Differential DNA methylation in Pacific oyster reproductive tissue associated with ocean acidification

Running title: Low pH influence on oyster methylome

Yaamini R. Venkataraman[1,2,*], Samuel J. White[2], Steven B. Roberts[2]

[1]Biology Department, Woods Hole Oceanographic Institution, 266 Water Street, Woods Hole, MA, 02543, USA

[2]School of Aquatic & Fishery Sciences, University of Washington, 1122 NE Boat St, Seattle, WA, USA

Abstract

Background: There is a need to investigate mechanisms of phenotypic plasticity in marine invertebrates as negative effects of climate change are experienced by coastal ecosystems. Environmentally-induced changes to the methylome can regulate gene expression, but methylome responses may be tissue- and species-specific. To assess how ocean acidification affects the Pacific oyster (*Crassostrea gigas*) epigenome, we exposed oysters to either low pH (7.31 ± 0.02) or ambient pH (7.82 ± 0.02) conditions for seven weeks. Whole genome bisulfite sequencing was used to identify methylated regions in female oyster gonad samples. C->T single nucleotide polymorphisms were identified and removed to ensure accurate methylation characterization.

Results: Analysis of gonad methylomes revealed a total of 1,284 differentially methylated loci (DML) found primarily in genes, with several genes containing multiple DML. Gene ontologies for genes containing DML were involved in development and stress response, suggesting methylation may promote gonad growth homeostasis in low pH conditions. Additionally, several of these genes were associated with cytoskeletal structure regulation, metabolism, and protein ubiquitination — commonly-observed responses to ocean acidification. Comparison of these DML with other *Crassostrea* spp. exposed to ocean acidification demonstrates that similar pathways, but not identical genes, are impacted by methylation.

Conclusions: Our work suggests DNA methylation may have a regulatory role in gonad and larval development. Combined with existing molluscan methylome research, our work further supports the need for tissue- and species-specific studies to understand the potential regulatory role of DNA methylation.

Keywords: Pacific oyster, bivalve, ocean acidification, DNA methylation, gonad development

Background

There is great interest in elucidating how changes in the environment can impact marine invertebrate stress responses by examining DNA methylation in marine invertebrates [1, 2]. Environmental variation can alter the positions of methyl groups on CpG dinucleotides, potentially regulating gene expression by fine-tuning expression of housekeeping genes or providing additional transcriptional opportunities for environmental response genes [3]. Methylation regulation of gene expression could then influence plasticity in response to environmental stressors [1]. Ocean acidification [4–9], heat stress [10], upwelling conditions [11], disease [12], salinity (Johnson et al. 2021), pesticide exposure [13], and temperature and salinity [9, 14] have all elicited changes in marine invertebrate methylomes, demonstrating a potential for methylation to regulate organismal responses to environmental conditions.

Given the impact of ocean acidification on calcifying species like oysters [15–17], several environmental epigenetic studies have examined the influence of ocean acidification on molluscan methylomes (Table 1). For example, eastern oyster (*Crassostrea virginica*) mantle [6], *C. virginica* reproductive tissue [8], Hong kong oyster (*Crassostrea hongkongensis*) mantle [4], and *C. hongkongensis* larval methylomes [7] all display changes in DNA methylation after experimental ocean acidification exposure. Between all four studies, only Venkataraman et al. (2020) reports the percent of CpGs methylated: the stated 22% is consistent with the baseline 15% methylation observed in initial Pacific oyster (*Crassostrea gigas*) methylation studies [18], as well as examination of methylome responses to heat stress [10] and diuron exposure [13]. In *C. virginica*, the mantle methylome was hypomethylated after ocean acidification exposure [6], while there were no differences in global gonad methylation patterns [8]. Experimental ocean

acidification exposure yielded differential methylation predominantly in gene bodies, but concentration in exons [7, 8] versus introns [4] varies. Hyper- and hypomethylation occurs with equal frequency in response to ocean acidification. This is consistent with diuron exposure [13], but not with heat stress [10]. Gene functions enriched in differential methylation datasets were inconsistent between species and tissue types (Table 1). While aminotransferase complex and biosynthetic processes were found to be enriched in hypomethylated DML in *C. virginica* mantle tissue [6], there were no enriched GOterms associated with genic DML in *C. virginica* gonad samples. In contrast, several processes were enriched in methylation datasets for *C. hongkongensis*, including acetoacetylco-A reductase activity, dehydrogenase activity, cellular response to pH, protein xylosyltransferase activity, translation factor activity, RNA binding, and diacylglycerol kinase activity in the mantle [4] and cytoskeletal and signal transduction, oxidative stress, metabolic processes, and larval metamorphosis in larvae [7]. While considerable effort has been made to understand molluscan methylation responses to ocean acidification, it is clear that there are likely species- and tissue-specific responses.

In addition to being an important global aquaculture species, *C. gigas* has been used to further our understanding of methylation in marine invertebrates [3, 18–26]. Examination of diuron exposure [13] and heat stress on different *C. gigas* phenotypes [10] reveal that environmental conditions do influence the Pacific oyster methylome, and other oyster methylomes are responsive to ocean acidification (Table 1). However, studies examining the influence of ocean acidification on the *C. gigas* methylome are currently absent. Additionally, a recent study found that female *C. gigas* exposure to low pH conditions, followed by four months in ambient pH conditions prior to spawning, can still negatively impact larval survival 18 hours post-fertilization [27]. Since maturation stage was not different between female *C. gigas* after low pH exposure, this finding suggests environmental “memory” may be mediating the carryover effect.

We sought to understand impacts of ocean acidification on the Pacific oyster gonad methylome in order to not only understand DNA methylation as a potential mechanism for the observed negative maternal carryover effect previously reported in Venkataraman et al. [27], but to also examine the functional role of methylation in regulating biological responses to ocean acidification. Female oyster gonad methylome sequencing revealed distinct methylation changes associated with low pH exposure. These results provide the foundation for examining phenotypic plasticity within a generation, as well as larval methylomes and potential intergenerational effects.

Table 1. Studies exploring methylation responses to environmental stress in *Crassostrea* spp.

Results

Sequence Alignment

Whole genome bisulfite sequencing produced 1.38 x 10^9 total 150 bp paired-end reads for eight total libraries. After quality trimming, 1.35 x 10^9 total paired reads remained (Supplementary Table 1). Of the trimmed paired-end reads, 8.31 x 10^8 (61.7%) were uniquely aligned to the bisulfite-converted *C. gigas* genome with appended mitochondrial sequence (Supplementary Table 1).

SNP Identification

C->T SNP variants were identified from WGBS data to remove loci that might have been incorrectly characterized across individuals. A total of 13,234,183 unique SNPs were found in individual or merged BAM files, including 300,278 C->T SNPs. These C->T SNPs were used to exclude loci from downstream analyses (see DML Characterization and Enrichment Analysis).

Global Methylation Patterns

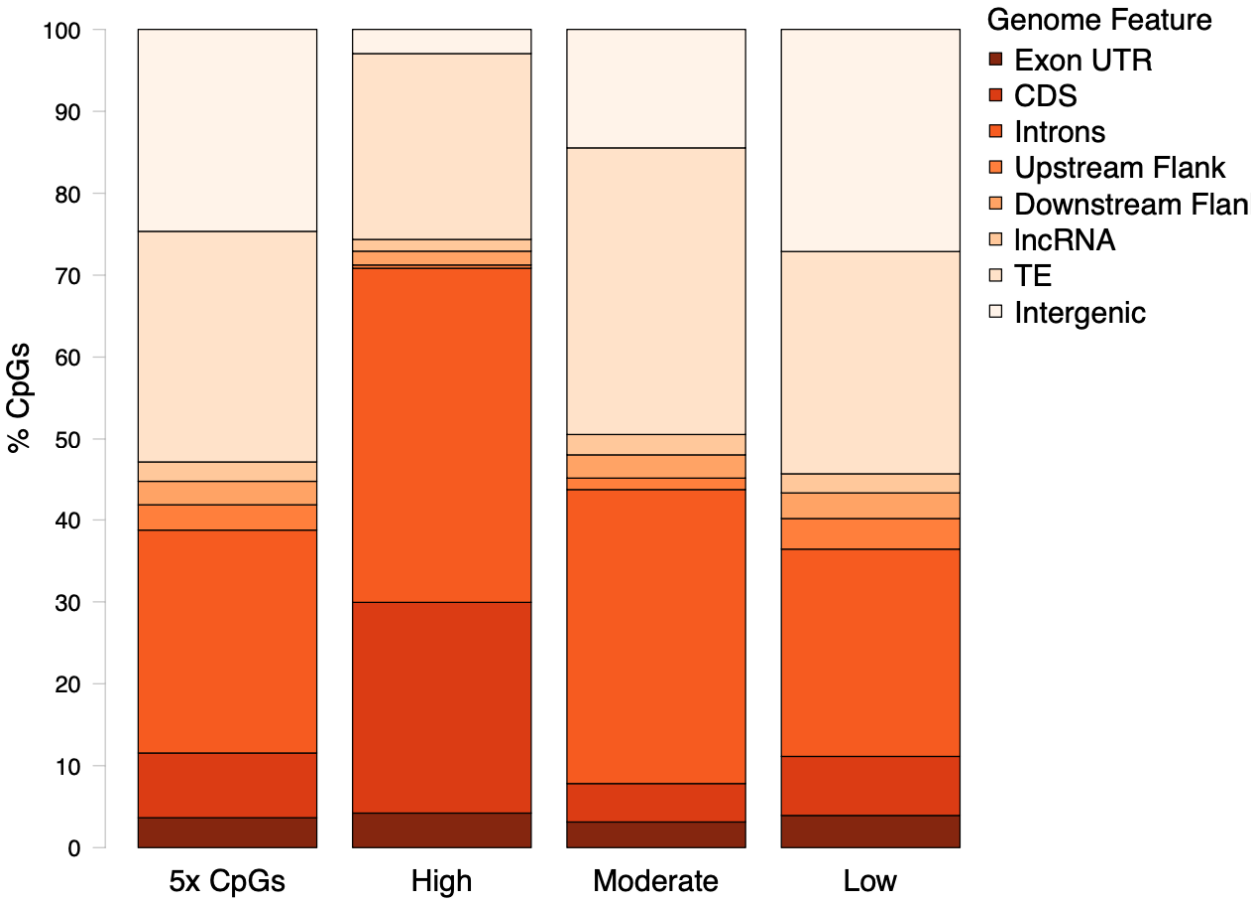


Figure 1. **Location of all CpGs with 5x coverage.** Percentage of highly methylated (≥ 50%), moderately methylated (10-50%), and lowly methylated (≤ 10%) CpGs in various genome features.

