DNA methylation modulates transcriptional noise in response to elevated pCO_2 in the eastern oyster (*Crassostrea virginica*)

Yaamini R. Venkataraman[1]*, Ariana S. Huffmyer[2,3], Samuel J. White[2], Alan Downey-Wall[4], Jill Ashey[3], Danielle M. Becker[3], Zachary Bengtsson[2], Hollie M. Putnam[3], Emma Strand[3,5], Javier A. Rodríguez-Casariego[6], Shelly A. Wanamaker[5], Kathleen E. Lotterhos[7], Steven B. Roberts[2]

[1]Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA, USA
[2]School of Aquatic and Fisheries Sciences, University of Washington, Seattle, WA, USA
[3]Department of Biological Sciences, University of Rhode Island, Kingston, RI, USA
[4]Embark Veterinary, Boston, MA, USA
[5]Gloucester Marine Genomics Institute, Gloucester, MA USA
[6]Environmental Epigenetics Laboratory, Institute of Environment, Florida International University, Miami, FL, US
[7]Northeastern University Marine Science Center, Nahant, MA, USA

*correspondence: yaamini.venkataraman@whoi.edu

Abstract

Ocean acidification significantly affects marine calcifiers like oysters, warranting the study of molecular mechanisms that contribute to adaptive plasticity in response to environmental change like DNA methylation. However, a consensus has not been reached on the extent to which methylation modules gene expression, and in turn plasticity, in marine invertebrates. To this end, we investigated the impact of pCO_2 on gene expression and DNA methylation in the eastern oyster, *Crassostrea virginica*. After a 30-day exposure to control (572 ppm) or elevated pCO_2 (2,827 ppm), whole genome bisulfite sequencing (WGBS) and RNA-Seq data were generated from adult female gonad and male sperm samples. Although differentially methylated loci (DML) were identified in female (89) and male (2,916) gonad tissues, there were no differentially expressed genes, and only one differentially expressed transcript in females. However, gene body methylation impacted other transcript-level processes in sperm, such as the maximum number of transcripts expressed per gene and shifts in the predominant transcript. Elevated pCO_2 exposure increased gene expression variability (transcriptional noise) in males but decreased noise in females, suggesting a sex-specific role of methylation in gene expression regulation. Functional annotation of genes with changes in transcript-level expression or containing DML revealed several enriched biological processes potentially involved in elevated pCO_2 response, including apoptotic pathways and signal transduction, as well as reproductive functions. Taken together, these results suggest that DNA methylation may regulate gene expression variability to maintain homeostasis in elevated pCO_2 conditions and could play a key role in environmental resilience in marine invertebrates.

Keywords: DNA methylation, gene expression, oysters, ocean acidification, reproduction

Abbreviations: OA: ocean acidification; DEG: differentially expressed gene; DET: differentially expressed transcript; DML: differentially methylated loci

Word Count (excluding Abstract and References): 11,204

Introduction

Ocean acidification (OA) is a pressing environmental issue resulting from the increased absorption of atmospheric carbon dioxide (CO_2) by the world's oceans (1). Increased concentrations of oceanic CO_2 leads to decreases in seawater pH, lower calcium carbonate concentrations, and reductions in aragonite saturation state (1). These environmental alterations pose a significant threat to marine ecosystems, particularly to calcifying organisms such as oysters [2–5]. Reductions in saturation state can compromise calcification and increase the energetic cost of building shells [6–9], and can impact broad physiological pathways such as protein synthesis, energy production, metabolism, antioxidant responses, and reproduction [10–14]. Since oysters are vital for biodiversity, aquaculture, and coastal economies [15–17], understanding the biological responses of these organisms to OA is critical for assessing ecological impacts and developing strategies to mitigate negative effects of OA in oyster restoration and aquaculture production.

The ability of organisms to respond to and cope with OA are dependent on complex changes in whole-organism physiology, behavior, and molecular mechanisms [4,18–21]. Additionally, plasticity, or the capacity for an organism to adjust their physiology to maintain function and performance in response to environmental change [22], is another important factor governing responses to OA. Plastic responses to environmental stressors can also be inherited, providing a mechanism for the parent to pass on “environmental memory,” through mechanisms such as alterations in reproduction and offspring provisioning [23]. There are several examples of adult oyster exposure to OA yielding beneficial plasticity in offspring. For example, Sydney rock oysters (*Saccostrea glomerata*) exposed to high pCO_2 during reproductive conditioning yielded larger and faster developing larvae compared to counterparts in ambient conditions [24,25], and Olympia oyster (*Ostrea lurida*) offspring with parental history of OA exposure had better performance than ambient counterparts when outplanted [26]. However, parental exposure to OA in the Pacific oyster (*Crassostrea gigas*), northern quahog (*Mercenaria mercenaria*), and bay scallop (*Argopecten irradians*) led to poorer larval survival [27,28]. These contrasting results demonstrate that the capacity for intergenerational plasticity is complex, and requires additional research. As adult reproductive tissue gives rise to gametes and offspring, it is critical to examine the molecular mechanisms impacting reproductive homeostasis in response to OA.

Epigenetic modifications are a potential mediator of plastic responses to environmental change. These modifications alter regulation of gene expression without altering the DNA sequence itself, and include multiple interacting mechanisms such as DNA methylation, chromatin remodeling, histone modifications, and non-coding RNAs [29]. DNA methylation — the addition of a methyl group to the cytosine adjacent to a guanine in a CpG dinucleotide [29,30] — is the most studied epigenetic mechanism in marine invertebrates [31]. DNA methylation can change expression of genes by altering binding of transcriptional machinery and can affect variability of transcription [30]. Changes in methylation are thought to be associated with phenotypic changes, with several examples of this connection illustrated in mammalian literature [32]. Methylomes (i.e. information of DNA methylation in a genome) are responsive to OA across a wide variety of taxa, such as pteropods [33], corals [34], sea urchins [35], and oysters [36–40]. In eastern oyster (*Crassostrea virginica*) mantle [36] and gonad [38], Hong kong oyster (*Crassostrea hongkongensis*) mantle [39] and larvae [37,41], and female *C.*

gigas reproductive tissue [40], genes with differential methylation were involved in protein ubiquitination, cytoskeletal regulation, and metabolism. These common processes impacted by DNA methylation across molluscan species and tissue types suggest a role for methylation in OA response.

Even though methylation is environmentally-responsive, direct connections between methylation and plasticity are not readily apparent in marine invertebrates. Where gene expression has been used as a proxy for plasticity, molluscan studies report no relationship between differential gene expression and genes with differential methylation in response to environmental stress [36,39,42–44]. This phenomenon extends beyond molluscan OA literature to other marine invertebrate taxa [32]. One exception is Dang et al. [41], which found significant overlap between differentially expressed genes and differentially methylated genes in juvenile *C. hongkongensis* after exposure to OA conditions as larvae. Differentially expressed and methylated genes were often methylated in introns, and had roles in signal transduction [41]. Studies that do not explore gene expression and methylation have found associations between methylation and phenotype, although causal relationships have not been established. While differences in mantle methylomes cannot be linked changes in calcification [36,39], there is evidence suggesting that differential methylation in response to OA promotes reproductive homeostasis in female *C. gigas* [40], which in turn could be linked to maladaptive plasticity in larval *C. gigas* [28]. Additionally, differential methylation of *C. hongkongensis* larvae exposed to elevated pCO_2 was associated with higher metamorphosis rates but poor substratum selection [37]. Taken together, this suggests that environmentally-responsive DNA methylation may play a role in modulating phenotypic plasticity in adult reproductive tissue, and could contribute to intergenerational carryover effects. Therefore, it is important to investigate knowledge gaps surrounding the role of methylation in altering gene expression and plasticity in oysters exposed to OA during reproductive conditioning.

This study explored the impact of OA on gene expression and DNA methylation in *C. virginica* adult reproductive tissue, and the relationship between gene expression and methylation. We used oysters from an accompanying study, which found that parental exposure to elevated pCO_2 increased growth rates in larvae, especially when those larvae were also reared in elevated pCO_2 conditions [45]. Female oysters in control and elevated pCO_2 conditions did not have any differences in egg quality or size, suggesting that epigenetic mechanisms may have mediated the observed carryover effect [45]. We build on previous work demonstrating environmentally-sensitive methylation in *Crassostrea* spp. reproductive tissue [38,40] by examining sex-specific effects of OA on reproductive tissue responses. As female and male oysters undergo distinct gametogenic processes, and egg and sperm have unique roles in reproduction, it follows that there are sex-specific methylation [46–48] and gene expression [49] patterns in reproductive tissue that would confound findings if analyzed together. We compared female gonad tissue and sperm exposed to control and elevated pCO_2 conditions to investigate 1) the influence of genetic variation on methylation and gene expression; 2) the impact of OA on changes in gene activity and methylation; and 3) the functional role of DNA methylation in regulating gene activity. We define gene activity as encompassing gene-level expression, but also transcript-level processes such as transcript expression, changes in the number of unique transcripts expressed per gene, shifts in the predominant isoform expressed, transcriptional noise, and alternative splicing. By incorporating

transcript-level analyses, we can further elucidate the functional role of methylation in regulating gene expression, and therefore plasticity, in response to environmental change. Through this comprehensive approach, we seek to uncover the molecular basis of oyster responses to OA, providing insights into potential impacts on reproduction and the molecular mechanisms involved in marine invertebrate stress response.

Results

Seawater Chemistry

As previously reported in McNally et al. [45], seawater of control and elevated pCO_2 conditions were 572 ppm and 2827 ppm, respectively. The $\Omega_{aragonite}$ of seawater differed significantly between control and elevated pCO_2 conditions (Welch's 2-sample t-test: t = 12.33, df = 7, p-value < 0.0001).

Genetic Variation

The full dataset of all samples ($female_{control}$ = 8, $female_{exposed}$ = 8, $male_{control}$ = 4, $male_{exposed}$ = 6) was used to examine genetic variation, and the influence of that variation of gene methylation and expression. Even though the full dataset contains different cell types (*i.e.* female gonad tissue and sperm), we wanted to determine the relatedness between female and male oysters. After filtering, we kept 2,343,637 single nucleotide polymorphisms (SNPs) out of 144,873,997 possible sites to generate a pairwise relatedness matrix and calculate genetic distance. There was no significant correlation between genetic distance and methylation (R = -0.16, *P*-value = 0.99) (**Supplementary S1 A**) nor genetic distance and gene expression (R = -0.053, *P*-value = 0.76) (**Supplementary S1 B**). Therefore, genetic distance was not incorporated in downstream analyses of methylation and gene expression.

Gene Activity

Gene and Transcript Expression

Across all female and male samples, 67.57% of RNA reads aligned to the *C. virginica* genome. Differentially expressed genes (DEG) and differentially expressed transcripts (DET) were identified between control and OA-exposed samples in female gonad tissue and male sperm samples separately. There were no DEG in females, but a single DET was identified (**Supplementary File S19**). The transcript was associated with gene ID LOC111134855 (uncharacterized). No DEG nor DET were found in males.

Changes in the maximum number of transcripts expressed per gene between control and OA-exposed samples were characterized as another measure of gene activity. A total of 3,966 genes (about 10% of total expressed genes) showed differences in the maximum number of transcripts expressed between female control and OA-exposed samples (**Table 1**), which was significantly less than what would be expected by random chance (two-sided 1-sample z-test: χ^2 = 24847.36, *P*-value = 0). Under OA conditions, 1,529 (38.6%) genes had a greater number of transcripts expressed per gene, while 2,437 had fewer transcripts expressed per gene (**Table 1**;

Supplementary File S2). A total of 211 biological processes were enriched in genes with more transcripts under elevated pCO_2, including establishment of protein localization and cell adhesion (**Supplementary File 18**). Genes with less transcripts expressed under elevated pCO_2 conditions had 469 enriched biological processes, including import into cell, anatomical structure development, and signaling (**Supplementary File 18**).

In sperm, 3,123 (72.2%) of genes had a greater number of transcripts expressed between control and OA-exposed samples, while 1,204 (27.8%) had fewer transcripts expressed in elevated pCO_2 conditions (**Table 1**; **Supplementary File S2**). Genes with more transcripts expressed in elevated pCO_2 conditions had 391 overrepresented biological processes, including signaling and cytoskeletal organization (**Supplementary File 18**). Genes with fewer transcripts expressed in elevated pCO_2 conditions had 238 enriched biological processes, including organelle organization and intracellular signal transduction (**Supplementary File 18**).

Table 1. Gene and transcript expression patterns attributable to elevated pCO_2 conditions. Changes in maximum number of transcripts are described with respect to elevated pCO_2 exposure, with increases in maximum number of transcripts representing more transcripts expressed in elevated pCO_2-exposed samples, and decreases representing more transcripts expressed in control samples. Genes with multiple (≥ 2 transcripts) were considered for predominant transcript analyses.

Sex	Total Genes Expressed	Genes with More Transcripts Expressed	Genes with Less Transcripts Expressed	Genes with ≥ 2 Transcripts	Genes with a Shift in Predominant Transcript
Female	39,047	1,529	2,437	38,575	2,252
Male	39,299	3,123	1,204	10,559	1,768

Predominant Transcript Identification

Low pH-induced shifts in the predominant transcript expression were assessed separately for female reproductive tissue and sperm, where the predominant transcript was defined as the most highly expressed transcript within a gene. Of the 38,575 genes that had expression data for multiple transcripts in females, 2,252 genes (5.83%) had shifts in the predominant transcript (**Table 1**; **Supplementary File S3**). The number of genes with this shift was significantly less than what would be expected due to chance alone (two-sided 1-sample z-test: $\chi^2 = 30091.12$, *P*-value = 0). A total of 226 biological process GO terms were overrepresented in genes where the predominant transcript shifted, including regulation of small GTPase mediated signaling, regulation of Ras protein signal transduction, and reproduction (**Supplementary File 18**).

For sperm, 1,768 (16.7%) out of 10,559 genes with data for multiple transcripts had shifts in the predominant transcript (**Table 1**; **Supplementary File S3**). Similar to the female oysters, the number of genes with a shift in the predominant isoform was significantly less than what would be expected due to random chance (two-sided 1-sample z-test: $\chi^2 = 4669.81$, *P*-value = 0). Thirty-three biological processes were enriched in genes with shifts in the

predominant transcript, including chemotaxis and epigenetic regulation of gene expression (**Supplementary File 18**).

Alternative Splicing

For each sex, we described patterns of alternative splicing by testing for multivariate variation across exon position and treatment using ANOVA-simultaneous components analyses (ASCA). The multivariate response measured was the expression exons 2 through 6 relative to expression of the first exon for every gene in each sample. The patterns identified by ASCA (i.e., principal components, or PCs) represent patterns of expression across exons. Of the ten PCs identified for each sex, PCs 1-6 explained > 1% of the variation in relative exon expression (**Supplementary File S5 AB**). In both sexes, PC1-4 described alternative splicing patterns that did not differ by treatment (**Figure 1**, **Supplementary File S6**) and PC5-6 described alternative splicing patterns that differed between treatments (**Figure 2, Supplementary File S7**). Due to low variance explained by PC 6 in both sexes, 95% confidence intervals could not be estimated for the PC scores but were successfully estimated for each individual gene loading.

