Invertebrate methylomes provide insight into mechanisms of environmental tolerance and reveal methodological biases

Shelly A. Trigg*[1], Yaamini R. Venkataraman*[1], Mackenzie R. Gavery[2], Steven B. Roberts[1], Debashish Bhattacharya[3], Alan Downey-Wall[4], Jose M. Eirin-Lopez[5], Kevin M. Johnson[6,7], Katie E. Lotterhos[4], Jonathan B. Puritz[8] and Hollie M. Putnam[8+]

[1] University of Washington, School of Aquatic and Fishery Sciences 1122 NE Boat St. Seattle, WA, 98195, USA
[2] Environmental and Fisheries Sciences Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 2725 Montlake Blvd E, Seattle, WA, 98112, USA
[3] Department of Biochemistry and Microbiology, Rutgers University, New Brunswick, NJ 08901 USA
[4] Department of Marine and Environmental Sciences, Northeastern University, 430 Nahant Road, Nahant, MA 01908
[5] Florida International University, Environmental Epigenetics Laboratory, Institute of Environment 3000 NE 151 St. North Miami, FL, 33181, USA
[6] Center for Coastal Marine Sciences, California Polytechnic State University, San Luis Obispo, CA, 93407, USA
[7] California Sea Grant, University of California San Diego, La Jolla, CA, 92093
[8] Department of Biological Sciences, University of Rhode Island, Kingston, RI 02881, USA

[+] corresponding author: hputnam@uri.edu

* joint first authors/equal contribution

Keywords: bisulfite sequencing, coral, epigenetics, marine invertebrate

Abstract

There is a growing focus on the role of DNA methylation in the ability of marine invertebrates to rapidly respond to changing environmental factors and anthropogenic impacts. However, genome-wide DNA methylation studies in non-model organisms are currently hampered by limited understanding of methodological biases. Here we compare three methods for quantifying DNA methylation at single base pair resolution — Whole Genome Bisulfite Sequencing (WGBS), Reduced Representation Bisulfite Sequencing (RRBS), and Methyl-CpG Binding Domain Bisulfite Sequencing (MBDBS) — using multiple individuals from two reef-building coral species with contrasting environmental sensitivity. All methods reveal substantially greater methylation in *Montipora capitata* (11.4%) than the more sensitive *Pocillopora acuta* (2.9%). The majority of CpG methylation in both species occurs in gene bodies and flanking regions. In both species, MBDBS has the greatest capacity for detecting CpGs in coding regions at our sequencing depth, however MBDBS may be limited by intra-sample methylation heterogeneity. RRBS yields robust information for specific loci albeit without enrichment of any particular genome feature and with significantly reduced genome coverage. Relative genome size strongly influences the number and location of CpGs detected by each method when sequencing depth is limited, illuminating nuances in cross-species comparisons. These findings reinforce the role and importance of DNA methylation underlying environmental sensitivity in critical marine invertebrate taxa, and provide a genomic resource for investigating the functional role of DNA methylation in environmental tolerance.

55 **Introduction**

56 Environmental stimuli interact with genomic content to drive variation in gene and
57 protein expression, resulting in phenotypic plasticity. This plasticity has the potential to
58 buffer against mortality under environmental change (Baldwin, 1902), or conversely be
59 maladaptive (Velotta et al., 2018). Furthermore, plasticity may enhance or diminish
60 evolutionary rates (Ghalambor et al., 2007), which is particularly relevant to plasticity-
61 evolution feedbacks (Ghalambor et al., 2007, 2015; Kronholm & Collins, 2016). This is of
62 particular concern in the Anthropocene (Lewis & Maslin, 2015), as global change
63 exacerbates the mismatch between phenotype and a rapidly changing environment.

64 The increase in negative global climate change consequences have prompted an
65 intensification of research into phenotypic plasticity, gene regulation, and epigenetic
66 mechanisms in non-model marine invertebrates (Eirin-Lopez & Putnam, 2019; Hofmann,
67 2017; Roberts & Gavery, 2012). Specifically, carryover effects and cross and multi-
68 generational plasticity in response to climate change (Byrne et al., 2020) may be
69 generated by epigenetic regulation of gene expression (Dixon et al., 2018; Liew et al.,
70 2018, 2020). As epigenetic research has increased, there has been a focus on DNA
71 methylation, or the addition of a methyl group on the Cytosine residues in the genome,
72 often in the Cytosine phosphate Guanine, or CpG context (Zemach et al., 2010). DNA
73 methylation has gene expression regulation capacity through the interaction of base
74 modification with transcriptional elements. Early bulk enzyme-based and fingerprinting
75 methods for quantifying DNA methylation in marine invertebrates provided initial insights
76 into DNA methylation and organismal phenotypic plasticity in response to environmental

77 changes (Dimond et al., 2017; Gavery & Roberts, 2010; Gonzalez-Romero et al., 2017;
78 Putnam et al., 2016; Rodriguez-Casariego et al., 2018; Suarez-Ulloa et al., 2018).
79 Non-sequencing approaches that quantify global or bulk methylation [e.g.,
80 colorimetric or fluorescent ELISAs (Dimond et al., 2017; Gavery & Roberts, 2010; Putnam
81 et al., 2016; Rodriguez-Casariego et al., 2018)] are low-cost, rapidly applied, and do not
82 require genomic resources to generate information on the responsiveness of the
83 methylome. These global estimates do not, however, fully capture local changes in DNA
84 methylation across different genome regions. Specifically, differences in the location and
85 amount of methylation in two samples or treatments could lead to an incorrect conclusion
86 when based on average percent methylation at the bulk level. Consequently, non-
87 sequencing methods are limited in their ability to elucidate specific mechanisms of
88 expression regulation and thus are unable to fully address the functional implications of
89 methylation-driven regulation within the genome. In contrast, the use of genome-wide
90 approaches that provide single base pair resolution allow the testing of hypotheses
91 regarding spurious transcription, alternative splicing, and exon skipping (Roberts &
92 Gavery, 2012). For example, the use of Whole Genome Bisulfite Sequencing (WGBS) to
93 investigate the role of DNA methylation in regulating genes involved in caste specification
94 in honeybees identified differential methylation in an exon of the *anaplastic lymphoma*
95 *kinase (ALK)* gene; this exon was differentially retained in a splice variant between queens
96 and workers (Foret et al., 2012). Thus there is a clear need for single base pair
97 assessment of DNA methylomes facilitated by next generation sequencing to more fully
98 elucidate the relationship of DNA methylation and gene expression in non-model
99 invertebrates.

100 Genome-wide levels of DNA methylation can be estimated by several bisulfite
101 conversion and sequencing approaches. Bisulfite conversion of DNA results in the
102 deamination of unmethylated Cytosine to Uracil, which leaves a base change signature in
103 the DNA that can be tracked via comparison of sequence between bisulfite-converted
104 samples and reference genomes. While the number of bisulfite sequencing approaches
105 are expanding [e.g., epiGBS (van Gurp et al., 2016)], the widely-used approaches are
106 WGBS, Reduced Representation Bisulfite Sequencing (RRBS), and more recently,
107 Methyl-CpG Binding Domain Bisulfite Sequencing (MBDBS). WGBS is considered to be
108 the gold-standard of bisulfite sequencing because it provides full coverage of the genome
109 and the capacity to detect the entire methylome at single base pair resolution.

110 While providing a comprehensive approach, the high cost of WGBS is juxtaposed
111 against the often very small fraction of methylated DNA common in invertebrate genomes
112 (Tweedie et al., 1997). Alternatively, approaches such as RRBS also use bisulfite
113 conversion to quantitatively assess DNA methylation with base pair resolution. RRBS
114 incorporates a restriction digestion of the genome to enrich for CpG rich regions, and is
115 designed to capture the majority of promoters and other genomic regions containing CpG
116 islands because they have important regulatory functions in vertebrates (Meissner et al.,
117 2008). This is a more cost-effective approach provided by sequencing only a small portion
118 of the genome, but requires restriction enzyme recognition sites near other CpGs to gather
119 high resolution data.Since DNA methylation in invertebrates is primarily limited to coding
120 regions (Dixon et al., 2018; Flores et al., 2012; Roberts & Gavery, 2012), it is less clear
121 whether enrichment of CG-rich DNA using RRBS will capture informative or regulatory

122 regions of invertebrate genomes, making the cost savings moot in the absence of
123 informative data.

124 In contrast to the CpG-rich, region-specific targeting of RRBS, MBDBS uses Methyl
125 Binding Domain Proteins to target and enrich methylated CpGs, then employs bisulfite
126 conversion to provide single base pair resolution of DNA fragments with methylated
127 regions. Many marine invertebrate genomes consist of highly methylated regions that are
128 distributed in predominantly unmethylated DNA in a mosaic pattern (Suzuki et al., 2007).
129 When invertebrate methylomes have been characterized, these highly methylated regions
130 overlap with gene bodies and have been shown to play a role in gene expression activity
131 (Roberts & Gavery, 2012). Therefore, using an enrichment approach such as MBDBS to
132 isolate gene body methylation can be a cost-effective and gene-body focused alternative
133 to WGBS or RRBS (Gavery & Roberts, 2013; Venkataraman et al., 2020). The base-pair
134 resolution offered by the combination of MBD enrichment and BS is an advantage
135 compared to MBD-seq alone (Dixon & Matz, 2020), as the latter assumes that methylation
136 level is proportional to read depth. In contrast to WGBS or RRBS, the quantification and
137 interpretation of MBDBS data can be complicated by individual variation in methylation
138 levels (e.g., one individual who has high methylation in a particular region would have
139 reads, whereas another individual who lacks methylation in that region would have
140 missing data).