Combined data with 5x coverage threshold provided 11,238,223 CpG loci (84.8% of 13,246,547 total CpGs in *C. gigas* genome). After C->T SNP removal, 10,939,918 CpG loci (82.6% of total *C. gigas* CpGs) were characterized. The majority of CpGs, 9,018,512 loci (82.4%), were lowly methylated, followed by 966,655 highly methylated (8.8%) and 954,751 moderately methylated

(8.7%) CpGs. Highly methylated CpGs were found primarily in genic regions (903,933 CpGs, 93.5%), with 53,951 (6.0%) in exon UTR, 329,117 (36.4%) in CDS, and 523,386 (7.9%) in introns (**Figure 1**). The distribution of highly methylated CpGs was significantly different than all 5x CpGs detected by WGBS (Supplementary Table 2). A Principal Component Analysis (PCA) did not demonstrate any clear separation of samples by treatment (Supplementary Figure 1), and consistently high Pairwise Pearson's correlation coefficients support the lack of global percent methylation differences between pH treatments (Supplementary Figure 2).

DML Characterization

Initially, 1,599 CpG loci were identified as potential DML using a 50% methylation difference threshold. Of these CpG loci, 315 (19.7%) overlapped with C->T SNPs, and thus were removed from downstream analysis. A total of 1,284 DML were used in all remaining analyses, which consisted of 654 (50.9%) hypomethylated and 630 (49.1%) hypermethylated DML (**Figure 2**). The DML were distributed amongst the ten main chromosomes, as well as scaffolds not placed in any chromosome (**Figure 3**, Supplementary Table 3).

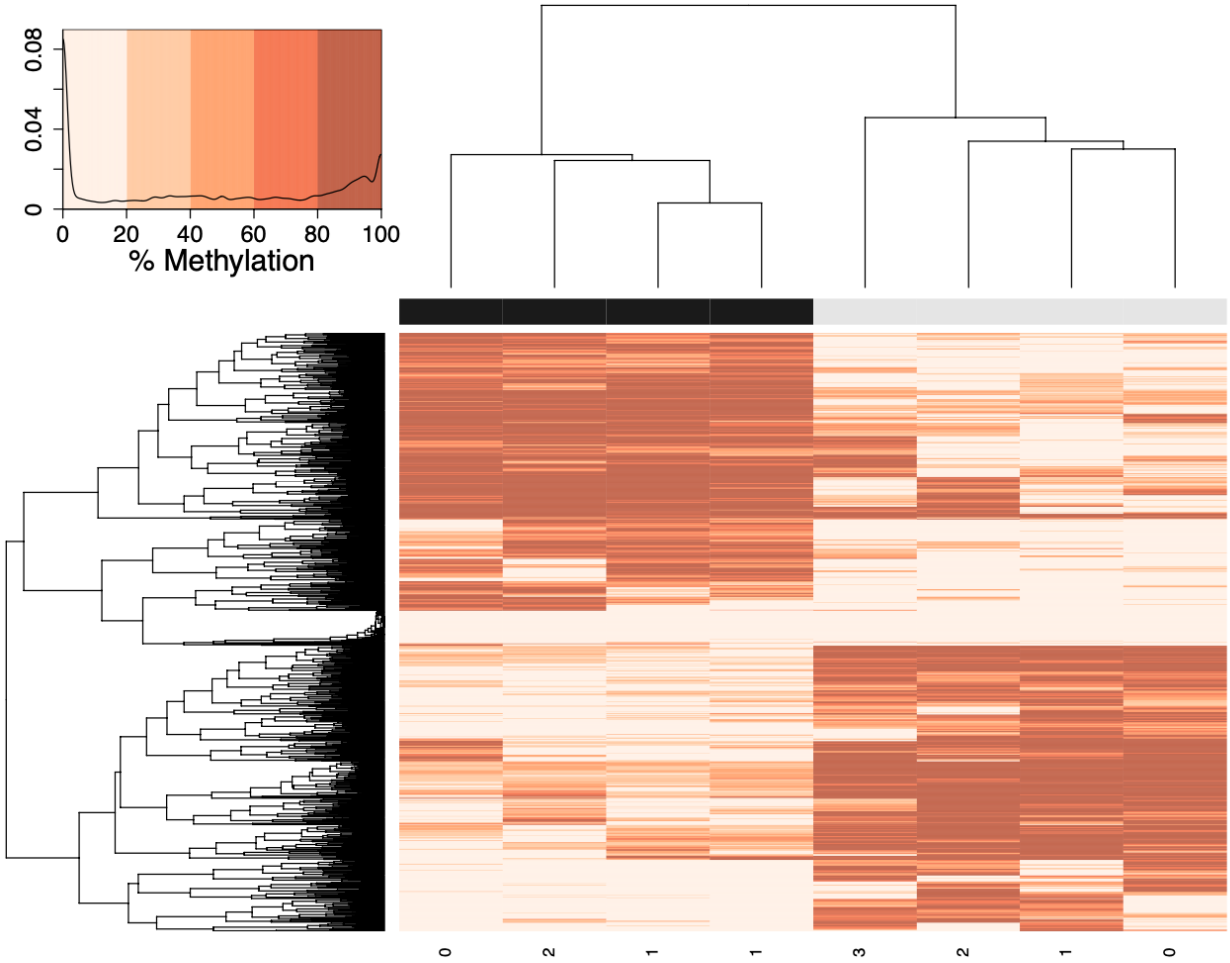


Figure 2. **Percent methylation values for DML created using an euclidean distance matrix**. Samples in low pH conditions are represented by black, and samples in ambient pH conditions are represented by gray, with maturation stage along the bottom (0 = indeterminate, 3 = spawn-ready mature female). Darker colors indicate higher percent methylation, and a density plot depicts the distribution of percent methylation values for a panel. After excluding C->T SNPs, 1,284 DML were identified using a logistic regression, using a chi-squared test and 50% methylation difference cut-off.

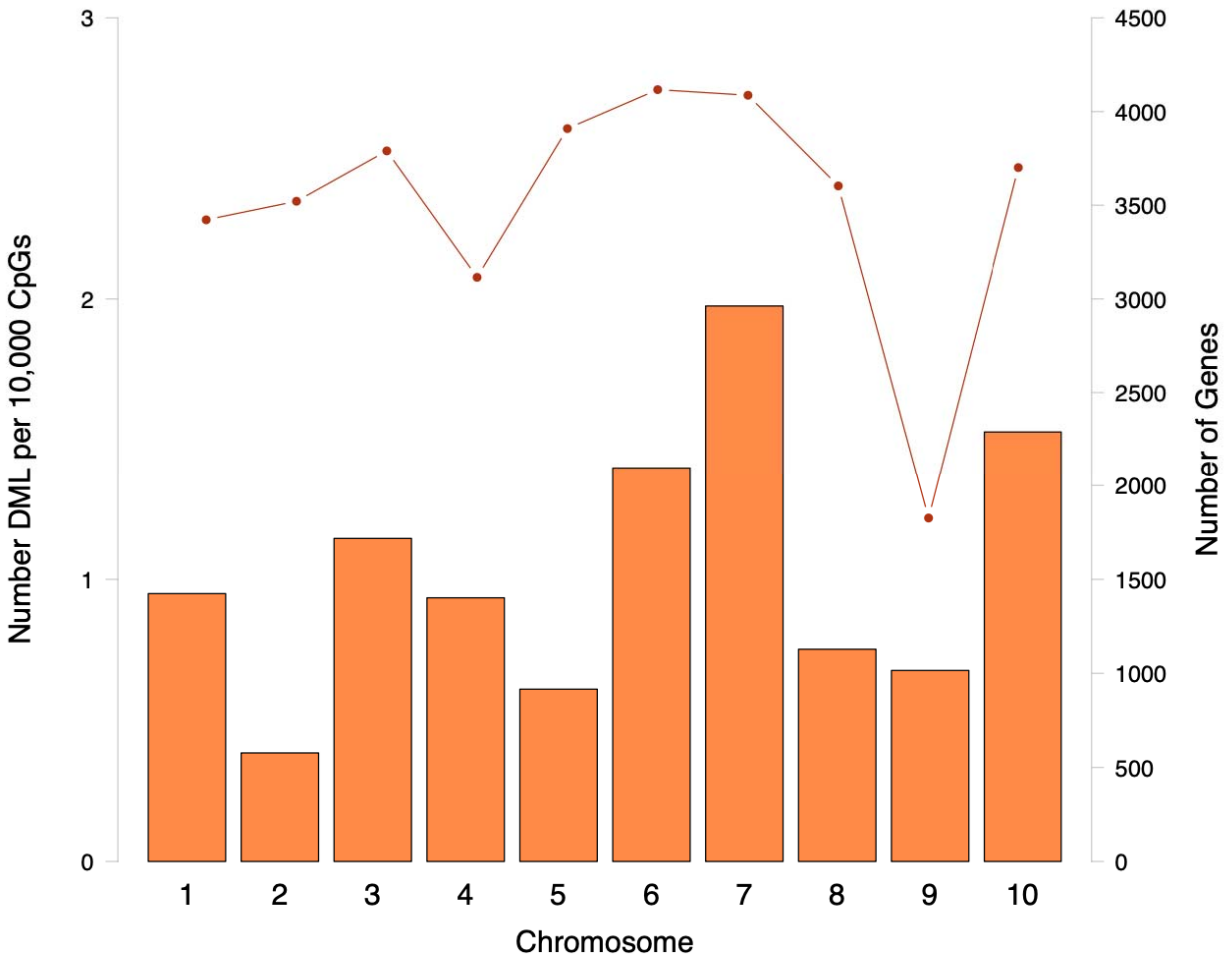


Figure 3. **Distribution of DML in main chromosomes**. Number of DML normalized by number of CpG in each chromosome (bars) and number of genes (line) in each chromosome. Additional DML were identified in scaffolds that were not mapped to any of the ten main linkage groups (Supplementary Table 3).

The majority of DML (1,181; 92.0%) were located in genic regions (**Figure 4**), consisting of 859 unique genes. Most genes only contained one DML; however, several genes contained multiple DML, ranging from 2-103 DML in one gene (Supplementary Table 4). The genic DML were primarily located in introns (783; 61.0%), followed by CDS (442; 34.4%) and exon UTR (77; 6.0%). Putative promoter regions upstream of transcription start sites contained DML (10;

0.8%), but more were located in downstream flanking regions (61; 4.8%). The DML were also found in transposable elements (434; 33.8%), lncRNA (9; 0.7%), and intergenic regions (38; 3.0%). The number of DML in all genome features differed significantly from all 5x CpGs (Supplementary Table 5).