The majority of variation in alternative splicing (91.77% in females; 86.72% in males) was explained by patterns that occurred in all samples irrespective of treatment. There were four distinct patterns of alternative splicing that were not affected by treatment and were consistent across sex (**Figure 1**). PC1 described genes with expression that changes linearly across exon position, which includes genes that have both increasing and decreasing expression across exon position (**Figure 1**A,E**, Supplementary File S6**). PC2 described genes that have either increased or decreased expression in middle exons (**Figure 1B,F, Supplementary File S6**) and PCs 3 and 4 described genes that have either higher or lower expression in other combinations of exon positions (**Figure 1**CDGH**, Supplementary File S6**).

Two patterns of alternative splicing differed between treatments (**Figure 2**) and described lower total variance (7.69% in females; 10.89% in males; **Supplementary File S5**). The general patterns of these PCs were the same between females and males, but the specific genes most strongly associated with these patterns differed by sex (**Figure 2**). PC5 described genes that show either higher or lower expression of exons 2-6 relative to exon 1 depending on treatment (**Supplementary File S7**). PC6 described genes that have different levels of expression in either the front or end exons depending on treatment (**Supplementary File S7**). For females, five of the top 20 genes from PC5 and PC6 had changes in the maximum number of transcripts expressed. There were 38 enriched biological process GO terms identified in the top female genes from PC5 and PC6 (**Supplementary File 18**). These processes included protein ufmylation and ion transport (**Supplementary File 18**). In males, four of the top 20 genes from PC5 and PC6 had changes in the maximum number of transcripts in males. A total of 26 overrepresented biological processes were identified, such as regulation of metallopeptidase activity and acetyl-CoA metabolic process (**Supplementary File 18**).

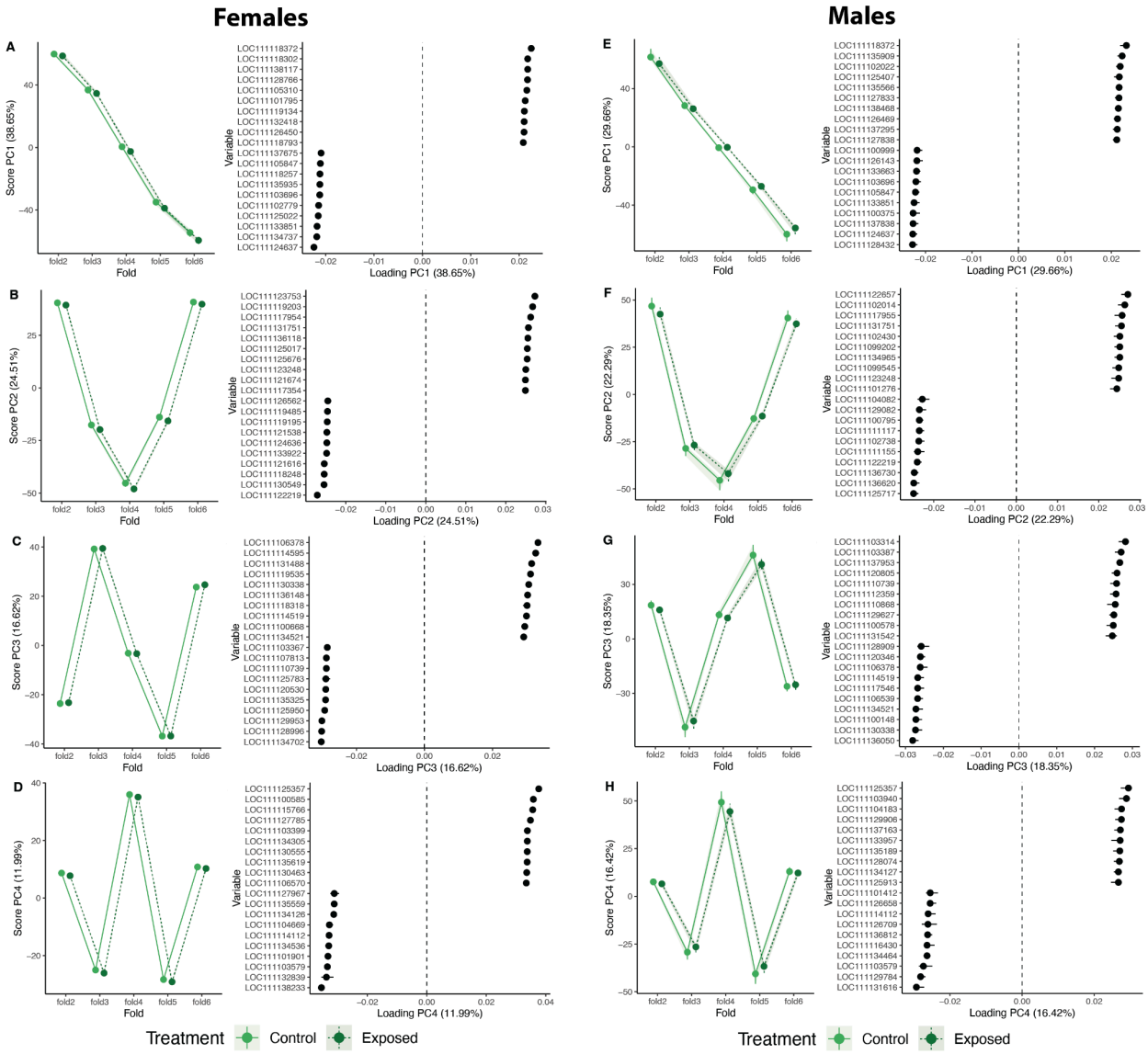


Figure 1. Principal components identified by ANOVA-simultaneous components analysis (ASCA) that describe patterns of alternative splicing that do not show differential effects of treatment. Each panel displays one PC with female PCs in the left column **(A-D)** and male PCs in the right column **(E-H)**. Each panel shows the PC score for each exon position (x-axis) with fold2 indicating change in expression at exon 2 relative to exon 1 and fold 3 indicating change in expression at exon 3 relative to exon 1, and so on. Note that PC scores show the association of each identified pattern in the PC plane and do not indicate absolute values of expression (i.e., a higher PC score does not necessarily indicate higher expression). The loading values of the 20 genes with the highest PC loading scores indicate the direction of expression change for that gene. For example, PC1 describes expression that changes linearly across exon position. In PC1, genes with loading values of <0 indicate genes that increase in expression across exon position whereas genes with loading values of >0 indicate genes that decrease in expression across exon position (see **Supplementary File S6**). Dark green indicates exposed treatment

and light green indicates control treatment. Shading indicates 95% confidence intervals. Confidence intervals resulted in NA values in PC6 and are therefore not shown.

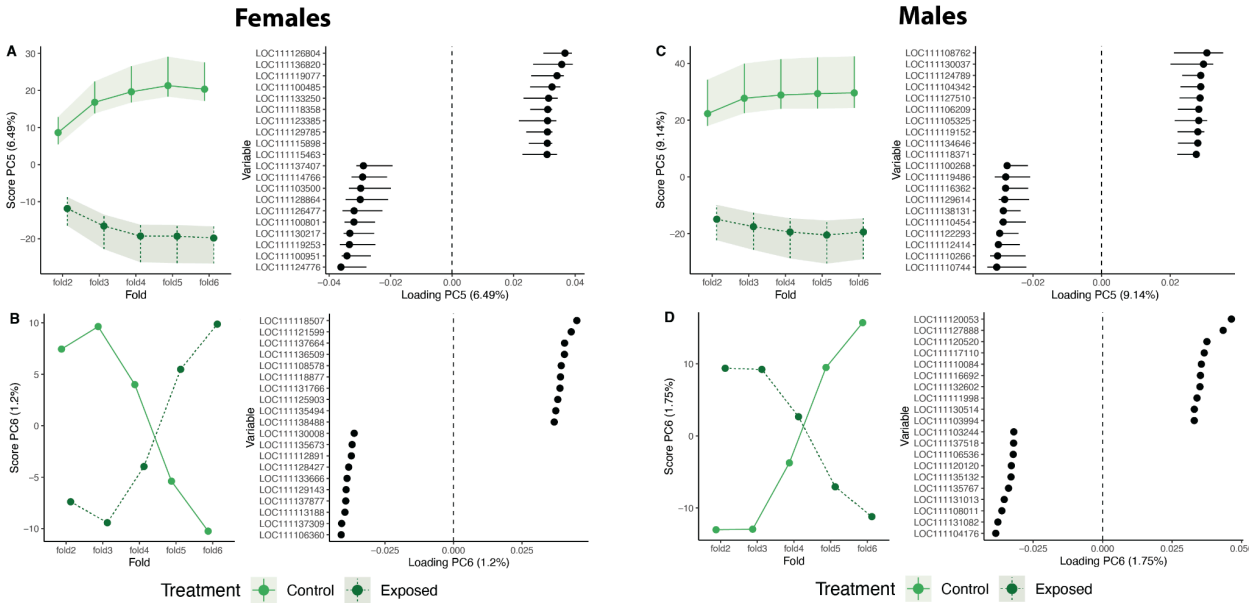


Figure 2. Principal components identified by ANOVA-simultaneous components analysis (ASCA) that describe patterns of alternative splicing that show differential effects of treatment. Each panel displays one PC with female PCs in the left column **(A-B)** and male PCs in the right column **(C-D)**. Each panel shows the PC score for each exon position (x-axis) with fold2 indicating change in expression at exon 2 relative to exon 1 and fold 3 indicating change in expression at exon 3 relative to exon 1, and so on. Note that PC scores show the association of each identified pattern in the PC plane and do not indicate absolute values of expression (i.e., a higher PC score does not necessarily indicate higher expression). The loading values of the 20 genes with the highest PC loading scores on each PC indicate the direction of expression change for that gene. PC5 describes genes that show either higher or lower expression of exons 2-6 relative to exon 1 depending on treatment. In PC5, genes with loading values of <0 indicate genes that have higher expression in exons 2-6 when exposed to treatment whereas genes with loading values of >0 indicate genes that have lower expression in exons 2-6 when exposed to treatment (see **Supplementary File S7**). PC6 describes genes that have different levels of expression in either front or end exons depending on treatment (see **Supplementary File S7**). Dark green indicates exposed treatment and light green indicates control treatment. Shading indicates 95% confidence intervals. Confidence intervals resulted in NA values in PC6 and are therefore not shown.

DNA Methylation Analysis

Approximately 57% of the quality trimmed reads aligned to the genome and on average 12.4% of the CpG loci were methylated. The average number of loci with a minimum of 10x coverage in the 26 individuals was 8.1M (56% of 14,458,703 total CpGs in the *C. virginica* genome). The baseline methylation landscape was characterized for each sex separately. A total of 12,678,572 and 11,991,497 non-C->T SNP CpGs had at least 10x coverage in one

female or male sample, respectively. For both sexes, a majority of CpGs were lowly methylated (females: 10,212,421 CpGs, 80.5%; males: 9,657,277 CpGs, 80.5%), followed by highly methylated (females: 1,254,583 CpGs, 9.9%; males: 1,467,485 CpGs, 12.2%), and moderately methylated (females: 1,211,567 CpGs, 9.6%; males: 866,734 CpGs, 7.2%).

In females, highly methylated CpGs were found primarily in genes, with 662,082 CpGs (52.8%) in introns, 416,646 CpGs in CDS (33.2%), and 62,784 CpGs in exon UTR (5.0%) (**Supplementary File S9**). More highly methylated CpGs were located in the 1000 kb flanks downstream of genes (40,697 CpGs; 3.2%) than in flanks upstream of transcription start sites (7,896 CpGs; 0.6%). A total of 75,013 CpGs (6.0%) were found in intergenic regions, 80,887 CpGs (6.4%) in transposable elements, and 21,399 CpGs (1.7%) in lncRNA. Similarly, highly methylated CpGs in males were concentrated in genes, with more CpGs located in introns (812,892 CpGs, 55.4%) than CDS (438,908 CpGs, 30.0%) or exon UTR (77,872 CpGs, 5.3%) (**Supplementary File S9**). Downstream flanks contained more CpGs (46,650 CpGs, 3.2%) than upstream flanks (10,088 CpGs, 0.7%). Males also had higher concentrations of CpGs in transposable elements (93,876 CpGs, 6.4%), followed by intergenic regions (92,758 CpGs, 6.3%) and lncRNA (27,706 CpGs, 1.9%). The distribution of highly methylated CpGs in both sexes in genome features was significantly different than the distribution of all CpGs in the genome (chi-squared contingency tests; **Supplementary File S10**, **Supplementary File S11**).

Differential Methylation at the CpG Locus-Level

There were 128 putative DML in females and 4,175 putative DML in males when comparing elevated pCO2-exposed and control oysters. After removing C->T SNPs, there were 89 and 2,916 female and male DML, respectively. Female DML were found predominantly in genes (74 DML; 83.1%), with 37 DML in introns (41.6%), 29 DML in CDS (32.6%), and 8 DML in exon UTR (9.0%) (**Figure 3A**). Seven DML (7.9%) were found in intergenic regions, four DML (4.5%) were found in both upstream and downstream flanks, and one DML (1.1%) was found in both lncRNA and TE. The proportion of female DML found in CDS and intergenic regions was significantly different from the proportion of 10x CpGs (**Supplementary File S12**). Gene ontology enrichment revealed 85 biological process GO terms enriched in genes containing DML (**Supplementary File S18**). These enriched processes include dolichol metabolic process, negative regulation of apoptotic process, and positive regulation of oocyte maturation.

Similar to females, the majority of male DML were found in gene bodies (2,322 DML; 79.6%) (**Figure 3B**). Within genes, 1,332 DML were found in introns (45.7%), 815 DML in CDS (27.9%), and 190 DML (6.5%) in exon UTR. More DML were found in downstream flanks (163 DML; 5.6%) than upstream flanks (60 DML; 2.1%). A total of 279 DML (9.6%) were found in intergenic regions, while 144 DML (4.9%) were in TE and 52 DML (1.8%) were in lncRNA. The proportion of male DML in exon UTR, CDS, introns, upstream flanks, downstream flanks, and intergenic regions differed significantly from the proportion of 10x CpGs in those genome features (**Supplementary File S14**). A total of 359 biological process GO terms such as organelle organization, regulation of GTPase activity, and macromolecule localization were enriched in genes containing DML in males (**Supplementary File S18**).

