Associations between DNA methylation and gene regulation depend on chromatin accessibility during transgenerational plasticity

Samuel N. Bogan[1*] , Marie E. Strader[1,2] and Gretchen E. Hofmann[1]

Abstract

Background Epigenetic processes are proposed to be a mechanism regulating gene expression during phenotypic plasticity. However, environmentally induced changes in DNA methylation exhibit little-to-no association with differential gene expression in metazoans at a transcriptome-wide level. It remains unexplored whether associations between environmentally induced differential methylation and expression are contingent upon other epigenomic processes such as chromatin accessibility. We quantified methylation and gene expression in larvae of the purple sea urchin *Strongylocentrotus purpuratus* exposed to different ecologically relevant conditions during gametogenesis (maternal conditioning) and modeled changes in gene expression and splicing resulting from maternal conditioning as functions of differential methylation, incorporating covariates for genomic features and chromatin accessibility. We detected significant interactions between differential methylation, chromatin accessibility, and genic feature type associated with differential expression and splicing.

Results Differential gene body methylation had significantly stronger effects on expression among genes with poorly accessible transcriptional start sites while baseline transcript abundance influenced the direction of this effect. Transcriptional responses to maternal conditioning were 4–13× more likely when accounting for interactions between methylation and chromatin accessibility, demonstrating that the relationship between differential methylation and gene regulation is partially explained by chromatin state.

Conclusions DNA methylation likely possesses multiple associations with gene regulation during transgenerational plasticity in *S. purpuratus* and potentially other metazoans, but its effects are dependent on chromatin accessibility and underlying genic features.

Keywords Epigenetics, DNA Methylation, Gene expression, Environment, Invertebrate, Chromatin

Background

Epigenetic modifications to the genome such as DNA methylation are one suite of regulatory factors that, in some cases, underpin phenotypic plasticity by affecting

changes in transcription [1] and subsequent phenotypes [2]. In particular, epigenetic variation is hypothesized to drive transgenerational plasticity (TGP): the influence of parental environment on the phenotype of offspring and latter generations via non-genetic means [3–5]. Connections between DNA methylation, plasticity, and acclimatization hinge on how and whether DNA methylation influences gene regulation. At a whole genome level, invertebrate DNA methylation exhibits negligible relationships with differential expression (DE) and other modes of gene regulation in environmental studies

*Correspondence:
Samuel N. Bogan
samuel.n.bogan@gmail.com
[1] Department of Ecology, Evolution and Marine Biology, University of California Santa Barbara, Santa Barbara, USA
[2] Department of Biology, Texas A&M University, College Station, USA

Bogan *et al. BMC Biology*   (2023) 21:149

Page 2 of 18

[6]. These poor associations are frequently observed in non-mammalian vertebrates as well [7–13], calling into question an epigenetic basis for acclimatization among many metazoans. However, DNA methylation is unlikely to influence gene expression independent of other epigenetic and genetic factors [14]. Using the purple sea urchin *Strongylocentrotus purpuratus* as a model invertebrate, we tested the hypothesis that the effects of differential methylation (DM) on gene regulation (differential expression and alternative splicing) are contingent upon additional epigenomic and genomic states such as chromatin accessibility and genic architecture (the frequency, length, and arrangement of features such as promoters, exons, and introns) by integrating larval RNA-seq and reduced representation bisulfite sequencing (RRBS) data from our study of transgenerational plasticity [15] with publicly available ATAC-seq data for *S. purpuratus* larvae.

Multiple lines of evidence support DNA methylation's association with plasticity in response to changing environments, but direct or indirect causal mechanisms remain unknown. Temporospatial environmental variation has been linked to modifications in invertebrate methylomes independent of genetic variation, demonstrating a potential role for DM during acclimatization [16–18]. DM of genes, gene modules, or whole genomes induced by environmental variation are associated with performance traits in stony corals, mollusks, crustaceans, and insects [18–21]. Causative tests of DNA methylation's effect on phenotype conducted in vertebrates, arthropods, and cnidarians have demonstrated significant effects of DNA methyltransferase inhibition on changes in performance, development, and survival across temperature [22–25]. By contrast, inducible changes in invertebrate gene expression in response to environmental variation have frequently possessed insignificant relationships with differential gene body methylation or GBM at a transcriptome-wide level [6, 15, 20, 26, 27]. This lack of relationship between transcriptome-wide DM and DE is also evident across environmental studies of non-mammalian vertebrates to date including fishes [7–10], birds [12], and reptiles [13]. If environmental variation in DNA methylation is associated with or influences phenotype in some invertebrates and metazoans, how does this occur in the absence of an association with gene expression?

DNA methylation's effects on gene regulation and its interactions with other epigenetic factors can both be numerous. Most invertebrate phyla exhibit sparsely methylated genomes punctuated by high levels of CpG methylation within gene bodies [28–31]. GBM positively correlates with baseline gene expression in cnidarians [6, 19, 29], bivalve mollusks of *Crassostrea sp.* [26,

27], arthropods [31–33], and vertebrates [34]. However, experimental demethylation of gene bodies in plants [35] and whole genomes in invertebrates [36] has resulted in upregulation and downregulation of corresponding genes, calling into question whether this association is obligately positive. Alternative splicing can be an important molecular response to environmental variation [37, 38] and is associated with baseline GBM in some invertebrates [31–35]. Here, baseline GBM refers to the percent methylation of a gene controlling against environmental effects: if differential GBM can be represented by a slope across environments, baseline GBM is that slope's intercept. Thus far, changes in alternative splicing in response to environmental variation have shown weak relationships with differential GBM [12]. Invertebrate DNA methylation is also associated with chromatin state [31, 39] and the suppression of spurious intragenic transcription [19]. For example, DE between cell types is more strongly associated with DM when accompanied by differential chromatin accessibility [40]. Relationships between DM and gene expression are also dependent on genic architecture such that enhanced or silenced expression can be driven by changes to DNA methylation at specific genic features such as promoters and/or exons while variation in methylation at other genic regions can yield lesser effects [33, 34, 41]. Determining the function of DM during transcriptional responses to the environment thus requires an integrated approach that considers genic architecture, additional epigenetic states such as chromatin accessibility, and multiple modes of gene regulation. Such interactions between chromatin state and DNA methylation remain untested in the context of environmental adaptation and acclimation.

The purple sea urchin *Strongylocentrotus purpuratus* is a uniquely poised model invertebrate in which to conduct an integrative test of DNA methylation's regulatory roles during phenotypic plasticity. *S. purpuratus* is an abundant herbivore distributed throughout North America's Pacific subtidal kelp forests and rocky intertidal [42]. Populations inhabiting environmental gradients or mosaics exhibit genetic evidence of local adaptation and interpopulation variation in performance and gene expression under ecologically relevant stress [43–45]. TGP linked to maternal effects has been observed in *S. purpuratus* for traits including egg protein content, larval body size, gene expression, and DNA methylation [15, 46–50]. Maternal conditioning of *S. purpuratus* to abiotic conditions mimicking coastal upwelling can induce 3–6×greater DM in offspring larvae relative to the effects of larval development under upwelling [15, 50]. These results suggest a function for DM in facilitating TGP's effects on gene expression, but negligible overlap between DM CpGs and DE genes has left that role ambiguous [15].

Bogan *et al. BMC Biology*   (2023) 21:149

Page 3 of 18

Accounting for interactions between DNA methylation and chromatin accessibility in *S. purpuratus,* which may better explain epigenetic effects on transcription, is made possible by the species' use as a model of deuterostome embryology, yielding developmental time series of ATAC-seq spanning the two-cell embryo to late prism larvae.

To elucidate the gene regulatory roles of DM during TGP, we quantified changes in DNA methylation, gene expression, and alternative splicing in prism larvae of *S. purpuratus* induced by maternal exposure to experimental upwelling during gametogenesis, an ecologically relevant, abiotic stress. Coastal upwelling increases pCO_2 and decreases temperature as a result of wind-driven upward movement of deep, cold seawater [51]. We then modeled DE and splicing as functions of DM, genic feature type, and chromatin accessibility to test the hypothesis that DNA methylation's regulatory role is contingent upon genic architecture and chromatin accessibility. To pursue these aims, we integrated data from Strader et al., initially exhibiting limited overlap between DM and DE genes and ATAC-seq measures of chromatin accessibility during the *S. purpuratus* prism stage [15, 52–54] and applied a Bayesian workflow that fitted multiple model iterations spanning low–to–high dimensionality before selecting the most likely model and evaluating effects of its parameters averaged across all iterations. Model selection and parameter evaluation was facilitated by Bayes factor tests comparing marginal likelihoods: the probability of a model or effect given one's data. Bayes factor tests have generally been shown to reduce error in model predictions compared to other model selection methods, improving false positive predictions and reproducibility [55–58].

Results

Our results demonstrated (i) that differential DNA methylation was associated with gene regulation during TGP in *Strongylocentrotus purpuratus* and (ii) that these effects were conditional upon chromatin accessibility and genic architecture. We observed positive correlations between baseline DNA methylation, transcript abundance, and the occurrence of alternative splicing (i.e., the average level of each value for a given gene, controlling for environmental variation) as shown in Fig. 1. With regard to plastic changes in DNA methylation and gene regulation induced by maternal environments (Fig. 2), associations between differential gene body methylation and differential gene expression or splicing were affected by interactions between differential methylation, chromatin accessibility, and genic architecture: the strength and direction of DM's effects were contingent upon these additional genomic and accessibility states (Figs. 3 and 4).

We describe these results in three sections below, focusing first on baseline relationships between DNA methylation and transcription followed by epigenetic and gene regulatory responses to experimental upwelling. Finally, we present the results of integrated epigenomic models of DNA methylation's association with gene regulation during TGP. We also quantified the intra- and interexperimental variation in ATAC-seq measures of chromatin accessibility in early-stage *S. purpuratus* (Additional file 1: Fig. S1) in order to ensure robust integration of functional genomic data between studies.

Associations between constitutive epigenomic states and transcription

GBM in *S. purpuratus* prism larvae measured via RRBS showed significant and positive correlations with gene expression level and the probability of alternative transcriptional variants. Mean promoter methylation demonstrated a significant, albeit weak, negative effect on expression (logCPM). Mean exon and intron methylation both exhibited stronger, positive effects on expression (Fig. 1A). Genes with high levels of intron or exon methylation were more likely to exhibit transcript variants consistent with alternatively spliced isoforms, alternative TSS, and/or exon skipping. Here, the presence of alternative transcript variants serves as a measure of baseline alternative splicing: a binary variable for whether a gene possesses annotations for multiple isoforms. This baseline measure of splicing is distinct from alternative splicing in response to environmental variation. The relationship between the probability of transcript variants and promoter methylation was insignificant (Fig. 1B). Significance was determined using a probability of direction test, a Bayesian corollary of the *p*-value, and is further described under the "Methods" section [59].