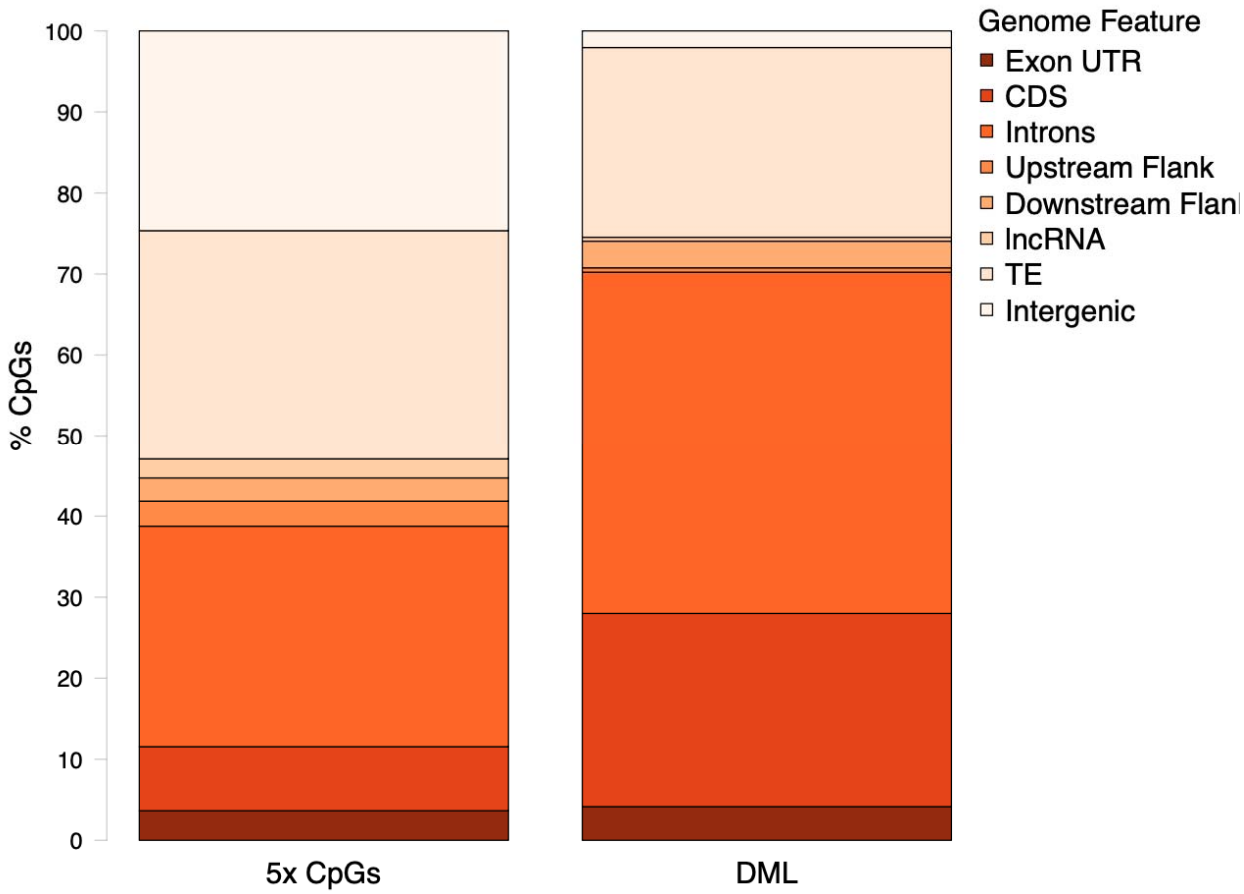


Figure 4. **Location of CpGs with 5x coverage and DML**. Percentage of 5x CpGs and DML found in various genome features.

Enrichment Analysis

Ten biological processes and 4 cellular component GOterms were significantly enriched in genes containing DML. A total of 339 genes with GOterm annotations containing 437 DML were

used for this analysis. The enriched biological process GOterms were involved in motility (ex. cilium-dependent cell motility, cilium movement) and development (ex. insta larval or pupal development, post-embryonic development) (Supplementary Table 6). Nine unique genes (37 associated transcripts) contained enriched biological process GOterms: *dynein heavy chain 5, axonemal, unconventional mysoin-VI, serine/threonine-protein kinase 36, helicase domino, protein neuralized, cytoplasmic aconitate hydratase, serine-protein kinase ATM, cation-independent mannose-6-phosphase receptor, uncharacterized LOC105324167*. These genes contained 16 DML, 13 of which were hypermethylated (Supplementary Table 6). Enriched cellular component GOterms were involved in microtubule-associated complexes and various acetyltransferase complexes (Supplementary Table 7). The six unique genes (25 associated transcripts) that contained enriched cellular component GOterms were *dynein heavy chain 5, axonemal, cytoplasmic dynein 2 light intermediate chain-1-like, unconventional myosin-VI, helicase domino, and kinesin-like protein KIF23* (Supplementary Table 7). Six of the seven DML contained in these genes were hypermethylated (Supplementary Table 7).

Discussion

Our work examined the potential role of environmentally-induced changes to the reproductive tissue methylome to mediate carryover effects in *C. gigas*. We identified a total of 1,284 differentially methylated loci, showing that the methylome is responsive to ocean acidification. We also found enriched biological processes associated with genic DML, implying that low pH may impact regulation of distinct processes, such as gonad development. Our findings contribute to the growing body of work examining epigenetic responses in molluscan species, and provide the foundation for future research examining intergenerational epigenetic inheritance and its potential to mediate plastic responses to environmental stressors.

Functions of genes containing DML with enriched GOterms suggest role for methylation regulation of gonad development

Enrichment of microtubule associated complex and cilium or flagellar motility GOterms imply methylation may control microtubule movement in response to low pH. Motor proteins like *dynein heavy chain 5, axonemal*, and *cytoplasmic dynein 2 light intermediate chain 1-like* form complexes to move flagella or cilia. Dyneins exhibit higher expression over the course of *C. gigas* spermatogenesis [48]. In female gonad tissue, hypermethylation may suppress expression of dyneins to promote oogenesis. Indeed, lack of significant sex ratio and maturation stage differences between low and ambient pH treatments suggests that oogenesis was maintained in low pH conditions [27]. Studies examining molluscan somatic tissues have found low pH impacts on dynein expression [49–51]. Similar to cytoplasmic dyneins, *unconventional myosin-VI* plays a role in organelle movement. Myosins generally exhibit high expression in immature gonads, which reduces as the gonad matures [48, 52]. The reduction in myosin expression may signify reorganization of female reproductive tissue to accommodate growing oocytes. Hyper- and hypomethylation of unconventional myosin may regulate cytoskeletal structure of cells to promote growth of reproductive over somatic tissue. Dynein and myosin genes exhibit differential methylation in response to low pH in *C. virginica* gonad as well [8]. Regulation of motility genes may be a conserved response to ocean acidification, and methylation of these genes may promote homeostasis during gametogenesis.

Several genes with DML and enriched GOterms have links to stress response and gonad development. In *C. gigas*, expression of cell cycle regulators *kinesin-like protein KIF23* and *serine-protein kinase ATM* changes in response to immune challenge or low pH and arsenic exposure, respectively [53, 54]. Kinesin-like proteins have higher expression in mature *C. gigas* gonads [48], and serine-protein kinase ATM is involved in razor clam *Sinonovacula constricta*

oogenesis [55]. Another kinase, *serine/threonine-protein kinase 36*, is involved in signaling pathways. Several serine/threonine-protein kinases are highly expressed in mature *C. gigas* female gonad, with certain homologues more expressed in females with mature gonad and during sex determination and differentiation events [48, 52]. Serine/threonine-protein kinases are also related to oocyte maturation in the king scallop *Pecten maximus* [56]. An investigation of the *C. gigas* kinome found various serine/threonine-protein kinases present in eggs, including *serine/threonine-protein kinase 36*, that were responsive to abiotic stressors [57]. Serine/threonine-protein kinase genes were also differentially methylated in *C. virginica* gonad [8]. As a catalyst of the TCA cycle, *cytoplasmic aconitate hydratase* is upregulated during spermatogenesis in *C. gigas* and downregulated during spermatogenesis in *P. maximus*, with differing consequences for energy metabolism [58, 59]. Early examination of *C. virginia* found *cytoplasmic aconitate hydratase* localized to eggs, with enzyme activity increasing during embryogenesis [60]. Therefore, it's possible *cytoplasmic aconitate hydratase* may regulate energy metabolism during oogenesis as well. Evidence of *cytoplasmic aconitate hydratase* abundance changing in response to low pH has been documented in *C. gigas*, which implies that low pH alters energy metabolism [54]. Energy metabolism would be important to regulate in developing oocytes. *Helicase domino* regulates transcription through histone phosphorylation, acetylation, and chromatin remodeling. Populations of the Olympia oyster (*Ostrea lurida*) contain an outlier SNP loci in the *helicase domino* gene, which is implicated in immune or stress response [61]. Upregulation of *helicase domino* was associated with oogenesis maintenance in *Acartia tonsa* shrimp [62]. Taken together, hypermethylation of *kinesin-like protein 23*, *serine-protein kinase ATM*, *cytoplasmic aconite hydratase*, and *helicase domino*, and hypomethylation of *serine/threonine-protein kinase 36*, suggests that methylation may regulate gonad development in stressful conditions, allowing the oyster to maintain homeostasis.

Similarities between differential methylation in gonads and larvae

Investigation of *C. hongkongensis* larvae exposed to ocean acidification revealed enrichment of cytoskeletal and signal transduction processes in differentially methylated genes [7]. Genes with DML in *C. gigas* gonad are enriched for similar processes, and have links to embryonic development. Unconventional myosins, important cytoskeletal components, were found to be important for fertilization, blastulation, and gastrulation in *S. purpuratus* [63]. Ocean acidification drove decreased expression of another cytoskeletal component, kinesin-like protein, in *S. purpuratus* larvae, which was related to measured reductions in calcification [64]. Expression of signal transduction kinases found in embryos changed in response to various abiotic stressors [57]. *Serine/threonine protein-kinase pim-3* gene was differentially methylated in *C. hongkongensis* larvae [7], and this study found a DML in *serine/threonine-protein kinase 36* in *C. gigas* gonad. Differential methylation in cytoskeleton and signal transduction genes in female gonad may influence larval responses to low pH.

