, Han M, Peng Z. Evolution and diversity of transposable elements in fish genomes. Sci Rep. 2019;9:15399.

[74] Volff JN. Genome evolution and biodiversity in teleost fish. Heredity (Edinb). 2005;94:280–294.

[75] Lerat E, Casacuberta J, Chaparro C, et al. On the Importance to acknowledge transposable elements in epigenomic analyses. Genes (Basel). 2019;10:258.

[76] Olsvik PA, Whatmore P, Penglase SJ, et al. Associations between behavioral effects of bisphenol

A and DNA methylation in zebrafish embryos. Front Genet. 2019;10:184.

[77] Vandenberg LN, Colborn T, Hayes TB, et al. Hormones and endocrine-disrupting chemicals: low-dose effects and nonmonotonic dose responses. Endocr Rev. 2012;33:378–455.

[78] Kamstra JH, Sales LB, Aleström P, et al. Differential DNA methylation at conserved non-genic elements and evidence for transgenerational inheritance following developmental exposure to mono(2-ethylhexyl) phthalate and 5-azacytidine in zebrafish. Epigenetics Chromatin. 2017;10:20.

[79] Srivastava RA. Estrogen-induced regulation of the ATP-binding cassette transporter A1 (ABCA1) in mice: a possible mechanism of atheroprotection by estrogen. Mol Cell Biochem. 2002;240:67–73.

[80] Wang H, Liu Y, Zhu L, et al. 17β-estradiol promotes cholesterol efflux from vascular smooth muscle cells through a liver X receptor α-dependent pathway. Int J Mol Med. 2014;33:550–558.

[81] Zhang S, Huang Y, Zhu YC, et al. Estrogen stimulates release of secreted amyloid precursor protein from primary rat cortical neurons via protein kinase C pathway. Acta Pharmacol Sin. 2005;26:171–176.

[82] Marmugi A, Ducheix S, Lasserre F, et al. Low doses of bisphenol A induce gene expression related to lipid synthesis and trigger triglyceride accumulation in adult mouse liver. Hepatology. 2012;55:395–407.

[83] Maunakea AK, Nagarajan RP, Bilenky M, et al. Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature. 2010;466:253–257.

[84] Feng S, Cokus SJ, Zhang X, et al. Conservation and divergence of methylation patterning in plants and animals. Proc Natl Acad Sci U S A. 2010;107:8689–8694.

[85] Yang X, Han H, De Carvalho DD, et al. Gene body methylation can alter gene expression and is a therapeutic target in cancer. Cancer Cell. 2014;26:577–590.

[86] Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012;13:484–492.

[87] Zemach A, McDaniel IE, Silva P, et al. Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science. 2010;328:916–919.

[88] Baccarelli A, Bollati V. Epigenetics and environmental chemicals. Curr Opin Pediatr. 2009;21:243–251.

[89] Gao D, Zhu B, Cao X, et al. Roles of NIPBL in maintenance of genome stability. Semin Cell Dev Biol. 2019;90:181–186.

[90] Zhu Z, Wang X. Roles of cohesin in chromosome architecture and gene expression. Semin Cell Dev Biol. 2019;90:187–193.

[91] Newkirk DA, Chen YY, Chien R, et al. The effect of Nipped-B-like (Nipbl) haploinsufficiency on genome-wide cohesin binding and target gene expression: modeling Cornelia de Lange syndrome. Clin Epigenetics. 2017;9:89.

[92] Meier M, Grant J, Dowdle A, et al. Cohesin facilitates zygotic genome activation in zebrafish. Development. 2018;145.

[93] Antony J, Dasgupta T, Rhodes JM, et al. Cohesin modulates transcription of estrogen-responsive genes. Biochim Biophys Acta. 2015;1849:257–269.

[94] Marin MG, Matozzo V. Vitellogenin induction as a biomarker of exposure to estrogenic compounds in aquatic environments. Mar Pollut Bull. 2004;48:835–839.

[95] Zhang Y-J, Wu H-C, Yazici H, et al. Global hypomethylation in hepatocellular carcinoma and its relationship to aflatoxin B1 exposure. World J Hepatol. 2012;4:169.

[96] Breton CV, Marsit CJ, Faustman E, et al. Small-magnitude effect sizes in epigenetic end points are important in children's environmental health studies: the children's environmental health and disease prevention research center's epigenetics working group. Environ Health Perspect. 2017;125:511–526.

[97] Silvestre F, Gillardin V, Dorts J. Proteomics to assess the role of phenotypic plasticity in aquatic organisms exposed to pollution and global warming. Integr Comp Biol. 2012;52:681–694.

[98] Diz AP, Martínez-Fernández M, Rolán-Alvarez E. Proteomics in evolutionary ecology: linking the genotype with the phenotype. Mol Ecol. 2012;21:1060–1080.