Indication of family-specific DNA methylation patterns in developing oysters

Claire E. Olson and Steven B. Roberts
School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA, USA

Correspondence: Steven B. Roberts, School of Aquatic and Fishery Sciences, 1122 Boat Street, Box 355020, Seattle, WA 98105-5020, USA. Email: sr320@u.washington.edu

Abstract

The roles of DNA methylation during invertebrate development remain enigmatic, especially regarding the inheritance and ontogenetic dynamics of methylation. Here, we characterized the genome-wide methylome of *Crassostrea gigas* sperm and larvae from two full-sib families nested within a maternal half-sib family across several developmental stages. Our data suggest that DNA methylation patterns are inherited, as methylation patterns were similar between the two sires and their offspring. Loci differing between the two paternal full-sib families (189) and among the developmental stages (160) were found throughout the genome but were concentrated in transposable elements and repeat regions. We suggest that the predominance of differentially methylated loci within transposable elements is a result of selection against changes in gene body methylation.

Introduction

DNA methylation is an epigenetic modification that is ubiquitous across many eukaryotes, with variable patterns and functions across taxa. This epigenetic mechanism involves the addition of a methyl group to cytosines, usually in CpG dinucleotides, catalyzed by DNA methyltransferases. Epigenetic modifications such as DNA methylation can alter gene expression without modifying the underlying nucleotide sequence, and functions in mammals to suppress transcription through increased

methylation in promoter regions (Bell and Felsenfeld, 2000). In mammals, DNA methylation is essential for development and differentiation of organs and tissues (Okano et al. 1999). Likewise, mutations of DNA methyltransferase in mammals may result in developmental delays and mortality (Li et al. 1992).

In contrast to the densely methylated mammalian genomes, several invertebrate species display low to intermediate levels of methylation. In invertebrates, it has been proposed that DNA methylation of genes may be associated with alternative splicing events (i.e. honey bee (Lyko et al. 2010) and *Nasonia* (Park et al. 2011)). Methylation of gene bodies has also been shown to have a positive relationship with transcriptional activity in oysters (Gavery and Roberts, 2013; Olson and Roberts, 2014a). Currently there is an incomplete understanding of the regulation of gene expression by DNA methylation in invertebrates, though it appears to be distinct from mechanisms observed in mammals and likely varies across species.

From the limited studies that have focused on invertebrates, research has shown that similar to mammals, DNA methylation has important regulatory functions during early development. For example, research on the honey bee *Apis mellifera* found DNA methylation to be abundant in the genome, with methylation being associated with altered gene expression resulting in bee caste differentiation (Elango et al. 2009; Kucharski et al. 2008). Furthermore, DNA methylation has been shown to regulate gene expression during *Octopus vulgaris* development, particularly during the first paralarval stage that includes significant morphological changes (Diaz-Freije et al. 2014). The first indication that methylation was an important regulator of development in *C. gigas* was by Riviere et

al. (2013), who found treatment with 5-Aza-cytidine leads to developmental alterations and abnormal phenotypes in oysters.

Despite the essential role of methylation in development, there is limited information on individual variation in DNA methylation patterns among invertebrates and particularly how any methylation information might be passed on to offspring. Furthermore, we do not have a full understanding of ontogenetic changes in DNA methylation. In order to better understand to what degree DNA methylation patterns are heritable, variable between individuals, and changing during *C. gigas* development, we analyzed genome-wide DNA methylation in gametes and larval oysters (72 and 120 hours post-fertilization) from two full-sib families.

Methods

Experimental Design

Oysters (two males and a single female) were collected from Oakland Bay, South Puget Sound, WA. Oysters were strip spawned by scoring the gonad tissue with a sterile razor blade and rinsing out gametes with sterile seawater. Oocytes were incubated for 30 minutes in sterile seawater and 2 million oocytes each were placed into two separate plastic containers. Sperm diluted in sterile seawater (1L) from each male were used to fertilize oocytes. Fertilization was confirmed by examining polar bodies in cells under a compound microscope.

Larvae were kept in static tanks (100L) up to 120 hours post-fertilization (hpf). Counts of oyster larvae were performed at 120 hpf to confirm normal development. Two samples for DNA methylation analyses were taken from sperm prior to fertilization, and larvae samples collected at days 72 hpf and 120 hpf. Larvae samples

were taken by filtering on a 20μm screen. All samples were preserved in 95% ethanol. For simplicity the sperm and corresponding larvae samples are referred to as family #1 and family #3 based on paternity.