Chromatin accessibility at TSS, exons, and introns was correlated with gene expression, but not with the probability of transcriptional variants. Chromatin accessibility was enriched proximal to TSS (Fig. 1C) but was greatest in introns, which exhibited a mean of 0.19 ± 0.11 ATAC-seq reads per bp compared to 0.043 ± 0.001 and 0.039 ± 0.043 in TSS and exons, respectively. Open chromatin regions showed ~30% less CpG methylation than inaccessible regions (Fig. 1D). Chromatin accessibility within ± 500 bp of TSS and gene bodies was significantly and positively correlated with gene expression, with TSS accessibility exhibiting an effect that was 32.98% stronger than gene body accessibility (Additional file 1: Fig. S2). TSS accessibility was also positively correlated with the length of first exons (e.g., distance of first intron to TSS) as shown in Additional file 1: Fig. S3 and uncorrelated with GBM (Additional file 1: Fig. S4). CpGs within -1 kb promoters, exons, and introns exhibited

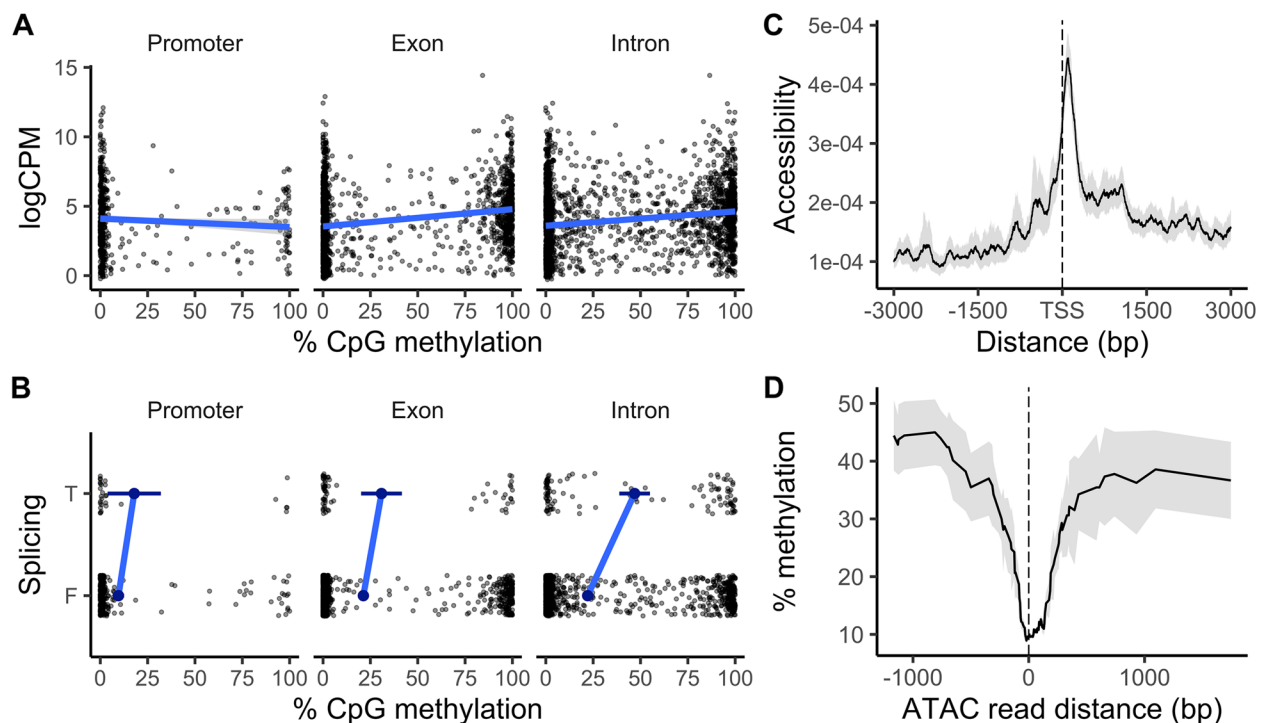Bogan *et al. BMC Biology*   (2023) 21:149

Page 4 of 18

Fig. 1 Relationships between constitutive DNA methylation, expression, and chromatin accessibility in larval *Strongylocentrotus purpuratus*. **A** Median percent methylation averaged across CpGs in the promoter, introns, and exons per gene is plotted against transcript abundance measured as logCPM. Each point depicts a single gene. **B** A binary value representing whether or not a gene was found to have transcript variants is plotted across averaged median methylation of genes' promoters, introns, and exons. Small black points depict single genes. Large blue points depict average DNA methylation among genes with or without transcript variants ± 95% CI. T/F values are spread across the *y*-axis to reduce overlap among points and better visualize their distribution across CpG methylation. **C** A loess trend of mean ATAC-seq read density (i.e., chromatin accessibility) ± 95% CI is plotted across distance to transcriptional start sites (TSS). **D** A 50-bp sliding window average of % CpG methylation ± 95% CI is plotted across distance to accessible chromatin regions for which all ATAC-seq replicates shared consensus for accessibility. *n* = 12 RNA-seq and RRBS replicate libraries; *n* = 3 ATAC-seq replicate libraries

mean methylation levels of 34.70%, 43.21%, and 44.64%, respectively. Although a strong decline in DNA methylation was observable at CpGs proximal to accessible chromatin regions, average gene-level intron and exon methylation showed no relationship with chromatin accessibility of either introns, exons, or TSS (Additional file 1: Figs. S2 and S4).

Regarding the effect of GBM on baseline gene expression, exon and intron methylation also exhibited a significant, antagonistic interaction such that genes with high methylation at both introns and exons were not more expressed than genes with high methylation at only introns or exons. Interestingly, accounting for TSS accessibility in models of logCPM resulted in the loss of a significant effect of intron methylation on gene expression. In addition to chromatin accessibility's correlation with gene expression, TSS accessibility also exhibited a significant and positive correlation with the probability of alternative transcriptional variants. This effect was insignificant however after accounting for GBM. Thus, GBM

at exon and introns was the only significant predictor of alternative splicing events.

Transcriptional and epigenetic responses to environmental variation

Maternal and developmental exposure to experimental upwelling induced changes to gene expression level (differential expression or "DE") as well as differential exon use (DEU), a measure of alternative splicing and exon skipping calculated by subtracting the logFC of DE in one exon from its corresponding transcript (Fig. 2). Developmental upwelling exposure induced 2263 upregulated and 2459 downregulated, differentially expressed genes (DEGs). Maternal exposure induced 1380 upregulated and 1025 downregulated DEGs (Fig. 2). After applying a $\log_2FC$ cutoff of > 1.0, 309 significant developmental DEGs were retained while 245 maternal DEGs were retained. Absolute logFCs of DE among maternal DEGs were significantly higher than developmental DEGs by 10.45%.
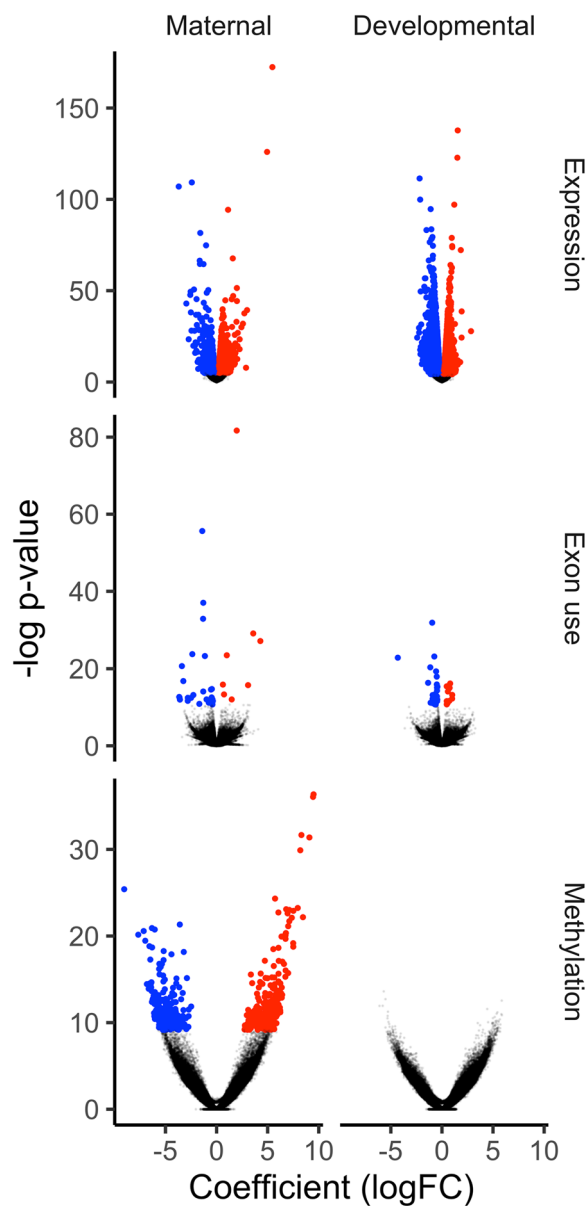
Bogan *et al. BMC Biology*   (2023) 21:149

Page 5 of 18

Fig. 2 Molecular responses to developmental and maternal upwelling exposure. Volcano plots depicting differential expression (DE), differential exon use (DEU), and differential methylation (DM). DE is depicted by measures of genewise $-\log_2 p$-values and $\log_2 FC$. DEU (e.g., splicing) is depicted as exon-level $-\log_2 p$-values, and $\Delta \log FC$ coefficients. DM is depicted as CpG-level $-\log_2 p$-values and $\log_2 FC$ of methylation. Red and blue points depict significant positive and negative coefficients, respectively (FDR < 0.05). $n = 12$ RNA-seq and RRBS replicate libraries

Significant DEU was evaluated using both gene- and exon-level tests. Developmental upwelling induced 78 alternatively spliced genes (ASGs) while maternal upwelling induced 121 ASGs, with 16 ASG genes shared between treatments (Additional file 1: Fig. S5). Significant DEU was detected among 44 exons in response to developmental upwelling: 14 upregulated or "included" exons and 30 downregulated or "dropped" exons. DEU induced by maternal upwelling occurred in 47 exons: 12 included and 35 dropped exons (Fig. 2). Maternal and developmental DE and DEU predictions are available in Additional file 2. Gene ontologies enriched among DE and DEU genes can be found in Additional file 3.

In response to maternal upwelling, 288 CpGs were hypermethylated and 233 were hypomethylated. 0 CpGs were differentially methylated in response to developmental upwelling unlike DE and DEU, which both exhibited greater or equal variation in response to developmental versus maternal exposure (Fig. 2). Among genic features, maternal upwelling induced hypermethylation at 6 promoters, 75 introns, and 32 exons and hypomethylation at 6 promoters, 57 introns, and 25 exons. Developmental upwelling exposure induced hypermethylation at 4 promoters, 19 introns, and 2 exons and hypomethylation at 0 promoters, 4 introns, and 0 exons. Because of the limited effect of developmental conditioning on DM, we exclusively focus on maternal effects for the remainder of the results section.

Forty-three and 49 genes were both differentially expressed and differentially spliced in response to developmental upwelling and maternal upwelling, respectively (Additional file 1: Fig. S5). The molecular function (MF) GO terms 'structural molecule activity' and 'structural constituent of ribosome' and the biological process terms 'obsolete GTP catabolic process,' 'small GTPase mediated signal transduction,' and 'cellular amide metabolic process' were enriched among differentially spliced exons in response to both maternal and developmental treatments. Exons differentially spliced under the maternal treatment were also enriched for the biological processes (BP) 'cellular localization' and 'nuclear transport' among others. All enriched gene ontologies among DE and DEU genes are described in Additional file 3. Additionally, functional enrichment among differentially methylated genes is extensively described in Strader et al. 2020 [15].