141 Given the need to assess plasticity mechanisms and the acclimatization potential
142 of a variety of marine taxa, it is critical to compare the potential of different approaches to
143 to detect, quantify, and assess DNA methylation with respect to specific biological
144 hypotheses of interest. To this end, we studied three DNA methylation quantification

145 approaches that provide single base pair resolution data using bisulfite conversion and
146 sequencing: WGBS, RRBS, and MBDBS. We applied these methods to two reef building
147 corals, *Montipora capitata* and *Pocillopora acuta*, which have different environmental
148 sensitivity, phenotypic plasticity, inducible DNA methylation (Putnam et al., 2016), and
149 genome sizes (Shumaker et al., 2019; Vidal-Dupiol et al., 2019). We assessed species-
150 specific differences in genome-wide methylation and contrasted percent methylation of
151 common loci and orthologous genes across methods. Then, we compared the coverage
152 and genomic location of CpG data generated from the three methods. Compared to
153 WGBS, both MBDBS and RRBS have advantages and potential limitations associated
154 with biology, genome characteristics, and experimental design, highlighting the need to
155 fully consider these aspects when evaluating DNA methylation for particular hypotheses
156 of function in invertebrates. As part of this effort, we described DNA methylation
157 differences in two coral species, providing valuable insights into the epigenetic
158 underpinnings of phenotypic plasticity.
159

160 **Materials and Methods**

161 *Sample collection*

162 The reef-building scleractinian coral species *Montipora capitata* and *Pocillopora*
163 *acuta* were collected from the reefs of Kaneʻohe Bay Hawaiʻi under SAP 2019-60 between
164 4 - 7 September 2018. Corals were transported to the Hawaiʻi Institute of Marine Biology
165 where they were held in tanks under ambient conditions for 15 days and then snap frozen
166 in liquid nitrogen and stored at -80°C until nucleic acid extraction was performed. For each

167 of the two coral species, fragments were collected from three different individuals collected
168 from ambient conditions.
169
170 *Nucleic Acid Extraction*
171 Samples were removed from -80°C and small tissue fragments were clipped
172 directly into a tube containing RNA/DNA shield (1 ml) and glass beads (0.5 mm). The
173 tissue clippings consisted of all coral cell types and their symbionts. Samples were
174 homogenized on a vortexer for 1 minute for the thin tissue imperforate coral *Pocillopora*
175 *acuta* and 2 minutes for the thick tissue perforate coral *Montipora capitata* at maximum
176 speed to ensure tissue extraction of all cell types. The supernatant was removed and DNA
177 was extracted using the Zymo Quick-DNA/RNA™ Miniprep Plus Kit and subsequently
178 checked for quality using gel electrophoresis on an Agilent 4200 TapeStation and
179 quantified using a Qubit. In summary, one DNA preparation was made from each of the
180 three individuals per coral species and was subsequently divided into three aliquots for
181 each of the three bisulfite sequencing methods (WGBS, MBDBS, and RRBS) to yield a
182 total of 18 libraries (**Figure 1**).
183
184 *Genome Information*
185 Previously sequenced and assembled coral genomes were used for mapping of
186 DNA methylation data. These include *Montipora capitata* (Shumaker et al., 2019) and
187 *Pocillopora acuta* (Vidal-Dupiol et al., 2019). Both of the coral genomes have a high and
188 similar number of predicted genes (63,227 in *M.capitata* and 64,558 in *P. acuta*). However,
189 *P. acuta* is much smaller in size (~352MB vs. ~886 MB in *M. capitata*), has less repetition,

190 a greater number of scaffolds (25,553 in *P. acuta* vs. 3,043 in *M.capitata*), and lower
191 genome assembly continuity (N50 is 171,375 in *P. acuta* and 540,623 in *M. capitata*).
192 Genome feature tracks for *M. capitata* and *P. acuta* were derived directly from the
193 published genomes for use in DNA methylation analyses. The *M. capitata* genome
194 annotation yielded gene (a combination of AUGUSTUS and GeMoMa predictions), coding
195 sequence, and intron tracks (Shumaker et al., 2019). Similarly, gene (AUGUSTUS
196 predictions), coding sequence, and intron information was obtained from the *P. acuta*
197 genome (Vidal-Dupiol et al., 2019). Flanking regions upstream and downstream of genes
198 were generated with BEDtools v2.29.2 for each genome separately (Quinlan & Hall,
199 2010); `flankBED` was used to generate 1000 bp flanks upstream and downstream of
200 annotated genes from each genome. Overlaps between genes and flanks were removed
201 from up- or down-stream flanking region tracks using `subtractBED`. Similarly, an
202 intergenic region track was created by finding the complement of genes with
203 `complementBED`, then removing any overlaps with flanking regions using `subtractBED`.
204 All tracks were verified with the Integrative Genomics Viewer (Thorvaldsdóttir et al., 2013).
205 Feature track files generated for both species are available in the project large file
206 repository (Putnam et al., 2020)
207
208 *MBD Enrichment*
209 Before enrichment, DNA (1 µg) in 80 µL Tris HCl was sheared to 500 bp using a
210 QSonica Q800R3. Samples were sonicated for 90 sec, with 15 sec on and 15 sec off
211 intervals at 25% amplitude. Fragment length was checked using a D5000 TapeStation

212 System (Agilent Technologies) and samples were sonicated an extra 15 sec to shear DNA
213 from 600 bp to 500 bp as needed.
214 The MethylMiner kit (Invitrogen; Cat. #ME10025) was used to enrich for methylated
215 DNA prior to MBDBS library generation, with 1µg of input DNA. Manufacturer's
216 instructions were adhered to with the following modifications: The capture reaction
217 containing the fragmented DNA and MBD beads was incubated with mixing at 4ºC
218 overnight, and enriched DNA was obtained with a single fraction elution using 2,000 mM
219 NaCl. Following ethanol addition, samples were centrifuged at 14,000 rcf at 1ºC for five
220 min. Pellets were resuspended in 25 µL ultra-pure water. Captured DNA was quantified
221 using a Qubit dsDNA HS Kit (Invitrogen).
222
223 *MBDBS and WGBS Library Preparation*
224 WGBS and MBDBS libraries were prepared using the Pico Methyl-Seq Library Prep
225 Kit (ZymoResearch Cat. # D5456). Manufacturer's instructions were followed with the
226 following modifications: For each sample, 1 ng of coral DNA and 0.05 ng of lambda phage
227 DNA (ZymoResearch Cat. # D5016) were used. Samples were always centrifuged at
228 12,000 rcf for 30 sec, however, samples were centrifuged for 12,000 rcf for 90 sec after
229 the second 200 µL addition of M Wash Buffer. Warmed elution buffer (56°C) was added
230 to each sample to increase DNA elution yield. During the second amplification cycle, 0.5
231 µL of PreAmp Polymerase was added. After initial clean-up with the DNA Clean-up and
232 Concentrator, the first amplification step was run for eight cycles. For amplification with i5
233 and i7 index primers, 1 µL of each primer was used to improve amplification. The volume

234 of the 2X LibraryAmp Master Mix was increased to 14 µL to match the increase in index
235 primers.

236 To remove excess primers from WGBS and MBDBS preparations, samples were
237 cleaned with 11 µL of KAPA pure beads (1X) (KAPA Cat # KK8000) and 80% ethanol.
238 Cleaned samples were resuspended in 12 µL of room-temperature DNA elution buffer
239 from the Pico Methyl-Seq Library Prep Kit. Samples were re-amplified with either two or
240 four cycles, depending on DNA concentration. Re-amplification was conducted with only
241 0.5 µL of each i5 and i7 index primer. After re-amplification, 26 µL of KAPA pure beads
242 (1X) and 80% ethanol were used for clean-up. Final samples were resuspended in 14 µL
243 of room-temperature elution buffer. Primer removal was confirmed by running samples on
244 a D5000 TapeStation System.

245

246 *RRBS Library Prep*

247 RRBS libraries were prepared with the EZ DNA RRBS Library Prep Kit
248 (ZymoResearch Cat. # D5460). Manufacturer's instructions were used with the following
249 modifications: For MspI digestion, 300 ng of input DNA and 15 ng of lambda phage DNA
250 were used. Digestions were carried out at 37°C for 4 hours. Adapter ligation was
251 performed overnight, with samples held at 4°C once cycling was completed. Similar to
252 WGBS and MBDBS library preparation, samples were always centrifuged at 12,000 rcf for
253 30 sec, with the exception of 90 sec centrifugation after the second 200 µL addition of M
254 Wash Buffer. Warmed elution buffer (56°C) was added to each sample to increase DNA
255 elution yield. Index primers were ligated using eleven cycles of the recommended
256 thermocycling protocol. Samples were cleaned using 50 µL of KAPA pure beads (1X) and

257 80% ethanol, then resuspended in 16 µL of the elution buffer. Primer removal was
258 confirmed by running samples on a D5000 TapeStation System.
259
260 *DNA Sequence Alignment*
261 All libraries (n = 18) were pooled in equimolar amounts and loaded at 250 pM onto
262 a single Illumina NovaSeq S4 flow cell lane for 2x150 bp sequencing at Genewiz (South
263 Plainfield, NJ). This was estimated to give 111-138 M reads per library and 99-123x
264 coverage of the *P. acuta* genome (3.3 M bp) and 38-47x coverage of the *M. capitata*
265 genome (8.8 M bp), assuming 100% even coverage (e.g., 150 bp read * 2 pairs * 111 M
266 reads/336,684,533 bp for *P. acuta*).
267 Sequence quality was checked by FastQC v0.11.8 and adapters from paired-end
268 sequences were trimmed using TrimGalore! version 0.4.5 (Krueger, 2012). Following
269 recommendations for methylation sequence analysis from the manufacturer's protocol
270 and from the Bismark User Guide, 10 bp were hard trimmed from the 5' and 3' end of each
271 read for WGBS and MBDBS samples, and RRBS samples were trimmed with `--`
272 `non_directional` and `--rrbs` options. Bisulfite-converted genomes were created in-
273 silico with Bowtie 2-2.3.4 [Linux x84_64 version; (Langmead & Salzberg, 2012)) using
274 `bismark_genome_preparation` through Bismark v0.21.0 (Krueger & Andrews, 2011).
275 Trimmed reads were aligned to the BS-converted *P. acuta* genome (Vidal-Dupiol et al.,
276 2019) and the BS-converted *M. capitata* genome (Shumaker et al., 2019) with Bismark
277 v0.21.0 with alignment stringency set by `-score_min L,0,-0.6` and the default MAPQ
278 score threshold of 20. To check mapping rates for endosymbionts and quantify percent

methylation, trimmed reads from *P. actua* libraries were also aligned to the *Cladicopium goreaui* genome [type C1, previously *Symbiodinium goreaui* (Liu et al., 2018)] using the same settings as specified above. Reads that mapped ambiguously were excluded and alignment files containing uniquely mapped reads were deduplicated with `deduplicate_bismark` for WGBS and MBDBS samples only. Methylation calls were extracted from sorted deduplicated alignment files using `bismark_methylation_extractor`. Cytosine coverage reports were generated using `coverage2cytosine` with the `--merge_CpG` option to combine data from both strand methylation. Resulting files include bedgraphs and Bismark coverage files (Putnam et al., 2020). MultiQC v1.8 (Ewels et al., 2016) was run on the trimmed reads, FastQC output, and Bismark reports to assess quality and summarize results.