Bisulfite treated DNA Sequencing (BS-Seq)

Genomic DNA was extracted using DNAzol according to the manufacturer's protocol (Molecular Research Center, Inc., Cincinnati, OH). High molecular weight genomic DNA (6 ug per sample), was used to prepare six libraries for whole-genome bisulfite sequencing. Briefly, DNA was fragmented to an average length of 250 bp in an Adaptive Focused Acoustics (AFA) microtube using a Covaris S2 (Covaris Inc Woburn, MA) with the following settings: duty cycle 20%, intensity of 4.0, cycles per burst 200, for 60 seconds. Fragment size resulting from DNA shearing was confirmed by gel electrophoresis. Libraries were constructed using the Paired-End DNA Sample Prep Kit (Illumina, San Diego, CA) with standard protocols. Unmethylated Lambda DNA (0.5%) (Promega Co. Madison, WI) was added to the each sample prior to fragmentation and library construction to serve as a measure of bisulfite conversion efficiency. DNA was treated with sodium bisulfite using the EpiTect Bisulfite Kit (Qiagen, Valencia, CA) and 72 bp paired-end sequencing was performed on the Illumina HiSeq 2000 system. Library construction and sequencing was performed by the High Throughput Genomics Center (htSEQ, Seattle, WA).

Bisulfite sequencing reads from the six libraries were quality filtered to remove adapter sequences and separately mapped to the *Crassostrea gigas* genome (version GCA_000297895.1; Zhang et al. 2012) using Bisulfite Sequencing Mapping Program BSMAP v2.74 (Xi and Li 2009). Resulting alignment from mapping bisulfite treated

reads was analyzed with *methratio*, a Python script that accompanies BSMAP. Parameters for *methratio* included reporting loci with zero methylation ratios (-z), combining CpG methylation ratios on both strands (-g) and only using unique mappings (-u). Only CpG loci covered by at least 3 sequenced reads were considered for further analysis. Data can be accessed and computational analysis performed as described via a GitHub repository (Olson and Roberts, 2014b). The IPython notebook in this repository includes all steps necessary for downloading and reproducing the analyses described in this manuscript.

Global DNA Methylation Comparison

Whole-genome DNA methylation correlation and clustering were performed using the program methylKit 0.9.2 (Akalin et al. 2012) in R v3.0.3. Pairwise Pearson's correlation coefficients scores were calculated based on the percent methylation profiles between all pairs of samples. Hierarchical clustering was performed using 1-Pearson's correlation distance of the six methylation profiles.

Differentially Methylated Loci

Differential methylation between the two full-sib families at each locus was determined using Fisher's exact test in methylKit. A CpG locus was determined to be different between families when the difference in methylation ratio between families was more than 25% and p-value <0.01.

Ontogenetic changes in DNA methylation patterns were tested by three pairwise comparisons (Fisher's exact tests in methylKit) between all developmental stages (sperm and 72 hpf larvae, sperm and 120 hpf larvae, and 72 hpf and 120 hpf larvae).

Differentially methylated loci were identified as any CpG with greater than 25% and *p*-value <0.01 for any comparison.

All loci with significantly different methylation ratios across families and ontogenetic stages were characterized with respect to genomic features (intron, exon, promoter region, transposable element) using Bedtools (i.e., intersectBed) (Quinlan and Hall, 2010). Genomic features were developed and reported elsewhere (Gavery and Roberts 2013). Putative promoters were defined as 1kb regions upstream from transcription start sites. Transposable elements were identified using RepeatMasker, a program that screens and annotates interspersed repeats (Smit et al., 1996-2010). Specifically, RepeatProteinMask, was used with repbase which contained 7,445 peptide sequences. RepeatProteinMask also uses Tandem Repeat Finder (Benson, 1999) to identify tandem repeats which were included in the genome feature track. A total of 119,786 features are in the transposable element genome feature file used for analysis in this paper including 61,319 tandem repeat regions and 58,467 transposable elements identified based on sequence similarity. This genome feature track along with intron, exon, and promoter region feature track are all available via the IPython notebook that provides code and data used in this manuscript (Olson and Roberts, 2014b). A Chi-squared test was performed to determine if the distribution of differentially methylated loci among genomic regions (intron, exon, promoter region, transposable elements) was significantly different to the distribution of all CpGs in the oyster genome.