Correlations between differential methylation and transcriptional responses to environmental variation were shaped by chromatin state and genic architecture

A Bayesian model fitting and selection workflow was employed to robustly test for effects of DM, chromatin accessibility, and genic architecture on DE and DEU induced by maternal upwelling exposure during gametogenesis. Described in greater detail under Materials and Methods, this workflow reduced overfitting, false positive predictions, and aimed to promote reproducibility of the results below. This was achieved by fitting multiple, iterative models of DE and DEU before performing
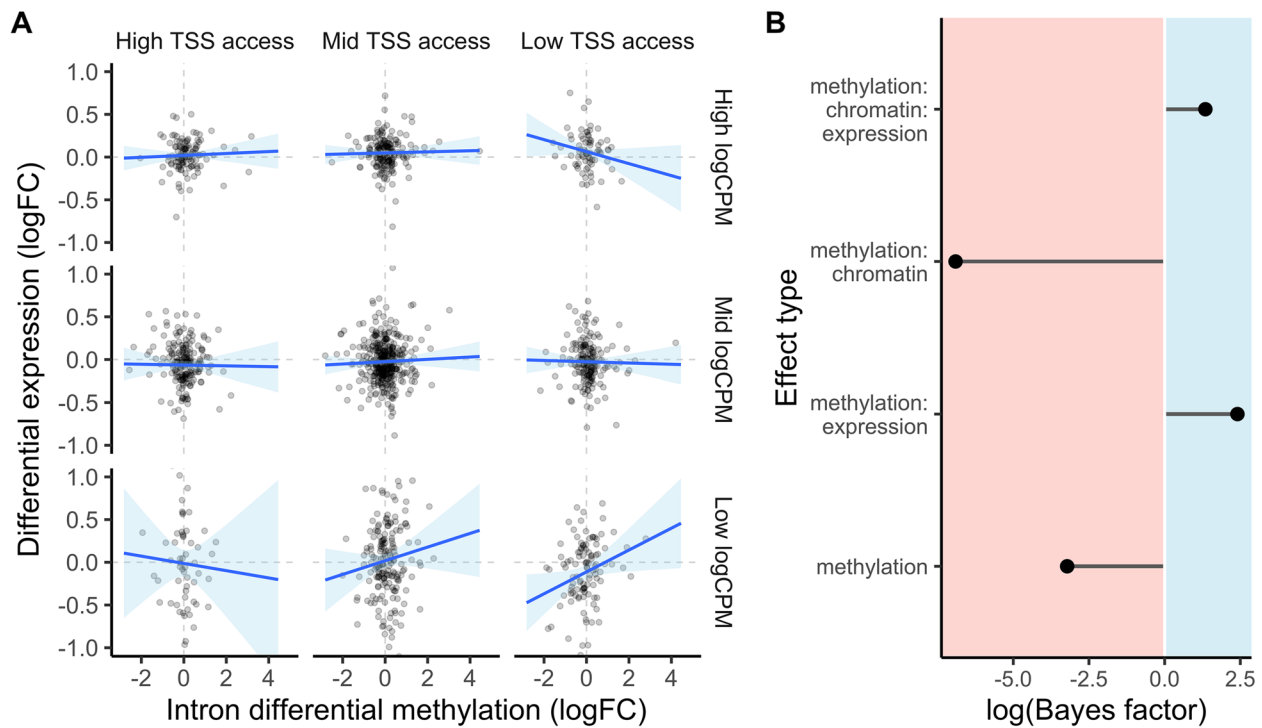
Bogan *et al. BMC Biology*   (2023) 21:149

Page 6 of 18

Fig. 3 Differential intron methylation affected expression conditional upon TSS accessibility and transcript abundance. **A** Differential gene expression in response to maternal upwelling is plotted against mean intron differential methylation. Rows and columns group data based on transcript abundance and transcriptional start site (TSS) accessibility quartiles. First and last rows/columns denote highest and lowest quartiles while the middle row/column denote second and third quartiles. **B** Log-scale inclusion Bayes factors depicting the probability of observed data under models with a parameter relative to those without it, including the interaction effect visualized in **A** ("methylation:chromatin:expression"). Positive Bayes factor values represent parameters that improved predictive strength across models and were jointly included in the selected model of differential expression responding to maternal conditioning. For parameters listed on the *y*-axis, "methylation"=intron differential methylation, "chromatin"=TSS accessibility, and "expression"=baseline gene expression (logCPM). *n* = 12 RNA-seq and RRBS replicate libraries; *n* = 3 ATAC-seq replicate libraries

model selection, parameter selection, and intensive quality checks of model predictions. For accurate interpretation of the results, it is important to state that models of DE and DEU incorporating ATAC-seq measures of chromatin accessibility were limited to testing for an effect of baseline accessibility on the outcome variable rather than differential accessibility. This is because ATAC-seq libraries were sourced from a publicly available dataset. Here, chromatin accessibility should be interpreted as a baseline measure as defined above: an expected level of chromatin accessibility, controlling for the effects of environmental variation.

The strength of associations between differential GBM and DE or DEU induced by maternal environment were dependent on chromatin accessibility within genic regions. Differential GBM across whole genes did not affect DE, while intron DM was significantly associated with DE. The selected model of DE as a function of intron DM under maternal upwelling included a significant three-way interaction between intron DM, TSS accessibility, and logCPM. Intron DM had higher

absolute effects on DE among genes with poorly accessible TSS. These effects were positive among genes with low expression and negative for those that were highly expressed, treating both expression level (logCPM) and TSS accessibility as continuous effects. For the sake of visualization, Fig. 3A shows this result using bins of logCPM and TSS accessibility rather than representing each variable as continuous. An inclusion Bayes factor test demonstrated that observed DE was 3.84×more likely under models that fit this interaction between intron DM, logCPM, and TSS accessibility relative to models lacking this parameter (Fig. 3B). Figure 3B visualizes the relative likelihood of observed DE across models with and without a given parameter including this three-way interaction and a two-way interaction between intron DM and logCPM, both of which were included in the selected model. The selected model also predicted a significant, negative, effect of constitutive intron methylation on DE (Additional file 1: Table S1). Model and parameter selection selected against singular, direct effects of intron DM (Fig. 3B). Models fit
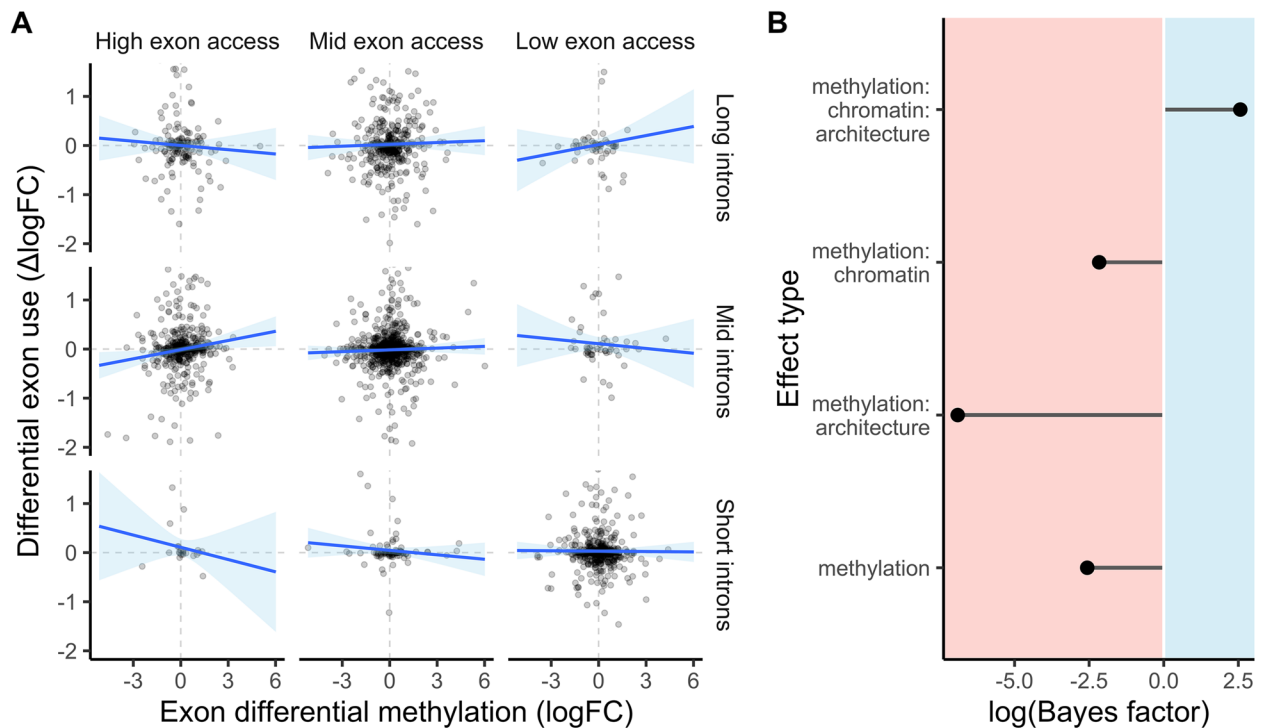
Bogan *et al. BMC Biology*   (2023) 21:149

Page 7 of 18

Fig. 4 Differential exon methylation affected splicing conditional upon exon accessibility and genic architecture. **A** Differential exon use in response to maternal upwelling is plotted against exon differential methylation. Rows and columns group data based on total genic intron length and exon accessibility quartiles. First and last rows/columns denote highest and lowest quartiles while the middle row/column denote second and third quartiles. **B** Log-scale inclusion Bayes factors depicting the probability of observed data under models fitting a parameter relative to those without it, including the interaction effect visualized in **A**. For parameters listed on the *y*-axis, "methylation" = exon differential methylation, "chromatin" = exon accessibility, and "architecture" = total genic intron length. *n* = 12 RNA-seq and RRBS replicate libraries; *n* = 3 ATAC-seq replicate libraries

with predictors for DM at promoters, first introns, and exons yielded null effects of methylation at these features.

Some model iterations in the top 25% of marginal likelihoods indicated that intron length interacted with intron DM to affect expression. Intron hypermethylation of genes in the top quartile of intron length silenced expression while genes in the middle-to-lowest quartiles of intron length showed enhanced expression as a result of intron hypermethylation (Additional file 1: Fig. S6), though the inclusion Bayes factor for this interaction was negligible (BF = 1.67). Information regarding the selected model of DE induced by maternal conditioning is available in Additional file 1: Table S1 and Additional file 1: Figs. S7–S12.

A Fisher's exact test demonstrated that genes in the lowest logCPM quartile and lowest quartile of TSS accessibility were enriched with MF GO terms that included 'nucleotidyltransferase activity' and 'cytoskeletal motor activity,' two MF terms that were also enriched among genes with CpGs that were differentially methylated under maternal upwelling [15]. The BP GO terms 'plasma

membrane bounded cell projection assembly' and 'movement of cell or subcellular component' were also enriched among genes with low TSS accessibility and expression, among others (Additional file 3). Genes in the lowest TSS accessibility quartile and highest logCPM quartile were enriched with the MF terms 'threonine-type endopeptidase activity,' 'transcription regulator activity,' and 'calcium ion binding.'

Differential methylation at singular exons induced by maternal upwelling interacted with gene body accessibility and genic architecture to affect DEU. Selected models of DEU yielded a significant three-way interaction between exon DM, exon accessibility, and total genic intron length. Positive correlations between exon DM and DEU were observed among exons from genes with poor exon accessibility while negative correlations were observed among exons from genes with high exon accessibility. The absolute effect strength of exon DM on DEU was stronger among genes with longer introns (Fig. 4A). The inclusion Bayes factor of the interaction between exon DM, exon accessibility, and intron length equaled 13.03 (Fig. 4B). The selected model of DEU also predicted

Bogan *et al. BMC Biology* (2023) 21:149

Page 8 of 18

a significant, positive effect of exon number on DEU and a singular, negative effect of exon length (Additional file 1: Table S2). Additional predictions and quality metrics for selected models of DE and DEU are available in Additional file 1: Figs. S7–S18.