Enrichment of metabolic processes in *C. gigas* gonad and *C. hongkongensis* larval methylation datasets [7] signify that methylation regulation of metabolism is important across life stages. In *S. purpuratus* with a history of low pH exposure, day-old embryos exhibited higher expression of *cytoplasmic aconitate hydratase* than embryos from populations without exposure history [65]. However, larvae from populations without exposure history had higher expression of *cytoplasmic aconitate hydratase* after seven days [65]. Modulation of the TCA cycle through *cytoplasmic aconitate hydratase* may be important for metabolic regulation in developing larvae, and inheritance of hypermethylation within this gene could prime larval metabolism for low pH exposure. *Cation-independent mannose-6-phosphate receptor* is a lysosomal enzyme associated with post-embryonic development (GO:0009791). Expression of this gene was upregulated during zebra mussel *Dreissena polymorpha* and shrimp *Marsupenaus japonicus* immune challenges [66, 67]. Although no current research has explored this gene's role in larval development or abiotic stress response, hypomethylation of *cation-independent mannose-6-*

phosphate receptor may impact lysosome metabolic activity and immune function during a critical growth period. *Protein neuralized* is part of a group of ubiquitin ligases. Regulation of ubiquitination is likely a conserved response to ocean acidification in oysters: elevated pCO_2 increased abundance of proteins involved in ubiquitination and decreased protein degradation in *C. gigas* posterior gill lamellae [49], and upregulated ubiquitination gene expression in adult *C. virginica* mantle [68]. Differential methylation of protein ubiquitination genes was also found in *C. virginica* gonad [8] and *C. hongkongensis* larvae [7]. Hypermethylated DML within the *protein neuralized* gene could regulate ubiquitination. Differential methylation in this gene may also regulate cell division and larval growth, as upregulation of protein neuralized has been found to suppress cell division and macromolecule synthesis in *Artemia sinica* during diapause [69].

Taken together, methylome similarities between *C. gigas* female gonad and *C. hongkongensis* larvae suggest that differential methylation in these genes could also impact larval calcification, cellular structure, and metabolism if patterns are inherited. Intergenerational methylation inheritance has been demonstrated previously in *C. gigas*: adult exposure to the pesticide diuron altered spat methylomes [13]. While we did not investigate *C. gigas* larval methylomes, we did find that prior exposure of female oysters to low pH conditions negatively impacted larval survival in ambient pH conditions 18 hours after fertilization [27]. It is possible that the methylation patterns responsible for maintaining homeostasis in female gonad tissue were maladaptive in larvae, especially when parent and larval environments were mismatched.

Considerations for marine invertebrate methylation research

The concentration of DML in genes is consistent with the literature examining molluscan methylome responses to ocean acidification [4, 6–8], disease [12], pesticide exposure [13], salinity (Johnson et al. 2021), and heat stress [10], as well as environmental responses in corals [9, 14], urchins [11], and pteropods [5]. This differs from the vertebrate model of methylation, in

which CpG islands are found in promoter regions upstream of transcription start sites [1]. In *C. gigas*, high levels of gene methylation are associated with increased gene expression and low expression variability, implying that methylation leads to transcriptional stability [18, 21, 26]. Similarly, exposure to the pesticide diuron modified the *C. gigas* methylome such that DNA methylation changes correlated with RNA abundance in a small group of highly methylated genes [13], and increased methylation was also found to reduce transcriptional noise in response to ocean acidification in corals [9]. While recent *Crassostrea* spp. studies have found no connection between differentially methylated and differentially expressed genes after exposure to low pH [4, 6], individual DML may play a role in shaping gene expression. Comparison of a single gene with differing methylation levels between *C. gigas* mantle and male gametes demonstrated that gene methylation was associated with alternative splicing and gene expression [26]. Therefore, it is critical to examine how individual DML can influence gene activity in response to environmental change.

The role of DML in shaping gene activity is likely dependent on their locations in genome features. We found 442 and 941 DML in introns and CDS, respectively, with roughly half of the genic DML being hypermethylated in treatment gonads. In *C. gigas*, exons included in the final transcript and longer exons generally have higher methylation [26]. Genes with higher methylation in exons or introns were highly expressed in *S. purpuratus*, and had a higher likelihood of alternative splicing, alternative start site, or exon skipping in response to upwelling conditions [70]. Differentially methylated positions in introns and CDS may have similar roles in *C. gigas* response to low pH. Additionally, presence of multiple DML in a gene may have conflicting effects on expression. Most genes examined in this study only contained one DML; however, we found several genes with multiple DML, ranging from 2-103 in one gene. High methylation in both exons and introns did not produce high expression levels in *S. purpuratus*, suggesting increased methylation in both of these genome features act antagonistically to

impact gene expression [70]. Multiple DML in CDS and introns may influence *C. gigas* expression similarly. Future work should examine how environmentally-influenced methylation impacts gene expression by considering individual DML, not averaged methylation for a gene.

Confounding elements may have influenced gonad methylome responses to ocean acidification, and our interpretation of those responses. Maturation stage is known to produce distinct baseline methylation patterns in *C. gigas* (Zhang et al., 2018). While the overall maturation stage was used as a covariate when identifying DML, it is possible that each sample methylome contains information from cells with various maturation stages or cell types. Single-cell sequencing efforts or investigations within a single maturation stage would clarify how these sources of variation influenced our findings. We also identified several C->T SNPs that overlapped with DML even though specimens were sibling oysters with similar genomic backgrounds. Recent examination of *C. virginica* methylomes suggest methylation is largely linked to genotype, emphasizing the need to account for genotypic differences in methylation calls [71]. Our work demonstrates how SNP identification from WGBS can be used to exclude loci with incorrectly characterized methylation due to genomic variation. While the lack of annotated genes used in the enrichment analysis limits our interpretation of which processes are most impacted by pH-sensitive methylation, removing SNP-DML overlaps prior to enrichment analysis provides a more biologically accurate understanding of how methylation may impact gonad growth.

Conclusion

Our work — which is the first to explore DNA methylation responses to ocean acidification in Pacific oysters — found that low pH treatment altered the *C. gigas* female gonad methylome after a seven-week exposure. We found that differential methylation occurs primarily in genic

regions, which is consistent with other studies examining oyster methylome responses to ocean acidification. In conjunction with histology data from Venkataraman et al. [27], enriched biological processes and cellular components in genes containing DML suggest that methylation may be used to maintain gonad growth in low pH conditions. Our work expands on previous molluscan methylation research by not only demonstrating a potential role for methylation in response to ocean acidification, but also showing the species- and tissue-specific nature of these responses.

As inheritance of environmentally-induced methylation — and any resulting phenotypic plasticity — is contingent on epigenetic marks being present in the germline, withstanding chromatin reorganization, and being packaged into gametes [1], the next step in this work is correlating germline methylation with offspring methylation, and examining gene expression, alternative splicing, chromatin regulation, and phenotype. Comparing parental and offspring phenotype will allow researchers to assess whether parental experience and resulting epigenetic responses are accurate, reliable cues for adaptive plasticity in offspring.

Methods

Experimental Design

The experimental design and seawater chemistry analysis with carbon chemistry parameters are described in [27]. Briefly, adult hatchery-raised *C. gigas* (average shell length = 117.46 ± 19.16 cm) were exposed to either low pH (7.31 ± 0.02) or ambient pH (7.82 ± 0.02) conditions from February 15, 2017 to April 8, 2017 at the Kenneth K. Chew Center for Shellfish Research and Restoration at the National Oceanic and Atmospheric Administration Manchester Field Station (47º 34’ 09.1” N 122º 33’ 19.0” W, Manchester, WA). Shell length, width, and oyster wet weight were obtained before and after pH exposure, and gonad tissue was extracted for

histological analysis at each sampling point. Prior to exposure, twenty *C. gigas* were randomly selected for histology, while ten oysters from each treatment were randomly selected from tanks to assess maturation after exposure. Gonad tissue in cassettes was fixed for histology using PAXgene Tissue Fix and Stabilizer and sent to Diagnostic Pathology Medical Group, Inc. (Sacramento, CA) for hematoxylin and eosin staining and slide preparation. Tissues exposed to ambient pH were confounded during processing, preventing any tank identification. Maturation state and organism sex was evaluated histologically at 40x magnification [28, 29].

DNA Extraction and Library Preparation

To determine how pH exposure altered maternal gonad methylation, DNA was extracted from histology blocks for four female (or presumed female) oysters for each treatment using the PAXgene Tissue DNA Kit (Qiagen) with modifications specified below at the end of the seven-week exposure. For each sample, up to 0.02 g of histology tissue embedded in paraffin was cut from the histology block and placed in a 2 mL round-bottomed tube. The tissue was then further homogenized within the tube. After adding 1 mL of xylene, each sample was vortexed for 20 seconds, incubated at room temperature for three minutes, then centrifuged at maximum speed for three minutes. To evaporate ethanol from samples after ethanol addition and vortexing steps, samples were placed on a heat block at 37ºC for 10-20 minutes.

The resuspended pellet was lysed using a TissueTearor at setting 1. Prior to lysis and between samples, the TissueTearor was run in a 10% bleach solution for 20 seconds, followed by two consecutive 20 second runs in DI water to clean the instrument. Lysates were transferred to clean, labeled 1.5 mL microcentrifuge-safelock tubes. Proteinase K (20 µL) was added to each sample, and pulse-vortexed for 15 seconds to mix. The sample-Proteinase K solution was incubated 56ºC for 60 minutes. Every ten minutes, the samples were briefly removed from the heat block to vortex at maximum speed for one minute. After 60 minutes, RNase A (4 µL, 100

mg/mL) was added to each sample. Samples were vortexed at maximum speed for 20 seconds to mix and incubated for two minutes at room temperature to obtain RNA-free genomic DNA and reduce possible interference for downstream enzyme reactions. Samples were then kept at 80ºC for 60 minutes, and vortexed at maximum speed for one minute in ten-minute intervals. After the incubation, Buffer TD2 and ethanol (200 µL) were added to each sample, vortexing thoroughly to mix.

DNA was isolated for each sample after lysis and incubation following manufacturer's instructions. To elute DNA, samples were loaded onto the spin column with Buffer TD5 (50 µL) and incubated at room temperature for five minutes to increase yield, then centrifuged at maximum speed for one minute. A Qubit™ 3.0 and dsDNA BR Assay (Invitrogen) was used to quantify sample yield using 1 µL of DNA for each sample. For two samples, DNA was further concentrated using an ethanol precipitation.

Bisulfite conversion was performed with the Zymo-Seq WGBS Library Kit (Cat. #D5465) using 100 ng of genomic DNA following manufacturer's instructions. Samples were spiked with a mixture of six unique double-stranded synthetic amplicons (180-200 bp) from the *Escherichia coli* K12 genome. Since each amplicon represents a different percent methylation, the spike-in was used to determine bisulfite conversion efficiency. Libraries were barcoded and pooled into a single lane to generate 150 bp paired-end reads in a single Illumina NovaSeq flowcell.