Results

Bisulfite treated DNA Sequencing (BS-Seq)

Bisulfite treated DNA sequence reads are available (NCBI Sequence Read Archive: Study Accession Number SRP028178 - Accession Numbers SRX795174-SRX795179). Sodium bisulfite conversion efficiency was estimated to be 99.9% based on analysis of lambda phage DNA. The number of sequenced cytosines ranged from 2.6×10^7 to 5.3×10^7 across libraries. Using a 3x coverage threshold, most cytosines (75-78%) were determined to be unmethylated (methylation ratio = 0), while 15-18% of the CpG dinucleotides were methylated (methylation ratio ≥ 0.5) (data not shown).

Genome-wide DNA Methylation Comparison

Relationships on a genome-wide scale were assessed by sample correlation and clustering. A total of 40,654 common loci were shared among all six samples and thus analyzed. Methylation ratios were highly correlated between sperm and respective progeny, with a pair-wise Pearson's correlation coefficient (r) of ≥ 0.84 for both families. These similarities within families are also evident in hierarchical clustering (Figure 1), where both male gamete samples were more similar in their methylation profiles to their respective 120 hpf larvae.

Family-specific and Developmental Differences

A total of 189 differentially methylated loci (DMLs) were identified between the two full-sib families. Of these, 102 were found to overlap with a defined genomic region (exon, intron, promoter region, transposable element) (Figure 2). Most CpG loci with different methylation ratios among oyster families were in introns. However, compared to the distribution of CpG dinucleotides in the oyster genome, the proportion of differentially methylated loci within transposable elements was significantly higher than expected ($\chi^2 = 18.84$, df= 1, $p < 0.0001$).

A total of 160 CpG loci showed differences in methylation ratios among developmental stages. Of these loci, 99 were within defined genomic regions (Figure 2). Again, the proportion of differentially methylated loci was significantly greater in transposable elements than in the rest of the oyster genome ($\chi^2 = 9.76$, df= 1, $p < 0.0018$).

Discussion

This study provides the first single-base pair resolution DNA methylomes for both oyster sperm and larval samples from multiple crosses. This research not only provides new information on DNA methylation patterns during oyster development, but also examines its inheritance and changes during early development. Interestingly, our research indicates that epigenetic patterns may differ among oyster families.

Methylation levels in oyster sperm and larvae samples ranged from 15-18% with interspersed regions of both methylated and unmethylated DNA in both male gamete and larval samples. This proportion of CpG methylation falls within the range of that previously reported for oyster male gonad tissue (Olson and Roberts, 2014a) and oyster gill tissue (Gavery and Roberts, 2013). Overall methylation levels are also comparable to those reported among multiple developmental stages of the Pearl oyster *Pinctada fucata* (Li et al. 2014). These findings indicate that overall genome methylation in *C. gigas* is at an intermediate level and suggests that DNA methylation levels do not significantly vary between multiple cell types and life history stages. This is similar to what has been described in global 5-methylcytosine content during different stages of *Ciona intestinalis* development (Suzuki et al. 2013). However, it contrasts with research on mammals and vertebrates, which exhibit the presence of tissue and developmental stage specific methylation profiles, as that seen in zebrafish (McGaughey et al. 2014).

Genome-wide comparisons indicated higher similarity of methylation patterns between oyster families than between developmental stages, suggesting that DNA methylation patterns are inherited. If epigenetic marks are indeed heritable, this mechanism has significant implications for selection. It has been proposed that epigenetic variation may compensate for a decrease in genetic variation in species such as sparrows (Shrey et al. 2012). While outside the scope of the current study, an assessment of relationships between genetic and epigenetic variation is critical. Several studies have examined epigenetic differentiation in vertebrate and plant populations experiencing different environments, indicating evidence for divergent selection in these species (Liu et al. 2012; Herrera and Bazaga 2010). Few studies have focused on invertebrates, though Jiang et al. (2013a) investigated the relationship between genetic and epigenetic variations in two groups of *C. gigas*, a base stock domesticated population and third generation mass selection population. This study demonstrated genetic differentiation between the base population and third generation mass selection populations of oysters, but did not find overall epigenetic variation (Jiang et al. 2013a). Nevertheless, a significant correlation was observed between genetic and epigenetic profiles, with few individuals having similar genetic but distinct epigenetic profiles (Jiang et al. 2013a). Regardless, if epigenetic variation is independent of genotype, mechanisms involved in epigenetic inheritance are not fully understood.