Genes in the lowest quartile of exon accessibility and highest intron length quartile were enriched with the MF GO terms "calcium ion binding," "cytoskeletal motor activity," and "ATPase activity," among others, and the BP terms "purine-containing compound metabolic process", "microtubule-based movement," and "cell adhesion." Genes with long introns and high exon accessibility were also enriched with the MF term "calcium ion binding," as well as "transporter activity," "small molecule binding", and others and enriched BP terms including "cell adhesion," "localization", "regulation of intracellular signal transduction", and others (Additional file 3). Information regarding the selected model of DEU induced by maternal conditioning is available in Additional file 1: Table S2 and Additional file 1: Figs. S13–S18.

Discussion

We sought to characterize relationships between differential methylation and transcriptional plasticity in the purple urchin *Strongylocentrotus purpuratus*, a species for which DNA methylation appears to play a role in transgenerational plasticity [15, 48, 50]. Differential gene body methylation in *S. purpuratus* larvae induced by maternal upwelling exposure (low temperature and high $p\mathrm{CO}_2$) exhibited significant and strong effects on differential expression and differential exon use among subsets of genes contingent upon chromatin accessibility and genic architecture. These results support the hypotheses that differential methylation induced during TGP elicits multiple gene regulatory effects in *S. purpuratus* and, secondly, that these effects are conditional upon the chromatin state and genic feature at which DM occurs. Here we interpret our results in the context of *S. purpuratus* and invertebrate physiological ecology before highlighting questions and approaches to be pursued in future studies of ecological epigenomics in metazoans. Because developmental conditioning to upwelling induced limited differential methylation, we restrict our discussion of environmental changes in DNA methylation and its relationship with differential gene expression to the effects of maternal conditioning.

Relationships between DNA methylation, chromatin accessibility, and gene expression in *Strongylocentrotus purpuratus* and other invertebrates

Baseline patterns of genomic methylation, chromatin accessibility, transcription, and the relationships between these processes in *S. purpuratus* were largely consistent with those typically observed in other invertebrates, particularly our observation that GBM correlated with transcript abundance [19, 26, 27, 31–33, 60–65]. GBM inhibition in metazoans can reduce gene expression, with some evidence of causal links between GBM and expression [36, 66, 67]. One potential link between invertebrate GBM and gene expression could be that when intragenic DNA methylation is bound by methyl-DNA-binding domain protein 2/3, acetyltransferases are recruited to promote H3K27 acetylation and transcriptional elongation as evidenced by empirical study in insects [68]. A non-competing hypothesis suggests that GBM correlates with gene expression to support sequence conservation and transcriptional homeostasis. For example, genes with intragenic hypermethylation are less accessible in at least some invertebrates [31], which can protect them from mutation [69] and slow evolutionary rates across invertebrate lineages [70–73].

Intron methylation was also positively correlated with the prevalence of transcript variants among *S. purpuratus* genes. This is consistent with associations between GBM and alternative splicing in other metazoans. Intragenic hypermethylation recruits methyl CpG binding protein 2 (MeCP2) to splice junctions, which promotes exon recognition and, in some cases, intron retention via alterations to elongation rate [74, 75]. Our results shed additional light on patterns observed in honeybees by Flores et al. that methylated genes exhibit more alternative splicing events and those by Libbrecht et al. in ants demonstrating that included introns have higher levels of methylation than expected [60, 62]. Further investigation into the strength of intron methylation's relationship with transcript variants in *S. purpuratus* relative to other genic features may better reveal the mechanisms underpinning GBM's association with alternative splicing.

Models of logCPM that included TSS or exon accessibility as predictors demonstrated significant, positive correlations with gene expression. However, variables related to chromatin accessibility were not selected for inclusion in models of logCPM, indicating their low predictive power. Chromatin accessibility was not associated with the occurrence of transcriptional variants and splice isoforms, but it is difficult to compare this result with other case studies because these two processes have not been frequently studied together in invertebrates. Compared to its association with constitutive levels of expression and splicing, ATAC-seq estimates of chromatin accessibility provided greater predictive power in models of DE and DEU induced by environmental variation.

Bogan *et al. BMC Biology*   (2023) 21:149

Page 9 of 18

Differential methylation's associations with gene regulation depend on chromatin state and genic architecture

Our finding that DM induced by maternal environment was associated with DE and DEU conditional upon chromatin accessibility and genic architecture supports the hypotheses that (i) DM is associated with gene regulation in invertebrates during plastic responses to the environment and (ii) that these correlations are contingent upon chromatin accessibility and genic architecture. DE genes exhibited little overlap with DEU (Additional file 1: Fig. S5), but both molecular responses to environmental stress were predicted to correlate with interacting epigenomic processes. This degree of complexity underlying differential GBM's functions juxtaposes the simple relationship between promoter DM and expression exhibited across vertebrates [76]. Such complexity is expected, however [77], and our evidence helps to resolve null one-to-one correlations between differential GBM and gene expression observed in prior studies: differential GBM bears multiple associations with gene regulatory processes in *S. purpuratus*, but these associations only exist among subsets of genes based on chromatin accessibility and the architecture of gene bodies. To our knowledge, our study is the first to detect that associations between differential expression and methylation induced by abiotic environments depend on chromatin accessibility. While a great deal remains to be uncovered about the gene regulatory roles of GBM, our results aid in understanding several key areas regarding epigenomic regulation of phenotypic plasticity and molecular responses to environmental stress.

Models of invertebrate environmental adaptation or acclimatization may not be improved by accounting for singular, independently acting chromatin modifications. Studies that aim to link DNA methylation to gene expression or phenotype across experimental treatments, natural populations, or generations via inheritance may suffer from poor predictive ability without accounting for the genomic features and chromatin states at which DM is occurring. While bisulfite sequencing has rapidly increased in application to experimental and natural studies of environmental adaptation [78, 79], methods that permit for the simultaneous quantification of chromatin accessibility and DNA methylation such as ATAC-Me [80] may offer greater promise for identifying functional epigenomic patterns when combined with RNA-seq. The need for integrated epigenomic studies in metazoans, and ecology and evolution at large, should not come as a surprise. Even in plants and mammals for which DNA methylation's regulatory roles are better understood, its associations with gene expression can be tenuous at whole-genome levels [1, 81] and can depend on chromatin state [14]. For example, correlations between CpG DM and DE between mammalian cell types are stronger among genes that also exhibit differential chromatin accessibility between cells [40]. The exact process underpinning interactions between DNA methylation and chromatin accessibility shaping expression in *S. purpuratus* during TGP remain unclear, but several results related to mechanism are worth highlighting.

Associations between differential GBM and DE were strongest when intron DM occurred in genes with low TSS accessibility (Fig. 3A). Introns play a significant role in enhancing gene expression and can be essential to transcriptional initiation at canonical 5′ TSS [82, 83] or spur alternative, downstream initiation sites when first exons are distally positioned [83]. In vertebrates, first intron methylation is often negatively correlated with gene expression [34]. Introns' contributions to gene regulation can also depend on chromatin modifications in vertebrates: length of the first exon (e.g., distance of first intron from TSS) is negatively correlated with activating chromatin marks at promoters such as H3K4me3 and H3K9ac [84]. In *S. purpuratus*, however, we found that (i) methylation of first introns did not silence expression and (ii) distance of the first intron from TSS was positively correlated with TSS accessibility (Additional file 1: Fig. S3), suggesting that different processes linking intron methylation to gene expression are likely at play compared to known mechanisms in vertebrates.

Introns possess both enhancing and silencing effects on gene expression across eukaryota [85]. Among genes with inaccessible TSS that exhibited correlations between intron DM and DE, the direction of this effect also depended on gene expression level: DE of highly expressed genes negatively correlated with DM while this correlation was positive for genes with low expression (Fig. 3A). Experimental demethylation of gene bodies [35] and whole genomes [36] has resulted in a mixture of significant up- and down-regulation in eukaryotes. It is plausible that differential GBM will induce upregulation or downregulation depending on the intragenic regulatory elements existing within gene bodies and how they interact with DNA methylation. The representation of enhancing and silencing regulatory elements among genes possessing positive or negative correlations between DM an DE should be further investigated.

Differential GBM in *S. purpuratus* larvae interacted with chromatin state and genic architecture to potentially influence alternative splicing and/or exon skipping. A three-way interaction between exon DM, exon accessibility, and genic intron length affected DEU such that (i) the absolute effect of exon DM was strongest among genes

Bogan *et al. BMC Biology*   (2023) 21:149

Page 10 of 18

with greater total intron lengths and (ii) exon accessibility altered the direction of this effect. To our knowledge, our results mark the first evidence in an invertebrate of a significant association between DM and exon inclusion responding to environmental variation. Associations between DM and DEU underscore the potential for DNA methylation's impact on environmental responses to be the additive result of associations with different modes of gene regulation.

The positive correlation between exon DM and DEU in genes with inaccessible exons was expected as hypermethylation at alternatively spliced exons has generally been associated with their inclusion in transcripts [60, 86]. The effect of total intron length on DEU is also consistent with alternative splicing's pervasiveness among longer genes [60, 87], whose size we found is largely attributable to intron length in *S. purpuratus*. Chromatin accessibility at exons may influence the direction of DM's association with DEU and splicing because chromatin state and DNA methylation are both known to influence exon inclusion and skipping [74, 86, 88]. Positive or negative associations between DM and exon use can manifest depending on the intragenic regulatory elements present within a gene, which are often associated with either accessible or inaccessible chromatin states, and their corresponding cofactors that may induce exon inclusion or skipping [89, 90].

Mixtures of positive and negative associations between DM and DE or DEU may be obscured by genome-wide regressions of DE and DM performed in prior studies. Such tests assume that a unidirectional effect of DM on DE should be evident across the whole transcriptome. Looking forward, methods permitting gene-wise rather than genome-wide tests of environmentally induced DM's effect on DE such as mediation analysis [91] and network-guided multivariate regressions [92, 93] may provide a more resolved understanding of DNA methylation's association with transcriptional responses to the environment. Mediation analysis has been performed using epigenetic and phenotypic data [94, 95] and can expand on traditional DE tests to estimate both the direct effect of an environmental variable on gene expression and indirect effect of environment on expression mediated by changes in DNA methylation. These methods can also incorporate indirect effects of chromatin accessibility measured across experimental replicates.

Correlated changes in DNA methylation and gene regulation could also be explained by mechanisms alternative to a causal effect of DNA methylation on expression. For example, evidence of DE preceding DM exists across plants and animals in a variety of contexts. Following bacterial infection in human dendritic cells, the majority of changes in DNA methylation occurred after DE of corresponding genes [96]. In rice, DM proximal to genes exhibiting DE under nutritional stress largely followed changes in gene expression [97]. If we assume the hypothesis that DE in *S. purpuratus* induced by maternal conditioning triggers changes in GBM, our evidence of interactions between chromatin state and DM should be interpreted differently. As discussed by Pacis et al. and others, DM that follows DE can potentially prime promoters or enhancers to block or be bound by *trans*-acting regulatory elements during secondary responses to environmental change that follow initial DE. DM may also serve to return DEGs to normal levels following acute responses or stabilize the expression of differentially regulated genes, consistent with GBM's role in transcriptional homeostasis [19, 31, 98]. Ecological epigenomic research in *S. purpuratus* and other metazoans may better resolve the temporal coupling of DNA methylation and gene expression by integrating bisulfite sequencing and RNA-seq with time series experiments. Assuming this study were replicated with incorporation of ATAC-seq in maternal upwelling and non-upwelling treatments, it is plausible if not likely that our reported results would vary due to the inclusion of differential accessibility as a fixed effect in addition to baseline accessibility, which we achieved here. It could be the case that variance explained by baseline accessibility is better predicted by differential accessibility as more accessible regions of the genome are more likely to vary in that accessibility [99]. Alternatively, both baseline and differential accessibility could have distinct singular effects or interactive effects associated with gene regulatory responses to environmental change.

