1 **Cytosine methylation within marine sediment microbial communities: potential epigenetic**
2 **adaptation to the environment**
3
4 **Running title:** Sediment epigenetics
5
6 **Authors:** Ian M. Rambo[1*], Adam Marsh[1], Jennifer F. Biddle[1#]
7 [1]School of Marine Science and Policy, University of Delaware, Lewes, Delaware USA
8
9
10 **Keywords:** epigenetics, methylation, metagenome, sediment, chitinase, transposase
11
12 [#]Corresponding author: Jennifer Biddle, 700 Pilottown Road, Lewes DE 19958; (302) 645-4267;
13 jfbiddle@udel.edu
14
15 [*]**Current address**: Department of Marine Science, University of Texas Marine Science Institute,
16 Port Aransas, Texas USA
17
18 **Abstract**
19   Marine sediments harbor a vast amount of Earth's microbial biomass, yet little is

20 understood regarding how cells subsist in this low-energy, presumably slow-growth

21 environment. Cells in marine sediments may require additional methods for genetic regulation,

22 such as epigenetic modification via DNA methylation. We investigated this potential

23 phenomenon within a shallow estuary sediment core spanning 100 years of age across its depth.

24 Here we provide evidence of dynamic community m5-cytosine methylation within estuarine

25 sediment metagenomes using a methylation-sensitive Illumina assay. The methylation states of

26 individual CpG sites were reconstructed and quantified across three depths within the sediment

27 core. A total of 6254 CpG sites were aligned for direct comparison of methylation states between

28 samples, with 4235 sites mapped to taxa and genes. Our results demonstrate the presence of

29 differential methylation within environmental CpG sites across an age/depth gradient of

30 sediment. We show that epigenetic modification can be detected within complex environmental

31 communities. The change in methylation state of environmentally relevant genes across depths

32 may indicate a dynamic role of DNA methylation in biogeochemical processes.

Introduction

Marine sediments are some of the largest reservoirs of microbial biomass on Earth (Whitman *et al.* 1998; Kallmeyer *et al.* 2012), and describing the relationships between community structure, activity, and ecosystem function in these habitats remains a challenge (Fuhrman 2009). The majority of sedimentary bacteria and archaea are unable to be successfully cultured in a laboratory setting, and if they are able to be cultivated, they likely do not exist in physiological states representative of those found within their natural habitats (Hoehler and Jørgensen, 2013). Next-generation sequencing technologies enable researchers to overcome the constraints of cultivation by directly analyzing environmental DNA and RNA. These technologies are employed in subsurface microbiology to provide information regarding community dynamics and ecological roles (Hua *et al.* 2014), classify rare or uncultured species (Albertsen *et al.* 2013; Seitz *et al.* 2016), and describe potential microbial activity (Orsi *et al.* 2013).

Determining the drivers that govern microbial activity in the subsurface is key to understanding the relationships between these communities and their environments. Models of the marine subsurface suggest that biomass turnover rates are on the scale of thousands of years and that many marine subsurface cells should be sporulated due to the low availability of energy (Lomstein *et al.* 2012), yet metagenomic analyses of deep-sea sediment communities exhibit low observed frequencies of endospore-specific genes (Kawai *et al.* 2015). While isolates obtained from the deep biosphere are phylogenetically similar to members of surface communities (Russell *et al.* 2016; Inagaki *et al.* 2015), cells adapted to the subsurface possibly suspend certain life processes though other functional strategies to subsist at low levels of activity. Epigenetic mechanisms offer potential microbial survival strategies within low-energy sediment, allowing

56   for cell maintenance and acclimation to environmental stressors (Bird 2002; Casadesús and Low

57   2006; Low and Casadesús 2008).

58   DNA methylation is a conserved epigenetic modifier in prokaryotes whose roles include

59   defense against invading foreign DNA and gene regulation (Kumar and Rao 2012; Wion and

60   Casadesus 2006; Low *et al.* 2001; Brunet *et al.* 2011), and involves the addition of a methyl

61   group via a DNA methyltransferase (MTase) to either the carbon 5 position of a cytosine

62   (resulting in 5-methylcytosine (m5C)), the nitrogen 4 position of a cytosine (resulting in N4-

63   methylcytosine (m4C)), or the nitrogen 6 position of an adenine (resulting in N6-methyladenine

64   (m6A)) within a specific nucleotide target sequence (Ratel *et al.* 2006). These modified bases

65   comprise an organism's methylome, and are generally formed by two different MTase activities.