Genome Information and Feature Tracks

Nuclear genome [30] (NCBI GenBank: GCA_902806645.1, RefSeq: GCF_902806645.1) and mitochondrial genome sequences (NCBI Reference Sequence: NC_001276.1) were used in downstream analysis. These sequences were combined for downstream alignment. The `fuzznuc` function within EMBOSS was used to identify CG motifs in the combined genomes.

The RefSeq annotation was used to obtain genome feature tracks, and create additional tracks. The annotation included a combination of Gnomon, RefSeq, cmsearch, and tRNAscan-SE predictions. Gene, coding sequence, exon, and lncRNA tracks were extracted from the RefSeq annotation. These tracks were then used to obtain additional genome feature information. To create an intron track, the complement of the exon track was generated with BEDtools `complementBed` v.2.26.0 to create a non-coding sequence track [31]. The overlap between the genes and the non-coding sequences was found with `intersectBed` and defined as introns. Similarly, untranslated regions of exons were obtained by subtracting coding sequences from exon information with `subtractBed`. Flanking regions were defined as the 1000 bp upstream or downstream of genes. Upstream flanks were generated by adding 1000 bp upstream of genes, taking into account strand, with `flankBed`. Existing genes were removed from flanking sequences using `subtractBed`. A similar process was performed to generate downstream flanking regions. Intergenic regions were isolated by finding the complement of genes with `complementBed`, then using `subtractBed` to remove any flanks. Transposable element locations were obtained from RepeatMasker using the NCBI RefSeq annotation (RefSeq: GCF_902806645.1) [32, 33]. The number of CG motifs in a given feature track was obtained using `intersectBed`. All genome feature tracks are available in the associated Open Science Framework repository (see *Availability of data and materials*).

Sequence Alignment

Reads were trimmed three times prior to alignment using TrimGalore! v.0.6.6 [34]. Known Illumina adapter sequences were removed (`--illumina`), along with 10 bp from both the 5' (`--clip_R1 10` and `--clip_R2 10`) and 3' ends (`--three_prime_clip_R1 10` and `--three_prime_clip_R2 10`) of 150 bp paired-end reads. Any remaining adapters were

removed in a second round of trimming. Finally, poly-G tails were trimmed out of samples by manually specifying adapter sequences (`--adapter GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG` and `--adapter2 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG`). Sequence quality was assessed with FastQC v.0.11.9 [35] and MultiQC [36] after each round of trimming.

Trimmed samples were then aligned to the combined nuclear and mitochondrial genomes for *C. gigas* [30]. The genome was bisulfite converted with Bowtie 2-2.3.4 (Linux x84_64 version [37]) using the `bismark_genome_preparation` function within Bismark v.0.22.3 [38]. Reads were aligned to the bisulfite-converted genome. Non-directional input was specified (`--non_directional`), and alignment scores could be no lower than -90 (`-score_min L,0,-0.6`). Resultant BAM files were deduplicated (`--deduplicate_bismark`), and methylation calls were extracted from deduplicated BAM files (`bismark_methylation_extractor`). Deduplicated BAM files were also sorted and indexed for downstream applications using SAMtools v.1.10 [39]. Using coverage files generated from methylation extraction, CpG information was consolidated between strands to report 1-based chromosomal coordinates (`coverage2cystosine`). Using `coverage2cytosine` coverage file output, genomic coordinates for CpGs with at least 5x coverage across all samples were written to BEDgraphs. Summary reports with alignment information for each sample were generated (`bismark2report` and `bismark2summary`) and concatenated with MultiQC. Code for all sequence alignment steps can be found in the associated Open Science Framework repository (see *Availability of data and materials*).

SNP Identification

During bisulfite conversion, cytosines methylated in a CpG context are retained as cytosines, while unmethylated cytosines are converted to uracils. These uracils are then converted to thymines during PCR amplification. As C->T single nucleotide polymorphisms (SNPs) would be interpreted by Bismark as an unmethylated cytosine, removing these loci is an important step to ensure accurate methylation characterization. SNP variants across all samples and in individuals were identified with `BS-Snper` [40]. This program takes into consideration nucleotide identity on the opposite strand: a thymine paired with an adenine indicates a C->T SNP, while a thymine paired with a guanine is a bisulfite-converted unmethylated cytosine. Sorted deduplicated BAM files generated with Bismark v.0.22.3 and SAMtools v.1.10 [38, 39] were merged (`samtools merge`) to identify SNPs. Variants were also identified in individual sorted deduplicated BAM files. Default `BS-Snper` settings were used, except for a minimum coverage of 5x to be consistent with methylation analyses. The C->T SNPs were used to exclude loci of interest.

Global Methylation Characterization

Global methylation patterns were characterized by averaging DNA methylation across all samples using loci with at least 5x coverage that did not overlap with C->T SNPs. A CpG locus was considered highly methylated if average percent methylation was greater than or equal to 50%, moderately methylated if percent methylation was between 10-50%, and lowly methylated if percent methylation was less than or equal to 10%. The genomic location of highly, moderately, and lowly methylated CpGs in relation to UTR, CDS, introns, up- and downstream flanking regions, transposable elements, and intergenic regions were determined by using `intersectBed`. We tested the null hypothesis that there was no association between the genomic location of CpG loci and methylation status (all CpGs versus highly methylated CpGs) with a chi-squared contingency test (`chisq.test` in R Version 3.5.3 [41]).

Identification of Differentially Methylated Loci

Identification of differentially methylated loci, or DML, was conducted with `methylKit v.1.17.4` [42] in R Version 3.5.3 [41]. Prior to DML identification, data were imported and processed. Data with at least 2x coverage were imported using `methRead` from merged CpG coverage files produced by `coverage2cytosine`. Imported data were filtered again (`filterByCoverage`) to require 5x coverage for each CpG locus (`lo.count = 5`) in each sample. Potential PCR duplicates in each sample were also removed by excluding CpG loci in the 99.9th percentile of coverage (`high.perc = 99.9`). Once filtered, data were normalized across samples (`normalizeCoverage`) to avoid over-sampling reads from one sample during downstream statistical analysis. Histograms of percent methylation (`getMethylationStats`) and CpG coverage (`getCoverageStats`) were used to confirm normalization. After filtering and normalization, data at each CpG locus with 5x were combined (`unite`) to create a CpG background, ensuring that each locus had at least 5x coverage in each sample.

Methylation differences were then identified for *C. gigas* gonad samples. A logistic regression (`calculateDiffMeth`) modeled the log odds ratio based on the proportion of methylation at each CpG locus. A covariate matrix with gonad stage information (`covariates`) were applied to the model for the *C. gigas* data to account for varying maturation stages in samples. Additionally, an overdispersion correction with a chi-squared test (`overdispersion = "MN", test = "Chisq"`) was applied.

Differentially methylated loci, or DML, were defined as CpG dinucleotides that did not overlap with C->T SNPs with at least a 50% methylation difference between pH treatments, and a q-value < 0.01 based on correction for false discovery rate with the SLIM method [43]. Hypermethylated DML were defined as those with significantly higher percent methylation in low

pH samples, and hypomethylated DML with significantly lower percent methylation in low pH samples. BEDfiles with DML locations were created and viewed with the Integrative Genomics Viewer (IGV) version 2.9.2 [44].

DML Characterization

The location of DML was characterized in relation to various genome feature tracks. Presence of DML in UTR, CDS, introns, up- and downstream flanking regions, transposable elements, and intergenic regions was discerned using `intersectBed`. A chi-squared contingency test was used to test the null hypothesis of no association between genomic location and differential methylation by comparing the genomic location of all CpGs with 5x data and DML. Additionally, the number of DML present in each chromosome was quantified to see if DML were distributed uniformly across the genome. Overlaps between DML and SNPs were identified to determine if DML were solely attributed to treatment differences, or if genetic differences contributed to differential methylation.

Enrichment Analysis

To determine if genes containing DML were associated with overrepresented biological processes, cellular component, or molecular function gene ontologies (GO), functional enrichment analysis was performed. Prior to this analysis, the *C. gigas* genome was annotated. A `blastx` alignment was performed against the Uniprot-SwissProt database (accessed June 01, 2021) to get Uniprot Accession information for each transcript [45]. The transcript nucleotide sequences were extracted from the genome (GCF_902806645.1_cgigas_uk_roslin_v1_rna_from_genomic.fna available in the NCBI Annotation Release 102). Transcript IDs from the `blastx` output were matched with GO terms from the Uniprot-Swissprot database using Uniprot Accession codes. These transcript IDs and corresponding GO terms were then used to create a gene ID-to-GO term database

(geneID2GO) for manual GO term annotation. The transcript IDs were filtered for genes that contained CpGs with at least 5x coverage in all samples, as the same parameters were used to generate the CpG background for differential methylation analysis. Each line of the database contained transcript ID in one column, and all corresponding GO terms in another column.

Gene enrichment analysis was conducted with topGO v.2.34.0 in R [46]. The geneID2GO database was used as the gene universe for enrichment, while transcript IDs for genes with DML were used as the genes of interest. These transcript IDs were filtered such that they did not include any transcripts associated with DML-SNP overlaps. A topGO object was generated for each DML list and GO category (biological process, cellular component, or molecular function), with GO term annotation performed using the geneID2GO database. A Fisher's exact test was used to identify GO terms in each DML list significantly enriched with respect to the gene background (*P*-value < 0.01). Keeping with topGO default settings, we did not correct for multiple comparisons [46]. Enriched GOterms were clustered by semantic similarity using simplifyEnrichment v.1.2.0 default settings [47]. A semantic similarity matrix was created from enriched GO IDs using the "Relevance" similarity measure and GO tree topology (GO_similarity). The semantic similarity matrix was clustered using the binary cut method, and visualized as a word cloud alongside a heatmap of semantic similarity (simplifyGO). Finally, the enriched GOterms and cluster information were appended to a list of genic DML and annotations to understand which genes with DML also had enriched GOterms.

List of Abbreviations

WGBS: whole genome bisulfite sequencing

DML: differentially methylated locus/loci

SNP: single nucleotide polymorphism

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data, genome feature tracks, scripts, and a supplementary materials list are available in the Oyster Gonad Methylation repository, doi.org/10.17605/OSF.IO/YGCTB. All raw data can be accessed at the NCBI Sequence Read Archive under BioProject accession number PRJNA806944.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was funded by National Science Foundation award 1634167 to SBR. The Hall Conservation Genetics Research Fund (YRV) supported sequencing for this project.

Authors' contributions

YRV conceived the experimental design and conducted the pH exposure experiment with assistance from SBR. YRV extracted DNA for sequencing and analyzed sequence data with SJW and SBR. YRV wrote the initial manuscript with input from SJW and SBR. All authors reviewed and approved the final manuscript.