Conclusions

Variation in DNA methylation appears to be a component of molecular responses by many metazoans to predicted global change including ecologically critical, threatened groups such as stony corals [21] and polar pteropods [100] or detrimental invasive species [101]. Given the potential heritability of DNA methylation in some clades [102, 103], exacting its transcriptional and phenotypic consequences is critical for understanding the mechanistic basis of TGP. Our findings (i) provide quantitative support for the hypothesis that gene regulation by differential gene body methylation in *S. purpuratus* is affected by chromatin accessibility and genic architecture and (ii) indicate that these effects influence both gene expression and mRNA splicing. The majority of ecological epigenomic studies in metazoans have focused on the singular effects of differential methylation on expression. However, DNA methylation is not a silver bullet to predict transcriptional changes by *S. purpuratus* in response

Bogan *et al. BMC Biology*   (2023) 21:149

Page 11 of 18

to environmental variation. Rather, it is likely one cog in the epigenomic machinery contributing to plasticity in gene expression and alternative splicing. A shift toward integrated studies combining DNA methylation, chromatin accessibility, and genomic/genic architecture may be necessary to accurately quantify non-genetic sources of transcriptional and phenotypic variation in invertebrates and other eukaryotes.

Methods

The methods applied in this study served to achieve three aims: (i) evaluating relationships between baseline DNA methylation (i.e., constitutive methylation level), chromatin accessibility, and gene expression by integrating RNA-seq, bisulfite-seq, and ATAC-seq data from prism-stage *S. purpuratus*, (ii) testing for effects of maternal and developmental conditioning on differential methylation, expression, and splicing, and (iii) modeling differential expression and splicing induced by maternal conditioning as a function of variation in methylation and chromatin accessibility.

Data sources

For RNA-seq and bisulfite sequencing datasets, a controlled transgenerational experiment was performed [15]. Briefly, adult urchins were conditioned to two treatments, non-upwelling (631 ± 106 μatm pCO$_2$ and 16.8 ± 0.2 °C) and upwelling (1390 ± 307 μatm pCO$_2$ and 12.7 ± 0.5 °C), mimicking variation in their natural environment [46]. Temperature and pCO$_2$ conditions were maintained by a flow-through CO$_2$ system [104] and described in detail by Strader et al. 2020 [15]. Treated seawater was evenly pumped from two reservoir tanks to conditioning tanks at a rate of 20 L/h. Adult urchins were induced to spawn and fertilizations were performed in ambient seawater conditions using 1 non-upwelling male and pooled eggs from 9 females per treatment. Embryos were reared in the same conditions as their parents or the reciprocal condition in triplicate using a flowthrough system with seawater treated as described above.

Once larval development progressed to the early prism stage, pooled samples from 12 different culturing replicates containing 6000 larvae per pool were collected for RNA-seq and RRBS and flash frozen in liquid nitrogen before storage at -80 °C. RRBS is less biased across genomic features relative to other reduced representation bisulfite sequencing methods [105]. Libraries for polyA-enriched RNA-seq and RRBS were constructed at the UC Davis genome center and sequenced on the Illumina HiSeq 4000 (BioProject: PRJNA548926). RRBS and RNA-seq libraries were separately sequenced on two lanes of Illumina HiSeq 4000 using single end 100 bp reads ($n = 12$ samples per lane). RNA-seq libraries averaged a size of

29.72 million raw reads ± 1.60 million SD and a mean mapping efficiency of $62.11\% \pm 1.89\%$ SD. RRBS libraries averaged a size of 24.79 million raw reads ± 4.65 million SD and a mean mapping efficiency of $38.13\% \pm 0.62\%$ SD. For more information on RNA-seq and RRBS libraries' quality and specifications, please see Additional file 4.

ATAC-seq data was obtained through the Gene Expression Omnibus (BioProject: PRJNA377768). For integration with the Strader et al. datasets, we chose ATAC-seq profiles for animals at 39 h post-fertilization, the closest developmental time point for early prism larvae, for which 3 pooled samples were sequenced (GSM2520650, GSM2520651, GSM2520652). ATAC-seq bed files were concatenated and summarized using the R package *ChIP-Seeker* v1.22.1 [106] to quantify chromatin accessibility as the mean density of Tn5 ATAC-seq reads per replicate per bp. In order to evaluate the applicability of the PRJNA377768 ATAC-seq dataset to our own experiment, we measured interexperimental variation in ATAC-seq peaks identified in early-stage *S. purpuratus* [107, 108]. We found that counts of ATAC-seq peaks detected between different experiments sequencing the same life history stages were strongly correlated as shown in Additional file 1: Fig. S1 (slope of 0.89; $R^2 = 0.80$; $p < 0.0001$) and comparable to within-experiment variation between replicates (slope $= 0.80$–0.93; $R^2 = 0.65$–0.87; $p < 0.0001$). These results suggest that integrating ATAC-seq reads from PRJNA377768 with RRBS and RNA-seq data from our own experiment is robust to interexperimental variation.

Mean chromatin accessibility of ± 500 bp transcriptional start sites (TSS), introns, and exons were each calculated in both gene- and feature-wise manners for downstream analyses by counting ATAC-seq reads from all three .bed replicates per genomic feature, dividing the total reads-per-feature by feature length to estimate ATAC-seq read density, and dividing the total density by 3 in order to calculate mean density.

The sequencing approaches used for data sourced in this study were suited to downstream analyses of gene regulation such as differential exon use and genomic feature-specific changes in DNA methylation. The use of polyA-enriched RNA-seq libraries is beneficial for analyzing alternative splicing as it mitigates the contribution of unprocessed RNA to quantification of differential exon use [109]. RRBS poses fewer biases on CpG representation across genomic feature type relative to other reduced representation BS-seq methods [105].

Evaluation of ATAC-seq integration

Although ATAC-seq datasets are commonly integrated from one experiment into another, interexperimental variation in chromatin accessibility measured via

Bogan *et al. BMC Biology*   (2023) 21:149

Page 12 of 18

ATAC-seq was assessed among early life stage *S. purpuratus* in order to assess the suitability of integrating publicly available ATAC-seq data within this study. Multiple prism-stage ATAC-seq datasets do not exist, prohibiting such a test on the stage examined in this study. Rather, chromatin accessibility peaks have been identified in 20 h post-fertilization (hpf) and 24 hpf *S. purpuratus* blastulas: GSE160461 and GSE96927 [107, 108, 110, 111]. Because of differences in sampling time point, sequencing parameters, and bioinformatic pipelines, these two datasets are likely to serve as an underestimation of variance in chromatin accessibility between experiments.

.bed files of called peaks for 2 biological replicates from 20 hpf blastulas (GSE160461) and 4 .bed files from 24 hpf blastulas representing isolated primary mesenchyme cells or all other cells (GSE96927) were input into R and converted to Genomic Ranges objects using GenomicRanges v1.38.0 [112]. The genomic locations of peaks from both GSE160461 20 hpf replicates and the 2 GSE96927 "other" 24 hpf blastula cell type replicates were annotated using ChIPSeeker v1.22.1 [106] before counting the number of peaks present within $\pm$ 500 bp TSS, exons, and introns of all CDS within the *S. purpuratus* v3.1 genome build. Using this dataset of genome-wide TSS, exon, and intron ATAC-seq peak counts from both experiments, a linear regression was fit to mean-standardized peak counts per feature in GSE160461 as a function of standardized peak counts in GSE96927. Accessibility measured as the standardized number of identified ATAC-seq peaks were strongly correlated between the two experiments (Additional file 1: Fig. S1). With a slope of 1.00 representing a 1:1 correlation, this regression fit a slope of 0.89; $R^2 = 0.80$; $p < 0.0001$. This between-experiment correlation is comparable to intraexperimental variation between replicates from GSE96927 (slope = 0.93; $R^2 = 0.87$; $p < 0.0001$) and is in fact stronger than intraexperimental correlation in GSE160461 (slope = 0.80; $R^2 = 0.65$; $p < 0.0001$).

Pairwise overlaps between genomic ranges of replicates from each experiment were then identified using the find-overlappairs() function from GenomicRanges [112] using two overlap distances: -1 bp and $+500$ bp. The number of overlaps was then calculated as the proportion of peaks in the smaller dataset (GSE96927) with at least one overlap in the opposite dataset. The number of overlaps was calculated across all reciprocal comparisons between replicates from the two experiments: 2 reps × 4 reps for 8 comparisons total. The mean number of overlaps across all 8 contrasts was found to be $72.16 \pm 6.15\%$ when overlap distance was set to $+500$ bp and $50.03 \pm 7.33\%$ with an overlap distance of -1 bp.

These results demonstrate that ATAC-seq measures of chromatin accessibility are reproducible at the scale of genomic features, with some but substantially less reproducibility at the scale of singular peaks. In fact, between-experimental correlations in feature-wide accessibility were comparable to within-experiment correlations between replicates. Therefore, the approach taken in this study to fit models of gene expression with parameters representing genewise levels of accessibility at TSS, exons, and introns taken from ATAC-seq data generated in a separate experiment is robust to potential interexperimental variation.

Gene expression analyses

RNA-seq reads were trimmed of adaptor sequences and filtered for quality using *TrimGalore*. Cleaned reads were mapped to the "Spur_3.1" genome assembly using *hisat2* [113]. Gene and exon counts were compiled with *featureCounts* [114]. Gene- and exon-level read counts were filtered to retain genes with > 0.5 counts per million (CPM) across ≥ 75% of samples and analyzed in *edgeR* v3.28.1 [61] in order to estimate differential expression and differential exon use, a measure of exon inclusion or exclusion attributable to skipping and splicing. DEU was selected as a measure of alternative splicing as opposed to differential isoform expression because this study leveraged 100 bp single end RNA-seq libraries rather than long read RNA-seq, which is necessary for performing isoform-level analyses of gene expression data [115, 116]. Furthermore, downstream analyses integrating DNA methylation and measures of alternative splicing were conducted at the exon rather than isoform level, and thus, DEU's robust measure of splicing and exon skipping [117] was appropriate for this study's scope and questions.

To test for differential expression and differential exon use, gene- and exon-level counts were modeled as a function of maternal environment, developmental environment and their interaction using the robust iteration of the *edgeR* glmQLfit function to fit negative binomial generalized linear models (GLMs). DEU was assessed by applying the edgeR function diffSpliceDGE to exon-level negative binomial GLMs, which outputs exon use coefficients denoted as ΔlogFC: exon logFC–gene logFC [118, 119]. To ensure that these tests were specific to alternative splicing and exon skipping, exon-level read count data were filtered to remove transcripts that exhibited DEU consistent with changes in spurious transcription marked by gradual reductions in exon use toward 5′ regions of a transcript. Significant DE and DEU was determined using FDR-adjusted *p*-values (alpha = 0.05). Enriched gene ontologies (GO) were identified among genes exhibiting DE or DEU with Mann–Whitney *U* tests input with signed, -log *p*-values using rank-based gene ontology analysis [120].

