

## HIGH-THROUGHPUT BISULFITE SEQUENCING REVEALS PATTERNS BETWEEN GENE EXPRESSION AND METHYLATION IN THE COMMERCIALLY IMPORTANT BIVALVE, *CRASSOSTREA GIGAS*

Gavery, M.<sup>1</sup>, and Roberts, S.B.<sup>1</sup>

<sup>1</sup>University of Washington School of Aquatic & Fishery Sciences, Seattle, USA

Email: mgavery@u.washington.edu

### SUMMARY

Epigenetic mechanisms have important functions in developmental gene regulation, response to environment, and natural variation in gene expression levels. As such, examining the role of these mechanisms in cultured aquatic organisms will greatly contribute to our understanding of how phenotypes are generated and maintained in these species (Figure 1). Specifically, our research is focused on understanding how DNA methylation, a well-studied epigenetic mark in mammals, is involved in gene regulation and variation in the economically important bivalve *Crassostrea gigas*.

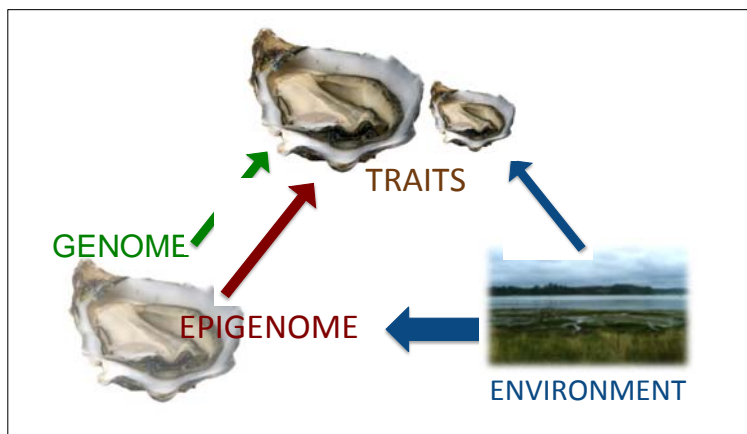


Figure 1. An organism's genotype and the environment it encounters contribute to traits (phenotypes). Increasingly it has been shown that an organism's epigenetic information also contributes to phenotypes. Unlike the genome, the epigenome can be influenced directly by the environment, providing a mechanism whereby an organism may respond or adapt to external changes, and even transmit these induced phenotypes to offspring.

DNA methylation refers to the addition of a methyl group to a cytosine residue, which occurs primarily in a CG dinucleotide context in animals. Previously, we have applied *in silico* approaches to characterize DNA methylation in the Pacific oyster (*Crassostrea gigas*). Our results suggest that there is a relationship between biological function of genes and their methylation patterns. Specifically, we have reported that genes associated with housekeeping functions (i.e. genes whose functions are required by a majority of cell types) have higher methylation than genes involved in inducible functions (i.e. genes involved in response to environment, embryonic development or tissue specific functions) indicating a role in the regulation of gene expression in *C. gigas*. One limitation to these early analyses was that only exonic regions were investigated due to limited availability genomic sequence information.

With the recent release of the *C. gigas* genome and the development of tools facilitating genome-wide epigenomic profiling, we have been able to assess DNA methylation patterns in additional regions of the genomic landscape including introns, promoter regions and repetitive elements. By utilizing high-throughput bisulfite sequencing approaches we have been able to assess the methylation status of 1,453,753 CG dinucleotides across the genome in gill tissue. A majority of these loci (69%) are located in intra-genic regions, including both exons and introns (Figure 2). The remaining 31% are distributed among promoters, repetitive elements and other inter-genic regions. In addition, we combined the DNA methylation data with publicly

available RNA-Seq datasets from *C. gigas* to investigate patterns between DNA methylation and gene expression. The results indicate that intra-genic methylation is significantly higher in genes that are abundantly expressed compared to those that are expressed at lower levels in gill tissue. This finding is consistent with an emerging view that intra-genic DNA methylation is associated with highly expressed genes, expanding the traditional view that DNA methylation is solely associated with the repression of gene expression.

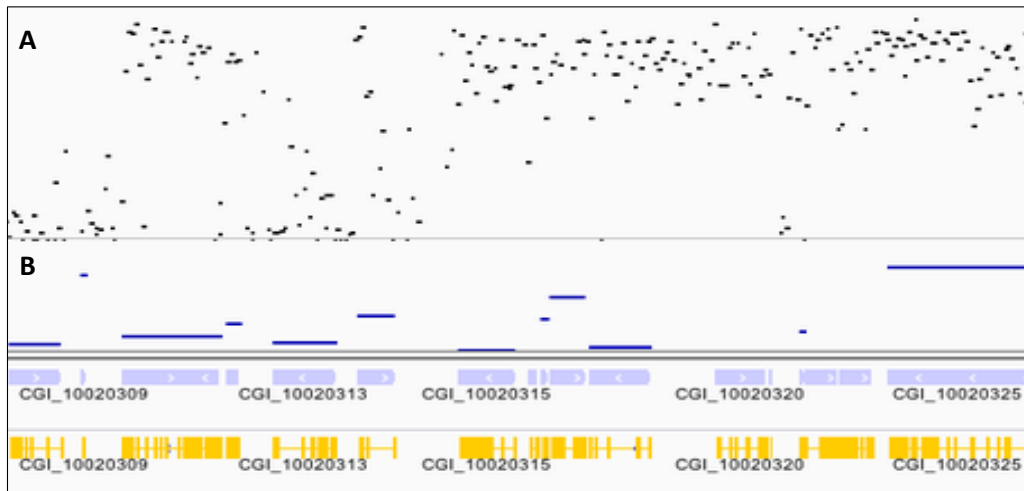


Figure 2. Visual representation of high-throughput bisulfite sequencing data from *C. gigas* gill tissue. DNA methylation ratios at CG loci are indicated by black dots (A) and gene expression values are indicated by blue bars (B). Genes (grey) and exons (yellow) are denoted below.

These data provide the first investigation into genome-wide patterns of methylation and relationships between spatial and functional gene attributes in a mollusc. These foundational studies will provide the basis for future work investigating the role of epigenetic mechanisms during physiological processes such as gametogenesis and stress responses as well as the determination of epigenetic variation among locally adapted populations. In order to facilitate epigenomic research in aquatic species, we are developing publicly available data repositories and genome browsers to promote data querying by the research community.