Acknowledgements

Grace Crandall performed histological analysis and Kaitlyn Mitchell assisted with DNA extractions. This work was facilitated through the use of advanced computational, storage, and networking infrastructure provided by the Hyak supercomputer system at the University of Washington.

Additional Files

Additional File 1 (.csv)

Sequencing summary information.

Descriptive statistics of sequencing reads and results of sequencing adapter trimming.

Additional File 2 (.csv)

Contingency test results for highly methylated CpGs.

Test statistics and P-values from chi-squared tests comparing genomic locations of highly methylated CpGs and all CpGs with 5x coverage detected by WGBS.

Additional File 3 (.pdf)

Principal components analysis (PCA) of CpG loci.

PCA of CpG methylation for loci covered at 5x read depth in all samples, with colors differentiating low and ambient pH conditions and shape indicating maturation stage. The PCA did not demonstrate any clear separation of samples by treatment.

Additional File 4 (.pdf)

Matrix of pairwise scatter plots for 5x CpG loci.

Data is presented for CpG covered at > 5x across all samples. Pearson correlation coefficients for each pairwise comparison are presented in the upper right boxes.

Additional File 5 (.csv)

Number of differentially methylated loci in each chromosome.

Number of DML, genes, and CpGs present in the ten *C. gigas* chromosomes and scaffolds not placed in any linkage group.

Additional File 6 (.csv)

Differentially methylated loci per gene.

Gene ID and the number of DML located in that gene.

Additional File 7 (.csv)

Contingency test results for DML.

Test statistics and P-values from chi-squared tests comparing genomic locations of DML and all CpGs with 5x coverage detected by WGBS.

Additional File 8 (.xlsx)

Biological process GOterm enrichment results.

Enrichment analysis p-values and annotation of genes containing significant biological process GOterms and differentially methylated loci.

Additional File 9 (.xlsx)

Cellular component GOterm enrichment results.

Enrichment analysis p-values and annotation of genes containing significant cellular component GOterms and differentially methylated loci

References

1. Eirin-Lopez JM, Putnam HM. Marine Environmental Epigenetics. Ann Rev Mar Sci. 2018. https://doi.org/10.1146/annurev-marine-010318-095114.

2. Hofmann GE. Ecological Epigenetics in Marine Metazoans. Frontiers in Marine Science. 2017;4.

3. Gavery MR, Roberts SB. A context dependent role for DNA methylation in bivalves. Brief Funct Genomics. 2014;13:217–22.

4. Chandra Rajan K, Yuan M, Yu Z, Roberts SB, Thiyagarajan V. Oyster biomineralisation under ocean acidification: from genes to shell. Glob Chang Biol. 2021. https://doi.org/10.1111/gcb.15675.

5. Bogan SN, Johnson KM, Hofmann GE. Changes in Genome-Wide Methylation and Gene Expression in Response to Future pCO_2 Extremes in the Antarctic Pteropod *Limacina helicina antarctica*. Frontiers in Marine Science. 2020;6:788.

6. Downey-Wall AM, Cameron LP, Ford BM, McNally EM, Venkataraman YR, Roberts SB, et al. Ocean Acidification Induces Subtle Shifts in Gene Expression and DNA Methylation in Mantle Tissue of the Eastern Oyster (*Crassostrea virginica*). Frontiers in Marine Science. 2020;7:828.

7. Lim Y-K, Cheung K, Dang X, Roberts SB, Wang X, Thiyagarajan V. DNA methylation changes in response to ocean acidification at the time of larval metamorphosis in the edible oyster, *Crassostrea hongkongensis*. Mar Environ Res. 2020;163:105214.

8. Venkataraman YR, Downey-Wall AM, Ries J, Westfield I, White SJ, Roberts SB, et al. General DNA methylation patterns and environmentally-induced differential methylation in the eastern oyster (*Crassostrea virginica*). Frontiers in Marine Science. 2020.

https://doi.org/10.3389/fmars.2020.00225.

9. Liew YJ, Zoccola D, Li Y, Tambutté E, Venn AA, Michell CT. Epigenome-associated phenotypic acclimatization to ocean acidification in a reef-building coral. Science Advances. 2018.

10. Arredondo-Espinoza R, Ibarra AM, Roberts SB, Sicard-Gonzalez MT, Escobedo-Fregoso C. Differentially methylated gene regions between resistant and susceptible heat-phenotypes of the Pacific oyster *Crassostrea gigas*. Aquaculture. 2021;543:736923.

11. Strader ME, Wong JM, Kozal LC, Leach TS, Hofmann GE. Parental environments alter DNA methylation in offspring of the purple sea urchin, *Strongylocentrotus purpuratus*. J Exp Mar Bio Ecol. 2019;517:54–64.

12. Johnson KM, Sirovy KA, Casas SM, La Peyre JF, Kelly MW. Characterizing the Epigenetic and Transcriptomic Responses to Perkinsus marinus Infection in the Eastern Oyster *Crassostrea virginica*. Frontiers in Marine Science. 2020;7.

13. Rondon R, Grunau C, Fallet M, Charlemagne N, Sussarellu R, Chaparro C, et al. Effects of a parental exposure to diuron on Pacific oyster spat methylome. Environ Epigenet. 2017;3.

14. Liew YJ, Howells EJ, Wang X, Michell CT, Burt JA, Idaghdour Y, et al. Intergenerational epigenetic inheritance in reef-building corals. bioRxiv. 2018;:269076.

15. Melzner F, Mark FC, Seibel BA, Tomanek L. Ocean Acidification and Coastal Marine Invertebrates: Tracking CO_2 Effects from Seawater to the Cell. Ann Rev Mar Sci. 2019. https://doi.org/10.1146/annurev-marine-010419-010658.

16. Parker LM, Ross PM, O'Connor WA, Pörtner HO, Scanes E, Wright JM. Predicting the response of molluscs to the impact of ocean acidification. Biology . 2013;2:651–92.

17. Kroeker KJ, Kordas RL, Crim RN, Singh GG. Meta-analysis reveals negative yet variable effects of ocean acidification on marine organisms. Ecol Lett. 2010;13:1419–34.

18. Gavery MR, Roberts SB. Predominant intragenic methylation is associated with gene expression characteristics in a bivalve mollusc. PeerJ. 2013;1:e215.

19. Roberts SB, Gavery MR. Is there a relationship between DNA methylation and phenotypic plasticity in invertebrates? Front Physiol. 2012;2.

20. Riviere G, Wu G-C, Fellous A, Goux D, Sourdaine P, Favrel P. DNA methylation is crucial for the early development in the oyster *C. gigas*. Mar Biotechnol . 2013;15:739–53.

21. Olson CE, Roberts SB. Genome-wide profiling of DNA methylation and gene expression in *Crassostrea gigas* male gametes. Front Physiol. 2014;5:224.

22. Olson CE, Roberts SB. Indication of family-specific DNA methylation patterns in developing oysters. bioRxiv. 2015;:012831.

23. Saint-Carlier E, Riviere G. Regulation of Hoxorthologues in the oyster *Crassostrea gigas* evidences a functional role for promoter DNA methylation in an invertebrate. FEBS Letters. 2015;589:1459–66.

24. Gavery MR, Roberts SB. Epigenetic considerations in aquaculture. PeerJ. 2017;5:e4147.

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

26. Song K, Li L, Zhang G. The association between DNA methylation and exon expression in the Pacific oyster *Crassostrea gigas*. PLoS One. 2017;12:e0185224.

27. Venkataraman YR, Spencer LH, Roberts SB. Larval Response to Parental Low pH

Exposure in the Pacific Oyster *Crassostrea gigas*. Journal of Shellfish Research. 2019;38:743.

28. Fabioux C, Huvet A, Le Souchu P, Le Pennec M, Pouvreau S. Temperature and photoperiod drive *Crassostrea gigas* reproductive internal clock. Aquaculture. 2005;250:458–70.

29. Enríquez-Díaz M, Pouvreau S, Chávez-Villalba J, Le Pennec M. Gametogenesis, reproductive investment, and spawning behavior of the Pacific giant oyster *Crassostrea gigas*: evidence of an environment-dependent strategy. Aquac Int. 2008;17:491.

30. Peñaloza C, Gutierrez AP, Eöry L, Wang S, Guo X, Archibald AL, et al. A chromosome-level genome assembly for the Pacific oyster *Crassostrea gigas*. Gigascience. 2021;10.

31. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26:841–2.

32. Tarailo-Graovac M, Chen N. Using RepeatMasker to Identify Repetitive Elements in Genomic Sequences. Current Protocols in Bioinformatics. 2009.

33. Smit AFA, Hubley R, Green P. 2013--2015. RepeatMasker Open-4.0. 2013.

34. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011;17:10–2.

35. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010.

36. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016;32:3047–8.

37. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–9.

38. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq

applications. Bioinformatics. 2011;27:1571–2.

39. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25:2078–9.

40. Gao S, Zou D, Mao L, Liu H, Song P, Chen Y, et al. BS-SNPer: SNP calling in bisulfite-seq data. Bioinformatics. 2015;31:4006–8.

41. R Core Team. R: A language and environment for statistical computing. 2019.

42. Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, et al. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. Genome Biol. 2012;13:R87.

43. Wang H-Q, Tuominen LK, Tsai C-J. SLIM: a sliding linear model for estimating the proportion of true null hypotheses in datasets with dependence structures. Bioinformatics. 2011;27:225–31.

44. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 2013;14:178–92.

45. UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 2019;47:D506–15.

46. Alexa A, Rahnenfuhrer J, Others. topGO: enrichment analysis for gene ontology. R package version. 2010;2:2010.

47. Gu Z, Hübschmann D. simplifyEnrichment: an R/Bioconductor package for Clustering and Visualizing Functional Enrichment Results. bioRxiv. 2021;:2020.10.27.312116.

48. Dheilly NM, Lelong C, Huvet A, Kellner K, Dubos M-P, Riviere G, et al. Gametogenesis in

the Pacific oyster Crassostrea gigas: a microarrays-based analysis identifies sex and stage specific genes. PLoS One. 2012;7:e36353.

49. Timmins-Schiffman E, Coffey WD, Hua W, Nunn BL, Dickinson GH, Roberts SB. Shotgun proteomics reveals physiological response to ocean acidification in *Crassostrea gigas*. BMC Genomics. 2014;15:951.

50. Goncalves P, Jones DB, Thompson EL, Parker LM, Ross PM, Raftos DA. Transcriptomic profiling of adaptive responses to ocean acidification. Mol Ecol. 2017;26:5974–88.

51. De Wit P, Durland E, Ventura A, Langdon CJ. Gene expression correlated with delay in shell formation in larval Pacific oysters (*Crassostrea gigas*) exposed to experimental ocean acidification provides insights into shell formation mechanisms. BMC Genomics. 2018;19:160.