Bogan *et al. BMC Biology* (2023) 21:149

Page 13 of 18

Robust negative binomial dispersion estimates were calculated using empirical Bayesian shrinkage with the *edgeR* function estimateGLMRobustDisp. $\log_2$ fold-changes (logFC), *F*-statistic scores, and *p*-values for gene-wise DE between maternal and developmental treatments were estimated using the *edgeR* function glmQLFTest to account for uncertainty of tagwise dispersion estimates and improve type I error control.

Quantifying DEU attributable to alternative splicing and exon skipping required the removal of genes exhibiting patterns of exon use consistent with spurious intragenic transcription and alternative TSS. Genes that are spuriously transcribed or exhibit alternative TSS possess exons with progressively lower inclusion toward 5′ ends [19]. Filtering out such genes from exon-level read counts used in DEU analysis required fitting linear models to exon-use data as a function of exon number and removing genes with positive slopes and a *y*-intercept of DEU > − 0.25. Without this filtering step, 56.0% of genes that exhibited significant DEU under maternal upwelling would likely have been attributed to alternative TSS or spurious transcription while such genes would have composed 64.9% of significant DEU under developmental upwelling. While this approach targeted DEU attributed to splicing and exon skipping, it likely removed genes with few exons for which a 5′ exon was removed during splicing. Plots of DEU trends demonstrative of spurious transcription or alternative TSS are available in the GitHub repository https://github.com/snbogan/Sp_RRBS_ATAC.

Mann–Whitney *U* tests used to test for GO enrichment were parameterized with an alpha value of 0.05 and a minimum GO-term group size of 5 genes for gene-level enrichment. Alpha equaled 0.01 and the minimum GO-term group size equaled 25 genes for exon-level enrichment to account for a mean exon count of ~ 5 per gene in the *S. purpuratus* genome [120].

DNA methylation analyses

RRBS sequences were trimmed and filtered with *Trim-Galore* specifying the –rrbs option. Trimmed RRBS reads were mapped to the "Spur_3.1" genome assembly using *Bismark* [121] and methylation was called using the bismark_methylation_extractor command with default settings and SNPs unmasked. This decision was made because no panel or database of known SNPs is available for *S. purpuratus* and the pooled, single end RNA-seq and RRBS data in our study are insufficient for SNP calling [122, 123]. As such, genetic variation is expected to have imposed random error in our estimates of methylation level and differential methylation that reduce statistical power [124]. Coverage files were used for subsequent differential methylation analysis

using an adapted *edgeR* workflow for RRBS data [125]. edgeR was selected for DM analysis to provide a statistical framework unified with estimations of DE and DEU such that methylation calls and RNA-seq counts underwent identical normalizations and were modeled using a negative binomial distribution fit with the same dispersion parameter. DM was tested among singular CpGs and genic features. To examine feature-specific responses by DNA methylation to environmental treatments, DM was estimated as logFC in percent methylation of all CpGs within the − 1 kb promoters, introns, and exons of a given gene or within singular exons and introns. Here, percent methylation was measured by *Bismark* from the proportion of methylated reads relative to all unmethylated plus methylated reads aligned to a CpG [125]. Methylation or DM within all introns/exons of a gene were used as a predictor in each reported model (e.g., models of baseline or differential gene expression) except for those predicting DEU. The DM of single exons was used as a predictor variable in models of DEU because this metric of splicing is estimated for singular exons. Counts summed across features were filtered to include genes represented by ≥ 10 reads across all samples. CpGs were also filtered to remove any loci aligned to fewer than 10 reads. Significant DM for both CpGs and genic features was assessed using an alpha value of FDR < 0.05. Functional enrichment of GO terms among differentially methylated genes was assessed using Mann–Whitney *U* tests input with signed, -log *p*-values using rank-based gene ontology analysis [120].

Modeling gene regulation as a function of epigenomic variation

A Bayesian model fitting and selection workflow was used to quantitatively test the hypothesis that associations between differential methylation and gene regulation induced by maternal environment were influenced by chromatin accessibility. This workflow fit multiple models of DE logFC and DEU values that were composed of different combinations of singular and interacting predictors. These predictors related to DM across genic features, chromatin accessibility across genic features, and genic architecture. The most likely model was then selected using marginal likelihoods. The explanatory power of significant parameters within the selected model was then measured using a model averaging approach by which the likelihood of observed DE or DEU was estimated across models including a predictor relative to their likelihood under models lacking that predictor. This workflow enabled objective effect estimation regarding how predictor variables should behave or interact and reduced spurious correlation, reduced false positive predictions, and promoted reproducibility.

Sixteen iterative models were fit to DE and DEU outcome variables each using weakly informative priors that assumed null effects of DM and its interactions on gene regulation. This null effect was assumed because of the frequent observation that DM does not correlate with plastic changes in gene expression. Model selection was used to compare each of the simple or complex iterative models and select the most likely iteration using Bayes factors, which compare marginal likelihoods, the likelihood of a model integrating over its parameter space, between models. Simply put, Bayes factors compare the relative likelihoods of models given the observed data. Bayes factors also reduce prediction error relative to other selection metrics [58]. A given parameter from the selected model was determined to have a significant effect if it passed a probability of direction test, a Bayesian corollary of the *p*-value, by which 95% of an effect's posterior probability must fall above or below 0. To avoid false positive predictions, significant effects were also evaluated using an inclusion Bayes factor, which averages the likelihood of observed outcome variables given all models that include a parameter and divides that value by the averaged likelihoods across all models that exclude it [126]. An inclusion Bayes factor ≥ 3 was considered evidence of significant explanatory power. We have reported effects as significant if they passed probability of direction and inclusion Bayes factor tests.

Studentized, gaussian linear models were fitted to measures of DE and DEU as functions of DM in -1 kb promoters, introns, and/or exons, as well as logCPM, chromatin accessibility across genic features, and components of genic architecture such as the total length of genic features. Models of DEU also included predictors related to singular exons at which DEU was measured: DM of the corresponding exon, exon number, and exon length. All models were fitted using the R package *brms* v2.14.0, an R interface to the Stan programming language for specifying Bayesian models [127]. Models of DE and DEU were specified with weakly informative normal priors (mean $=0$; SD $=0.5$) for both slope (β) and intercept parameters. Posterior distributions were sampled using 4 chains at 20,000 iterations each, including 5000 warmup iterations.

Models were fitted with scaled Z-score transformations of continuous variables in order to improve model convergence and allow for comparisons of posterior β parameter distributions for predictors of different dependent variables. Unlike models of plastic changes in gene regulation (DE and DEU), models of baseline

expression (logCPM) and the occurrence of transcript variants did not undergo model and parameter selection. logCPM was measured using a gaussian linear model. The binomial presence/absence of transcript variants per gene was measured using a generalized linear model set to a Bernoulli distribution. Both models were assigned weakly informative priors for all parameters (intercepts and $\beta = 0 \pm 0.5$). Linear models of DE and DEU were fit with studentized model families to reduce prediction of artificially high or low outcome variables. To account for variation in RRBS read coverage across the data, the selected model was then refit to include an error parameter for estimated methylation and differential methylation that equaled the inverse CpG coverage of each gene or feature in the dataset. RRBS CpG coverage per feature is described in Additional file 1 for reported models (Additional file 1: Figs. S12 and S18). Posterior predictive checks were used to evaluate selected model predictions according to observed data (Additional file 1: Figs. S10 and S16).

Model selection was performed using the "bayesfactor_models" function in *bayestestR* v0.9.0 [128] to compare the likelihoods of models fit with iterative combinations of predictor variables. Probability of direction tests were conducted using the "posterior_interval" function of *brms* [127]. Inclusion Bayes factors for parameters were measured across all iterative models using the "bayesfactor_inclusion" function of *bayestestR* v0.9.0 [128]. Additional validation was performed on selected models using leave one out cross-validation and posterior predictive checks. Information about validation and predictions of selected models are available in Additional file 1. The specifications and relative likelihoods of selected and unselected models are available in the following GitHub repository: https://github.com/snbogan/Sp_RRBS_ATAC.

Abbreviations

ASG	Differentially spliced gene
ATAC	Assay for transposase-accessible chromatin
BP	Base pairs
CDS	Coding domain sequence
CPM	Counts per million
DE	Differential expression
DEG	Differentially expressed gene
DEU	Differential exon use
FC	Fold change
GBM	Gene body methylation
GLM	Generalized linear model
GO	Gene ontology
HPF	Hours post-fertilization
MF	Molecular function
RRBS	Reduced representation bisulfite sequencing
SNP	Single-nucleotide polymorphism
TGP	Transgenerational plasticity
TSS	Transcriptional start site

Bogan *et al. BMC Biology*   (2023) 21:149

Page 15 of 18

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12915-023-01645-8.

Additional file 1: Tables S1–S2. Summaries of posterior probabilities for parameters from differential expression and differential exon use models. **Figs. S1–S6.** Plots of ATAC-seq quality checks and relationships among baseline or differential accessibility, DNA methylation, and gene expression. **Fig. S7–S12.** Figures plotting quality checks of differential expression model. **Fig. S13–S18.** Figures plotting quality checks of differential exon use model.

Additional file 2. File containing genewise parameter estimates for maternal and developmental effects on differential expression and differential exon use.

Additional file 3. File containing the results of Mann Whitney U and Fisher's exact tests for gene ontological enrichment among gene sets.

Additional file 4. File containing library quality and alignment metrics for RNA-seq and RRBS replicates.

Authors' contributions

GEH conceived the original design and scope of the experiment. SNB, MES, and GEH conceived hypotheses and analytical approaches. MES led experimental execution and prepared DNA and RNA for sequencing. SNB and MES performed bioinformatic analyses. SNB wrote the manuscript with contributions from MES and input from GEH. All authors have read and approved the final manuscript.

Availability of data and materials

All data generated or analyzed during this study are included in this published article, its supplementary information files, and publicly available repositories, which we describe below. Raw RNA-seq and RRBS fastq files associated with this study are available through the NCBI Short Read Archive under the accession PRJNA548926. Scripts associated with trimming, mapping, and counting RNA-seq reads and CpG methylation are available in the following GitHub repository: https://github.com/mariestrader/S.purp_RRBS_RNAseq_2019. Code and genomic data files (all data except for fastq or bed files available on SRA) corresponding to all analyses reported in this study, their intermediate files, and outputs can be found in the following GitHub repository: https://github.com/snbogan/Sp_RRBS_ATAC. This repository's DOI is as follows: https://doi.org/10.5281/zenodo.7729770.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References