Bisulfite conversion efficiency assessment

Trimmed sequence reads were aligned to the genome of *E. coli* strain K-12 MG1655 (Riley et al., 2006) using Bismark v0.21.0 with the `-non_directional` option and alignment stringency set by `-score_min L,0,-0.6`. Bisulfite conversion efficiency was also estimated from coral alignments as the ratio of the sum of unmethylated cytosines in CHG and CHH context to the sum of methylated and unmethylated cytosines in CHG and CHH. ANOVA was used to test for an effect of library preparation method on conversion efficiency within each species (conversion efficiency ~ library preparation method) for both estimated and lambda alignment calculated conversion efficiencies. A two-sample t-test was used to test if conversion

301 efficiency calculated from lambda alignments was the same as estimated conversion
302 efficiency for each library preparation method within each species.

303 *Genome-Wide Methylation*

304 General *M. capitata* and *P. acuta* methylation was characterized to describe any
305 species specific patterns. This was carried out by combining BEDgraphs derived from all
306 methods for each species using `unionBedGraphs`. Percent methylation for every CpG
307 locus with at least 5x coverage was averaged, irrespective of how many samples had
308 coverage for that locus. Loci with no data within a method were excluded from downstream
309 analysis. CpGs were classified as being either highly methylated (≥ 50% methylation),
310 moderately methylated (>10% and <50%), or lowly methylated (≤10% methylation).

311 *Percent Methylation of Shared CpG Loci*

312 Comparisons of percent DNA methylation at CpG loci analyzed by more than one
313 method were performed using the R-package `methylKit` (Akalin et al., 2012). A
314 minimum of 5x coverage was required across all samples for a CpG locus to be
315 considered in the analyses. The `unite` function in methylKit was used to identify CpG loci
316 that were covered across all 9 samples (3 individuals per method) per species.
317 Scatterplots and Pearson correlation coefficients were calculated using the function
318 `getCorrelation`. Additionally, differential methylation tests were performed on
319 pairwise comparisons between methods (WGBS versus RRBS, WGBS versus MBDBS,
320 and RRBS versus MBDBS). Discordant methylation was quantified using a logistic

321 regression model on CpG loci that were covered across all 6 samples (3 samples from
322 each method compared) in each pairwise comparison using the `calculateDiffMeth`
323 function with default parameters.
324
325 *CpG Coverage*
326 To assess average genome-wide CpG coverage, the number of Cytosines passing
327 different read depth thresholds (5x, 10x, 15x, 20x, 25x, 30x, 40x, and 50x) were totaled
328 from the CpG coverage reports output by the Bismark `coverage2cytosine` function
329 (detailed above) for each sample. These totaled CpGs were then relativized to the number
330 of CpGs in their respective genomes (*M. capitata,* 28,684,519 CpGs; *P. acuta,* 9,155,620
331 CpGs). Next, average and standard deviation of genome-wide CpG fractions were
332 calculated for each method within each species (n = 3), and these were plotted across
333 different read depth thresholds using ggplot2 (Gómez-Rubio, 2017).
334 To estimate overall genome-wide CpG coverage, a downsampling analysis was
335 performed by pooling all sample reads within a method and species. Briefly, trimmed fastq
336 files were concatenated for each method and species then randomly subsampled to 50,
337 100, 150, and 200 million reads. Next, alignment and methylation calling were carried out
338 as described above on each subset, and the number of cytosines passing with 5 or more
339 reads were totaled from CpG coverage reports from each subset. Sequencing saturation
340 was estimated from a Michaelis-Menten model with the 'drm' function from the R package
341 drc (Ritz et al., 2015) using CpG coverage reports from subsampled data as input. Both

342 observed CpG coverage from subsampled data and estimated CpG coverage were
343 plotted using the R package ggplot2 (Gómez-Rubio, 2017).
344
345 *Proportion of Detected CpGs for Orthologs*
346 To describe the differences in DNA methylation detected by each method at a more
347 functional level, and given the connection of gene body methylation and gene expression
348 in invertebrates (Roberts & Gavery, 2012) and corals specifically (Liew et al., 2018), the
349 presence of CpG data within all genes was calculated for each species, by method. First,
350 a CpG gff track was generated using EMBOSS (Rice et al., 2011) with the `fuzznuc`
351 command searching for the pattern CG. For each sample, `intersectBED` was used to
352 identify CpGs with 5x coverage that intersected with gene regions. This was also done for
353 the reference genome CpG gff track. CpG counts per gene were compiled for each sample
354 and the mean taken per method. The proportion of CpGs per orthologous gene was
355 calculated by dividing the mean number of CpGs with 5x coverage from the three samples
356 per method and dividing that by the number of CpG possible summed per gene from the
357 reference genome CpG gff track. The proportion of CpG data in a gene was then
358 visualized in heatmaps for all genes of *M. capitata* and *P. acuta*.

359 *Genomic Location of CpGs*

360 For both *M. capitata* and *P. acuta*, the overlap between genome feature tracks and
361 species-specific CpG data at 5x coverage was characterized with BEDtools v2.29.2 to
362 assess the presence CpGs in various regions by method (Quinlan & Hall, 2010). Since
363 only gene, coding sequence, intron, flanking regions, and intergenic region tracks were

364 common between species, these were the tracks used in downstream analyses. A
365 combination of PCoA, PERMANOVA and beta-dispersion tests, and chi-squared
366 contingency tests were used to determine if the library preparation method influenced the
367 proportion of CpGs detected in a specific genomic feature. A separate contingency test
368 was used for each genomic feature.
369

370 **Results**

371 To compare the performance of bisulfite sequencing methods in the reef-building
372 scleractinian corals *Montipora capitata* and *Pocillopora acuta*, we isolated DNA and
373 generated WGBS, RRBS, and MBDBS libraries for three individuals from each species to
374 yield a total of 18 libraries (**Figure 1**).

375 Sequencing of all 18 libraries resulted in 1.82 x 10^9 read pairs, of which 99.1%
376 remained after QC and trimming (Additional file 1: **Table ST1**). Individual libraries were
377 generally sequenced to the same depth (~7.5 x 10^7 reads) across library preparation
378 methods and species, with the exception of *P. acuta* RRBS libraries, which were
379 sequenced 2-4-fold deeper. The average mapping efficiencies for all *P. acuta* and all *M.*
380 *capitata* libraries were 45% and 39% read alignments, respectively (Additional file 2: **Table**
381 **ST2**). In comparison to other methods, MBDBS libraries had a larger proportion of reads
382 (73.1% ± 9.9%) that did not align to the coral genomes (Additional file 3: **Figure SF1**). To
383 investigate this we aligned *P. acuta* libraries to a known symbiont *Cladocopium goreaui*
384 genome (C1 (Liu et al., 2018)) for which the genome sequence was available at the time
385 of analysis. We found a sizable proportion of the MBDBS reads mapped to the symbiont
386 genome (23.6 ± 10.6%), while a much smaller proportion of RRBS and WGBS reads

387 mapped to the symbiont genome (5.04 ± 0.22% and 1.92 ± 0.3% respectively) (Additional
388 file 4: **Table ST3**).

389 Bisulfite conversion efficiency calculated from alignments of the unmethylated
390 lambda DNA spike-in ranged from 98.6 to 99.3% in *M. capitata* and from 98.3 to 99.1% in
391 *P. acuta* (Additional file 5: **Table ST4**), and this differed by library preparation method for
392 both *M. capitata* (F_{df}= X, P = 1.676e^{-05}) and *P. acuta* (F_{df}= X, P = 0.025) libraries. In
393 general, conversion efficiency calculated from the lambda alignments did not differ from
394 conversion efficiency estimates from CHG and CHH methylation (under the assumption
395 that non-CpG methylation does not occur in corals, see also (Liew et al., 2018)) from coral
396 alignments in *M. capitata* and *P. acuta*. (Additional file 6: **Figure SF2** and Additional file
397 7: **Table ST5**).

398 For each species, the general methylation landscape was characterized for CpG
399 loci with 5x coverage identified in any method. The *M. capitata* genome was more
400 methylated than *P. acuta* (**Figure 2**). Using a cutoff of ≥ 50% methylation to define
401 methylated CpGs, of the 13,340,268 CpGs covered by the *M. capitata* data, 11.4% were
402 methylated. In contrast, only 2.9% of the 7,326,297 CpGs in *P. acuta* were methylated.
403 Both genomes were predominantly lowly methylated (≤ 10% methylated): 79.6% CpGs in
404 *M. capitata* and 91.3% CpGs in *P. acuta* were lowly methylated. The remaining 9.0% of
405 CpGs in *M. capitata* and 5.8% of CpGs in *P. acuta* were moderately methylated (10-50%
406 methylation). The different methods captured varying proportions of highly, moderately,
407 and lowly methylated CpGs (Additional file 8: **Figure SF3**).

408 For quantitative comparison of method performance, we reduced the dataset to loci
409 covered at 5x read depth across all methods and samples for each species, referred to
410 here as 'shared loci'. The number of shared loci was 4,666 CpG for *M. capitata* and 93,714
411 CpG for *P. acuta*. A PCA of CpG methylation for loci covered at 5x read depth showed
412 that libraries tended to cluster in PC space by preparation method, rather than by
413 individual (Additional file 9: **Figure SF4**). Variation in methylation levels of the shared loci
414 across all *M. capitata* samples was lower within a method than between methods
415 (Additional file 9: **Figure SF4A**). In contrast, for *P. acuta*, RRBS and WGBS methods
416 showed similar methylation levels of shared loci but these were different from the
417 methylation level of loci identified in MBDBS (Additional file 9: **Figure SF4B**). To further
418 explore the variation in methylation observed by method, we directly correlated
419 quantitative methylation calls for the shared loci (**Figure 3**). For *M. capitata*, correlations
420 among biological replicates within a method were higher, whereas for *P. acuta*,
421 correlations were lower and more variable. Correlations between pairs of methods for *M.*
422 *capitata* ranged on average from 0.75-0.82, whereas correlations for *P. acuta* ranged from
423 0.40-0.64. For *M. capitata*, WGBS versus MBDBS had the highest correlation. For *P.*
424 *acuta*, WGBS versus RRBS had the highest correlation.