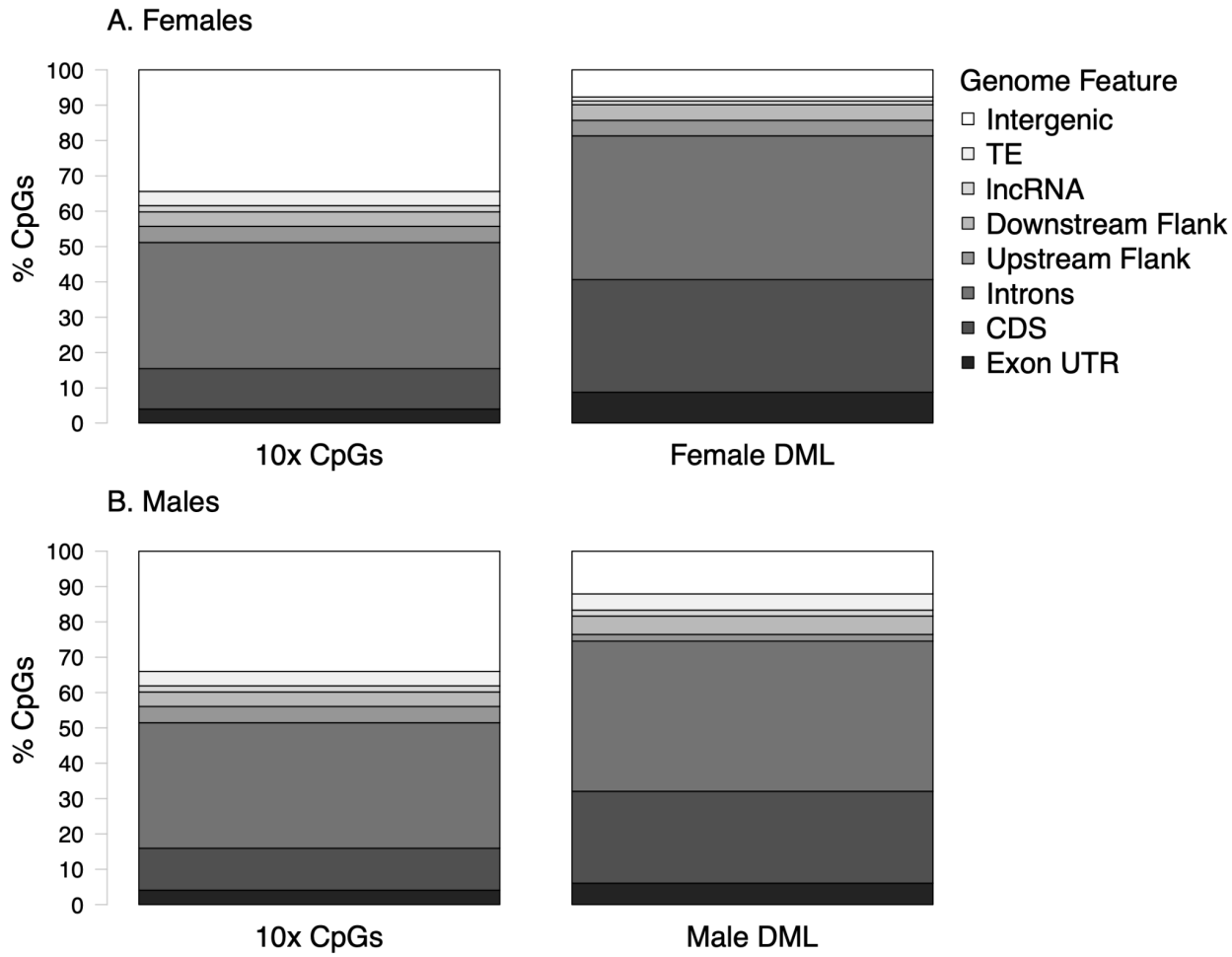


Figure 3. Genomic locations of all 10x CpGs with data versus DML for a) females and b) males.

Methylation Influence on Gene Activity

A significant and positive correlation was found between gene methylation and expression distance matrices, indicating that samples that were highly similar in gene expression were also highly similar in methylation (R = 0.843, *P*-value = 0.01) (**Supplementary File S1**). To investigate the reason for this higher-level correlation, methylation's influence on gene activity was evaluated by examining its effect on 1) gene and transcript expression, 2) predominant transcript shifts, 3) transcriptional noise, and 4) alternative splicing.

Influence of Methylation on Gene and Transcript Expression

The one transcript differentially expressed between control and elevated pCO_2-exposed female oysters did not contain DML. Since a one-to-one relationship was not identified between DML and DET, the influence of methylation on other aspects of gene activity was explored. We first examined connections between methylation and change in the maximum number of transcripts expressed. The number of DML in genes with changes in the maximum number of expressed transcripts ranged from zero to two in females (**Supplementary File S15**) and zero to 25 in

males (**Supplementary File S15**). A binomial model was used to evaluate the impact of methylation on whether or not there was a change in the maximum number of transcripts expressed due to elevated pCO_2 exposure. The number of DML in a gene was not a significant predictor of a change in maximum transcripts expressed (**Table 2**, **Supplementary File S16**) in female or male gonad tissue. Change in gene body methylation had a marginally significant negative effect on changes in the maximum number of transcripts male oysters (**Table 2**). For both sexes, change in gene expression and gene length were significant predictors (**Table 2**).

Table 2. Summary of binomial model for female and male oysters examining the influence of gene body methylation, DML count, gene length, and gene expression on changes to the maximum number of transcripts expressed due to elevated pCO_2 exposure. **P*-value < 0.01; ***P*-value < 0.05

Predictor	Females			Males		
	β Estimate	**z-value**	***P*-value**	**β Estimate**	**z-value**	***P*-value**
Intercept	0.24	0.67	0.51	2.66	6.88	6.00×10^{-12}**
Change in Gene Body Methylation	0.007	0.44	0.66	-0.01	-1.76	0.08*
DML Count	0.35	0.23	0.82	-0.05	-1.07	0.29
$\log_2$(Change in Gene Expression)	0.65	11.31	1.13×10^{-29}**	-0.54	-5.86	4.63×10^{-9}**
$\log_{10}$(Gene Length)	-0.19	-2.14	0.03**	0.18	6.27	3.67×10^{-10}**

Influence of Methylation on Predominant Transcript Shifts

Several DML were found in genes with a shift in the predominant transcript expressed. In females, zero to three DML were found in these genes (**Supplementary File S15**), with zero to 25 DML in males (**Supplementary File S15**). Binomial models were used to quantify the impact of gene body methylation, DML count, gene expression, and gene length on whether or not there was a shift in the predominant transcript expressed due to low pH exposure. In female oysters, change in gene body methylation and number of DML in genes did not have a significant effect (**Table 3, Supplementary File S17**). However, change in gene body methylation had a significant positive effect on a shift in the predominant transcript in males (**Table 3, Supplementary File S17**). Gene expression and gene length were significant predictors of predominant transcript shift in both females and males (**Table 3**).

Table 3. Summary of binomial model for female and male oysters examining the influence of gene body methylation, DML count, gene length, and gene expression on a shift in the predominant transcript due to low pH exposure. **P*-value < 0.01; ***P*-value < 0.05

Predictor	Females			Males		
	β Estimate	**z-value**	***P*-value**	**β Estimate**	**z-value**	***P*-value**
Intercept	-8.49	-35.93	9.54×10^{-283}**	-3.04	-9.44	3.71×10^{-21}**

Predictor	Females			Males		
	β Estimate	z-value	*P*-value	β Estimate	z-value	*P*-value
Change in Gene Body Methylation	-0.005	-0.58	0.56	0.01	2.03	0.04**
DML Count	-11.64	-0.10	0.92	0.07	1.61	0.11
$\log_2$(Change in Gene Expression)	-0.21	-11.10	7.20×10^{-140}**	0.05	2.16	0.03**
$\log_{10}$(Gene Length)	1.45	25.18	1.32×10^{-28}**	0.33	4.28	1.88×10^{-5}**

Transcriptional Noise Analysis

Models demonstrate both gene body methylation and expression have a significant influence on gene expression variability, or transcriptional noise (**Figure 4**). All predictors were included in the best-fit model for each sex. For female oysters (model adjusted $R^2 = 0.60$), transcriptional noise was negatively correlated with treatment (*P*-value = 0.002; **Figure 4A**), gene expression (*P-value* = 0; **Figure 4C**), gene body methylation (*P*-value = 0; **Figure 4E**), and the interaction between treatment and expression, but positively correlated with methylation variation (*P*-value = 1.87×10^{-272}; **Figure 4G**), gene expression squared, the interaction between treatment and gene expression squared, and gene length (**Table 4**). Inclusion of polynomial terms for gene expression reflects that gene expression is most variable at very low and very high levels of gene expression. Methylation had the highest partial R^2 (partial $R^2 = 0.47$; **Table 4**), followed closely by model residuals (partial $R^2 = 0.40$). For male oysters, transcriptional noise was positively correlated with treatment (*P*-value = 6.19×10^{-15}; **Figure 4B**), negatively correlated with gene expression (*P*-value = 0; **Figure 4D**) and gene body methylation (*P*-value = 8.45×10^{-50}; **Figure 4F**), and positively correlated with methylation variation (*P*-value = 6.40×10^{-163}; **Figure 4H**). Transcriptional noise was also negatively correlated with the interaction between treatment and gene expression and gene length, but positively correlated with gene expression squared, the interaction between treatment and gene expression squared, and the interaction between treatment and gene body methylation (**Table 4**). The model residuals had the highest partial R^2 (partial $R^2 = 0.62$) followed by methylation (partial $R^2 = 0.16$; **Table 4**).

Table 4. Summary of multiple linear regression examining transcriptional noise, or the CV of gene expression in female (adjusted $R^2 = 0.60$) and male oysters (adjusted $R^2 = 0.38$). Coefficients of determination, or partial R^2, were calculated separately for each predictor.
**P*-value < 0.05

Predictor	Females				Males			
	β Estimate	t-value	*P*-value	Partial R^2	β Estimate	t-value	*P*-value	Partial R^2
Intercept	0.14	19.71	3.04×10^{-86}*	-	0.43	19.22	1.22×10^{-81}*	-
Treatment	-8.62×10^{-3}	-3.17	0.002*	0.02	0.15	7.81	6.19×10^{-15}*	3.96×10^{-3}

Predictor	Females				Males			
	β Estimate	t-value	*P*-value	Partial R^2	β Estimate	t-value	*P*-value	Partial R^2
$\log_2$(Gene Expression)^2	0.01	35.85	1.10×10^{-278}*	0.02	0.07	46.84	0*	2.98×10^{-3}
$\log_2$(Gene Expression)	-0.10	-51.19	0*	0.06	-0.41	-47.87	0*	0.15
Gene Methylation	-3.85×10^{-3}	-56.32	0*	0.47	-1.40×10^{-3}	-14.88	8.45×10^{-50}*	0.16
$\log_{10}$(CV Methylation)	0.10	35.44	1.87×10^{-272}*	0.02	0.14	27.46	6.40×10^{-163}*	0.05
Treatment * $\log_2$(Gene Expression)^2	6.18×10^{-3}	14.89	4.81×10^{-50}*	2.80×10^{-3}	0.02	7.25	4.40×10^{-13}*	2.29×10^{-6}
Treatment * $\log_2$(Gene Expression)	-0.07	-26.15	6.46×10^{-150}*	4.29×10^{-3}	-0.10	-7.50	6.73×10^{-14}*	1.77×10^{03}
Treatment * Gene Methylation	6.38×10^{-4}	7.43	1.13×10^{-13}*	3.06×10^{-3}	7.03×10^{-4}	7.06	1.67×10^{-12}*	1.32×10^{-3}
$\log_{10}$(Gene Length)	7.14×10^{-3}	3.80	0.001*	9.90×10^{-5}	-0.03	-6.38	1.87×10^{-10}*	1.14×10^{-3}

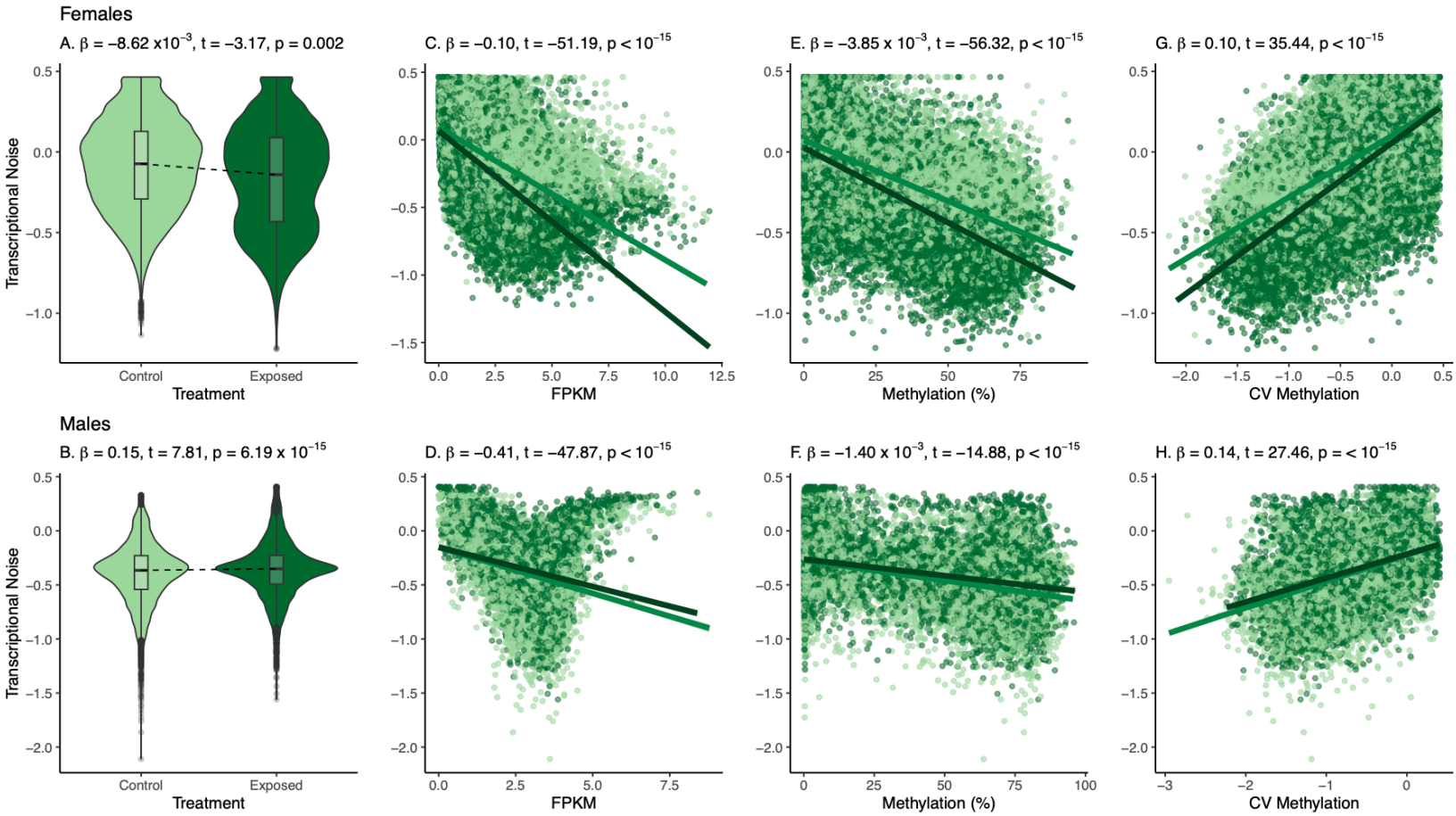


Figure 4. Relationship between transcriptional noise and A) and B) elevated pCO_2 treatment; C) and D) FPKM; E) and F) gene methylation (%); and G) and H) CV of methylation defined by the best-fit linear model for female (adjusted R^2 = 0.60) and male (adjusted R^2 = 0.38) oysters. Transcriptional noise and FPMK are log_{10} and log_2 transformed, respectively, for visualization. Beta statistics, t-values, and *P*-values are included for each explanatory variable of the linear model. For panels C-H, each point represents values for a single gene, with associated control values in light green and elevated pCO_2 values in dark green. Linear model fits are included for control and elevated pCO_2 treatments. Additional model predictors, including polynomial terms for gene expression, can be found in Table 3.