52. Broquard C, Saowaros S-A, Lepoittevin M, Degremont L, Lamy J-B, Morga B, et al. Gonadal transcriptomes associated with sex phenotypes provide potential male and female candidate genes of sex determination or early differentiation in *Crassostrea gigas*, a sequential hermaphrodite mollusc. BMC Genomics. 2021;22:609.

53. de Lorgeril J, Zenagui R, Rosa RD, Piquemal D, Bachère E. Whole transcriptome profiling of successful immune response to Vibrio infections in the oyster *Crassostrea gigas* by digital gene expression analysis. PLoS One. 2011;6:e23142.

54. Moreira A, Figueira E, Mestre NC, Schrama D, Soares AMVM, Freitas R, et al. Impacts of the combined exposure to seawater acidification and arsenic on the proteome of *Crassostrea angulata* and *Crassostrea gigas*. Aquat Toxicol. 2018;203:117–29.

55. Yao H, Lin Z, Dong Y, Kong X, He L, Xue L. Gonad Transcriptome Analysis of the Razor Clam (*Sinonovacula constricta*) Revealed Potential Sex-Related Genes. Frontiers in Marine Science. 2021;8:1246.

56. Pauletto M, Milan M, Huvet A, Corporeau C, Suquet M, Planas JV, et al. Transcriptomic features of *Pecten maximus* oocyte quality and maturation. PLoS One. 2017;12:e0172805.

57. Epelboin Y, Quintric L, Guévélou E, Boudry P, Pichereau V, Corporeau C. The Kinome of Pacific Oyster *Crassostrea gigas*, Its Expression during Development and in Response to Environmental Factors. PLoS One. 2016;11:e0155435.

58. Kingtong S, Kellner K, Bernay B, Goux D, Sourdaine P, Berthelin CH. Proteomic identification of protein associated to mature spermatozoa in the Pacific oyster *Crassostrea gigas*. J Proteomics. 2013;82:81–91.

59. Boonmee A, Heude Berthelin C, Kingtong S, Pauletto M, Bernay B, Adeline B, et al. Differential protein expression during sperm maturation and capacitation in an hermaphroditic bivalve, *Pecten maximus* (Linnaeus, 1758). J Molluscan Stud. 2016;82:575–84.

60. Black RE. The Concentrations Of Some Enzymes Of The Citric Acid Cycle And Electron Transport System In The Large Granule Fraction Of Eggs And Trochophores Of The Oyster, *Crassostrea virginica*. Biol Bull. 1962;123:71–9.

61. Silliman K. Population structure, genetic connectivity, and adaptation in the Olympia oyster (*Ostrea lurida*) along the west coast of North America. Evol Appl. 2019;12:923–39.

62. Nilsson B, Jepsen PM, Bucklin A, Hansen BW. Environmental stress responses and experimental handling artifacts of a model organism, the copepod *Acartia tonsa* (Dana). Front Mar Sci. 2018;5.

63. Sirotkin V, Seipel S, Krendel M, Bonder EM. Characterization of sea urchin unconventional myosins and analysis of their patterns of expression during early embryogenesis. Mol Reprod Dev. 2000;57:111–26.

64. Padilla-Gamiño JL, Kelly MW, Evans TG, Hofmann GE. Temperature and CO_2 additively regulate physiology, morphology and genomic responses of larval sea urchins, *Strongylocentrotus purpuratus*. Proceedings of the Royal Society B: Biological Sciences. 2013;280:20130155.

65. Evans TG, Pespeni MH, Hofmann GE, Palumbi SR, Sanford E. Transcriptomic responses to seawater acidification among sea urchin populations inhabiting a natural pH mosaic. Mol Ecol. 2017;26:2257–75.

66. Zhang K-Y, Yuan W-J, Xu J-D, Wang J-X. Cation-dependent mannose-6-phosphate receptor functions as a pattern recognition receptor in anti-bacterial immunity of *Marsupenaeus japonicus*. Dev Comp Immunol. 2018;89:122–30.

67. Leprêtre M, Almunia C, Armengaud J, Salvador A, Geffard A, Palos-Ladeiro M. The immune system of the freshwater zebra mussel, *Dreissena polymorpha*, decrypted by proteogenomics of hemocytes and plasma compartments. J Proteomics. 2019;202:103366.

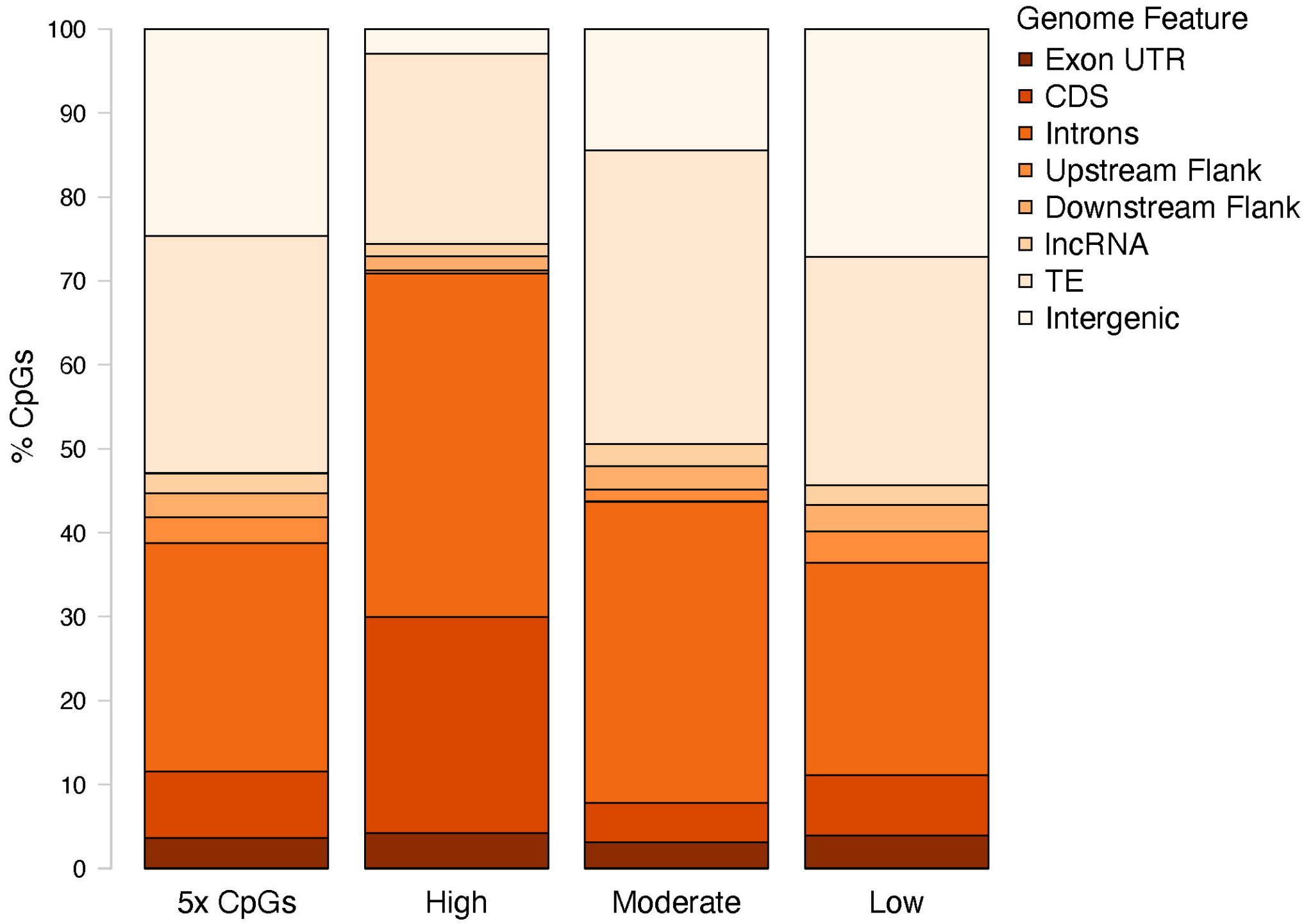
68. Tomanek L, Zuzow MJ, Ivanina AV, Beniash E, Sokolova IM. Proteomic response to elevated pCO_2 level in eastern oysters, *Crassostrea virginica*: evidence for oxidative stress. J Exp Biol. 2011;214 Pt 11:1836–44.

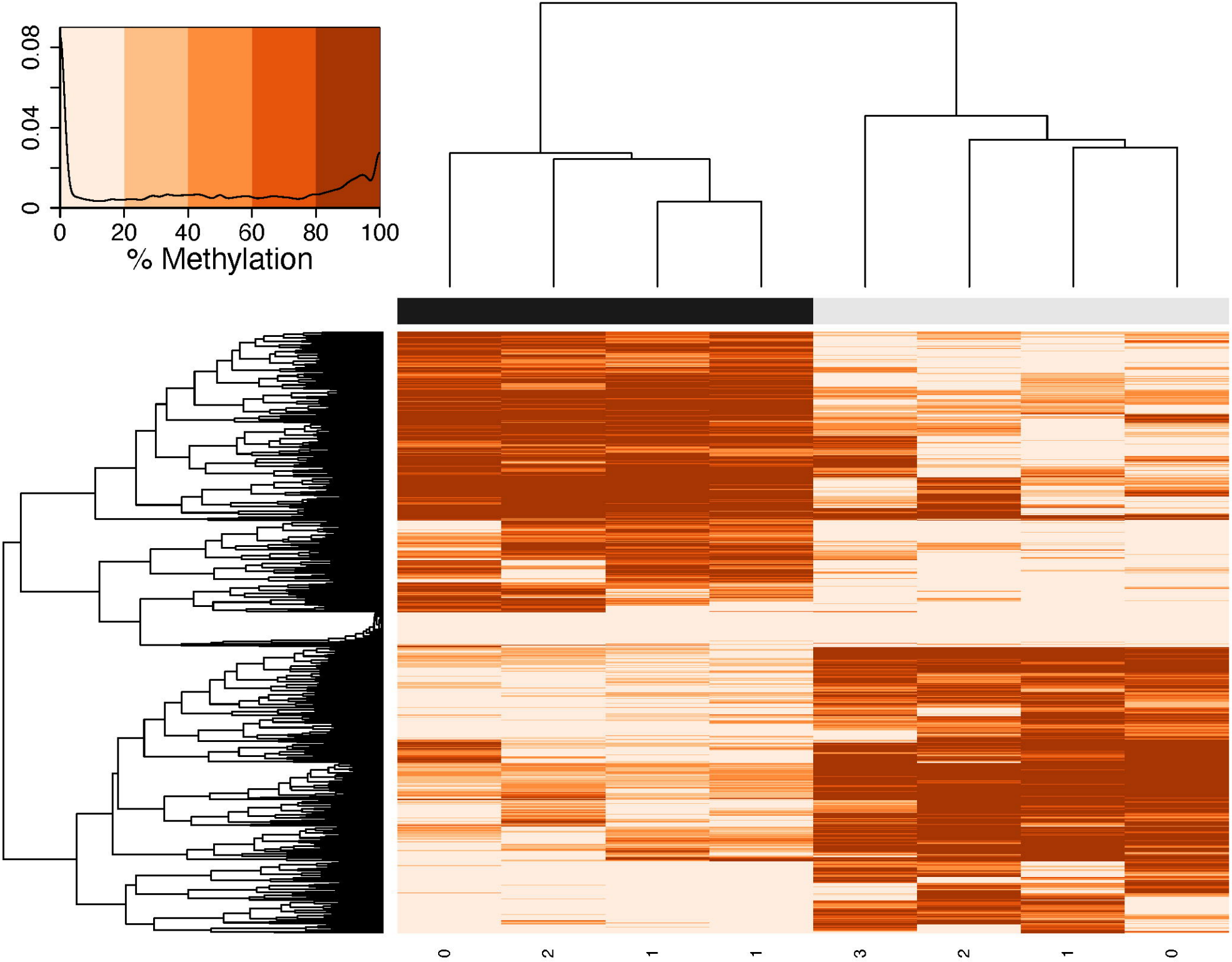
69. Jiang G, Xu X, Jing Y, Wang R, Fan T. Comparative studies on sorting cells from Artemia sinica at different developmental stages for in vitro cell culture. In Vitro Cell Dev Biol Anim. 2011;47:341–5.