Differences in genome-wide methylation ratios between full-sib families nested within a maternal half-sib family suggested paternal inheritance of DNA methylation patterns. These results are similar to what has been seen in zebrafish where embryos inherit the methylation profile of sperm rather than oocyte (Jiang et al. 2013b, Potok et al.

2013). In contrast, methylation ratios in Pearl oysters are mainly influenced by oocytes, rather than sperm (Li et al. 2014), possibly due to maternal influences on DNA methylation patterns in early larvae, while later stages attain methylation patterns similar to sperm.

Differentially methylated loci across families were distributed throughout the genome, through a higher proportion was found in transposable elements. This concentration of methylation in transposable elements may be due to selection against methylation in functionally important parts of the genome. For instance, many differentially methylated loci in gene bodies could be lethal or deleterious as they would alter gene expression. It should be noted that the role of DNA methylation in regulating genome activity in *C. gigas* is still unclear. However, it has been suggested that elevated methylation decreases spurious transcription of housekeeping genes and limited methylation in inducible genes facilitates multiple transcriptional opportunities (Roberts and Gavery, 2012). In other words, DNA methylation patterns in gene bodies may have evolved over time based on gene function to fit the needs of organisms in highly variable environments, and random changes in these patterns could be detrimental. Furthermore, we suggest that random variations in methylation within transposable elements may have a relatively higher chance of persisting than elsewhere in the genome. Transposable elements are mobile DNA sequences that may be methylated in many species to silence activity (Yoder et al. 1997; Liu and Schmid 1993). Limited information is available about the methylation status of transposable elements in other invertebrate species, but the available studies suggest that transposons are generally unmethylated and contain similar levels of methylation to neighboring DNA (Suzuki and Bird, 2008). This is in agreement

with our previous research, which showed limited DNA methylation in transposable elements in oyster male gamete tissue (Olson and Roberts, 2014a). Assuming that transposable element activity is less critical to survival than coding gene activity, differentially methylated loci in transposable elements may be less likely to have negative selective effects. On the other hand, differentially methylated loci may also provide advantageous phenotypic variation by increasing transposable element mobility. However, such selection hypotheses assume that methylation is introduced randomly, something we do not have evidence for.

A significant proportion of loci differentially methylated during oyster development were found within transposable elements, relative to the distribution of CpG dinucleotides in the oyster genome. Interestingly, we did not observe a high proportion of differentially methylated loci among promoter regions, as would be expected if promoter methylation was regulating gene expression to play a role in oyster development. Recent research has found that DNA methylation of promoter regions specifically reduces expression of *Hox* genes during oyster development (Riviere et al. 2013). Considerable stage-specific differences in total methylation levels during oyster early development indicated that DNA methylation plays a crucial role in oyster embryogenesis (Riviere et al. 2013). We previously found variation in expression levels depending on the level of promoter region methylation (Olson and Roberts, 2014a). Surprisingly, we did not observe any dramatic differences in overall methylation levels during oyster development, nor higher methylation of promoter regions. This discrepancy is likely due to the analysis of different ontogenetic stages, as Riviere et al. (2013) examined the first 24 hours post-fertilization, and our first larval sample was taken at 72 hpf. It is also

possible that only a subset of genes are transcriptionally controlled via DNA methylation and our global approach masked the ability to see differences.

In conclusion, this research suggests epigenetic inheritance as DNA methylation patterns were similar between males and their offspring and differed between oyster families. Interestingly, we found a high proportion of family-specific and stage-specific methylation patterns within transposable elements. Future research should focus on the relationship between epigenetic and genetic variation, and explore the possible relationship of DNA methylation and transposable element activity.

Acknowledgements

This work was supported by an National Science Foundation award (Grant Number 1158119) to SR.

Works cited

Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F. E., Figueroa, M. E., Melnick, A., and Mason, C. E. (2012). methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology* 13, R87.

Bell, A. C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* 405, 482-485.

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic acids research* 27:2.

Diaz-Freije, E., Gestal, C., Castellanos-Martinez, S., and Moran, P. (2014) The role of DNA methylation on *Octopus vulgaris* development and their perspectives. *Frontiers in Physiology* 5: 1-7.