Influence of Methylation on Alternative Splicing

The influence of methylation was examined in the top 20 genes from PC5 and PC6, which exhibited changes in splicing from elevated pCO_2 exposure. There were no DML found in top female genes, while six male genes contained DML. Average gene body methylation in control and exposed samples were investigated as predictors of alternative splicing using sex-specific binomial GLM. In both sexes, average gene body methylation in control samples, methylation in exposed samples, and gene length had no significant impact on alternative splicing status (**Table 4**; **Figure 5**).

Table 4. Binomial GLM results evaluating the influence of gene body methylation and gene length on alternative splicing status for the top 20 genes from each ASCA PC. No explanatory variables were significant predictors of alternative splicing status.

Predictor	Females			Males		
	β Estimate	z-value	*P*-value	β Estimate	z-value	*P*-value
Intercept	4.80	1.67	0.09	-2.91	-1.25	0.21
Gene Body Methylation in Control Samples	-0.02	-0.44	0.65	-0.04	-1.03	0.30
Gene Body Methylation in Control Samples	-0.02	-0.37	0.71	0.04	1.14	0.25
$\log_{10}$(Gene Length)	-0.85	-1.26	0.2	0.49	0.85	0.39

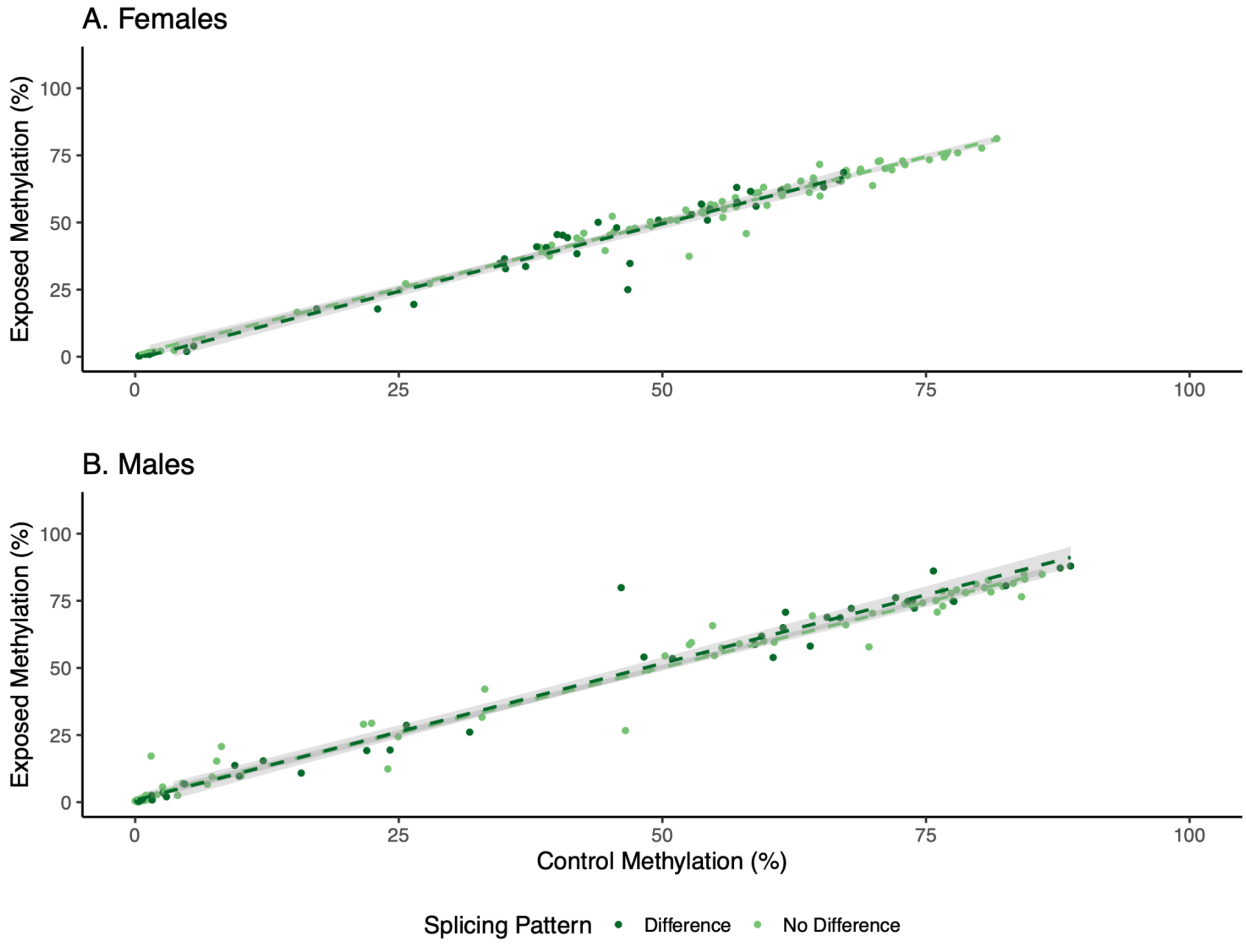


Figure 5. Average gene body methylation in control and exposed samples for genes that were alternatively spliced (difference) and not alternatively spliced (no difference) based on elevated pCO_2 exposure in A) females and B) males. Each point represents a gene that was one of the

top 20 drivers of expression patterns in a PC. Regression lines with 95% confidence intervals are included for each gene category.

Discussion

As elevated pCO_2 continues to threaten marine invertebrates, it is critical to understand the role epigenetic and transcriptomic mechanisms play in stress resilience and potential for environmental memory. Impacts on adult reproduction can influence intergenerational plasticity, so disentangling how various molecular mechanisms influence reproductive tissue is key for understanding offspring responses to stress. McNally et al. [45] used adult oysters from this study to produce offspring and understand intergenerational plasticity in elevated pCO_2 conditions. After three days of exposure, larvae from parents exposed to elevated pCO_2 had higher shell growth rates and larger shell sizes, and these effects were pronounced when larvae were also reared in elevated pCO_2 conditions [45]. Interestingly, there was no impact of elevated pCO_2 on egg size or shape [45]. Since maternal provisioning may not have been responsible for larval phenotypes, exploring methylation and gene expression responses can elucidate regulatory mechanisms underlying observed carryover effects. By better understanding these mechanisms, we can better predict how environmental stress will influence intergenerational responses, and therefore ecosystem outcomes.

Given that the functional role of DNA methylation in marine invertebrates remains unclear [32], we leveraged whole genome, transcript-level data to conduct the first extensive assessment of DNA methylation-gene activity relationships in response to elevated pCO_2 in *Crassostrea* spp. gonad tissues to our knowledge. We found clear evidence of DNA methylation reducing gene expression variability, or transcriptional noise, in a sex-specific manner in response to elevated pCO_2. These findings align with the growing consensus that gene body methylation may not have a direct regulatory role in gene-level expression, but instead functions to broadly reduce transcriptional variation across the genome [32,50,51]. Additionally, differences in gene activity and methylation patterns between female and male oysters highlight the need to examine how OA may affect reproduction differently in females and males [21], and underscore the cell- and tissue-specificity of DNA methylation [52,53].

Methylation does not impact gene-level responses in gonads to elevated pCO_2

In this study there was no substantial impact of elevated pCO_2 on gene-level expression in female and male gonadal tissues. This is not surprising when considering similar studies on elevated pCO_2 on gene expression in *C. virginica* mantle tissue [36]. There are several studies that do see an impact on gene expression [20], with predominant function of influenced genes associated with stress response, acid-base regulation, metabolic processes, and calcification. While we did not see substantial gene-level impacts, we did see significant changes to transcript-level activity — changes in the maximum number of transcripts expressed per gene, shifts in the predominant transcript expressed, and patterns of alternative splicing — and to DNA methylation. These changes were associated with biological processes such as reproduction (*ex.* gamete development, regulation of female receptivity, sperm motility) various signaling pathways, and protein ubiquitination. These data suggest the importance of these molecular mechanisms in response to elevated pCO_2 conditions.

Beyond understanding the response of oyster reproductive tissue to elevated pCO_2, a primary impetus for this work was to address hypotheses surrounding the role of methylation in environmental response and gene expression regulation. Following mammalian studies, it was previously considered that environmentally-sensitive changes in methylation drive differential expressed genes on a gene-by-gene basis in marine invertebrates. However, to date, numerous molluscan studies have failed to find any significant relationships between differential DNA methylation and differential gene expression [36,39,444], or have found minimal overlaps in differentially expressed genes and differentially methylated regions [43,54]. As described above, we could not examine the association between differential methylation and differential gene expression since we did not observe any of the latter. We also did not see a concordant change in gene expression at the gene level where differential methylation occurred. Taken together, our data supports the lack of a one-to-one relationship between methylation and gene expression reported in the literature. Further, these results demonstrate that this relationship does not exist in gonad tissues, similar to what has been previously reported in somatic tissues.

Methylation alters some aspects of transcript-level gene activity in gonad tissues

Although there is limited evidence for DNA methylation driving differential gene expression, it is possible that methylation may change transcriptional opportunities. Sparse methylation may allow for exon skipping or access to alternative start sites, thereby modifying gene expression and phenotype [31,32,55]. We observed that elevated pCO_2 elicits alternative splicing, although environmentally-responsive splicing accounted for a small proportion for the total splicing variance. There were minimal overlaps with DML and alternatively spliced genes, and gene body methylation did not significantly influence splicing. This contrasts with work in *C. gigas* and upwelling-exposed *Strongylocentrotus purpuratus* urchin larvae showing that high levels of methylation in exons [56,57] and/or introns [57] are associated with alternative splicing. In *C. gigas*, alternative splicing was explored in one gene (CGI_10021620) with higher methylation and gene expression mantle tissues as opposed to male gametes [56]. Lower levels of methylation in male gametes may have opened access to an alternative stop codon, producing a shorter protein product when compared to the product from mantle tissues [56]. Since *C. gigas* were not exposed to any stressors, the role of methylation in regulating alternative splicing could be distinct for cell type differentiation versus environmental response. Significant methylation regulation of alternative splicing was found in *S. purpuratus* larvae from female urchins exposed to upwelling, but only when considering transcriptional start site accessibility [57]. It is possible that incorporating additional epigenetic mechanisms like chromatin state would elucidate connections between methylation and alternative splicing in *C. virginica*.

The lack of correlation between methylation and alternative splicing in our study could indicate that these molecular mechanisms are not intrinsically linked in *C. virginica*. In both sexes, we observed distinct biological functions associated with alternatively spliced genes and genes containing DML. For example, in female oysters alternatively spliced genes were involved in metabolic and developmental processes such as the glycolysis pathway and DNA ligation, while genes with DML had prominent roles in apoptosis, reproductive processes, and cell organization and biogenesis. The latter processes have been previously implicated in oyster response to elevated pCO_2 [11,37], including studies examining gonad methylome responses

[38,40]. Therefore, it is possible that DNA methylation and alternative splicing have separate, but essential, roles in moderating reproductive tissue response to OA.

Although methylation and alternative splicing patterns were not correlated, the influence of gene body methylation on other aspects of gene activity suggests a potential regulatory role. In male oysters, methylation significantly impacted changes in the maximum number of transcripts expressed per gene, and had a marginal effect on shifts in the predominant transcripts. Interestingly, genes with DML and alterations to gene activity had similar enriched biological processes involved in signal transduction and cell organization and biogenesis. These commonalities further support the idea methylation may have a role in gene activity beyond regulating alternative splicing patterns for specific pathways.

Methylation regulates gene activity differently in female and male reproductive tissue

A critical regulatory role for DNA methylation is the reduction of gene expression variability, or transcriptional noise. We speculate that this genome-wide role may be more important than direct modulation of gene expression. The inverse relationship between methylation and transcriptional noise found in this study is consistent with previous research in molluscs [36,39,44,58] and across invertebrate taxa [32,50,51]. We also demonstrate that methylation variability is positively correlated with transcriptional noise. In other words, larger changes in gene body methylation lead to higher expression variability. The general relationships described above were also consistent in response to elevated pCO_2, further supporting the idea that methylation has a broad regulatory role in response to stress. In this study, we show that these responses are also seen in gonad tissues. We also identified a sex-specific pattern: elevated pCO_2 decreases transcriptional noise in females but increases noise in males with respect to control conditions. The difference in methylation-noise relationships between sexes could suggest that sperm cells require additional methylation to achieve similar levels of transcriptional noise reductions to females. As egg quality is essential to successful reproduction, reducing transcriptional noise may be necessary to maintain egg quality, while gene expression variability may not be maladaptive for sperm. Given the small effect size in males, future work should investigate the phenotypic implications of these shifts. Combined with the subtle gene expression differences between control and OA-exposed oysters, our work suggests that reductions in general gene variability are an integral component of OA response in gonad tissues of *C. virginica*.