1. Kribelbauer JF, Lu X-J, Rohs R, Mann RS, Bussemaker HJ. Toward a mechanistic understanding of DNA methylation readout by transcription factors. J Mol Biol. 2019;S0022–2836(19):30617–25.
2. Xu J, Chen G, Hermanson PJ, Xu Q, Sun C, Chen W, et al. Population-level analysis reveals the widespread occurrence and phenotypic consequence of DNA methylation variation not tagged by genetic variation in maize. Genome Biol. 2019;20:243.
3. Galloway LF, Etterson JR. Transgenerational plasticity is adaptive in the wild. Science. 2007;318:1134–6.
4. Donelson JM, Salinas S, Munday PL, Shama LNS. Transgenerational plasticity and climate change experiments: where do we go from here? Glob Chang Biol. 2018;24:13–34.
5. Donelan SC, Hellmann JK, Bell AM, Luttbeg B, Orrock JL, Sheriff MJ, et al. Transgenerational plasticity in human-altered environments. Trends Ecol Evol. 2020;35:115–24.
6. Dixon GB, Matz MV. Changes in gene body methylation do not correlate with changes in gene expression in Anthozoa or Hexapoda. BMC Genom. 2022;23:234.
7. Anastasiadi D, Shao C, Chen S, Piferrer F. Footprints of global change in marine life: Inferring past environment based on DNA methylation and gene expression marks. Mol Ecol. 2021;30:747–60.
8. Wang S, et al. Hypoxia causes transgenerational impairments in reproduction of fish. Nat Commun. 2016;7:12114.
9. Jones E, Griffitt R. Oil and hypoxia alter DNA methylation and transcription of genes related to neurological function in larval *Cyprinodon variegatus*. Aquat Toxicol. 2022;251: 106267.
10. Skjærven KH, Jakt LM, Fernandes JMO, Dahl JA, Adam A-C, Klughammer J, et al. Parental micronutrient deficiency distorts liver DNA methylation and expression of lipid genes associated with a fatty-liver-like phenotype in offspring. Sci Rep. 2018;8:3055.
11. Lai KP, Wang SY, Li JW, Tong Y, Chan TF, Jin N, et al. Hypoxia causes transgenerational impairment of ovarian development and hatching success in fish. Environ Sci Technol. 2019;53:3917–28.
12. Lindner M, Verhagen I, Viitaniemi HM, Laine VN, Visser ME, Husby A, et al. Temporal changes in DNA methylation and RNA expression in a small song bird: within- and between-tissue comparisons. BMC Genomics. 2021;22:36.
13. Ruhr I, Bierstedt J, Rhen T, Das D, Singh SK, Miller S, et al. Developmental programming of DNA methylation and gene expression patterns is associated with extreme cardiovascular tolerance to anoxia in the common snapping turtle. Epigenetics Chromatin. 2021;14:42.
14. Lea AJ, Vockley CM, Johnston RA, Del Carpio CA, Barreiro LB, Reddy TE, et al. Genome-wide quantification of the effects of DNA methylation on human gene regulation. eLife. 2018;7:e37513.
15. Strader ME, Kozal LC, Leach TS, Wong JM, Chamorro JD, Housh MJ, 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. Front Mar Sci. 2020;7:205.
16. Dimond JL, Roberts SB. Convergence of DNA methylation profiles of the reef coral *Porites astreoides* in a novel environment. Front Mar Sci. 2020;6:792.
17. Rodríguez-Casariego JA, Mercado-Molina AE, Garcia-Souto D, Ortiz-Rivera IM, Lopes C, Baums IB, et al. Genome-Wide DNA methylation analysis reveals a conserved epigenetic response to seasonal environmental variation in the staghorn coral *Acropora cervicornis*. Front Mar Sci. 2020;7:560424.
18. Clark MS, Thorne MAS, King M, Hipperson H, Hoffman JI, Peck LS. Life in the intertidal: Cellular responses, methylation and epigenetics. Funct Ecol. 2018;32:1982–94.

Bogan *et al. BMC Biology*   (2023) 21:149

Page 16 of 18

19. Li Y, Liew YJ, Cui G, Cziesielski MJ, Zahran N, Michell CT, et al. DNA methylation regulates transcriptional homeostasis of algal endosymbiosis in the coral model *Aiptasia*. Sci Adv. 2018;4:142.

20. Arsenault SV, Hunt BG, Rehan SM. The effect of maternal care on gene expression and DNA methylation in a subsocial bee. Nat Commun. 2018;9:3468.

21. Putnam HM, Davidson JM, Gates RD. Ocean acidification influences host DNA methylation and phenotypic plasticity in environmentally susceptible corals. Evol Appl. 2016;9:1165–78.

22. Dai T-M, Zhang Y-K, Zhou X-P, Liu S-S. Molecular characterizations of DNA methyltransferase 3 and its roles in temperature tolerance in the whitefly Bemisia tabaci Mediterranean. Insect Mol Biol. 2018;27:123–32.

23. Fuchs B, Wang W, Graspeuntner S, et al. Regulation of polyp-to-jellyfish transition in *Aurelia aurita*. Curr Biol. 2014;24:263–73.

24. McCaw BA, Leonard AM, Stevenson TJ, Lancaster LT. Epigenetic control of temperature-dependent female reproductive life history trade-offs in seed beetles, *Callosobruchus maculatus*. bioRxiv. 2021. https://doi.org/10.1101/2021.10.08.463711.

25. Loughland I, Little A, Seebacher F. DNA methyltransferase 3a mediates developmental thermal plasticity. BMC Biol. 2021;19:11.

26. Downey-Wall AM, Cameron LP, Ford BM, McNally EM, Venkataraman YR, Roberts SB, et al. Ocean acidification induces subtle shifts in gene expression and DNA methylation in mantle tissue of the eastern oyster (*Crassostrea virginica*). Front Mar Sci. 2020;7:566419.

27. Johnson K, Sirovy K, Kelly M. Differential DNA methylation across environments has no effect on gene expression in the eastern oyster. J Anim Ecol. 2021. https://doi.org/10.1111/1365-2656.13645.

28. Keller TE, Han P, Yi SV. Evolutionary transition of promoter and gene body DNA methylation across invertebrate–vertebrate boundary. Mol Biol Evol. 2015;33:1019–28.

29. Zemach A, McDaniel IE, Silva P, Zilberman D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science. 2010;328:916–9.

30. Suzuki MM, Kerr ARW, De Sousa D, Bird A. CpG methylation is targeted to transcription units in an invertebrate genome. Genome Res. 2007;17:625–31.

31. Gatzmann F, Falckenhayn C, Gutekunst J, Hanna K, Raddatz G, Carneiro VC, et al. The methylome of the marbled crayfish links gene body methylation to stable expression of poorly accessible genes. Epigenetics Chromatin. 2018;11:57.

32. Kvist J, Gonçalves Athanàsio C, Shams Solari O, Brown JB, Colbourne JK, Pfrender ME, et al. Pattern of DNA methylation in *Daphnia*: evolutionary perspective. Genome Biol Evol. 2018;10:1988–2007.

33. Glastad KM, Gokhale K, Liebig J, Goodisman MAD. The caste- and sex-specific DNA methylome of the termite *Zootermopsis nevadensis*. Sci Rep. 2016;6:37110.

34. Anastasiadi D, Esteve-Codina A, Piferrer F. Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species. Epigenetics Chromatin. 2018;11:37.

35. Bewick AJ, Zhang Y, Wendte JM, Zhang X, Schmitz RJ. Evolutionary and experimental loss of gene body methylation and its consequence to gene expression. Genes Genomes Genetics. 2019;9:2441.

36. Lindeman LC, Thaulow J, Song Y, Kamstra JH, Xie L, Asselman J, et al. Epigenetic, transcriptional and phenotypic responses in two generations of *Daphnia magna* exposed to the DNA methylation inhibitor 5-azacytidine. Environ Epigenet. 2019;5:016.

37. Jakšić AM, Schlötterer C. The interplay of temperature and genotype on patterns of alternative splicing in *Drosophila melanogaster*. Genetics. 2016;204:315–25.

38. Chan SKN, Suresh S, Munday P, Ravasi T, Bernal MA, Schunter C. The alternative splicing landscape of a coral reef fish during a marine heatwave. Ecol Evol. 2022;12: e8738.

39. Nanty L, Carbajosa G, Heap GA, Ratnieks F, van Heel DA, Down TA, et al. Comparative methylomics reveals gene-body H3K36me3 in Drosophila predicts DNA methylation and CpG landscapes in other invertebrates. Genome Res. 2011;21:1841–50.

40. Rizzardi LF, Hickey PF, Rodriguez DiBlasi V, Tryggvadóttir R, Callahan CM, Idrizi A, et al. Neuronal brain-region-specific DNA methylation and chromatin accessibility are associated with neuropsychiatric trait heritability. Nat Neurosci. 2019;22:307–16.

41. Ehrlich M, Lacey M. DNA methylation and differentiation: silencing, upregulation and modulation of gene expression. Epigenomics. 2013;5:553–68.

42. Pearse JS. Ecological role of purple sea urchins. Science. 2006;314:940–1.

43. Kelly MW, Padilla-Gamiño JL, Hofmann GE. Natural variation and the capacity to adapt to ocean acidification in the keystone sea urchin *Strongylocentrotus purpuratus*. Glob Chang Biol. 2013;19:2536–46.

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

45. Pespeni MH, Chan F, Menge BA, Palumbi SR. Signs of adaptation to local pH conditions across an environmental mosaic in the California Current Ecosystem. Integr Comp Biol. 2013;53:857–70.

46. Hoshijima U, Hofmann GE. Variability of seawater chemistry in a kelp forest environment Is linked to in situ transgenerational effects in the purple sea urchin Strongylocentrotus purpuratus. Front Marine Sci. 2019;6:62.

47. Wong JM, Johnson KM, Kelly MW, Hofmann GE. Transcriptomics reveal transgenerational effects in purple sea urchin embryos: adult acclimation to upwelling conditions alters the response of their progeny to differential pCO2 levels. Mol Ecol. 2018;27:1120–37.

48. Wong JM, Kozal LC, Leach TS, Hoshijima U, Hofmann GE. Transgenerational effects in an ecological context: conditioning of adult sea urchins to upwelling conditions alters maternal provisioning and progeny phenotype. J Exp Mar Bio Ecol. 2019;517:65–77.

49. Strader ME, Wolak ME, Simon OM, Hofmann GE. Genetic variation underlies plastic responses to global change drivers in the purple sea urchin *Strongylocentrotus purpuratus*. Proc Biol Sci. 2022;289:20221249.

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

51. Gruber N, Hauri C, Lachkar Z, Loher D, Frölicher TL, Plattner G-K. Rapid progression of ocean acidification in the California Current System. Science. 2012;337:220–3.

52. Kudtarkar P, Cameron RA. Echinobase: an expanding resource for echinoderm genomic information. Database. 2017;2017:bax074.

53. Strader ME. Epigenomics and transcriptomics of *Stronglyocentrotus purpuratus* in a transgenerational context. 2019.

54. Cameron RA. Whole genome developmental time course of Tn5 transposon accessible chromatin patterns in sea urchin embryos. 2017.

55. Berger JO, Pericchi LR. The intrinsic Bayes factor for model selection and prediction. J Am Stat Assoc. 1996;91:109–22.

56. Hooten MB, Hobbs NT. A guide to Bayesian model selection for ecologists. Ecol Monogr. 2015;85:3–28.

57. Etz A, Wagenmakers E-J. JBS haldane's contribution to the Bayes factor hypothesis test. Stat Sci. 2017;32:313–29.

58. Ward EJ. A review and comparison of four commonly used Bayesian and maximum likelihood model selection tools. Ecol Model. 2008;211:1–10.

59. Makowski D, Ben-Shachar MS, Chen SHA, Lüdecke D. Indices of effect existence and significance in the Bayesian framework. Front Psychol. 2019;10:2767.

60. Flores K, Wolschin F, Corneveaux JJ, Allen AN, Huentelman MJ, Amdam GV. Genome-wide association between DNA methylation and alternative splicing in an invertebrate. BMC Genom. 2012;13:480.

61. Li-Byarlay H, Li Y, Stroud H, Feng S, Newman TC, Kaneda M, et al. RNA interference knockdown of DNA methyl-transferase 3 affects gene alternative splicing in the honey bee. Proc Natl Acad Sci U S A. 2013;110:12750–5.

62. 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.

63. Gao F, Liu X, Wu X-P, Wang X-L, Gong D, Lu H, et al. Differential DNA methylation in discrete developmental stages of the parasitic nematode *Trichinella spiralis*. Genome Biol. 2012;13:R100.