Elango, N., Hunt, B. G., Goodisman, M. A. D., and Yi, S. V. (2009). DNA methylation is widespread and associated with differential gene expression in castes of the honey bee, *Apis mellifera*. *Proceedings of the National Academy of Sciences* 106: 11206-11211.

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

Gavery, M. R., and Roberts, S. B. (2014). A context dependent role for DNA methylation in bivalves. *Briefings in functional genomics* elt054.

Herrera, C. M and Bazaga, P. (2010). Epigenetic differentiation and relationship to adaptive genetic divergence in discrete populations of the violet *Viola cazorlensis*. *New Phytologist* 187: 867-876.

Jiang, Q., Li, Q., Yu, H., and Kong, L.F. (2013). Genetic and epigenetic variation in mass selection populations of Pacific oyster *Crassostrea gigas*. *Genes Genomics* 35, 641–647.

Jiang, L., Zhang, J., Wang, J. J., Wang, L., Zhang, L., Li, G., Yang, X., Ma, X., Sun, X., Cai, J., et al. (2013). Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. *Cell* 153: 773-84.

Kucharski, R., Maleszka, J., Foret, S., and Maleszka, R. (2008). Nutritional Control of Reproductive Status in Honey bees via DNA Methylation. *Science* 319: 1827-1830.

Li, E., Bestor, T. H., Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69: 915-26.

Li, Y., Guan, Y., and He, M. (2014). Analysis of DNA methylation in tissues and development stages of pearl oyster *Pinctada fucata*. *Genes and Genomics* 1-8. doi: 10.1007/s13258-014-0246-1

Liu, S., Sun, K., Jiang, T., Ho, J. P., Liu, B., and Feng, J. (2012). Natural epigenetic variation in the female great roundleaf bat (*Hipposiderous arminger*) populations. *Molecular Genetics and Genomics* 287: 643-50.

Liu, W. M., and Schmid, C. W. (1993). Proposed roles for DNA methylation in Alu transcriptional repression and mutational inactivation. *Nucleic Acid Resources* 21: 1351-9.

Lyko, F., Foret, S., Kucharski, R., Wolf, S., Falckenhayn, C., and Maleszka,R. (2010). The honey bee epigenomes: differential methylation of brain DNA in queens and workers. *PLoS ONE* 8:1000506.

McGaughey, D. M., Abaan, H. O., Miller, R. M., Kropp, P. A., and Brody, L. C. (2014). Genomic of CpG methylation in developing and developed zebrafish. *Genes Genomes Genetics* 4: 861-9.

Okano, M., Bell, D. W., Harber, D. A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99: 247-57.

Olson, C. E. and Roberts, S. B. (2014). Genome-wide profiling of DNA methylation and gene expression in *Crassostrea gigas* male gametes. *Frontiers in Physiology* 5: 1-7.

Olson, C. and Roberts, S. (2014). Olson-ms-nb: v1.0. *Figshare*. doi: http://dx.doi.org/10.6084/m9.figshare.1271360

Park, J. Peng, Z., Zeng, J., Elango, N., Park, T., Wheeler, D., Werren, J. H., and Yi, S. V. (2011). Comparative Analyses of DNA Methylation and Sequence Evolution Using *Nasonia* Genomes. *Molecular Biology and Evolution* 28: 3345-3354.

Potok, M. E., Nix, D. A., Parnell, T. J., and Cairns, B. R. (2013). Reprogramming the Maternal Zebrafish Genome after Fertilization to Match the Paternal Methylation Pattern. *Cell* 153: 759-772.

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842.

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

Roberts, S. B., Gavery, M. R. (2012). Is there a relationship between DNA methylation and phenotypic plasticity in invertebrates? *Frontiers in Physiology,* 2.

Shrey, A. W., Coon, C. A. C., Grispo, M. T., Awad, M., Imboma, T., McCoy, E. D., Mushinsky, H. R., Richards, C. L., and Martin, L. B. (2012). Epigenetic Variation May Compensate for Decreased Genetic Variation with Introductions: A Case Study Using House Sparrows (*Passer domesticus*) on Two Continents. *Genetics Research International* 2012: 1-7.