66   The two major forms of MTase activity are "maintenance" and "de novo" methylation.

67   Maintenance methylation (MM) provides cells with a means of propagating DNA methylation

68   patterns across generations. Daughter DNA strands with methylated parent strands are modified

69   by a maintenance MTase after replication (Bird, 2002). Similarly, non-methylated parent strands

70   normally produce non-methylated daughter strands. Unlike MM, which propagates existing

71   methylation patterns, de novo methylation adds methyl groups to previously unmethylated bases

72   (Kuhlmann *et al.* 2005).

73   While DNA methylation is an integral part of restriction-modification (RM) systems

74   involved in the recognition of self vs. non-self DNA for cellular defense, growing evidence

75   indicates that prokaryotes utilize both adenine and cytosine methylation as a means of regulating

76   gene expression (Reisenauer and Shapiro 2002; Srikhanta *et al.* 2005; Srikhanta *et al.* 2009;

77   Collier 2009; Brunet *et al.* 2011; Low and Casadesús 2008; Marinus and Casadesus 2009;

78   Løbner-Olesen *et al.* 2005; Wion and Casadesus 2006; Low *et al.* 2001; Kahramanoglou *et al.*

79   2012; Gonzalez *et al.* 2014; Blow *et al.* 2016). RNA polymerase, transcription factors and

80   binding proteins are able to recognize the methylated states of modified bases within target sites,

81   and this discrimination of differentially methylated DNA acts as a method for determining which

82   genes are transcribed at specific stages in the cell cycle (Low and Casadesús 2008; Gonzalez *et*

83   *al.* 2014; Collier 2009).

84   Compared to an organism's genome which generally remains static, the modified bases

85   of the methylome exist in a dynamic system exhibiting plasticity outside of binary "methylated"

86   or "non-methylated" states (Ichida *et al.* , 2007; Chernov *et al.* , 2015). A system of genetic

87   "switches" (Hernday *et al.* 2004) regulated by dynamic DNA methylation could be a viable

88   mechanism for both long-term and short-term transcriptional silencing for microbes inhabiting

89   marine sediments. To better understand the prevalence and behavior of DNA methylation within

90   marine sediment microbe communities, we utilized an Illumina sequencing-based assay to

91   identify dynamic shifts in CpG methylation within sediment metagenomes from the Broadkill

92   River estuary system. We opted to utilize this assay due to the presence of cytosine methylation

93   at CpG sites in prokaryotes and the anticipation of a heterogeneous, dynamic community. Since

94   adenine methylation is considered to be more widespread than cytosine methylation in

95   prokaryotes (Ratel *et al.* 2006), this choice of motif also serves to reduce signal saturation in this

96   mixed community. To the best of our knowledge, this is the first report on DNA methylation

97   within metagenomic sequence data, and is the first to utilize this method of CpG methylation

98   analysis in an environmental application.

99

100   **Materials and Methods**

101   *Core collection*

4

102   Sediment cores were sampled from the Oyster Rocks site of the Broadkill River, Milton,

103   DE, USA (38.802161, -75.20299) at low tide in July 2012 and 2014. The 2012 core was

104   sectioned into 3 cm sections and immediately frozen at -80°C for subsequent processing and

105   DNA extraction. The sediment collection from 2012 was depleted to extract sufficient DNA for

106   sequencing. Three cores were extracted from the same site in 2014 ~5 m from the riverbank: a 32

107   cm radionuclide dating core (R), and 25 cm (S) and 30 cm (L) cores for pore water ion

108   chromatography, methane flame ionization gas chromatography, and porosity measurements.

109   Cores L and S were sliced into 3 cm depth samples and immediately frozen at -80 °C, while Core

110   R was immediately processed.

111   *Radionuclide dating*

112   Core R was sectioned into 1 cm thick intervals from 0-10 cm, and 2 cm thick intervals

113   from 10-32 cm. Samples were weighed, dried at 60 °C for 48 hours, reweighed, and transferred

114   to a 25 °C desiccation chamber for storage until further processing. Dried samples were crushed

115   with a mortar and pestle, and ground into a fine powder with an IKA Werke M20 mill (IKA

116   Werke, Staufen, Germany). Powdered samples were transferred to 60 ml plastic jars and

117   compressed at 3.4×10^3 kPa with a manual hydraulic press. Radionuclide counting of compressed

118   samples was performed for 24 hours on a Canberra Instruments Low Energy Germanium

119   Detector (Canberra Industries, Meriden, CT, USA). Levels of ^{7}Be ($t_{1/2}$ = 53.22 days), ^{210}Pb ($t_{1/2}$ =

120   22.20 years), and ^{137}Cs ($t_{1/2}$ = 30.17 years) activity were measured by gamma spectroscopy of the

121   478, 46.5, and 662 kEV photopeaks, respectively (Igarashi *et al.* 1998; Cutshall *et al.* 1983;

122   Wallbrink *et al.* 2002).

123   *Porewater ion chromatography*

124   Porewater was extracted from 50 mL sediment samples by centrifugation at 13,000 G for

125   30 minutes. Porewater ions were measured with a Metrohm 850 Professional ion chromatograph

126   (Metrohm, Herisau, Switzerland). Dilutions were measured to determine a standard curve.