425 Discordance in methylation quantification between methods was evaluated by
426 identifying the number of CpG loci with large differences (>50%) in methylation for each
427 species. WGBS versus RRBS showed the lowest discordance in both species (0.4% for
428 *M. capitata* and 0.5% for *P. acuta*). The highest discordance in methylation was found in
429 comparisons with MBDBS for *P. acuta*, with 11% and 15% of CpG sites being called at
430 least 50% different for comparisons with WGBS and RRBS, respectively. In contrast, only

431 0.4% and 5% of common CpG sites were at least 50% difference between MBDBS versus
432 WGBS and MBDBS versus RRBS, respectively, for *M. capitata*. A majority of the
433 discordance was due to higher methylation calls in MBDBS compared to WGBS or RRBS
434 (**Figure 3B**).

435 Consistent with what would be expected based on genome size, *P. acuta* libraries
436 have higher genome-wide CpG coverage than *M. capitata* regardless of library
437 preparation method (**Figure 4A-C**). For both species, WGBS and MBDBS libraries
438 covered more CpGs than RRBS libraries, whereas RRBS libraries tended to show greater
439 read depth for the CpGs that it did cover. In other words, at >20x read depth, RRBS
440 libraries covered more CpGs than either WGBS or MBDBS (**Figure 4** insets). Modelling
441 increased sequencing depth for RRBS or MBDBS libraries showed little impact on the
442 fraction of genome-wide CpGs covered in *M. capitata*, while increasing sequencing depth
443 from 50 M to 200 M for WGBS libraries in both species and for MBDBS in *P. acuta* showed
444 a substantially larger fraction of CpGs covered (Additional file 10: **Figure SF5**).

445 In order to assess the potential for cross-species comparisons using an equivalent
446 dataset we quantified CpG data available across one-to-one orthologous genes. For *M.*
447 *capitata*, WGBS yielded the highest proportion of CpGs, followed by RRBS, and then
448 MBDBS (Additional file 11: **Figure SF6A**). This differed in *P. acuta* with WGBS yielding
449 the highest proportion of CpGs on average across orthologs, followed by MBDBS, and
450 then RRBS (Additional file 11: **Figure SF6B**).

451 In order to compare locations of CpG data between genomic features for each
452 species and method, all CpGs with 5x coverage were characterized based on genomic
453 feature location (**Figure 5)**. Global PERMANOVA tests found significant differences

454 between library preparation methods for CpG coverage in various genome features for *M.*
455 *capitata* and *P. acuta* (Additional file 12: Table ST6). Although post-hoc pairwise
456 PERMANOVA tests did not reveal differences between sequencing methods, power for
457 these was probably low (due to low sample size). Pairwise chi-squared tests indicated
458 there are differences in CpG location for both species. In particular, CpGs in gene bodies
459 were significantly enriched over other genomic features with MBDBS (**Figure 5**; Additional
460 file 13: Table ST7). Visual inspection of PCoA also revealed the proportion of CpGs
461 captured in coding sequences (CDS) drove differences between MBDBS and the other
462 methods in both species (**Figure 5C-D**).
463

464 **Discussion**

465 We evaluated the performance of three approaches that use bisulfite-treated library
466 preparation to enable single base pair resolution quantification of DNA methylation in
467 corals. Our results demonstrate that the methylation landscape can vary significantly
468 across species, which is a critical consideration for both interpreting environmental
469 response capacity, and therefore for experimental design. We report significant
470 differences in DNA methylation in two coral species that may contribute to their differential
471 environmental sensitivity of these organisms (Gibbin et al., 2015; Putnam et al., 2016).
472 Whereas WGBS is the gold standard for studying methylation, it comes at a high cost.
473 MBDBS enriches for gene regions, which may be useful for taxa with gene body
474 methylation. On the other hand, RRBS provides greater coverage depth for a smaller
475 fraction of the genome, but lacks specificity for genomic features, or DNA methylation.
476 Taken together, our findings indicate biology, genome architecture, regions of interest,

477 and depth of coverage are critical considerations when choosing methods for high
478 resolution quantification of DNA methylation profiles in invertebrates.

479 *M. capitata* has a relatively high environmental tolerance (Bahr et al., 2016; Gibbin
480 et al., 2015; Putnam et al., 2016), which has previously been attributed to its symbiont
481 composition (Cunning et al., 2016), genome characteristics (Shumaker et al., 2019),
482 perforate tissue-skeletal architecture and tissue thickness, and heterotrophic capacity
483 (Rodrigues & Grottoli, 2007). Of particular relevance to DNA methylation are genomic
484 aspects such as gene family duplication and high repeat content in *M. capitata* (Shumaker
485 et al., 2019). Given genetic-epigenetic correlations, particularly in the case of DNA
486 methylation and the requirement for a CpG sequence target site (Dimond & Roberts, 2020;
487 Johnson et al., 2020), variation in genome architecture, gene number, and content will
488 impact the presence and use of DNA methylation as a mechanism of gene expression
489 regulation. We found overall DNA methylation was higher in *M. capitata* than in *P. acuta*,
490 supporting early bulk analyses of DNA methylation in these species (Putnam et al., 2016).
491 While the predicted number of genes is similar, the genome size of *M. capitata* is over
492 twice that of *P. acuta* (Shumaker et al., 2019; Vidal-Dupiol et al., 2019). One explanation
493 for the higher methylation in *M. capitata* is that with greater energy availability — through
494 translocation from high density Symbiodiniaceae populations and energy stores in
495 perforate tissues — there is greater capacity for maintenance methyltransferase to
496 maintain high methylation, and thus reduce gene expression variability and spurious
497 expression (Liew et al., 2018; Li et al., 2018). High constitutive methylation could allow
498 “frontloading” of stress response genes (e.g., (Barshis et al., 2013)), providing greater
499 stress tolerance. Another possible explanation is that the higher level of methylation

500 contributes to the silencing of repeated genetic elements. In contrast, with a small and
501 non-repetitive genome, imperforate thin tissues, and low energy reserves, *P. acut*a may
502 be more energetically limited. Thus, *P. acuta* may be expected to show lower DNA
503 methylation across the genome as we demonstrate here, as well as a higher propensity
504 for inducible methylation in the presence of stressors (Putnam et al., 2016).

505 Another striking contrast in DNA methylation in these species is the lack of
506 concordance in the percent methylation values for *P. acuta* among methods compared to
507 *M. capitata* (**Figure 3**). The potential for chimerism in corals (Oury et al., 2020;
508 Schweinsberg et al., 2015) and differences in tissue structure (e.g., perforate or
509 imperforate) between species could contribute to differences in concordance across
510 methods for quantifying DNA methylation. One possibility is that Pocilloporids are chimeric
511 and multiple genotypes exist (Oury et al., 2020; Schweinsberg et al., 2015). Although
512 percent DNA methylation concordance across methods was generally high, in *P. acuta*
513 there was approximately a 10% higher level of discordance in percent methylation
514 quantification when compared WGBS to RRBS or MBDBS (**Figure 3**). This discordance
515 could have resulted from differences in *P. acuta* and *M. capitata* tissue structure. There is
516 the potential to homogenize and extract DNA from all cell types from the thin, imperforate
517 tissues of *P. acuta*, as opposed to the thick, perforate tissues in *M. capitata* (Putnam et
518 al., 2017), likely contributing to a greater number of cell types, and thus methylation
519 differences, captured in our *P. acuta* samples. Furthermore, the microhabitats created in
520 the tissues of these two species likely differ substantially spatially (Putnam et al., 2017),
521 creating cell-to-cell variability in methylation content. Since the likelihood of capturing
522 multiple cell types in bulk DNA extractions varies with tissue structure, future studies

523 should consider methods such as fluorescence-activated cell sorting (Hu et al., 2020;
524 Rosental et al., 2017), or laser microdissection (e.g., (Liew et al., 2018)), to target specific
525 tissues or cell types and reduce cell-to-cell methylation variability. Whereas this does not
526 necessarily indicate a bias in our methods, it highlights the needs to account for the
527 biological characteristics of a species when designing an experiment and evaluating
528 results.

529 The gold standard for bisulfite sequencing, WGBS, can be cost prohibitive
530 particularly if comparing multiple species and treatments. As expected, we found that
531 WGBS performed well, particularly for *P. acuta,* which has a smaller genome. Focusing
532 on gene orthologs, WGBS performed the best in terms of data for CpGs per gene. Based
533 on the gene ortholog comparisons, MBDBS provided more information than RRBS for *P.*
534 *acuta*, however the opposite held true for *M. capitata.* This is likely attributable to the
535 different genome size and inherent differences in methylation that result in differential
536 enrichment.

537 For both species, WGBS and MBDBS libraries covered more CpGs than RRBS
538 libraries; however, RRBS libraries showed greater read depth for CpGs. This is because
539 RRBS subsampled a specific, smaller portion of the genome than MBDBS or WGBS,
540 allowing more read coverage. Hence, CpG coverage did not largely increase when deeper
541 sequencing was modeled using RRBS data (Additional file 10: **Figure SF5**). RRBS was
542 designed to enrich for CpG islands, short stretches of DNA with higher levels of CpGs (~1
543 CpG per 10bp), that are typically found in mammalian promoters and enhancer regions
544 and thought to play a role in gene regulation (Gu et al., 2011). We found RRBS yielded a
545 well-covered reduced representation of the genome, which is important for bisulfite data

546 where high read depth is desired, and locus methylation levels were concordant with
547 WGBS for both species. However, RRBS did not enrich for promoters or other particular
548 genomic regions compared to the other bisulfite sequencing methods (**Figure 5**), and in
549 fact tended to identify unmethylated regions. For this reason, RRBS is not the best choice
550 for gene-focused methylation studies in corals and other invertebrates.