Additional differences in female and male reproductive tissue responses to elevated pCO_2 underscore the importance of disentangling biological and environmental effects on gene expression and DNA methylation. It is well-known that OA can have sex-specific impacts on physiology [21], and previous work in *C. gigas* has documented distinct gene expression patterns associated with oogenesis and spermatogenesis [49]. We extend this work by demonstrating contrasting OA effects between female and male gonad tissues. We found that sperm had more genes with higher number of transcripts expressed in elevated pCO_2, as opposed to female gonad tissue that had more genes with a smaller number of transcripts expressed in the same conditions. Even though OA-responsive patterns of alternative splicing were similar between sexes, the top 20 gene drivers of those patterns differed and were involved in sex-specific reproductive processes. For example, protein UFMylation genes were

alternatively spliced in female gonad tissue, while genes involved in regulation of metallopeptidase activity. Protein UFMylation is involved in broader protein ubiquitination processes, which are conserved across oyster responses to pCO_2 [11,37] and characteristic of female gonad methylome responses [40]. Metallopeptidases are important for sperm motility in humans [59,60], and have been found to increase in expression in *C. gigas* during spermatogenesis [61]. Differences in female and male reproductive tissue responses extend to methylation as well. Sperm had a more dramatic methylation response to elevated pCO_2, with 2,916 DML compared to 89 in female gonad tissue oysters, even though a similar number of CpGs with 10x coverage were detected in both sexes. Sex-specific differences in methylation [46–48] are well-documented in *C. gigas*. Male *C. gigas* have higher levels of gene body methylation than females in both somatic [48] and reproductive tissue [46,47]. Our results confirm that these patterns are present in *C. virginica* gonads as well. These gene expression and methylation differences can be attributed to cell type: oocytes typically have higher levels of maternal RNA as opposed to sperm cells with more DNA. Enrichment of processes such as oocyte development in females and spermatogenesis in males in genes with altered activity or with DML align with the specific biological roles of these cell types. In our study, female samples were likely a mix of oocytes and connective tissue, which further complicates the observed patterns. Previous studies have documented environmentally-sensitive methylation in response to elevated pCO_2 in molluscan reproductive tissue [38,40], none examine sex-specific patterns or have data for both sexes. As the role of methylation is cell-, tissue-, and sex-specific [52,53], future investigations should compare molecular and phenotypic responses to elevated pCO_2 between tissues and sexes when appropriate.

Conclusion

Elucidating the relationship between epigenetic and transcriptomic mechanisms will allow for a holistic understanding of phenotypic responses to climate stressors. We highlight the importance of transcript-level data for parsing the relationship between these mechanisms. Future research should integrate transcript-level data along with exon- and intron-specific methylation and expression analyses to further clarify the relationship between methylation and gene activity within and across generations. As DNA methylation is a potential mechanism for intergenerational carryover effects, it is critical to assess environmentally-responsive methylation in adult reproductive tissue. Our study suggests that methylation has a genome-wide regulatory role, effectively maintaining gene expression homeostasis in reproductive tissues under elevated pCO_2 by reducing transcriptional noise. Importantly, we found that methylation regulates transcriptional noise differently in female reproductive tissue and sperm: male oysters required higher levels of methylation to achieve similar reductions in transcriptional noise when compared to female oysters. An accompanying study using the same adult *C. virginica* found no difference in female egg quality after OA exposure during reproductive conditioning, and no delays in early larval development [45]. Since this signal was observed even with significant genotype effects [45], it suggests that epigenetic maintenance of reproduction may also confer intergenerational resilience to elevated pCO_2. We demonstrate that OA conditions lead to molecular changes in oyster gonads and future work should evaluate

specific connections between these changes and offspring fitness to understand whether inheritance of environmental memory can occur.

Materials and Methods

Experimental Design and Seawater Chemistry Analysis

The experimental design and seawater chemistry manipulations have been published by McNally et al. [45]. Briefly, adult *C. virginica* (mean ± SD shell length: 7.92 ± 1.80 cm) were acclimated to ambient pCO_2 conditions (mean $pCO_2 \pm SD = 632 \pm 64$ ppm) for one week prior to random placement in experimental pCO_2 conditions (either control (572 ± 107 ppm) or elevated $4pCO_2$ (2827 ± 360 ppm)) during reproductive conditioning at 20ºC. Oysters were exposed to experimental conditions for 30 days, which is known to elicit differences in *C. virginica* gonad methylation [37]. Temperature, salinity, and pH_T were recorded three times per week during the experimental period, and seawater samples were collected during the first and third weeks of the experiment for dissolved inorganic carbon and total alkalinity measurements. Oysters were fed in accordance with best practices from Helm and Bourne [62]. Gonad tissue was collected from 16 female oysters (eight control and eight exposed to elevated pCO_2), and sperm from ten male oysters (four control and six exposed to elevated pCO_2), after the 30 day exposure period for WGBS and RNA-Seq.

Nucleic Acid Extraction and Library Preparation

Simultaneous DNA and RNA extraction was conducted for female gonad and sperm samples using the Zymo Quick-DNA/RNA Microprep Plus kit (Cat # D7005). A total of 10-15 mg of frozen gonad tissue and 50 µL of sperm in seawater solution were used as inputs for females and males, respectively. Manufacturer's instructions were used, with the following exceptions. After addition of 300 µL DNA/RNA Shield (1X), 30 µL PK Digestion Buffer, and 15 µL Proteinase K, female samples were incubated at 55ºC overnight. Four volumes of DNA/RNA Lysis Buffer (200 µL) were added to sperm samples. All centrifugation steps were conducted at 16,000 rcf. DNase I digestion was performed for RNA. Concentration of eluted DNA and RNA were verified with Qubit dsDNA High Sensitivity or RNA High Sensitivity assays (Thermo Fisher Scientific), respectively, and quality confirmed on a BioAnalyzer.

All samples were submitted to ZymoResearch for library preparation and sequencing. Whole genome bisulfite sequencing (WGBS) libraries were created from 100 ng of DNA with the Zymo-Seq WGBS Library Kit (Cat#: D5465) following manufacturer's protocol and PCR performed with Illumina Unique Dual Indices. Library quality was checked with an Agilent 2200 TapeStation. Libraries were sequenced on an Illumina NovaSeq instrument to generate 150 bp paired-end reads. Prior to sequencing, total RNA-Seq libraries were created using 250 ng of RNA. For all samples, rRNA was removed following methods described in Bogdanova et al. [62] with some modifications. Libraries were constructed using the Zymo-Seq RiboFree Total RNA Library Prep Kit (Cat # R3000) according to the manufacturer's instructions. Libraries were sequenced on an Illumina NovaSeq to generate at least 30 million read pairs for 150 bp paired-end sequences per sample. Resulting sequences were submitted to the NCBI Sequencing Read Archive (SRA) underBioProject accession number PRJNA1089235.

Genome Information and Feature Tracks

The *C. virginica* genome was used for several analyses [64]. Gene, coding sequence, exon, and lncRNA tracks were provided by the *C. virginica* genome NCBI RefSeq annotation (RefSeq: GCF_002022765.2) on NCBI [63]. These tracks were used to generate remaining feature tracks with BEDtools v2.26.0[64]. The complement of the exon track, or non-coding sequences, was created with BEDtools `complementBed` v.2.26.0. Overlaps between these non-coding sequences and genes — introns — were obtained using `intersectBed`. Untranslated regions of exons were identified by subtracting the coding sequence track from the exon track with `subtractBed`. Upstream flanking regions, or putative promoters, were defined as the 1000 bp upstream of a transcription start site. These flanks were created by adding 1000 bp upstream of genes with `flankBed`, taking strandedness into account. Other genes that overlapped with these flanks were removed using `subtractBed`. Flanks 1000 bp downstream of genes were created as well. To identify intergenic regions, the complement of the gene track was created with complementBed, then `subtractBed` was used to remove flanking regions from the intergenic region track. Transposable element information from RepeatMasker was obtained from the NCBI RefSeq annotation [66,67]. CG motifs in the *C. virginica* genome were identified using the `fuzznuc` function from EMBOSS. The number of CG motifs in a given track were obtained using `intersectBed` between the CG motifs and a feature track. The genome features are available along with corresponding code [68].

Genetic Variation

SNPs were identified in the BS-seq data using the Nextflow [69] EpiDiverse/snp pipeline [70] to characterize genetic differences between samples. Sorted, deduplicated BAM files, generated through methylation analysis with Bismark [71], were used as inputs, along with the National Center for Biotechnology Information (NCBI) *C. virginica* genome FastA (`GCF_002022765.2_C_virginica-3.0_genomic.fa`) provided as the reference. The options `--variants` and `--clusters` were provided when running the EpiDiverse/snp pipeline with default settings (`--min-coverage 0`, `--min-repeat-entropy 1`) . After variant identification, NgsRelate was used to estimate pairwise relatedness using VCF output from EpiDiverse/snp [72]. Variants were filtered to include only bi-allelic sites (`--min-alleles 2 --max-alleles 2`) and minor allele counts of at least two (`--mac 2`). Variants were included if data was missing for more than 50% of the samples (`--max.missing 0.5`).

Recent studies have shown that genetic population structure can influence methylation [73]. In order to determine if genetic variation had a significant impact on gene expression and methylation, we first performed correlation analyses correlated genetic distance with methylation and gene expression distance between individuals to determine whether genetic distance correlated with methylation and gene expression distance between individuals. In the case of a significant correlation, genetic distance would need to be included in downstream model analyses of differential methylation and gene expression. To perform these correlations, we generated euclidean distance matrices of genetic data, expression data, and DNA methylation data. For genetic data SNPs were derived from the Nextflow EpiDiverse SNP Pipeline as described above. Gene expression matrices were derived from DESeq2 output. DNA

methylation data from methylkit output using a 10x coverage minimum threshold. For analysis, Mantel tests (`mantel.rtest`) in R were used with 99 replicates. We visualized these relationships using Pearson correlation matrices. Corresponding code is available for all genetic variation analyses [68].

Gene Activity

Gene and Transcript Expression

FastQ files were quality trimmed using two rounds of `fastp` [74]. The initial round used default settings for paired end adapter trimming. The second round was performed without adapter trimming and removed an additional 20bp from the five prime end of all reads, followed by visualization with `MultiQC` [75]. To identify isoforms, the *C. virginica* genome (NCBI accession: GCF 002022765.2) was indexed using HISAT2 [76]. A pipeline detailed by Pertea et al. [77] was used as a guide to identify isoforms and analyze differential expression. Trimmed reads were aligned to the genome using HISAT2 (v2.1.0), followed by StringTie [78] to identify isoforms and generate files for downstream analysis using the R [79] packages `ballgown` [80] and `tidyverse` [81]. Differentially expressed genes (DEGs) and differentially expressed transcripts (DETs) (p-value and q-value < 0.05) were identified between pCO_2 treatments within a given sex using `ballgown` [80], with '`FPKM`' as the expression measurement.

Elevated pCO_2 may influence finer-scale expression processes. To explore this, we quantified differences in the maximum number of transcripts expressed in each gene for each sex separately. The maximum number of transcripts for each gene was identified across all samples. The difference in maximum number of transcripts expressed in each gene was calculated with respect to the elevated pCO_2-exposed samples. Positive differences signify more unique transcripts expressed in elevated pCO_2 conditions, while the opposite holds true for negative differences. A two-sided 1-sample z-test was used to determine if a change in the maximum number of transcripts expressed happened significantly less or more than what would be expected by chance.

An enrichment test with `topGO` [82] was conducted to understand if certain biological processes were overrepresented in genes with differences in the maximum number of transcripts expressed. First, a gene ID-to-GO term database (`geneID2GO`) was created for manual GO term annotation. Each line of the database contained gene ID in one column, and all corresponding GO terms in another. The list of gene IDs from this database was used as the gene universe for enrichment. Two separate lists were compiled for genes of interest: one for genes with increased transcript counts, and another for genes with decreased transcript counts due to elevated pCO_2 exposure. A `topGO` object was generated for each sex- and direction-specific list of genes and biological process GO terms, with GO term annotation performed using the `geneID2GO` database. A Fisher's exact test was used to identify GO terms significantly enriched with respect to the gene background (P-value < 0.01). Corresponding code is available for all gene activity analyses [68].

Predominant Transcript Identification

In addition to quantifying transcript expression, we examined shifts in the predominant transcript — the transcript with the highest mean level of expression in an experimental group — due to low pH exposure within male and female samples separately. Gene expression data from `ballgown` was filtered to retain genes with multiple transcripts where gene expression was greater than 0. Differences in the identity of the predominant transcript were determined for genes remaining after filtering. A two-sided 1-sample z-test was used to determine if a shift in predominant isoform identity happened significantly less or more than what would be expected by chance. An enrichment test with `topGO` [80] was used to identify significantly enriched biological process GO terms (*P*-value < 0.01) in genes where the predominant transcript changed using similar methods as described above (see *Gene and Transcript Expression*).

Alternative Splicing

The final way we characterized gene activity was through an alternative splicing analysis. We hypothesized that OA exposure would induce different patterns of exon expression, or alternative splicing. To characterize these patterns, expression of exons 2-6 were calculated as fold change values relative to expression of exon 1 for every gene in each sample. To do this, `bedtools::coverage` was used with alignment files (BAMs) from HISAT2 and the genome feature file with a minimum read count coverage of 10x was used. These values were generated for females and males separately. We then used ANOVA simultaneous component analysis (ASCA) using the ALASCA package (v1.0.14) [83] to perform longitudinal multivariate analysis across exon position. ASCA is useful for the analysis of both fixed and random effects on high-dimensional multivariate data when other multivariate ANOVA analyses (i.e., MANOVA) are not able to include more variables than there are observations [84]. Each sex-specific ASCA model analyzed multivariate expression as a function of exon position, treatment, and their interaction, with the sample as a random effect to account for repeated measures:

```
expression relative to exon 1 ~ exon position * treatment +
                        (1|sample)
```

Genes that were detected in all samples were included (11,270 genes in females and 12,645 genes in males). The model was validated with bootstrapping with replacement for 100 iterations [85]. We extracted and plotted principal components (PCs) that explained >1% of the variance to visualize expression patterns across exons. We categorized PCs as those that either 1) show no difference in exon expression patterns between treatments or 2) show differences between treatments. For each PC, we extracted the top 20 genes with the highest PC scores. These gene lists were combined into the two previously described categories for downstream analyses. For these top genes, the number of genes with changes in the maximum number of transcripts expressed were quantified for females and males.

A functional enrichment analysis was conducted with `topGO` [82] to understand if alternatively spliced genes contained any overrepresented biological processes. Enrichment was conducted using similar methods as described above (see *Gene and Transcript Expression*). The top 20 genes with the highest PC scores, irrespective of alternative splicing status, were used as the gene background for this analysis. Biological process GO terms with *P*-values < 0.01 were considered significant.

DNA Methylation Analysis

CpG locus-level Methylation Quantification

Quality-trimmed sequence reads were aligned to the *C. virginica* genome using Bismark v.0.22.3 using `-score_min L,0,-0.6` to dictate alignment specificity. In order to characterize CpG methylation for individual samples, bedGraph files were generated based on Bismark `coverage2cytosine` output, where CpG data from each strand was merged (`--merge_CpG)`, filtered to retain CpG loci with at least 10x coverage, and sorted. Since C->T SNPs can be interpreted as unmethylated cytosines, they were identified in bedGraphs using default BS-Snper settings [86], then the methylation values were changed to 0.