127   Samples were diluted to ensure signal within the standard curve.

128   *Methane flame ionization gas chromatography*

129   Methane concentrations were determined for Core L and S subsamples (volume = 305

130   cm^3) extracted from each core slice using a 5 ml syringe whose top had been removed with a

131   sterile razor blade. Core subsamples were transferred into 20 mL amber glass vials, and 1 mL 1

132   M NaOH was added to each vial to halt microbial activity. Vials were crimped, shaken, and

133   stored for 10 days at 25°C. A standard curve was calculated from 500, 1000, and 5000 ppm

134   standards. Mean headspace methane concentrations were determined by running 100 μL gas

135   extractions in triplicate via flame ionization gas chromatography using a 5890 Series II gas

136   chromatograph equipped with a flame ionization detector (Hewlett-Packard, Palo Alto,

137   California, USA).

138   *Metagenome library preparation and sequencing*

139   Metagenome libraries were prepared from the 2012 sediment core sections. Genomic

140   DNA (gDNA) was extracted from 0.5 g of sediment with a MoBio PowerSoil (MoBio, Valencia,

141   CA) kit per the manufacturer's protocol. A 10 μg aliquot of purified gDNA was digested with

142   the methylation-sensitive restriction endonuclease HpaII, which cleaves at the unmodified

143   internal cytosine of a 5'-CCGG-3' motif. Digested DNA was cleaned with a QIAquick PCR

144   purification kit (Qiagen, Hilden, Germany), sheared to a median size of 300 bp using a Covaris

145   focused-ultrasonicator (Covaris, Woburn, MA, USA), and cleaned again with QIAquick.

146   Digested extracts were immediately transferred to -20°C until library preparation. Illumina

147   libraries were prepared using the NEBNext Ultra Library Prep Kit for Illumina (New England

6

148   BioLabs, Ipswich, MA, USA) and sequenced with an Illumina Hi-Seq 2500 (Illumina, San

149   Diego, California, USA) at the Delaware Genomics and Biotechnology Institute (Newark, DE,

150   USA). Single-read sequencing was performed for all samples, with 150-cycle sequencing for the

151   3-6 cm and 12-15 cm samples, and 50-cycle sequencing for the 24-27 cm sample. All sequence

152   reads are deposited in GenBank under the study PRJEB11699.

153   *16S rRNA gene amplicon sequencing and analysis*

154   DNA was extracted, purified, and digested using the previously described method for

155   Illumina libraries. Purified DNA was quantified and tested for successful PCR reactions for the

156   bacterial 16S rRNA gene. Amplicon library preparation and sequencing of 16S rRNA genes

157   were performed by Molecular Research, LP (Clearwater, Texas, USA).

158   Analysis of 16S rRNA gene sequences was performed with QIIME 1.8.0 (Caporaso *et al.*

159   , 2010). Dereplication, abundance sorting, and discarding reads less than 2 bp was performed

160   with the USEARCH7 algorithm (Edgar, 2013). Chimeras were filtered with UCHIME (Edgar *et*

161   *al.* , 2011) using the RDP Gold Classifier training database v9 (Cole *et al.* , 2014). Operational

162   taxonomic unit (OTU) picking was performed at 97% similarity with UCLUST (Edgar, 2010).

163   Non-chimeric sequences were chosen as the representative set of sequences for taxonomic

164   assignment and alignment. Taxonomic assignments were performed with UCLUST (Edgar 2010)

165   using the Greengenes V13.8 database for 97% OTUs (DeSantis *et al.* 2006). OTU tables were

166   rarefied from 2000 to 9500 sequences per sample by steps of 100, with 10 iterations performed at

167   each step.