551 A critical consideration in deciding to perform MBDBS in corals is the amount of
552 DNA methylation present in any symbiont. If any non-target organisms have substantially
553 more DNA methylation than the target organism, MBDBS data could become saturated
554 by methylated DNA from non-target organisms, lowering sampling of the target species.
555 We observed this in *P. acuta*, for which we had the genome of its Symbiodiniaceae which
556 has ~90% genome-wide methylation (de Mendoza et al., 2018; Lohuis & Miller, 1998).
557 When compared to RRBS and WGBS data, we found a 4 to 10-fold enrichment of
558 Symbiodiniaceae DNA in *P. acuta* MBDBS data. Separation of host and symbionts is
559 therefore recommended to obtain the greatest read counts for the organism of interest,
560 but this comes at the cost of not being able to obtain RNA from the same nucleic acid
561 pool. For example, physical separation of the host and symbiont in living cells impacts
562 gene expression, and attempts at physical separation after freezing can degrade the host
563 RNA. Simultaneous extraction of holobiont RNA and DNA from the same nucleic acid pool
564 provides the optimal approach for detecting interactions between DNA methylation and
565 epigenetic regulation of gene expression. This comes at the cost of generating excess
566 reads to overcome highly methylated Symbiodiniaceae DNA.

567 MBDBS can enrich for gene regions in species where methylation is primarily found
568 in gene bodies such as in corals (reviewed in (Eirin-Lopez & Putnam, 2019)), and can thus

569 provide insight into mechanisms underlying physiological or organismal responses. We
570 found that MBDBS significantly enriched for gene bodies, specifically CDS and introns,
571 when compared to RRBS and WGBS in both *M. capitata* and *P. acuta* (**Figure 5**). While
572 MBDBS may be a good choice to examine gene body methylation at a reduced cost,
573 species differences in CpG coverage within orthologous genes with MBDBS (Additional
574 file 10: **Figure SF6**) may complicate cross-species comparisons by reducing the amount
575 of data available for analysis. Additionally, we found high discordance between MBDBS
576 and non-enrichment methods, WGBS and RRBS, for *P. acuta*. MBDBS is the only method
577 we evaluated that can non-randomly sub-sample genomes present in a DNA sample
578 through preferential pull-down of methylated DNA. Differences in methylation across the
579 sampled genomes could result from cell-to-cell heterogeneity in methylation or cell-type
580 (e.g., methylation of calcifying cells may differ from symbiont hosting cells). In other words,
581 MBDBS data may represent only a subpopulation of highly methylated cells, while WGBS
582 and RRBS represent the average methylation across all cells in the sample. Using a
583 consistent tissue type is important to limit potential methylation heterogeneity, and caution
584 should be taken when comparing MBDBS data directly to that of non-enrichment bisulfite
585 sequencing approaches.

586 Although MBDBS did enrich for methylated regions of the genome, 80% of CpGs
587 in *M. capitata* and 82% of CpGs in *P. acuta* interrogated with MBDBS were lowly
588 methylated (<10% methylated) (Additional file 8: **Figure SF3**). This is expected and is
589 consistent with previous reports applying MBDBS in other marine invertebrates where
590 unmethylated CpGs actually represent the highest proportion of loci in the data,
591 attributable to the nature of the methylation landscape and enrichment protocol (e.g.

592 (Gavery & Roberts, 2013; Venkataraman et al., 2020)). The base pair resolution of
593 methylation revealed by MBDBS is a benefit over MBD-Seq alone because it enables a
594 fine-scale examination of specific genomic features (e.g. exon-intron boundaries) that may
595 not be possible with the regional resolution of MBD-Seq. Without complete knowledge of
596 the relative importance of a single loci compared to a region, it is difficult to compare trade-
597 offs between MBDBS and MBD-seq. However, bisulfite sequencing requires significant
598 coverage to quantify DNA methylation.

599 MBDBS may have potential biases that should be considered when interpreting
600 results. If a treatment, population comparison, or developmental change results in a given
601 region (~500bp) going from being highly methylated to fully unmethylated, then it is likely
602 that this region would not be interrogated by MBDBS, due to an absence of data in the
603 unmethylated condition. This is a potential source of bias in MBDBS data and may
604 contribute to important differentially methylated regions being overlooked: for example if
605 one treatment results in high methylation and is captured by MBDBS and another
606 treatment results in no methylation and is not captured by MBDBS, this region would be
607 filtered out of the analysis because of missing data in some individuals. Further, the
608 potential of MBDBS to provide limited information for unmethylated genes may introduce
609 bias in studies that seek to draw relationships between methylation level and gene
610 expression. Just as with many interpretations of key findings we present, a more complete
611 understanding of the mechanistic functional role DNA methylation plays in genome
612 regulation in the species of interest is needed.

613 There is a greater capacity to gain mechanistic insight when using methods that
614 have single base pair resolution compared to methylation enrichment without bisulfite

615 treatment or bulk percent methylation approaches. For example, hypotheses such as the
616 linkage between DNA methylation and alternative splicing (Roberts & Gavery, 2012) are
617 more accurately tested with bisulfite sequencing approaches. We acknowledge the cost
618 of generating genomic resources and bisulfite sequencing data can be higher than other
619 approaches. While WGBS is supported here as the gold standard for DNA methylation
620 quantification, consideration should be given to specific study hypotheses in light of the
621 pros and cons of the enrichment or reduced representation approaches presented here
622 and in other comparative works (Dixon & Matz, 2020). Our results suggest that it would
623 be unwise to use multiple different library preparation methods for comparing individuals
624 within a study, especially for studies in which familial relationships are to be compared. As
625 technology advances, it would be ideal to move away from harsh bisulfite conversion to
626 assess DNA methylation with single base pair resolution across whole genomes in the
627 absence of DNA treatment (e.g., Oxford Nanopore).

628 Our results provide a quantitative comparative assessment that can be used to
629 inform the choice of sequencing DNA methylation in corals and other non-model
630 invertebrates. Together these metrics enable comparative capacity for three common
631 methods in two coral taxa that vary in their phylogeny, genome size, symbiotic unions,
632 and environmental performance, and thus provide the community with a more
633 comprehensive foundation upon which to build laboratory and statistical analyses of DNA
634 methylation, plasticity, and acclimatization.

635

636 **Abbreviations**

637 CpG: cytosine and guanine separated by a phosphate group

638 ELISA: Enzyme-Linked Immunosorbent Assay

639 MBD-seq: Methyl-CpG Binding Domain Sequencing

640 MBDBS: Methyl-CpG Binding Domain Bisulfite Sequencing

641 RRBS: Reduced representation bisulfite sequencing

642 WGBS: Whole genome bisulfite sequencing

643

644 **Acknowledgements**

645 We thank Putnam Lab members E. Strand and M. E. Schedl for their technical support.
646 Data analyses were facilitated through the use of advanced computational, storage, and
647 networking infrastructure provided by the Hyak supercomputer system at the University of
648 Washington.

649

650 **Funding**

651 This work was partially supported by the USDA: National Institute of Food and Agriculture,
652 Hatch projects RI0019-H020 to HMP, NJ01180 to DB, and the NRSP-8 National Animal
653 Genome Research Program to HMP and SBR. This work was also partially supported by
654 funding from the National Science Foundation: 1756623 to HMP, 1756616 to DB, 1921149
655 to SBR, 1921402 to JEL, and 1635423 to KEL.

656

657 **References**

658 Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F. E., Figueroa, M. E., Melnick, A.,
659 & Mason, C. E. (2012). methylKit: a comprehensive R package for the analysis of
660 genome-wide DNA methylation profiles. *Genome Biology*, *13*(10), R87.
661 https://doi.org/10.1186/gb-2012-13-10-r87
662 Bahr, K. D., Jokiel, P. L., & Rodgers, K. S. (2016). Relative sensitivity of five Hawaiian

coral species to high temperature under high-pCO2 conditions. *Coral Reefs* , *35*(2), 729–738. https://doi.org/10.1007/s00338-016-1405-4

Baldwin, J. M. (1902). Development and evolution, including psychophysical evolution, evolution by orthoplasy, and the theory of genetic modes. https://doi.org/10.1037/13707-000

Barshis, D. J., Ladner, J. T., Oliver, T. A., Seneca, F. O., Traylor-Knowles, N., & Palumbi, S. R. (2013). Genomic basis for coral resilience to climate change. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(4), 1387–1392. https://doi.org/10.1073/pnas.1210224110

Byrne, M., Foo, S. A., Ross, P. M., & Putnam, H. M. (2020). Limitations of cross-and multigenerational plasticity for marine invertebrates faced with global climate change. *Global Change Biology*, *26*(1), 80–102. https://doi.org/10.1111/gcb.14882

Cunning, R., Ritson-Williams, R., & Gates, R. D. (2016). Patterns of bleaching and recovery of *Montipora capitata* in Kāne ‘ohe Bay, Hawai ‘i, USA. *Marine Ecology Progress Series*, *551*, 131–139. https://www.int-res.com/abstracts/meps/v551/p131-139/

de Mendoza, A., Bonnet, A., Vargas-Landin, D. B., Ji, N., Li, H., Yang, F., Li, L., Hori, K., Pflueger, J., Buckberry, S., Ohta, H., Rosic, N., Lesage, P., Lin, S., & Lister, R. (2018). Recurrent acquisition of cytosine methyltransferases into eukaryotic retrotransposons. *Nature Communications*, *9*(1), 1–11. https://doi.org/10.1038/s41467-018-03724-9

Dimond, J. L., Gamblewood, S. K., & Roberts, S. B. (2017). Genetic and epigenetic insight into morphospecies in a reef coral. *Molecular Ecology*, *26*(19), 5031–5042. https://doi.org/10.1111/mec.14252

Dimond, J. L., & Roberts, S. B. (2020). Convergence of DNA Methylation Profiles of the Reef Coral *Porites astreoides* in a Novel Environment. *Frontiers in Marine Science*, *6*, 792. https://doi.org/10.3389/fmars.2019.00792