Sex-specific baseline methylation landscapes were characterized using methylation information for CpG loci with at least 10x coverage in a sample. Individual bedGraph files (described above) were combined using `bedtools::unionBedGraphs` v2.30.0 [65] to create sex-specific union bedGraphs. Percent methylation was averaged across all samples for each CpG locus. Highly (> 50%), moderately (10-50%), and lowly (≤ 10%) methylated CpGs were identified in sex-specific union bedGraphs. Overlaps between highly methylated CpGs and various genome features were characterized using `bedtools::intersectBed`. A chi-squared contingency test (`chi.sq`) was used to determine if the distribution of highly methylated CpGs was significantly different from the distribution of all CpGs in the *C. virginica* genome, with *P*-values < 0.05 indicating a significant difference.

Deduplicated BAM alignment files from Bismark were used for sex-specific differential methylation analysis with `methylKit` [87]. The threshold for comparison was a minimum of 10x coverage (`lo.count = 10`) and maximum coverage set at 98% (`hi.perc=98`), using the destranded setting (`destrand = TRUE`). Putative differentially methylated loci were determined based on a 50% difference in comparison group and q-value threshold of 0.01. Finally, C->T SNPs were removed from putative female and male DML lists using `bedtools::subtractBed`. Genomic location of female and male DML were determined by overlapping with annotated features using `bedtools::intersectBed`. A chi-squared contingency test was used to test the null hypothesis of no association between methylation and specific genome features using all CpGs with 10x data and DML for females and males separately. A *P*-value < 0.05 signified a significantly different proportion of DML in a specific genome feature when compared to 10x CpG data. An enrichment test with `topGO` [82] was used to identify significantly enriched biological process GO terms (*P*-value < 0.01) in genes containing DML The gene universe for the enrichment was a sex-specific list of genes containing 10x CpGs. The enrichment was conducted using similar methods as described above (see *Gene and Transcript Expression*).

Gene-level Methylation Quantification

Gene level methylation was determined by taking the mean value of methylation at loci within a gene where coverage was at least 10x. Specifically, bedgraph files generated as part of the Bismark alignment described above were used in conjunction with `bedtools::intersectBed` to extract CpG loci for each sample, for each gene.

Methylation Influence on Gene Activity

To broadly examine methylation and gene expression signatures of OA treatment for each sex, we used a similar approach to the genetic variation analysis. We used Monte Carlo tests (`mantel.rtest`) to calculate correlations between bray-curtis distance matrices derived from gene methylation and expression datasets. Separate subsequent analyses were conducted to examine if methylation influenced specific aspects of gene activity.

Influence of Methylation on Gene and Transcript Expression

The first set of analyses were used to determine the relationship between gene methylation and expression. First, the number of DML in DEG and DET were quantified. Next, we investigated if gene methylation influenced differences in the maximum number of transcripts expressed between control and exposed samples using a binomial GLM (`glm`). For each sex separately, we examined if a change in the maximum number of transcripts expressed due to OA exposure ("success") could be explained by change in gene body methylation, number of DML in a gene, change in gene expression, or gene length:

```
occurrence of a change in the maximum number of expressed transcripts
 ~ change gene methylation + number of DML in the gene + log2(change
            in gene expression + 1) + log10(gene length)
```

Change in gene body methylation was calculated by subtracting average gene body methylation in the control treatment from the OA treatment. A similar calculation was done to obtain change in gene expression. Explanatory variables were considered significant if *P*-values < 0.05. All code and analyses are available on GitHub (https://github.com/sr320/ceabigr/blob/main/code/75-max-transcript-enrichment.qmd)

Influence of Methylation on Predominant Transcript Shifts

The second set of analyses examined if previously-identified transcript shifts were due in part to gene methylation. Similar sex-specific binomial GLM (`glm`) to those described above (see *Influence of Methylation on Gene and Transcript Expression*) were constructed to examine if occurrence of a predominant transcript shift due to OA exposure could be explained by change in gene body methylation, number of DML in a gene, change in gene expression, or gene length:

```
occurrence of predominant transcript shift ~ change gene methylation
 + number of DML in the gene + log2(change in gene expression + 1) +
                        log10(gene length)
```

Predictors were significant if *P*-values < 0.05.

Transcriptional Noise Analysis

Following Wu et al. [88], transcriptional noise, or the CV of gene expression, was used as a response variable in a linear model (`lm`) with average gene methylation, average gene

expression, CV of gene methylation, gene length, OA treatment, and the interaction between treatment and methylation or expression:

$$\log_{10}(\text{CV of gene expression}) \sim \text{gene methylation} + \log_2(\text{gene expression} + 1)^2 + \log_2(\text{gene expression}) + \log_{10}(\text{CV of gene methylation}) + \log_{10}(\text{gene length}) + \text{treatment} + \text{treatment}*\text{gene methylation} + \text{treatment}*\log_2(\text{gene expression} + 1)^2 + \text{treatment}*\log_2(\text{gene expression} + 1)$$

The polynomial terms for gene expression were included after initial data visualization, and represent a relationship between gene expression and transcriptional noise where gene expression variability is highest at very low or high levels of gene expression. A backwards deletion approach was used for model construction, with explanatory variables removed from the model based on a *P*-value < 0.05 threshold. Competing models were evaluated with the Akaike information criterion (AIC). The best model fit was defined by the lowest AIC and all input variables having a *P*-value < 0.05. Residuals of the best fit model were evaluated for normality and homoscedasticity. Coefficients of determination, or partial R^2, were calculated for all significant explanatory variables. To calculate partial R^2, the sum of squares from `anova` output was divided by the total sum of squares for all input variables and residuals.

Influence of Methylation on Alternative Splicing

Finally, methylation's influence on gene activity was evaluated by understanding methylation patterns in alternatively spliced genes. The number of DML were counted in genes that were and were not alternatively spliced. A sex-specific binomial GLM (`glm`) was used to determine if average gene body methylation in control or exposed samples and gene length were significant predictors of alternative splicing:

$$\text{whether or not the gene was alternatively spliced} \sim \text{average gene body methylation in control samples} + \text{average gene body methylation in exposed samples} + \text{log10(gene length)}$$

Terms were considered significant predictors of alternative splicing if *P*-values < 0.05.

Acknowledgements

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. This work was funded by the National Science Foundation awards 1634167 to SBR and 1635423 to KEL. We are grateful for insight from Drs. Shayle Matsuda, Katherine Silliman, Laura H. Spencer, and Samuel N. Bogan on manuscript ideation and earlier drafts of this manuscript.

Data Availability

Raw WGBS and RNA-Seq data can be accessed at the NCBI Sequence Read Archive under BioProject accession number PRJNA1089235. Additional data, genome feature tracks, scripts, and supplementary materials are available in the *C.virginica* methylation and gene expression repository on Open Science Framework (https://osf.io/xuy2f/) and on Github (https://github.com/sr320/ceabigr) [68].

Author Contributions

AD-W and KEL conceived and ran the experiment. YRV performed DNA and RNA extractions. All authors contributed to initial analyses. YRV, ASH, SJW, and SBR conducted final analyses. AD-W and KEL contributed to analysis. YRV, ASH, SJW, and SBR wrote the initial manuscript draft. All authors reviewed and approved the manuscript.

References

1. IPCC, 2023: Sections. In: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland, pp. 35-115, DOI: 10.59327/IPCC/AR6-9789291691647

2. Gazeau F, Parker LM, Comeau S *et al.* Impacts of ocean acidification on marine shelled molluscs. *Mar Biol* 2013;**160**:2207–45.

3. Ducker J, Falkenberg LJ. How the pacific oyster responds to ocean acidification: Development and application of a meta-analysis based adverse outcome pathway. *Front Mar Sci* 2020;**7**, DOI: 10.3389/fmars.2020.597441.

4. Parker LM, Ross PM, O'Connor WA *et al.* Predicting the response of molluscs to the impact of ocean acidification. *Biology* 2013;**2**:651–92.

5. Lemasson AJ, Fletcher S, Hall-Spencer JM *et al.* Linking the biological impacts of ocean acidification on oysters to changes in ecosystem services: A review. *J Exp Mar Bio Ecol* 2017;**492**:49–62.

6. Gazeau F, Quiblier C, Jansen JM *et al.* Impact of elevated CO_2 on shellfish calcification. *Geophys Res Lett* 2007;**34**:L07603.

7. Kurihara H, Kato S, Ishimatsu A. Effects of increased seawater pCO_2 on early development of the oyster *Crassostrea gigas*. *Aquat Biol* 2007;**1**:91–8.

8. Beniash E, Ivanina A, Lieb NS *et al.* Elevated level of carbon dioxide affects metabolism and shell formation in oysters *Crassostrea virginica*. *Mar Ecol Prog Ser* 2010;**419**:95–108.

9. Ries JB. A physicochemical framework for interpreting the biological calcification response to CO_2-induced ocean acidification. *Geochim Cosmochim Acta* 2011;**75**:4053–64.

10. Tomanek L, Zuzow MJ, Ivanina AV *et al.* Proteomic response to elevated pCO_2 level in eastern oysters, *Crassostrea virginica*: evidence for oxidative stress. *J Exp Biol* 2011;**214**:1836–44.

11. Timmins-Schiffman E, Coffey WD, Hua W *et al.* Shotgun proteomics reveals physiological response to ocean acidification in *Crassostrea gigas*. *BMC Genomics* 2014;**15**:951.

12. Dineshram R, Chandramouli K, Ko GWK *et al.* Quantitative analysis of oyster larval proteome provides new insights into the effects of multiple climate change stressors. *Glob Chang Biol* 2016;**22**:2054–68.

13. Boulais M, Chenevert KJ, Demey AT *et al.* Oyster reproduction is compromised by acidification experienced seasonally in coastal regions. *Sci Rep* 2017;**7**:13276.

14. Omoregie E, Mwatilifange NSI, Liswaniso G. Futuristic Ocean Acidification Levels Reduce Growth and Reproductive Viability in the Pacific Oyster (*Crassostrea gigas*). *J Appl Sci Environ Manage* 2019;**23**:1747–54.

15. Coen LD, Brumbaugh RD, Bushek D *et al.* Ecosystem services related to oyster restoration. *Mar Ecol Prog Ser* 2007;**341**:303–7.

16. Smith RS, Cheng SL, Castorani MCN. Meta-analysis of ecosystem services associated with oyster restoration. *Conserv Biol* 2023;**37**:e13966.

17. Grabowski JH, Brumbaugh RD, Conrad RF *et al.* Economic Valuation of Ecosystem Services Provided by Oyster Reefs. *Bioscience* 2012;**62**:900–9.

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

19. Melzner F, Mark FC, Seibel BA *et al.* Ocean Acidification and Coastal Marine Invertebrates: Tracking CO_2 Effects from Seawater to the Cell. *Ann Rev Mar Sci* 2019, DOI: 10.1146/annurev-marine-010419-010658.

20. Strader ME, Wong JM, Hofmann GE. Ocean acidification promotes broad transcriptomic responses in marine metazoans: a literature survey. *Front Zool* 2020;**17**:7.

21. Padilla-Gamiño JL, Alma L, Spencer LH *et al.* Ocean acidification does not overlook sex: Review of understudied effects and implications of low pH on marine invertebrate sexual reproduction. *Frontiers in Marine Science* 2022;**9**, DOI: 10.3389/fmars.2022.977754.

22. Gienapp P, Teplitsky C, Alho JS *et al.* Climate change and evolution: disentangling environmental and genetic responses. *Mol Ecol* 2008;**17**:167–78.

23. Byrne M, Foo SA, Ross PM *et al.* Limitations of cross- and multigenerational plasticity for marine invertebrates faced with global climate change. *Glob Chang Biol* 2020;**26**:80–102.

24. Parker LM, Ross PM, O'Connor WA *et al.* Adult exposure influences offspring response to ocean acidification in oysters. *Glob Chang Biol* 2012;**18**:82–92.

25. Parker LM, O'Connor WA, Raftos DA *et al.* Persistence of Positive Carryover Effects in the Oyster, *Saccostrea glomerata*, following Transgenerational Exposure to Ocean Acidification. *PLoS One* 2015;**10**:e0132276.

26. Spencer LH, Venkataraman YR, Crim R *et al.* Carryover effects of temperature and pCO_2 across multiple Olympia oyster populations. *Ecol Appl* 2019:e02060.

27. Griffith AW, Gobler CJ. Transgenerational exposure of North Atlantic bivalves to ocean acidification renders offspring more vulnerable to low pH and additional stressors. *Sci Rep* 2017;**7**:11394.

28. 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.

29. Deans C, Maggert KA. What Do You Mean, "Epigenetic"? *Genetics* 2015;**199**:887–96.

30. Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev* 2002;**16**:6–21.

31. Eirin-Lopez JM, Putnam HM. Marine Environmental Epigenetics. *Ann Rev Mar Sci* 2018, DOI: 10.1146/annurev-marine-010318-095114.

32. Bogan SN, Yi SV. Potential Role of DNA Methylation as a Driver of Plastic Responses to the Environment across Cells, Organisms, and Populations. *Genome Biol Evol* 2024, DOI: 10.1093/gbe/evae022.

33. 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.

34. Liew YJ, Zoccola D, Li Y *et al.* Epigenome-associated phenotypic acclimatization to ocean acidification in a reef-building coral. *Science Advances* 2018.

35. Strader ME, Kozal LC, Leach TS *et al.* Examining the Role of DNA Methylation in Transcriptomic Plasticity of Early Stage Sea Urchins: Developmental and Maternal Effects in a Kelp Forest Herbivore. *Frontiers in Marine Science* 2020;**7**:205.

36. Downey-Wall AM, Cameron LP, Ford BM *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.

37. Lim Y-K, Cheung K, Dang X *et al.* 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.

38. Venkataraman YR, Downey-Wall AM, Ries J *et al.* General DNA methylation patterns and environmentally-induced differential methylation in the eastern oyster (*Crassostrea virginica*). *Frontiers in Marine Science* 2020, DOI: 10.3389/fmars.2020.00225.

39. Chandra Rajan K, Yuan M, Yu Z *et al.* Oyster biomineralisation under ocean acidification: from genes to shell. *Glob Chang Biol* 2021, DOI: 10.1111/gcb.15675.

40. Venkataraman YR, White SJ, Roberts SB. Differential DNA methylation in Pacific oyster reproductive tissue in response to ocean acidification. *BMC Genomics* 2022;**23**:556.

41. Dang X, Lim Y-K, Li Y *et al.* Epigenetic-associated phenotypic plasticity of the ocean acidification-acclimated edible oyster in the mariculture environment. *Mol Ecol* 2023;**32**:412–27.

42. Rondon R, Grunau C, Fallet M *et al.* Effects of a parental exposure to diuron on Pacific oyster spat methylome. *Environ Epigenet* 2017;**3**, DOI: 10.1093/eep/dvx004.