Smit, A. F. A., Hubley, R., and Green, P. (1996–2010). *RepeatMasker Open-3.0*. Available online at http://www.repeatmasker.org

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

Suzuki, M. M., Yoshinari, A., Obara, M., Takuno, S., Shigenobu, S., Sasakura, Y., et al. (2013). Identical sets of methylated and nonmethylated genes in *Ciona intestinalis* sperm and muscle cells. *Epigenetics Chromatin* 6:38. doi: 10.1186/1756-8935-6-38

Xi, Y., and Li, W. (2009). BSMAP: whole genome bisulfite sequence MAPping program. *BMC Bioinformatics* 10:232.

Yoder, J. A., Walsh, C. P., and Bestor, T. H. (1997). Cytosine methylation and the ecology of intragenomic parasites. *Trends in Genetics* 13: 335-40.

Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., Xu, F., et al. (2012).The oyster genome reveals stress adaptation and complexity of shell formation. *Nature* 490, 49–54.

Figures

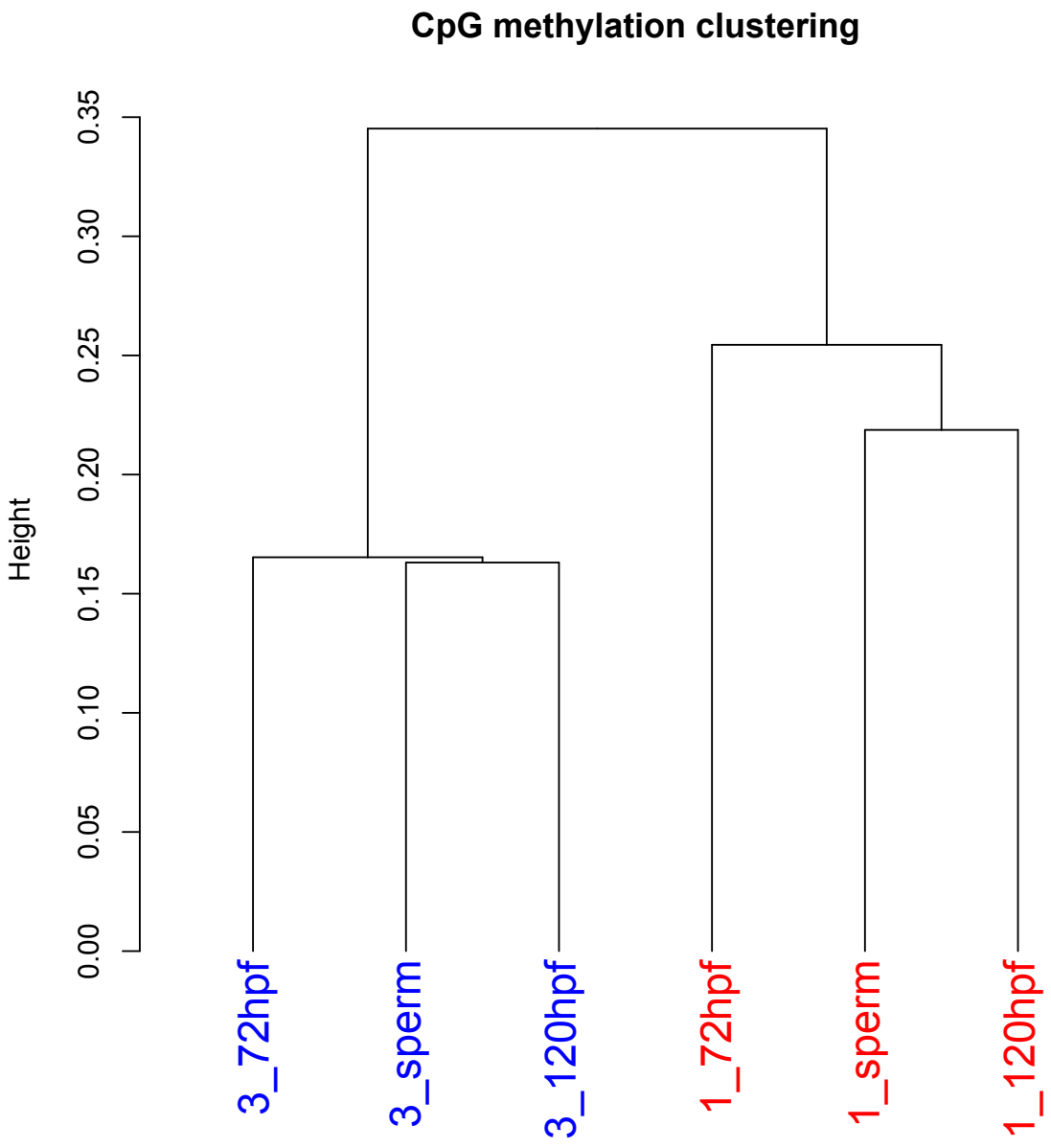

Figure 1. *Dendogram of the male sperm and oyster larvae genome-wide methylation profiles using Pearson's correlation distance*. Numeric prefix refers to family.