Dixon, G., Liao, Y., Bay, L. K., & Matz, M. V. (2018). Role of gene body methylation in acclimatization and adaptation in a basal metazoan. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(52), 13342–13346. https://doi.org/10.1073/pnas.1813749115

Dixon, G., & Matz, M. (2020). Benchmarking DNA methylation assays in a reef-building coral. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.13282

Eirin-Lopez, J. M., & Putnam, H. M. (2019). Marine Environmental Epigenetics. *Annual Review of Marine Science*, *11*, 335–368. https://doi.org/10.1146/annurev-marine-010318-095114

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* , *32*(19), 3047–3048. https://doi.org/10.1093/bioinformatics/btw354

Flores, K., Wolschin, F., Corneveaux, J. J., Allen, A. N., Huentelman, M. J., & Amdam, G. V. (2012). Genome-wide association between DNA methylation and alternative splicing in an invertebrate. *BMC Genomics*, *13*, 480. https://doi.org/10.1186/1471-2164-13-480

Foret, S., Kucharski, R., Pellegrini, M., Feng, S., Jacobsen, S. E., Robinson, G. E., & Maleszka, R. (2012). DNA methylation dynamics, metabolic fluxes, gene splicing,

and alternative phenotypes in honey bees. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(13), 4968–4973. https://doi.org/10.1073/pnas.1202392109

Gavery, M. R., & Roberts, S. B. (2010). DNA methylation patterns provide insight into epigenetic regulation in the Pacific oyster (*Crassostrea gigas*). *BMC Genomics*, *11*(1), 483. https://doi.org/10.1186/1471-2164-11-483

Gavery, M. R., & Roberts, S. B. (2013). Predominant intragenic methylation is associated with gene expression characteristics in a bivalve mollusc. *PeerJ*, *1*, e215. https://doi.org/10.7717/peerj.215

Ghalambor, C. K., Hoke, K. L., Ruell, E. W., Fischer, E. K., Reznick, D. N., & Hughes, K. A. (2015). Non-adaptive plasticity potentiates rapid adaptive evolution of gene expression in nature. *Nature*, *525*(7569), 372–375. https://doi.org/10.1038/nature15256

Ghalambor, C. K., McKay, J. K., Carroll, S. P., & Reznick, D. N. (2007). Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology*, *21*(3), 394–407. https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2435.2007.01283.x

Gibbin, E. M., Putnam, H. M., Gates, R. D., Nitschke, M. R., & Davy, S. K. (2015). Species-specific differences in thermal tolerance may define susceptibility to intracellular acidosis in reef corals. *Marine Biology*, *162*(3), 717–723. https://doi.org/10.1007/s00227-015-2617-9

Gómez-Rubio, V. (2017). ggplot2 - Elegant Graphics for Data Analysis (2nd Edition). In *Journal of Statistical Software* (Vol. 77, Issue Book Review 2). https://doi.org/10.18637/jss.v077.b02

Gonzalez-Romero, R., Suarez-Ulloa, V., Rodriguez-Casariego, J., Garcia-Souto, D., Diaz, G., Smith, A., Pasantes, J. J., Rand, G., & Eirin-Lopez, J. M. (2017). Effects of Florida Red Tides on histone variant expression and DNA methylation in the Eastern oyster *Crassostrea virginica*. *Aquatic Toxicology* , *186*, 196–204. https://doi.org/10.1016/j.aquatox.2017.03.006

Gu, H., Smith, Z. D., Bock, C., Boyle, P., Gnirke, A., & Meissner, A. (2011). Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nature Protocols*, *6*(4), 468–481. https://doi.org/10.1038/nprot.2010.190

Hofmann, G. E. (2017). Ecological Epigenetics in Marine Metazoans. *Frontiers in Marine Science*, *4*, 4. https://doi.org/10.3389/fmars.2017.00004

Hu, M., Zheng, X., Fan, C.-M., & Zheng, Y. (2020). Lineage dynamics of the endosymbiotic cell type in the soft coral *Xenia*. In *Nature* (Vol. 582, Issue 7813, pp. 534–538). https://doi.org/10.1038/s41586-020-2385-7

Johnson, K. M., Sirovy, K. A., Casas, S. M., La Peyre, J. F., & Kelly, M. W. (2020). Characterizing the Epigenetic and Transcriptomic Responses to *Perkinsus marinus* Infection in the Eastern Oyster *Crassostrea virginica*. *Frontiers in Marine Science*, *7*. https://doi.org/10.3389/fmars.2020.00598

Kronholm, I., & Collins, S. (2016). Epigenetic mutations can both help and hinder adaptive evolution. *Molecular Ecology*, *25*(8), 1856–1868.

https://doi.org/10.1111/mec.13296

Krueger, F. (2012). Trim Galore: a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files, with some extra functionality for MspI-digested RRBS-type (Reduced Representation Bisufite-Seq) libraries. *URL Http://www. Bioinformatics. Babraham. Ac. Uk/projects/trim_galore/. (Date of Access: 28/04/2016)*.

Krueger, F., & Andrews, S. R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* , *27*(11), 1571–1572. https://doi.org/10.1093/bioinformatics/btr167

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Lewis, S. L., & Maslin, M. A. (2015). Defining the anthropocene. *Nature*, *519*(7542), 171–180. https://doi.org/10.1038/nature14258

Liew, Y. J., Howells, E. J., Wang, X., Michell, C. T., Burt, J. A., Idaghdour, Y., & Aranda, M. (2020). Intergenerational epigenetic inheritance in reef-building corals. *Nature Climate Change*, *10*(3), 254–259. https://doi.org/10.1038/s41558-019-0687-2

Liew, Y. J., Zoccola, D., Li, Y., Tambutté, E., Venn, A. A., Michell, C. T., Cui, G., Deutekom, E. S., Kaandorp, J. A., Voolstra, C. R., Forêt, S., Allemand, D., Tambutté, S., & Aranda, M. (2018). Epigenome-associated phenotypic acclimatization to ocean acidification in a reef-building coral. *Science Advances*, *4*(6), 188227. https://doi.org/10.1126/sciadv.aar8028

Liu, H., Stephens, T. G., González-Pech, R. A., Beltran, V. H., Lapeyre, B., Bongaerts, P., Cooke, I., Aranda, M., Bourne, D. G., Forêt, S., Miller, D. J., van Oppen, M. J. H., Voolstra, C. R., Ragan, M. A., & Chan, C. X. (2018). Symbiodinium genomes reveal adaptive evolution of functions related to coral-dinoflagellate symbiosis. *Communications Biology*, *1*, 95. https://doi.org/10.1038/s42003-018-0098-3

Li, Y., Liew, Y. J., Cui, G., Cziesielski, M. J., Zahran, N., Michell, C. T., Voolstra, C. R., & Aranda Lastra, M. (2018). DNA methylation regulates transcriptional homeostasis of algal endosymbiosis in the coral model *Aiptasia. Sci Adv*, *4*.

Lohuis, M. R., & Miller, D. J. (1998). Hypermethylation at CpG-Motifs in the dinoflagellates *Amphidinium carterae* (*Dinophyceae*): Evidence from restriction analyzes, 5-azacytidine and ethionine treatment. In *Journal of Phycology* (Vol. 34, Issue 1, pp. 152–159). https://doi.org/10.1046/j.1529-8817.1998.340152.x

Meissner, A., Mikkelsen, T. S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B. E., Nusbaum, C., Jaffe, D. B., Gnirke, A., Jaenisch, R., & Lander, E. S. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*, *454*(7205), 766–770. https://doi.org/10.1038/nature07107

Oury, N., Gélin, P., & Magalon, H. (2020). Together stronger: Intracolonial genetic variability occurrence in *Pocillopora* corals suggests potential benefits. *Ecology and Evolution*, *10*(12), 5208–5218. https://doi.org/10.1002/ece3.5807

Putnam, H. M., Barott, K. L., Ainsworth, T. D., & Gates, R. D. (2017). The Vulnerability and Resilience of Reef-Building Corals. *Current Biology: CB*, *27*(11), R528–R540. https://doi.org/10.1016/j.cub.2017.04.047

Putnam, H. M., Davidson, J. M., & Gates, R. D. (2016). Ocean acidification influences host DNA methylation and phenotypic plasticity in environmentally susceptible

corals. *Evolutionary Applications*, *9*(9), 1165–1178. https://doi.org/10.1111/eva.12408

Putnam, H. M., Roberts, S. B., & Venkataraman, Y. (2020). *Coral Methylation Methods Comparison*. Open Science Framework. https://doi.org/10.17605/OSF.IO/X5WAZ

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* , *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Rice, P. M., Bleasby, A. J., & Ison, J. C. (2011). *EMBOSS User's Guide: Practical Bioinformatics with EMBOSS*. Cambridge University Press. https://play.google.com/store/books/details?id=zC5WGJxDNtcC

Riley, M., Abe, T., Arnaud, M. B., Berlyn, M. K. B., Blattner, F. R., Chaudhuri, R. R., Glasner, J. D., Horiuchi, T., Keseler, I. M., Kosuge, T., Mori, H., Perna, N. T., Plunkett, G., 3rd, Rudd, K. E., Serres, M. H., Thomas, G. H., Thomson, N. R., Wishart, D., & Wanner, B. L. (2006). *Escherichia coli* K-12: a cooperatively developed annotation snapshot--2005. *Nucleic Acids Research*, *34*(1), 1–9. https://doi.org/10.1093/nar/gkj405

Ritz, C., Baty, F., Streibig, J. C., & Gerhard, D. (2015). Dose-Response Analysis Using R. *PloS One*, *10*(12), e0146021. https://doi.org/10.1371/journal.pone.0146021

Roberts, S. B., & Gavery, M. R. (2012). Is There a Relationship between DNA Methylation and Phenotypic Plasticity in Invertebrates? *Frontiers in Physiology*, *2*, 116. https://doi.org/10.3389/fphys.2011.00116

Rodrigues, L. J., & Grottoli, A. G. (2007). Energy reserves and metabolism as indicators of coral recovery from bleaching. *Limnology and Oceanography*, *52*(5), 1874–1882. https://doi.org/10.4319/lo.2007.52.5.1874