43. Johnson KM, Sirovy KA, Casas SM *et al.* Characterizing the Epigenetic and Transcriptomic Responses to Perkinsus marinus Infection in the Eastern Oyster *Crassostrea virginica*. *Frontiers in Marine Science* 2020;**7**, DOI: 10.3389/fmars.2020.00598.

44. 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, DOI: 10.1111/1365-2656.13645.

45. McNally EM, Downey-Wall AM, Titmuss FD *et al.* Parental exposure of Eastern oysters (*Crassostrea virginica*) to elevated pCO_2 mitigates its negative effects on early larval shell growth and morphology. *Limnol Oceanogr* 2022;**67**:1732–45.

46. Sun D, Li Q, Yu H. DNA methylation differences between male and female gonads of the oyster reveal the role of epigenetics in sex determination. *Gene* 2022:146260.

47. Sun D, Yu H, Kong L *et al.* The role of DNA methylation reprogramming during sex determination and sex reversal in the Pacific oyster *Crassostrea gigas*. *Int J Biol Macromol* 20424:128964.

48. Teng W, Xu C, Liu S *et al.* DNA methylation in oysters is influenced by genetics and sex. *Research Square* 2024, DOI: 10.21203/rs.3.rs-3866716/v1.

49. Broquard C, Saowaros S-A, Lepoittevin M *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.

50. Sarda S, Zeng J, Hunt BG *et al.* The evolution of invertebrate gene body methylation. *Mol Biol Evol* 2012;**29**:1907–16.

51. Dixon G, Matz M. Changes in gene body methylation do not correlate with changes in gene expression in Anthozoa or Hexapoda. *BMC Genomics* 2022;**23**:234.

52. Feil R, Fraga MF. Epigenetics and the environment: emerging patterns and implications. *Nat Rev Genet* 2012;**13**:97–109.

53. Yi SV. Insights into Epigenome Evolution from Animal and Plant Methylomes. *Genome Biol Evol* 2017;**9**:3189–201.

54. Sol Dourdin T, Guyomard K, Rabiller M *et al.* Ancestors' Gift: Parental Early Exposure to the Environmentally Realistic Pesticide Mixture Drives Offspring Phenotype in a Larger Extent Than Direct Exposure in the Pacific Oyster, *Crassostrea gigas*. *Environ Sci Technol* 2024, DOI: 10.1021/acs.est.3c08201.

55. Roberts SB, Gavery MR. Is there a relationship between DNA methylation and phenotypic plasticity in invertebrates? *Front Physiol* 2012;**2**, DOI: 10.3389/fphys.2012.00116/abstract.

56. 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.

57. Bogan SN, Strader ME, Hofmann GE. Associations between DNA methylation and gene regulation depend on chromatin accessibility during transgenerational plasticity. *BMC Biol* 2023;**21**:149.

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

59. Belardin LB, Antoniassi MP, Camargo M *et al.* Semen levels of matrix metalloproteinase (MMP) and tissue inhibitor of metallorproteinases (TIMP) protein families members in men with high and low sperm DNA fragmentation. *Sci Rep* 2019;**9**:903.

60. Cabral-Pacheco GA, Garza-Veloz I, Castruita-De la Rosa C *et al.* The Roles of Matrix Metalloproteinases and Their Inhibitors in Human Diseases. *Int J Mol Sci* 2020;**21**, DOI:

10.3390/ijms21249739.

61. Riviere G, Fellous A, Franco A *et al.* A crucial role in fertility for the oyster angiotensin-converting enzyme orthologue CgACE. *PLoS One* 2011;**6**:e27833.

62. Helm MM, Bourne N. *Hatchery Culture of Bivalves: A Practical Manual*. Food and Agriculture Organization of the United Nations, 2004.

63. Bogdanova EA, Barsova EV, Shagina IA *et al.* Normalization of full-length-enriched cDNA. *Methods Mol Biol* 2011;**729**:85–98.

64. Gómez-Chiarri M, Warren WC, Guo X *et al.* Developing tools for the study of molluscan immunity: The sequencing of the genome of the eastern oyster, *Crassostrea virginica*. *Fish & Shellfish Immunology* 2015;**46**:2–4.

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

66. Tarailo-Graovac M, Chen N. Using RepeatMasker to Identify Repetitive Elements in Genomic Sequences. *Current Protocols in Bioinformatics* 2009, DOI: 10.1002/0471250953.bi0410s25.

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

68. White S, Roberts S. *C.virginica* methylation and gene expression. 2022, DOI: 10.17605/OSF.IO/XUY2F.

69. Di Tommaso P, Chatzou M, Floden EW *et al.* Nextflow enables reproducible computational workflows. *Nat Biotechnol* 2017;**35**:316–9.

70. Nunn A, Can SN, Otto C *et al.* EpiDiverse Toolkit: a pipeline suite for the analysis of bisulfite sequencing data in ecological plant epigenetics. *NAR Genom Bioinform* 2021;**3**:lqab106.

71. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* 2011;**27**:1571–2.

72. Korneliussen TS, Moltke I. NgsRelate: a software tool for estimating pairwise relatedness from next-generation sequencing data. *Bioinformatics* 2015;**31**:4009–11.

73. Silliman K, Spencer LH, White SJ *et al.* Epigenetic and Genetic Population Structure is Coupled in a Marine Invertebrate. *Genome Biol Evol* 2023;**15**, DOI: 10.1093/gbe/evad013.

74. Chen S, Zhou Y, Chen Y *et al.* fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 2018;**34**:i884–90.

75. Ewels P. *MultiQC*. Github

76. Kim D, Paggi JM, Park C *et al.* Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 2019;**37**:907–15.

77. Pertea M, Kim D, Pertea GM *et al.* Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat Protoc* 2016;**11**:1650–67.

78. Pertea M, Pertea GM, Antonescu CM *et al.* StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol* 2015;**33**:290–5.

79. R Core Team. R: A Language and Environment for Statistical Computing. 2021.

80. Fu J, Frazee AC, Collado-Torres L *et al.* ballgown: Flexible, isoform-level differential expression analysis. 2021.

81. Wickham H, Averick M, Bryan J *et al.* Welcome to the tidyverse. *Journal of Open Source Software* 2019;**4**:1686.

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

83. Jarmund AH, Madssen TS, Giskeødegård GF. ALASCA: An R package for longitudinal and cross-sectional analysis of multivariate data by ASCA-based methods. *Front Mol Biosci* 2022;**9**:962431.

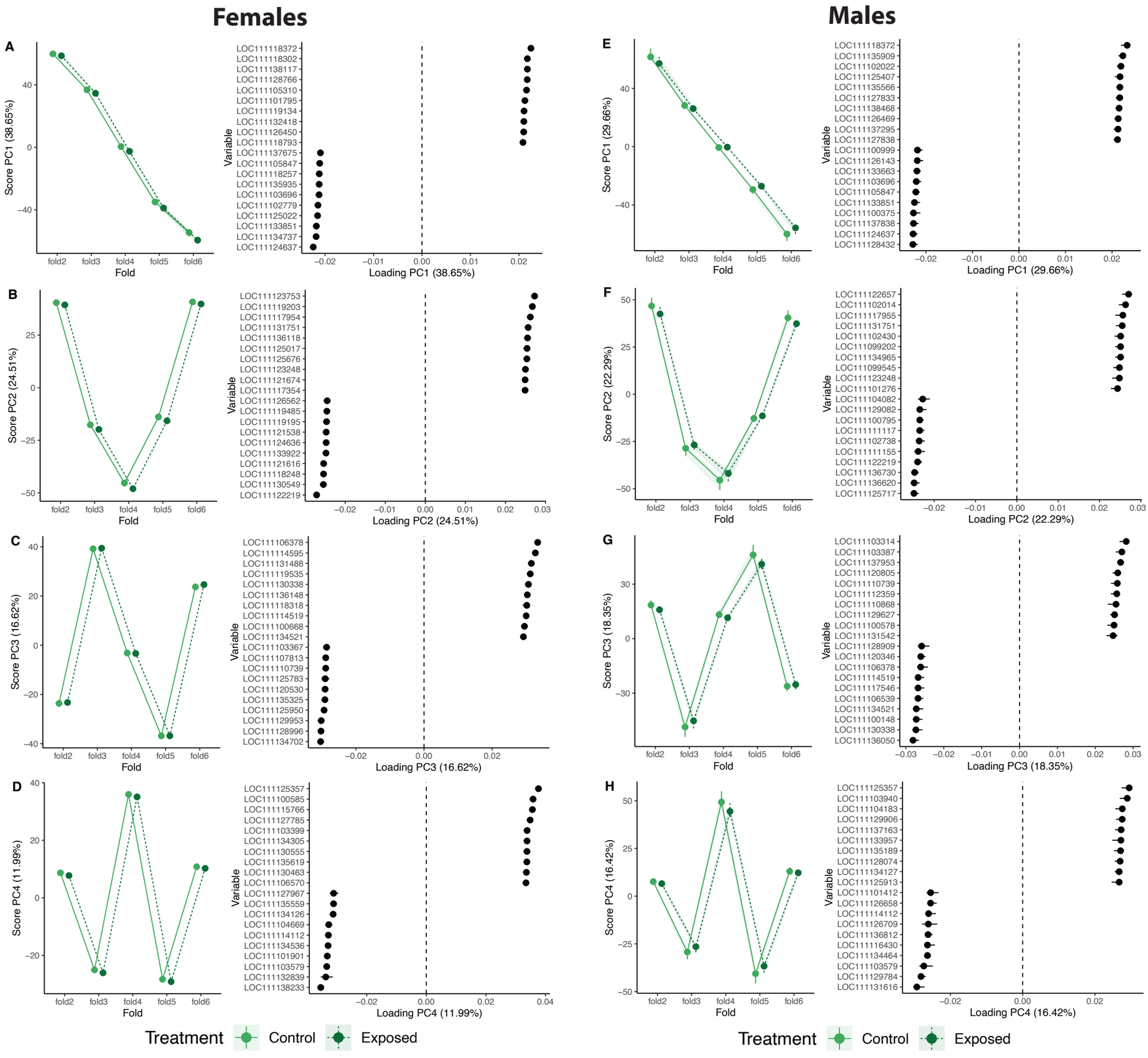
84. Bertinetto C, Engel J, Jansen J. ANOVA simultaneous component analysis: A tutorial review. *Anal Chim Acta X* 2020;**6**:100061.

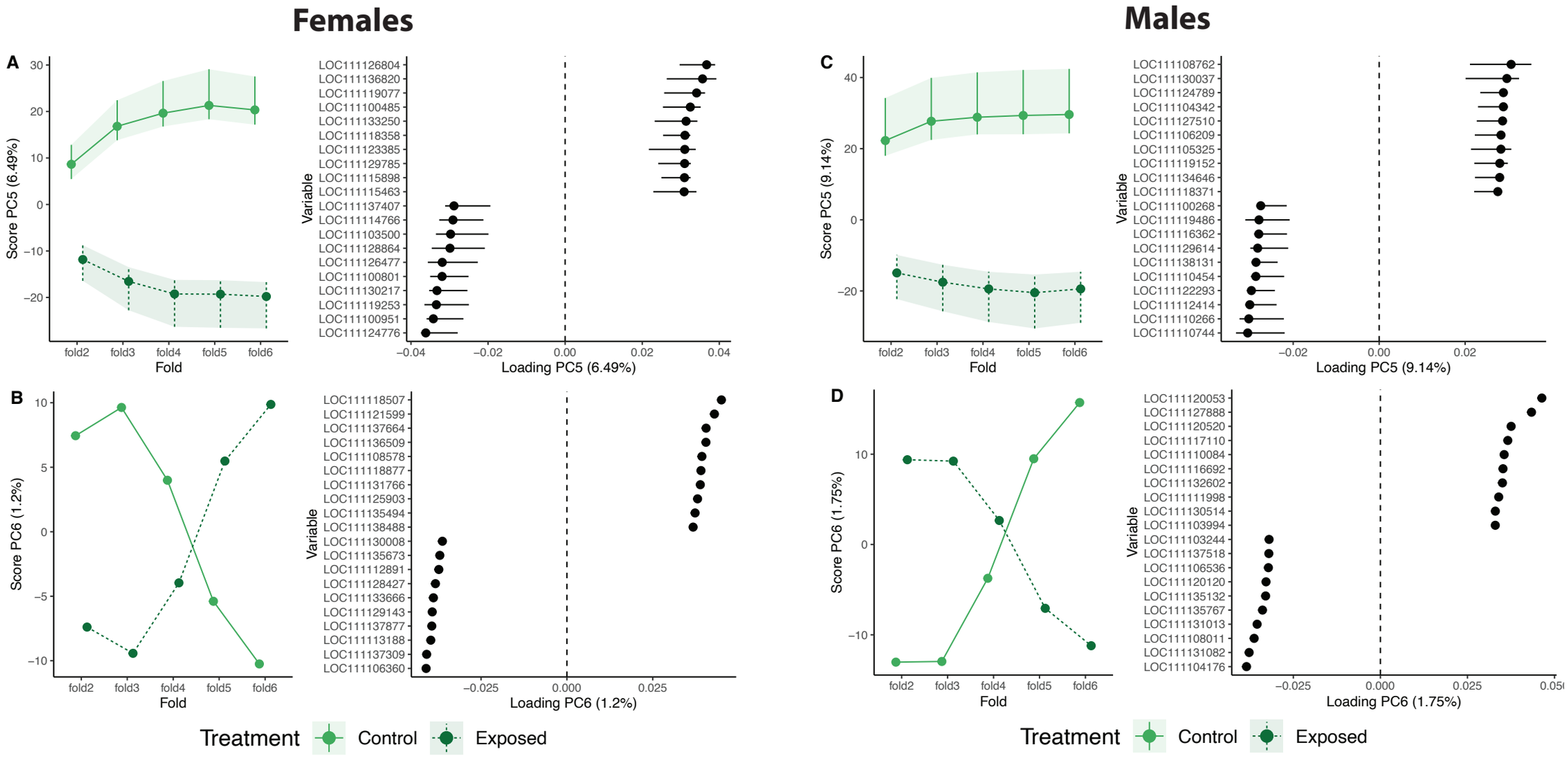
85. Pattengale ND, Alipour M, Bininda-Emonds ORP *et al.* How many bootstrap replicates are necessary? *J Comput Biol* 2010;**17**:337–54.