168   *Metagenome assembly and annotation*

169   Metagenome sequence reads were trimmed to 51 bp and quality controlled to only

170   include those with Phred nucleotide confidence scores greater than or equal to 95%. Quality-

171 controlled reads were assembled in IDBA (Peng *et al.* , 2010) with parameters –mink 18 –maxk

172 36 –step 2 –similar 0.97 –min_count 2 (Table S1). Phylogenetic classification of IDBA-

173 assembled contigs was performed with PhymmBL (Brady and Salzberg, 2011) and Kraken

174 (Wood and Salzberg, 2014). A PhymmBL identity confidence score threshold of 65% was

175 imposed to designate higher-confidence Order-level assignments. Comparative taxonomic

176 classifications were performed with Kraken (Wood and Salzberg, 2014) using the standard

177 database comprised of complete RefSeq bacterial, archaeal, viral, and fungal genomes. Contigs

178 assigned to viral or fungal genomes in Kraken were removed from downstream analyses. Marker

179 gene annotation of filtered contigs was performed with Phylosift (Darling *et al.* , 2014).

180 Open reading frame (ORF) prediction was performed in six reading frames with

181 MetaGene (Noguchi *et al.* 2006). ORFs were annotated for KEGG Orthology (KO) families

182 (Kanehisa *et al.* , 2016) in HMMER 3.0 (Eddy, 2011) using the Functional Ontology

183 Assignments for Metagenomes (FOAM) database (Prestat *et al.* , 2014) and an e-value

184 acceptance threshold of 1e-4. In the case of multiple KO assignments per contig, the result with

185 the best e-value and bitscore was chosen to represent that contig. Contigs that did not receive a

186 protein annotation from these software were aligned with BLASTX (Altschul *et al.* , 1997)

187 against the NCBI non-redundant protein database and scored with the BLOSUM62 substitution

188 matrix (Henikoff and Henikoff, 1992), with a maximum expectation value of 1e-4 and a word

189 size of 3.

190 *Metagenome CpG methylation quantification and statistical analysis*

191 CpG methylation was calculated using a custom bioinformatic pipeline and software

192 platform (Genome Profiling LLC; Marsh and Pasqualone 2014). Platform workflow is managed

193 down a decision tree from assembled contigs, and performs the following tasks: 1) isolation of

194   informative target contigs, 2) sequence compression to reduce complexity, 3) contig re-assembly

195   and mapping to reference metagenomes, 4) CpG quantification for m5C site distributions (%

196   methylation across all gDNA copies in the sample), and 5) methylation profiling comparison

197   between samples based on quantitative statistics. The methylation score metrics recovered from

198   assembled contigs are based on independent characteristics of DNA fragmentation via HpaII

199   restriction digest and random shearing.

200   Computational reconstruction of CpG methylation is based on a null selection model,

201   where the distribution of m5C modifications at any single CpG site is expected to be 50% in a

202   large population of cells (genome copies) for CpG sites that are non-functional or silent. Where

203   CpG methylation status is important for cellular fitness and thus there is a selection force

204   pushing the m5C distributions away from a 50:50 equilibrium, these scoring algorithms are

205   focused on quantifying this departure from the null expectation, and the degree of departure

206   measured is proportional to overall methylation status of that CpG site among all the cellular

207   genome copies being sampled.

208   Statistical analyses of methylation scores were performed using R statistical package.

209   Modalities were tested with Hartigans' dip test for unimodality. Methylation score bootstrap

210   standard errors (SE) and coefficients of variation (CV) were estimated (n = 10,000). Score

211   variances were tested with a Brown-Forsythe Levene-type test. Two-tailed Jonckheere-Terpstra

212   trend tests were performed with 10,000-permutation reference distributions.

213

214   **Results**

215   *Sediment properties*

9

216   Radionuclide dating constraints show that the Oyster Rocks site is comprised of a top

217 layer of recently deposited tidally mixed or bioturbated sediment (~ 4 cm, sediment age < 106

218 days) situated above older sediment established 50-100+ years ago (Figure S1). Sulfate

219 concentrations were more varied between 0-3 cm and 3-6 cm (Figure S1 A) for Core L, but

220 concentrations were higher in deeper samples from 6-9 cm to 27-30 cm. Methane concentrations

221 of Core L were shown to increase with depth, with higher variance between 0-12 cm and lower

222 variance from 15-30 cm (Figure S1 B). Porosity for Core R was shown to be far lower within

223 older sediments (Figure S1 C). The higher variability of SO_4^{-2} and CH_4 in more recently

224 deposited sediments could be due to tidal forcing. While cores from separate years were used to

225 generate sequence and geochemical data, the ages of all sediments are consistent, in that the

226 shallowest sequenced sample is less than a year old and the deeper sequenced samples are

227 significantly older (50+ years). In general, the deepest samples are anoxic, with a transition

228 occurring past 9 cm depth. This follows the trends established in earlier sampling in this area

229 (Cheng 2013), showing that generalities can be drawn over time.