Rodriguez-Casariego, J. A., Ladd, M. C., Shantz, A. A., Lopes, C., Cheema, M. S., Kim, B., Roberts, S. B., Fourqurean, J. W., Ausio, J., Burkepile, D. E., & Eirin-Lopez, J. M. (2018). Coral epigenetic responses to nutrient stress: Histone H2A.X phosphorylation dynamics and DNA methylation in the staghorn coral *Acropora cervicornis*. *Ecology and Evolution*, *8*(23), 12193–12207. https://doi.org/10.1002/ece3.4678

Rosental, B., Kozhekbaeva, Z., Fernhoff, N., Tsai, J. M., & Traylor-Knowles, N. (2017). Coral cell separation and isolation by fluorescence-activated cell sorting (FACS). In *BMC Cell Biology* (Vol. 18, Issue 1). https://doi.org/10.1186/s12860-017-0146-8

Schweinsberg, M., Weiss, L. C., Striewski, S., Tollrian, R., & Lampert, K. P. (2015). More than one genotype: how common is intracolonial genetic variability in scleractinian corals? *Molecular Ecology*, *24*(11), 2673–2685. https://doi.org/10.1111/mec.13200

Shumaker, A., Putnam, H. M., Qiu, H., Price, D. C., Ehud, Z., Harel, A., Wagner, N. E., Gates, R. D., Yoon, H. S., & Bhattacharya, D. (2019). Genome analysis of the rice coral *Montipora capitata*. *Scientific Reports*, *9*(2571). https://doi.org/10.1038/s41598-019-39274-3

Suarez-Ulloa, V., Rivera-Casas, C., Michel, M., & Eirin-Lopez, J. M. (2018). Seasonal variation and epigenetic changes in the flat tree oyster *Isognomon alatus* during a 2-year study of a mangrove ecosystem in North Biscayne Bay, Florida. *Journal of Shellfish Research*, 38(1):79-88 (2019). https://doi.org/10.2983/035.038.0108

Suzuki, M. M., Kerr, A. R. W., De Sousa, D., & Bird, A. (2007). CpG methylation is

843 targeted to transcription units in an invertebrate genome. *Genome Research*, *17*(5),
844 625–631. https://doi.org/10.1101/gr.6163007
845 Thorvaldsdóttir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics
846 Viewer (IGV): high-performance genomics data visualization and exploration.
847 *Briefings in Bioinformatics*, *14*(2), 178–192. https://doi.org/10.1093/bib/bbs017
848 Tweedie, S., Charlton, J., Clark, V., & Bird, A. (1997). Methylation of genomes and
849 genes at the invertebrate-vertebrate boundary. *Molecular and Cellular Biology*,
850 *17*(3), 1469–1475. https://www.ncbi.nlm.nih.gov/pubmed/9032274
851 van Gurp, T. P., Wagemaker, N. C. A. M., Wouters, B., Vergeer, P., Ouborg, J. N. J., &
852 Verhoeven, K. J. F. (2016). epiGBS: reference-free reduced representation bisulfite
853 sequencing. *Nature Methods*, *13*(4), 322–324. https://doi.org/10.1038/nmeth.3763
854 Velotta, J. P., Ivy, C. M., Wolf, C. J., Scott, G. R., & Cheviron, Z. A. (2018). Maladaptive
855 phenotypic plasticity in cardiac muscle growth is suppressed in high-altitude deer
856 mice. *Evolution; International Journal of Organic Evolution*, *72*(12), 2712–2727.
857 https://doi.org/10.1111/evo.13626
858 Venkataraman, Y. R., Downey-Wall, A. M., Ries, J., Westfield, I., White, S. J., Roberts,
859 S. B., & Lotterhos, K. E. (2020). General DNA methylation patterns and
860 environmentally-induced differential methylation in the eastern oyster (*Crassostrea*
861 *virginica*). *Frontiers in Marine Science*. https://doi.org/10.3389/fmars.2020.00225
862 Vidal-Dupiol, J., Chaparro, C., Pratlong, M., Pontaroti, P., Grunau, C., & Mitta, G. (2019).
863 Sequencing, de novo assembly and annotation of the genome of the scleractinian
864 coral, *Pocillopora acuta*. In *bioRxiv* (p. 698688). https://doi.org/10.1101/698688
865 Zemach, A., McDaniel, I. E., Silva, P., & Zilberman, D. (2010). Genome-wide
866 evolutionary analysis of eukaryotic DNA methylation. *Science*, *328*(5980), 916–919.
867 https://doi.org/10.1126/science.1186366

Data accessibility

The datasets supporting the conclusions of this article are available in the Coral Methylation Methods Comparison repository, http://doi.org/10.17605/OSF.IO/X5WAZ, and included included within the article (and its additional file(s)). All raw data can be accessed under NCBI Bioproject PRJNA691891.

Author contributions

HMP designed the study and collected the samples. YRV, SAT, MRG, HMP, and SBR performed data analyses and drafted the manuscript. DB, AD-W, JME-L, KMJ, KEL, and JBP provided analytical insight and manuscript revisions. All authors read and approved the final manuscript.

Ethics declarations

Ethics approval and consent to participate

The coral samples were collected under Hawai‘i Department of Aquatic Resources Special Activity Permit SAP 2019-60.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Figures and Tables

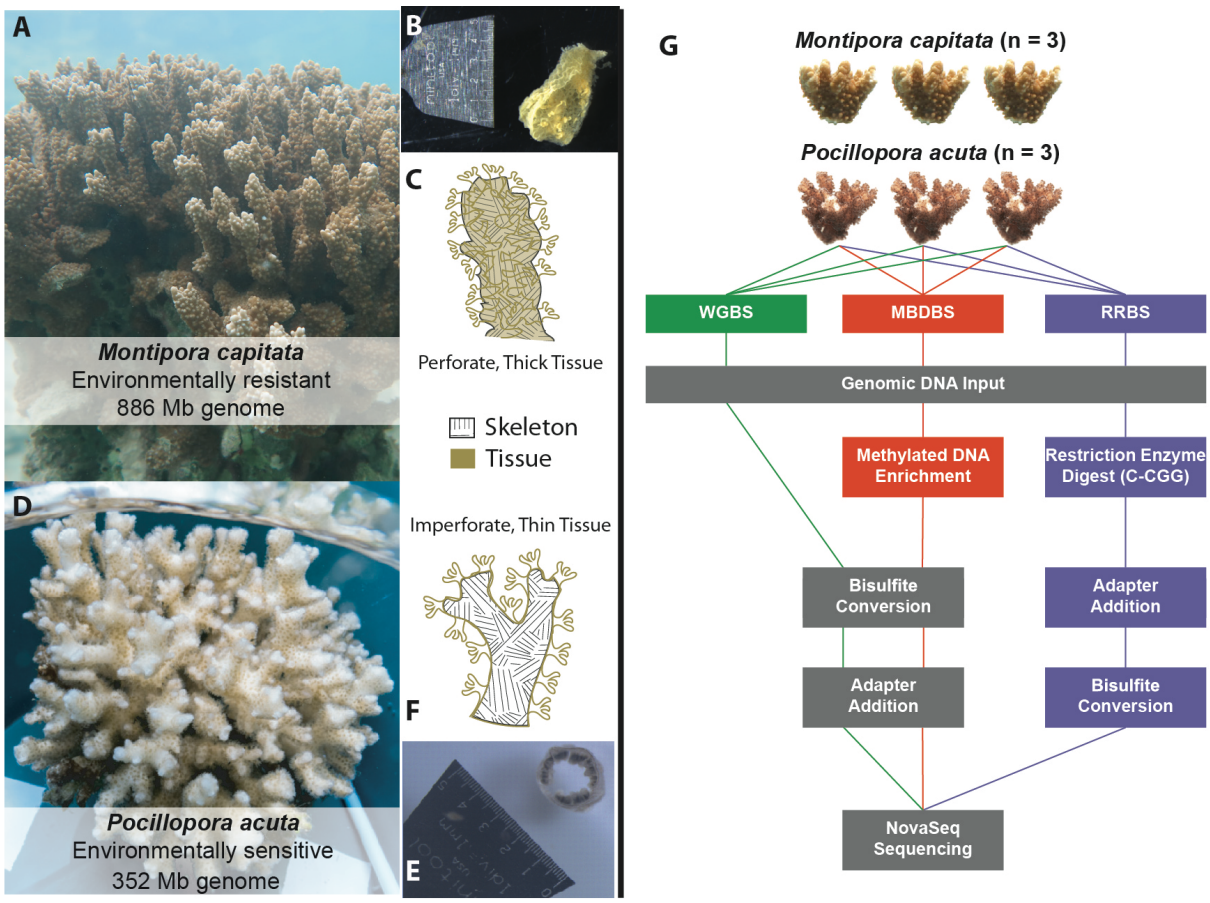

Figure 1. Experimental design. Three biological replicate coral samples were obtained from both coral species. A) *M. capitata*, where B) a cross section of a decalcified fragment reveals thick tissue, and C) a perforate tissue skeletal interaction. In contrast in D) *P. acuta*, a E) a cross section of a decalcified fragment reveals thin tissue, and F) an imperforate tissue skeletal interaction. DNA was extracted from each coral sample and split for use in Whole Genome Bisulfite Sequencing (WGBS), Reduced Representation Bisulfite Sequencing (RRBS), and Methyl-CpG Binding Domain Bisulfite Sequencing (MBDBS) library preparation methods. Three libraries were generated for each of the three methods, yielding nine libraries for each species and 18 libraries total.