64. Libbrecht R, Oxley PR, Keller L, Kronauer DJC. Robust DNA methylation in the clonal raider ant brain. Curr Biol. 2016;26:391–5.

Bogan *et al. BMC Biology*   (2023) 21:149

Page 17 of 18

65. Dixon G, Liao Y, Bay LK, Matz MV. Role of gene body methylation in acclimatization and adaptation in a basal metazoan. Proc Natl Acad Sci U S A. 2018;115:13342–6.

66. Yang X, Han H, De Carvalho DD, Lay FD, Jones PA, Liang G. Gene body methylation can alter gene expression and is a therapeutic target in cancer. Cancer Cell. 2014;26:577–90.

67. Biergans SD, Giovanni Galizia C, Reinhard J, Claudianos C. Dnmts and Tet target memory-associated genes after appetitive olfactory training in honey bees. Sci Rep. 2015;5:16223.

68. Xu G, Lyu H, Yi Y, Peng Y, Feng Q, Song Q, et al. Intragenic DNA methylation regulates insect gene expression and reproduction through the MBD/Tip60 complex. iScience. 2021;24:102040.

69. Shi Y, Su X-B, He K-Y, Wu B-H, Zhang B-Y, Han Z-G. Chromatin accessibility contributes to simultaneous mutations of cancer genes. Sci Rep. 2016;6:35270.

70. Hunt B, Brisson J, Yi S, Goodisman M. Functional conservation of DNA methylation in the pea aphid and the honeybee. Genome Biol Evol. 2010;2:719–28.

71. Park J, Peng Z, Zeng J, Elango N, Park T, Wheeler D, et al. Comparative analyses of DNA methylation and sequence evolution using *Nasonia* genomes. Mol Biol Evol. 2011;28:3345–54.

72. Sarda S, Zeng J, Hunt B, Yi S. The evolution of invertebrate gene body methylation. Mol Biol Evol. 2012;29:1907–16.

73. Dixon GB, Bay LK, Matz MV. Evolutionary consequences of DNA methylation in a basal metazoan. Mol Biol Evol. 2016;33:2285–93.

74. Maunakea AK, Chepelev I, Cui K, Zhao K. Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. Cell Res. 2013;23:1256–69.

75. Wong JJ-L, Gao D, Nguyen TV, Kwok C-T, van Geldermalsen M, Middleton R, et al. Intron retention is regulated by altered MeCP2-mediated splicing factor recruitment. Nat Commun. 2017;8:15134.

76. Boyes J, Bird A. Repression of genes by DNA methylation depends on CpG density and promoter strength: evidence for involvement of a methyl-CpG binding protein. EMBO J. 1992;11:327–33.

77. Zilberman D. An evolutionary case for functional gene body methylation in plants and animals. Genome Biol. 2017;18:87.

78. Husby A. Wild epigenetics: insights from epigenetic studies on natural populations. Proc Biol Sci. 2022;289:20211633.

79. Eirin-Lopez JM, Putnam HM. Marine environmental epigenetics. Ann Rev Mar Sci. 2019;11:335–68.

80. Guerin LN, Barnett KR, Hodges E. Dual detection of chromatin accessibility and DNA methylation using ATAC-Me. Nat Protoc. 2021;16:5377–97.

81. Barnett KR, Decato BE, Scott TJ, Hansen TJ, Chen B, Attalla J, et al. ATAC-Me captures prolonged DNA methylation of dynamic chromatin accessibility loci during cell fate transitions. Mol Cell. 2020;77:1350-1364.e6.

82. Gallegos JE, Rose AB. Intron DNA sequences can be more important than the proximal promoter in determining the site of transcript initiation. Plant Cell. 2017;29:843–53.

83. Shaul O. How introns enhance gene expression. Int J Biochem. 2017;91:145–55.

84. Bieberstein NI, Carrillo Oesterreich F, Straube K, Neugebauer KM. First exon length controls active chromatin signatures and transcription. Cell Rep. 2012;2:62–8.

85. Rose AB. Introns as gene regulators: a brick on the accelerator. Front Genet. 2018;9:672.

86. Shayevitch R, Askayo D, Keydar I, Ast G. The importance of DNA methylation of exons on alternative splicing. RNA. 2018;24:1351–62.

87. Grishkevich V, Yanai I. Gene length and expression level shape genomic novelties. Genome Res. 2014;24:1497–503.

88. Brown SJ, Stoilov P, Xing Y. Chromatin and epigenetic regulation of pre-mRNA processing. Hum Mol Genet. 2012;21:R90–6.

89. Yearim A, Gelfman S, Shayevitch R, Melcer S, Glaich O, Mallm J-P, et al. HP1 is involved in regulating the global impact of DNA methylation on alternative splicing. Cell Rep. 2015;10:1122–34.

90. Lev Maor G, Yearim A, Ast G. The alternative role of DNA methylation in splicing regulation. Trends Genet. 2015;31:274–80.

91. VanderWeele TJ. Mediation analysis: a practitioner's guide. Annu Rev Public Health. 2016;37:17–32.

92. Wang Z, Curry E, Montana G. Network-guided regression for detecting associations between DNA methylation and gene expression. Bioinformatics. 2014;30:2693–701.

93. Ma X, Liu Z, Zhang Z, Huang X, Tang W. Multiple network algorithm for epigenetic modules via the integration of genome-wide DNA methylation and gene expression data. BMC Bioinform. 2017;18:72.

94. Hong X, Hao K, Ladd-Acosta C, Hansen KD, Tsai H-J, Liu X, et al. Genome-wide association study identifies peanut allergy-specific loci and evidence of epigenetic mediation in US children. Nat Commun. 2015;6:6304.

95. Fujii R, Sato S, Tsuboi Y, Cardenas A, Suzuki K. DNA methylation as a mediator of associations between the environment and chronic diseases: a scoping review on application of mediation analysis. Epigenetics. 2022;17:759–85.

96. Pacis A, Mailhot-Léonard F, Tailleux L, Randolph HE, Yotova V, Dumaine A, et al. Gene activation precedes DNA demethylation in response to infection in human dendritic cells. Proc Natl Acad Sci U S A. 2019;116:6938–43.

97. Secco D, Wang C, Shou H, Schultz MD, Chiarenza S, Nussaume L, et al. Stress induced gene expression drives transient DNA methylation changes at adjacent repetitive elements. eLife. 2015;4:e09343.

98. Neri F, et al. Intragenic DNA methylation prevents spurious transcription initiation. Nature. 2017;543:72–7.

99. Su Y, Shin J, Zhong C, Wang S, Roychowdhury P, Lim J, et al. Neuronal activity modifies the chromatin accessibility landscape in the adult brain. Nat Neurosci. 2017;20:476–83.

100. Bogan SN, Johnson KM, Hofmann GE. Changes in genome-wide methylation and gene expression in response to future pCO2 extremes in the Antarctic pteropod *Limacina helicina antarctica*. Front Mar Sci. 2020;6:788.

101. Hawes NA, Tremblay LA, Pochon X, Dunphy B, Fidler AE, Smith KF. Effects of temperature and salinity stress on DNA methylation in a highly invasive marine invertebrate, the colonial ascidian Didemnum vexillum. PeerJ. 2018;6: e5003.

102. Liew YJ, Howells EJ, Wang X, Michell CT, Burt JA, Idaghdour Y, et al. Intergenerational epigenetic inheritance in reef-building corals. Nat Clim Chang. 2020;10:254–9.

103. Wang X, Werren JH, Clark AG. Allele-specific transcriptome and methylome analysis reveals stable inheritance and cis-regulation of DNA methylation in *Nasonia*. PLoS Biol. 2016;14: e1002500.

104. Fangue NA, O'Donnell MJ, Sewell MA, Matson PG, MacPherson AC, Hofmann GE. A laboratory-based, experimental system for the study of ocean acidification effects on marine invertebrate larvae. Limnol Oceanogr Methods. 2010;8:441–52.

105. Trigg SA, Venkataraman YR, Gavery MR, Roberts SB, Bhattacharya D, Downey-Wall A, et al. Invertebrate methylomes provide insight into mechanisms of environmental tolerance and reveal methodological biases. Mol Ecol Resour. 2022;22:1247–61.

106. Yu G, Wang L-G, He Q-Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics. 2015;31:2382–3.

107. Shashikant T, Khor JM, Ettensohn CA. Global analysis of primary mesenchyme cell cis-regulatory modules by chromatin accessibility profiling. BMC Genom. 2018;19:206.

108. Arenas-Mena C, Miljovska S, Rice EJ, Gurges J, Shashikant T, Wang Z, et al. Identification and prediction of developmental enhancers in sea urchin embryos. BMC Genom. 2021;22:1–5.

109. Sultan M, Amstislavskiy V, Risch T, Schuette M, Dökel S, Ralser M, et al. Influence of RNA extraction methods and library selection schemes on RNA-seq data. BMC Genom. 2014;15:675.

110. Arenas-Mena C. PRO-seq, ATAC-seq, and Pol II ChIP-seq in 20 hour sea urchin embryos (ATAC-Seq) (purple sea urchin). 2020.

111. Shashikant T. Chromatin accessibility profiling identifies cis-regulatory modules in an early embryonic cell lineage. 2017.

112. Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, et al. Software for computing and annotating genomic ranges. PLoS Comput Biol. 2013;9: e1003118.

113. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12:357–60.

114. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014; 30:923–30.

115. Engström PG, Steijger T, Sipos B, Grant GR, Kahles A, Rätsch G, et al. Systematic evaluation of spliced alignment programs for RNA-seq data. Nat Methods. 2013;10:1185–91.

116. De Paoli-Iseppi R, Gleeson J, Clark MB. Isoform age - splice isoform profiling using long-read technologies. Front Mol Biosci. 2021;8: 711733.

117. Mehmood A, Laiho A, Venäläinen MS, McGlinchey AJ, Wang N, Elo LL. Systematic evaluation of differential splicing tools for RNA-seq studies. Brief Bioinformatics. 2020;21:2052–65.

118. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26:139–40.

119. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. 2012;40:4288–97.

120. Wright RM, Aglyamova GV, Meyer E, Matz MV. Gene expression associated with white syndromes in a reef building coral Acropora hyacinthus. BMC Genom. 2015;16:371.

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

122. Nielsen R, Paul JS, Albrechtsen A, Song YS. Genotype and SNP calling from next-generation sequencing data. Nat Rev Genet. 2011;12:443–51.

123. Lea AJ, Vilgalys TP, Durst PAP, Tung J. Maximizing ecological and evolutionary insight in bisulfite sequencing data sets. Nat Ecol Evol. 2017;1:1074–83.

124. Laine VN, Sepers B, Lindner M, Gawehns F, Ruuskanen S, van Oers K. An ecologist's guide for studying DNA methylation variation in wild vertebrates. Mol Ecol Resour. 2022;00:1–21.

125. Chen Y, Pal B, Visvader JE, Smyth GK. Differential methylation analysis of reduced representation bisulfite sequencing experiments using edgeR. F1000Res. 2017;6:2055.

126. Hinne M, Gronau QF, van den Bergh D, Wagenmakers E-J. A conceptual introduction to Bayesian model averaging. Adv Meth Pract Psychol Sci. 2020;3:200–15.

127. Bürkner P-C. brms: an R package for Bayesian multilevel models using Stan. J Stat Softw. 2017;80:1–28.

128. Makowski D, Ben-Shachar M, Lüdecke D. BayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. J Open Source Softw. 2019;4:1541.

Publisher's Note