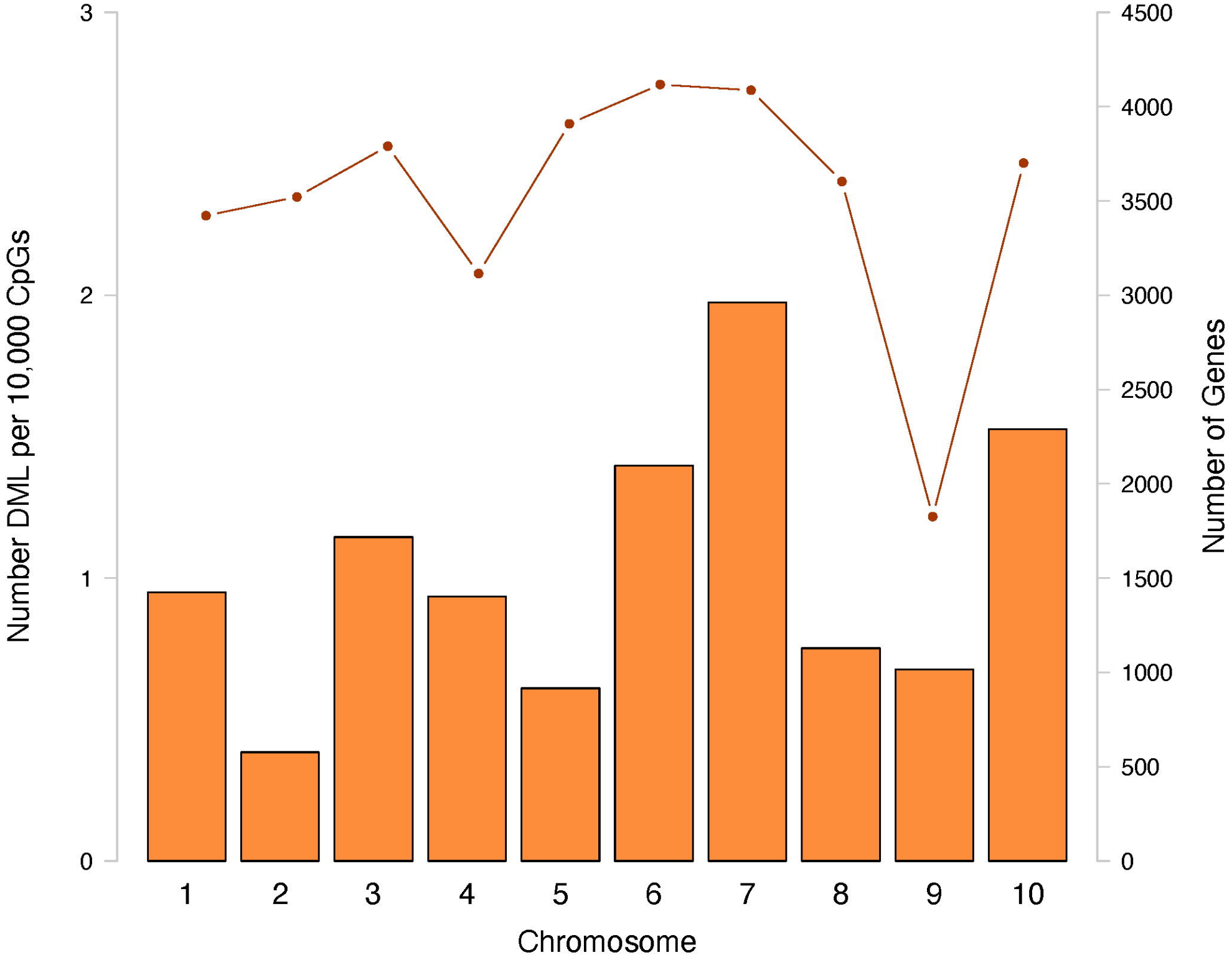
70. Bogan SN, Strader ME, Hofmann GE. Gene regulation by DNA methylation is contingent on chromatin accessibility during transgenerational plasticity in the purple sea urchin. bioRxiv. 2021;:2021.09.23.461091.

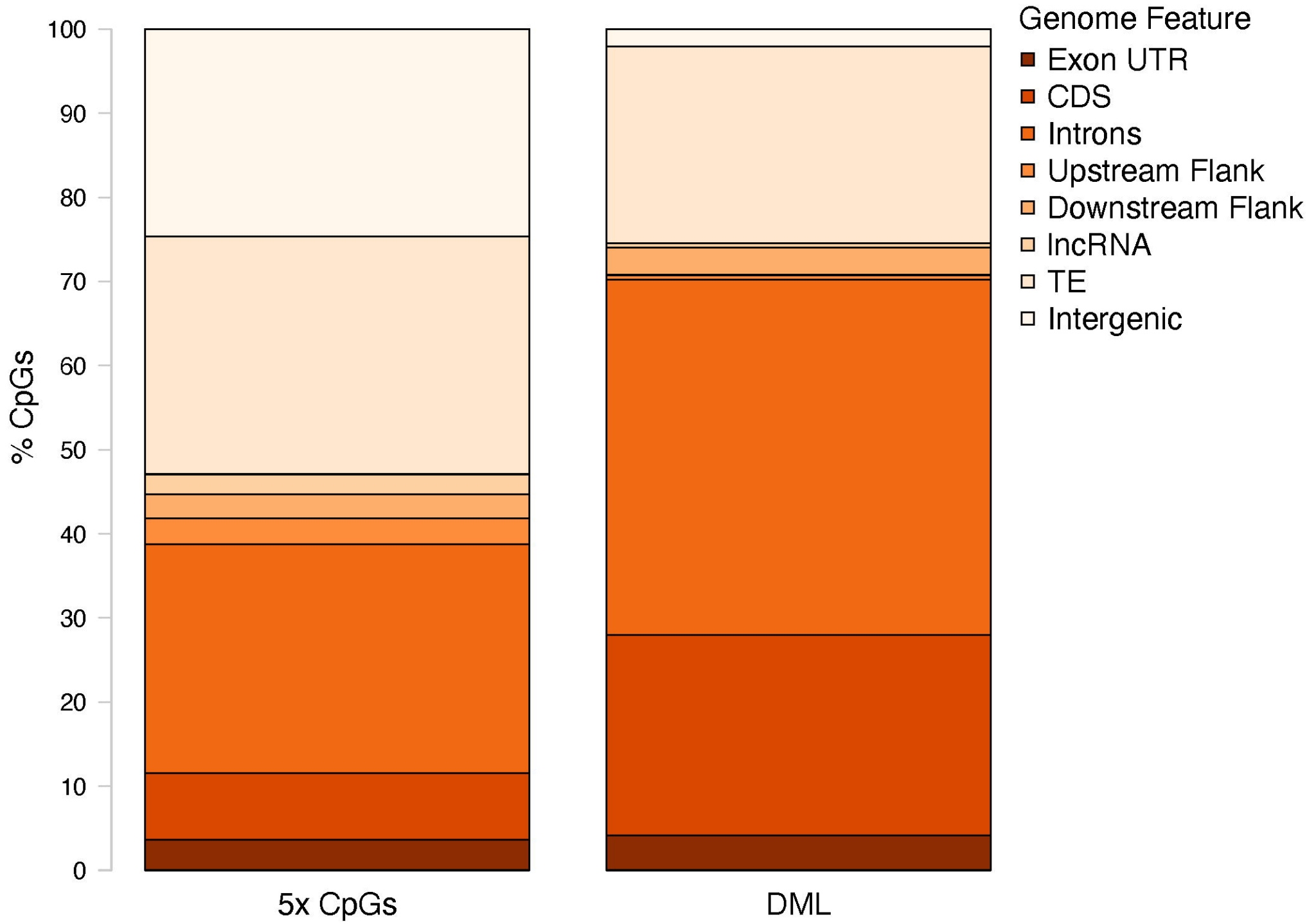
71. Johnson KM, Sirovy KA, Kelly MW. Differential DNA methylation across environments has

no effect on gene expression in the eastern oyster. J Anim Ecol. 2021. https://doi.org/10.1111/1365-2656.13645.


Genome Feature
Exon UTR
CDS
Introns
Upstream Flank
Downstream Flank
lncRNA
TE
Intergenic
% CpGs
100
90
80
70
60
50
40
30
20
10
0
5x CpGs
High
Moderate
Low


0.08
0.04
0
0
20
40
60
80
100
% Methylation
0
2
1
1
3
2
1
0


Number DML per 10,000 CpGs
Number of Genes
Chromosome
0
1
2
3
0
500
1000
1500
2000
2500
3000
3500
4000
4500
1
2
3
4
5
6
7
8
9
10


Genome Feature
Exon UTR
CDS
Introns
Upstream Flank
Downstream Flank
lncRNA
TE
Intergenic
% CpGs
100
90
80
70
60
50
40
30
20
10
0
5x CpGs
DML