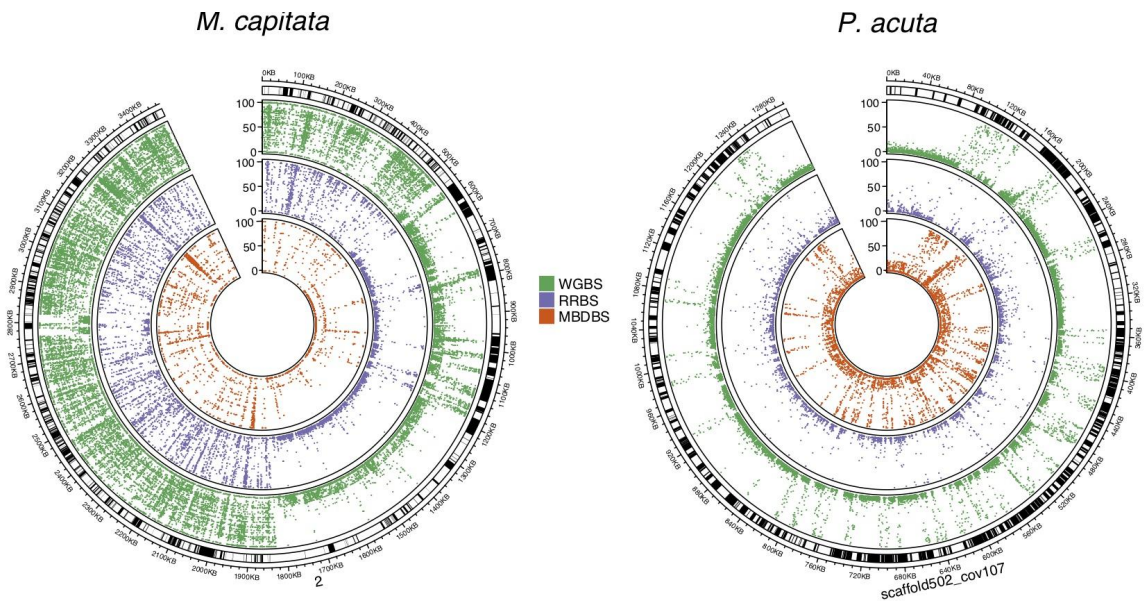

Figure 2. **Mean percent methylation of CpGs**. Data is presented for CpGs with 5x coverage for each method on the largest scaffolds of each genome. The outer track shows the scaffold locations and dots indicate the percent methylation as indicated by the y-axes from 0-100% for each of the inner tracks.

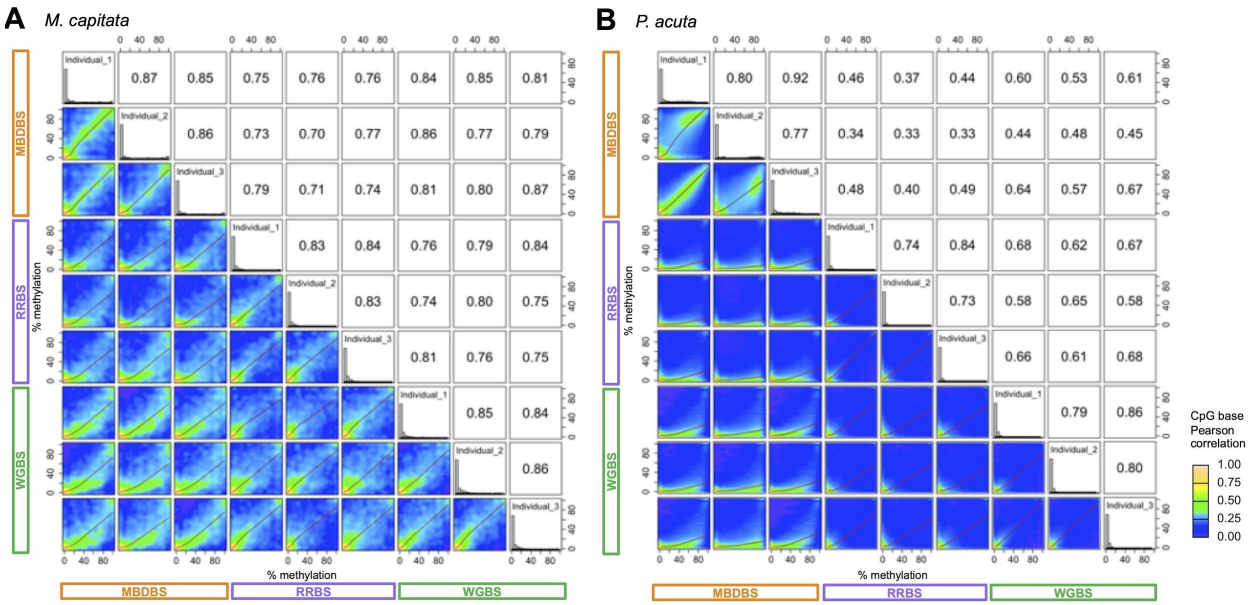

Figure 3. Matrix of pairwise scatter plots for shared CpG loci. Data is presented for CpG covered at ≥ 5x across all samples) for A) *M. capitata* (n=4,666 common loci) and B) *P. acuta* (n=93,714 common loci). The red lines represent linear regression fits and the

green lines are polynomial regression fits. Pearson correlation coefficients for each pairwise comparison are presented in the upper right boxes. Methods are color coded on the X and Y axes (WGBS = green, MBDBS = purple, and RRBS = orange) and replicate samples are indicated on the diagonal along with histograms of % CpG methylation.

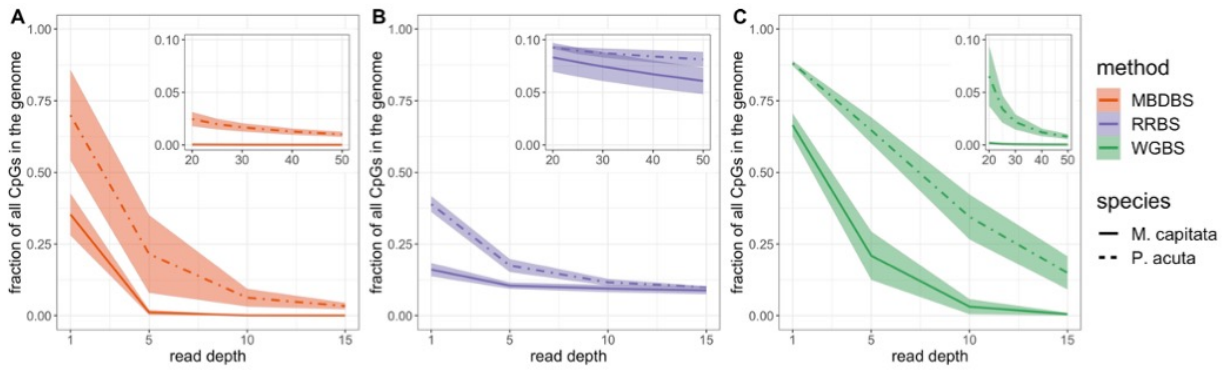

Figure 4. CpG site coverage across library preparation methods. Mean fraction of CpG sites in the genome covered at different sequencing depths (read depths) by (**A**) MBDBS libraries, (**B**) RRBS libraries, and (**C**) WGBS libraries with standard deviations shown by shaded areas (see Additional file 2: **Table ST2** for number of reads in each sample).

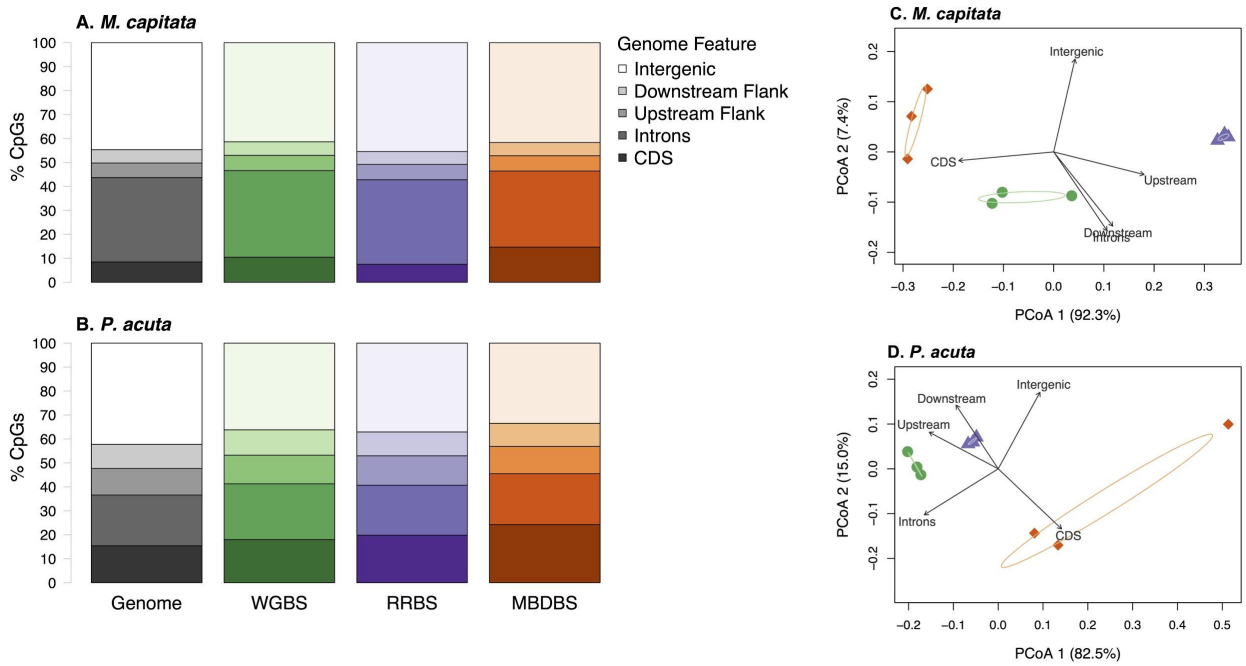

Figure 5. Percent of CpGs detected by sequencing methods in genome features A) for *M. capitata* and B) *P. acuta*. Genome features considered were coding sequences

(CDS), introns, 1 Kb flanking regions upstream (Upstream Flank) or downstream of genes (Downstream Flank), and intergenic regions Principal Coordinate Analyses associated with PERMANOVA and beta-dispersion tests related to Additional file 12: **Table ST6** that show differences in proportion of CpGs in various genomic locations (CDS, introns, upstream flanks, downstream flanks, and intergenic regions) for **C**) M. capitata and **D**) P. acuta. WGBS is represented by green circles, RRBS by purple triangles, and MBDBS by orange diamonds. Percent variation explained by each PCoA axis is included in the axis label. Ellipses depict 95% confidence intervals for each sequencing method. All eigenvectors are significant at the $\alpha = 0.05$ level.

Supplementary Information