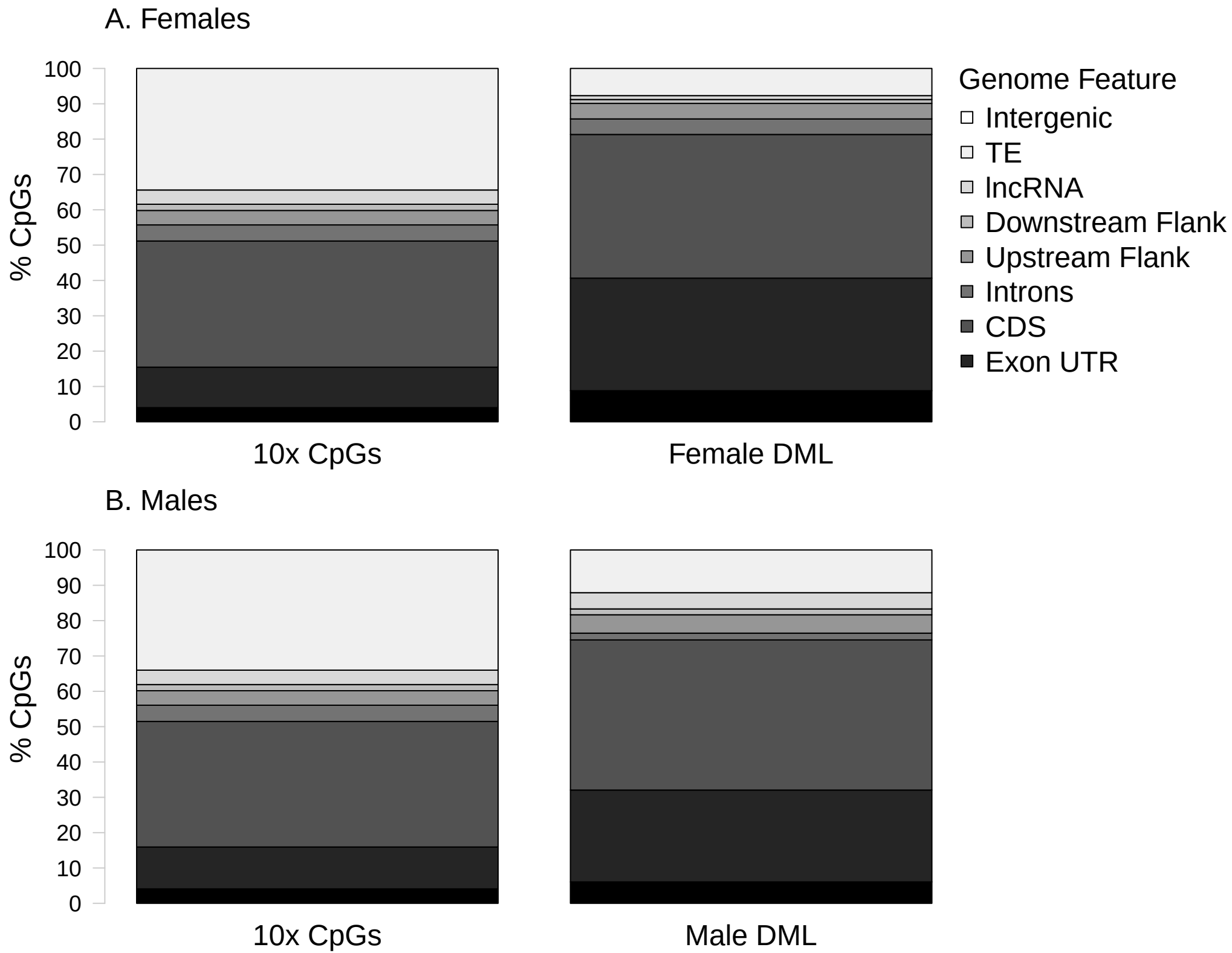
86. Gao S, Zou D, Mao L *et al.* BS-SNPer: SNP calling in bisulfite-seq data. *Bioinformatics* 2015;**31**:4006–8.

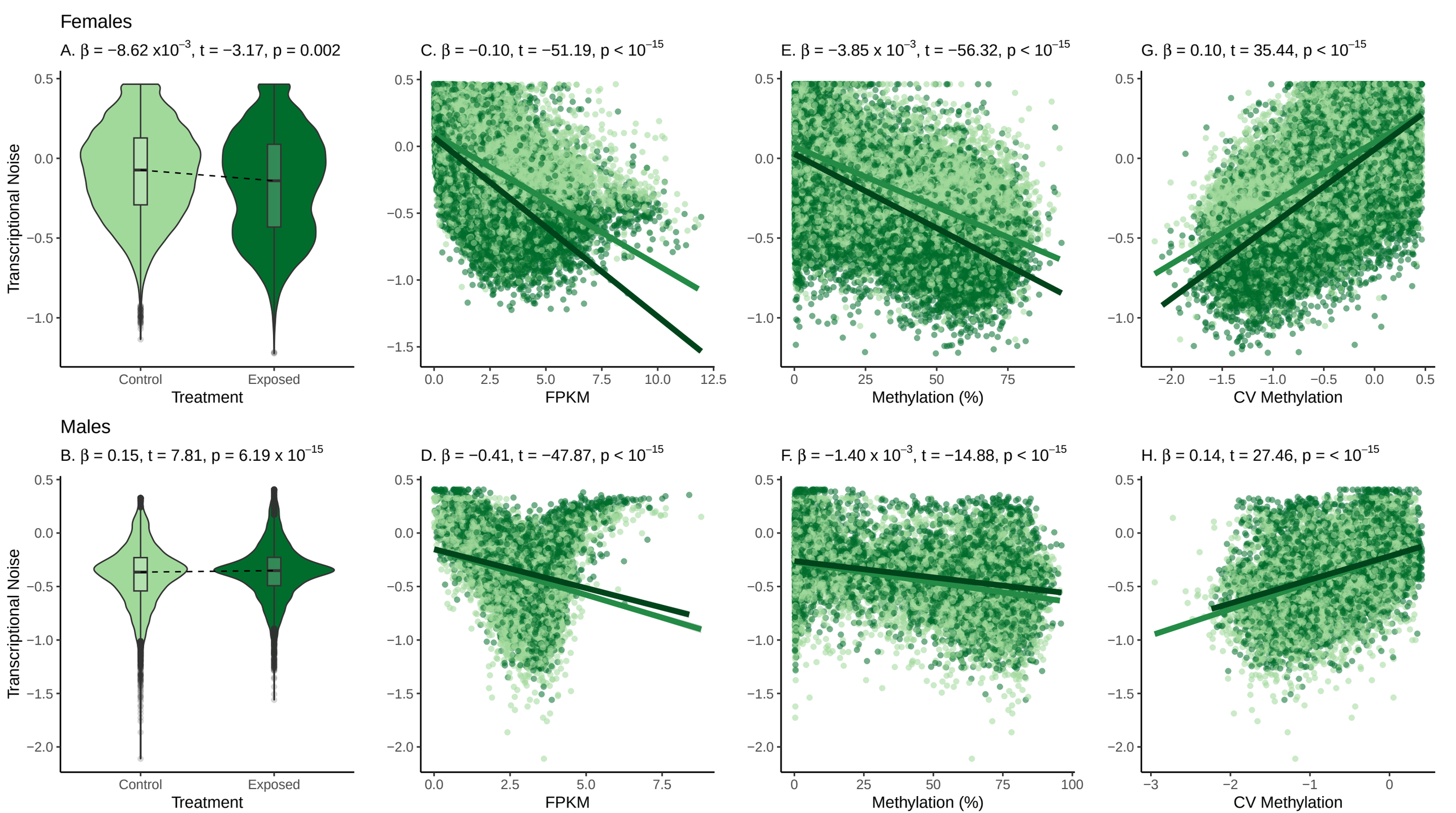
87. Akalin A, Kormaksson M, Li S *et al.* methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol* 2012;**13**:R87.

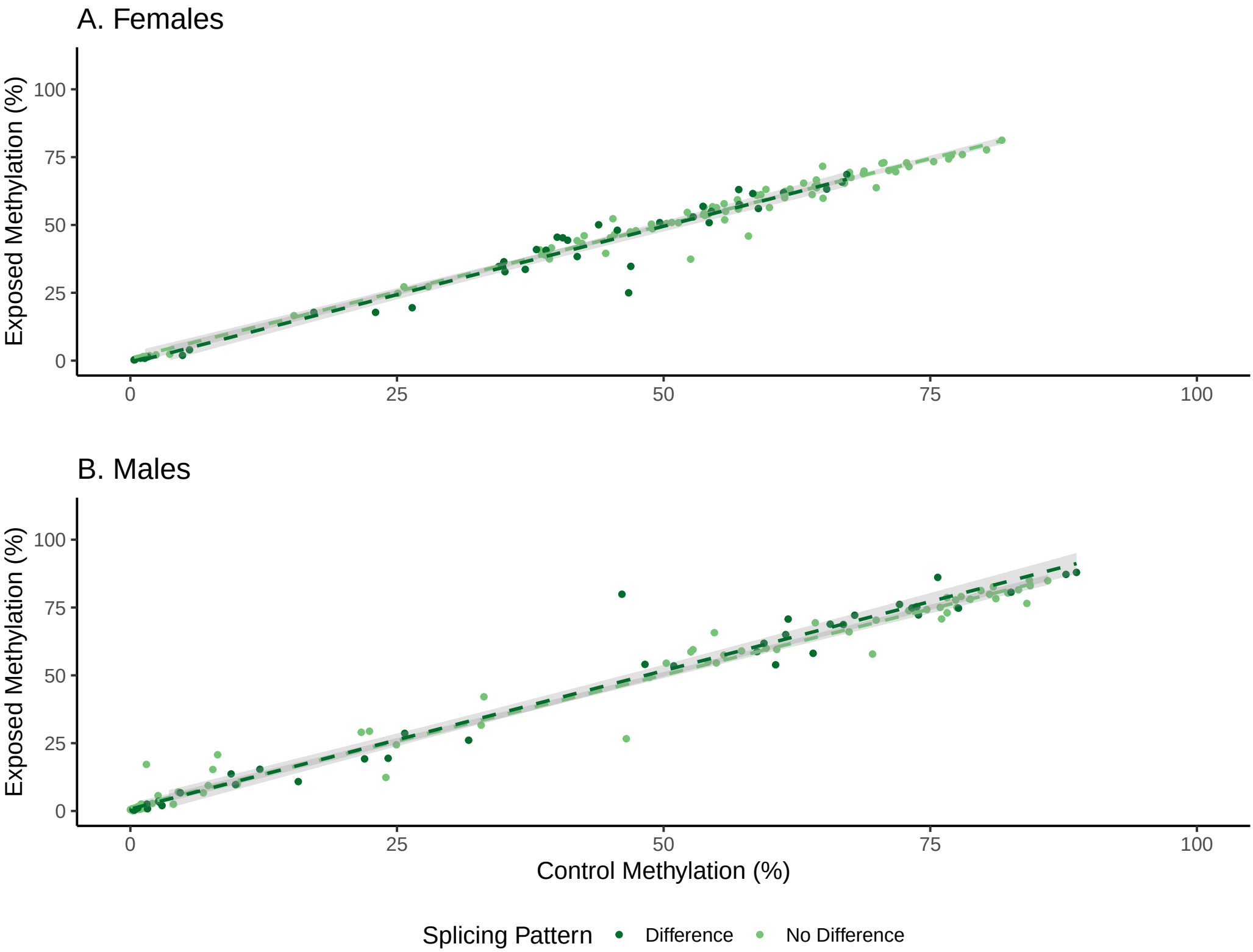
88. Wu X, Lindsey ARI, Chatterjee P *et al.* Distinct epigenomic and transcriptomic modifications associated with Wolbachia-mediated asexuality. *PLoS Pathog* 2020;**16**:e1008397.


Females
Males
A
Score PC1 (38.65%)
Fold
Variable
Loading PC1 (38.65%)
B
Score PC2 (24.51%)
Loading PC2 (24.51%)
C
Score PC3 (16.62%)
Loading PC3 (16.62%)
D
Score PC4 (11.99%)
Loading PC4 (11.99%)
E
Score PC1 (29.66%)
Loading PC1 (29.66%)
F
Score PC2 (22.29%)
Loading PC2 (22.29%)
G
Score PC3 (18.35%)
Loading PC3 (18.35%)
H
Score PC4 (16.42%)
Loading PC4 (16.42%)
fold2
fold3
fold4
fold5
fold6
Treatment
Control
Exposed


Females
Males
A
B
C
D
Score PC5 (6.49%)
Score PC6 (1.2%)
Score PC5 (9.14%)
Score PC6 (1.75%)
Fold
fold2
fold3
fold4
fold5
fold6
Variable
LOC111126804
LOC111136820
LOC111119077
LOC111100485
LOC111133250
LOC111118358
LOC111123385
LOC111129785
LOC111115898
LOC111115463
LOC111137407
LOC111114766
LOC111103500
LOC111128864
LOC111126477
LOC111100801
LOC111130217
LOC111119253
LOC111100951
LOC111124776
Loading PC5 (6.49%)
LOC111118507
LOC111121599
LOC111137664
LOC111136509
LOC111108578
LOC111118877
LOC111131766
LOC111125903
LOC111135494
LOC111138488
LOC111130008
LOC111135673
LOC111112891
LOC111128427
LOC111133666
LOC111129143
LOC111137877
LOC111113188
LOC111137309
LOC111106360
Loading PC6 (1.2%)
LOC111108762
LOC111130037
LOC111124789
LOC111104342
LOC111127510
LOC111106209
LOC111105325
LOC111119152
LOC111134646
LOC111118371
LOC111100268
LOC111119486
LOC111116362
LOC111129614
LOC111138131
LOC111110454
LOC111122293
LOC111112414
LOC111110266
LOC111110744
Loading PC5 (9.14%)
LOC111120053
LOC111127888
LOC111120520
LOC111117110
LOC111110084
LOC111116692
LOC111132602
LOC111111998
LOC111130514
LOC111103994
LOC111103244
LOC111137518
LOC111106536
LOC111120120
LOC111135132
LOC111135767
LOC111131013
LOC111108011
LOC111131082
LOC111104176
Loading PC6 (1.75%)
Treatment
Control
Exposed


A. Females
100
90
80
70
60
50
40
30
20
10
0
% CpGs
10x CpGs
Female DML
Genome Feature
Intergenic
TE
lncRNA
Downstream Flank
Upstream Flank
Introns
CDS
Exon UTR
B. Males
100
90
80
70
60
50
40
30
20
10
0
% CpGs
10x CpGs
Male DML


Females
A. β = −8.62 x10−3, t = −3.17, p = 0.002
C. β = −0.10, t = −51.19, p < 10−15
E. β = −3.85 x 10−3, t = −56.32, p < 10−15
G. β = 0.10, t = 35.44, p < 10−15
Males
B. β = 0.15, t = 7.81, p = 6.19 x 10−15
D. β = −0.41, t = −47.87, p < 10−15
F. β = −1.40 x 10−3, t = −14.88, p < 10−15
H. β = 0.14, t = 27.46, p = < 10−15
Transcriptional Noise
Control
Exposed
Treatment
FPKM
Methylation (%)
CV Methylation


A. Females
Exposed Methylation (%)
100
75
50
25
0
0
25
50
75
100
B. Males
Exposed Methylation (%)
100
75
50
25
0
0
25
50
75
100
Control Methylation (%)
Splicing Pattern
Difference
No Difference