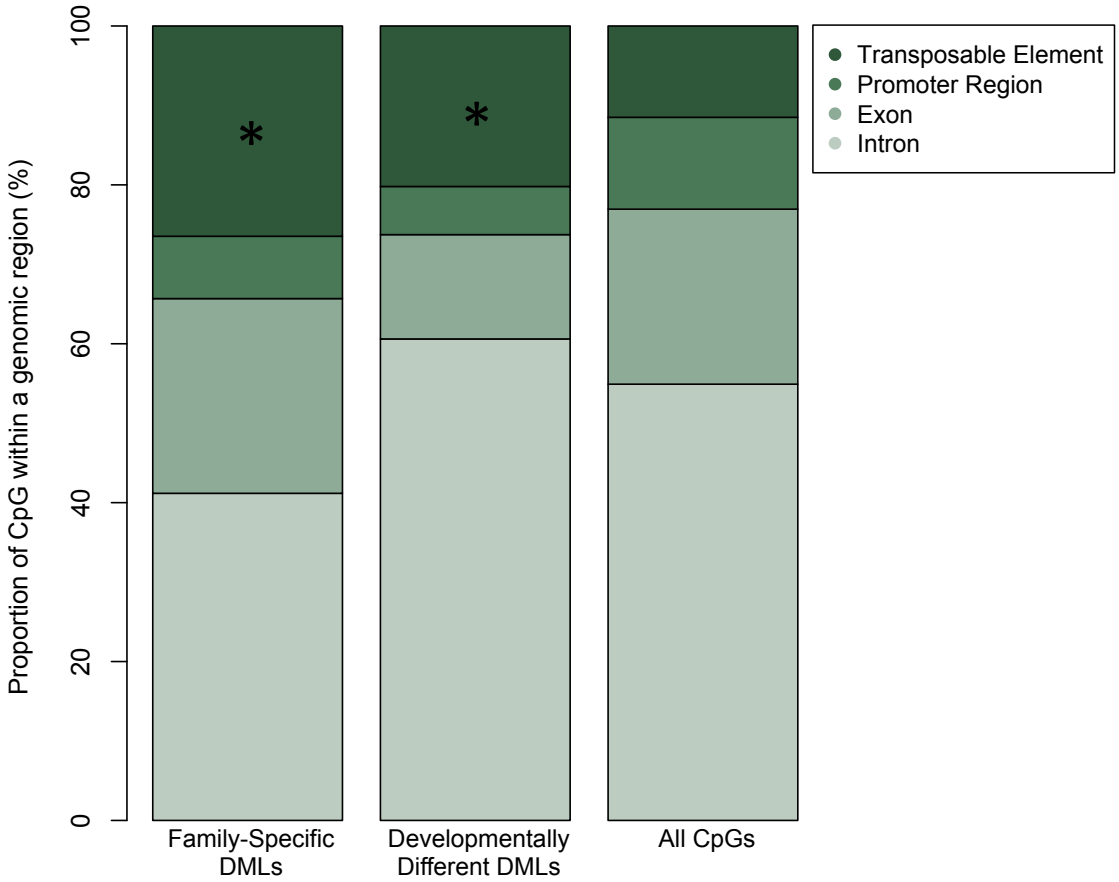

Figure 2. *Proportion of family-specific differentially methylated loci, developmentally different differentially methylated loci, and all CpGs in the oyster genome based on genomic region*. The proportion of loci located within a genomic region (Intron, Exon, Promoter Region, Transposable Element) for differentially methylated loci between families (n= 102), differentially methylated loci among the developmental stages (n= 99), and all CpGs in the oyster genome (n= 10035701) are displayed. An asterisk indicates a statistically different distribution relative to the distribution of all CpGs in the oyster genome.