230 *16S rRNA gene analysis*

231   The diversity of 16S rRNA genes was generally higher at 3-6 cm than the 12-15 cm and

232 24-27 cm samples (Figure S2). The 12-15 cm and 24-27 cm samples had similar profiles for

233 rarefied Chao1 diversity and observed OTU counts. These deeper samples also had a higher

234 presence of Dehalococcoidetes and sulfate-reducing Deltaproteobacteria, as well as Marine

235 Crenarchaoetal Group and Marine Hydrothermal Vent Group archaea (Figure S3). OTUs were

236 clearly shared between the three depths, and corresponding abundance changes suggest that

237 known anaerobic taxa were more abundant at depth.

238 *Metagenome taxonomic composition and function*

10

239   The most abundant taxonomic classes present in metagenomic and 16S rRNA gene data

240   across all depths were the Actinobacteria, Bacilli, Clostridia, Deinococci, and α-β-δ-γ-

241   proteobacteria (Figure 1, S4). The sediment community transitions to a mostly anaerobic

242   environment with depth based on taxonomic and functional annotations. Metagenome contigs

243   with both higher-scoring PhymmBL annotations and Kraken annotations further indicate a

244   prevalence of anaerobic taxa within the 12-15 and 24-27 cm samples. The presence of

245   methanogenic archaea (Methanomicrobiales and Methanosarcinales) and anaerobic

246   Dehalococcoidia and Deltaproteobacteria within the 12-15 and 24-27 cm samples (Figure 1, S4)

247   suggest these communities support anaerobic lifestyles (Oremland and Polcin, 1982). KO

248   annotations suggest that deeper communities have the potential for anaerobic metabolism (Figure

249   2), as greater abundances of genes involved in sulfate reduction (formate dehydrogenase,

250   adenylyl sulfate kinase, NADH dehydrogenase, and heterodisulfide reductase) and

251   methanogenesis (trimethylamine corrinoid protein co-methyltransferase) were present within the

252   24-27 cm sample.

253   *Metagenome CpG methylation*

254   From these metagenome data, we assessed the methylation states of CpG sites. A total of

255   6254 CpG sites that could be directly compared between all three samples were mapped to 3743

256   contigs (4.33% of all three unprocessed IDBA assemblies). Differential methylation states were

257   observed in 1173 sites, while the remaining 5081 had equivalent methylation states. Of these

258   CpG sites, 4235 (67.7%) were identified within contigs receiving higher-confidence PhymmBL

259   Order classifications.

260   The methylation shift behaviors of individual CpG sites are varied and highly dependent

261   upon their original states. Community-wide methylation distributions showed higher proportions

11

262   of CpG sites that remain in highly methylated states from 3-6 cm to 12-15 cm (Figure 3, A;

263   Figure 4 A), and can be associated with MM. A gradual overall loss in methylation and transition

264   into more equilibrated, binary states of high and low methylation was seen at 12-15 cm (Figure

265   3, B) and 24-27 cm (Figure 3, C). An apparent increase of methylation losses ranging from ~25-

266   50% was shown to account for this (Figure 4, B), with higher numbers of sites shifting from ~80-

267   90% methylated states to hemimethylated and non-methylated states (Figure 5, B). A greater

268   number of CpG sites experienced shifts from non-methylated states to fully methylated states

269   when transitioning from the 12-15 cm to 24-27 cm, potentially indicating *de novo* methylation of

270   these sites. It should be noted that many CpG sites were shown to remain in non-methylated

271   states between 12-15 cm and 24-27 cm.

272   The methylation dynamics of individual CpG sites were analyzed for taxa with higher

273   numbers of recovered sites. An overall trend of increasing methylation score standard error (SE)

274   and coefficient of variation (CV) with depth was seen in all analyzed phyla (Table S2). There is a

275   general trend of decreasing CV for methylation scores with depth, and this is influenced by an

276   overall trend towards bimodal score distributions. Hartigans' dip test results support a non-

277   unimodal distribution of methylation scores for analyzed phyla (Table S3), verifying mixed

278   methylation profiles. Brown-Forsythe tests suggest that CpG score variances across depths were

279   unequal for 70% of analyzed phyla ($p < 0.05$), supporting the presence of mixed methylation

280   profiles and dynamic shifts in methylation states (Table S4). Jonckheere-Terpstra trend test

281   results show that community methylation scores decrease overall with depth ($p = 2e-4$).

282   Methylation scores for the majority of phyla exhibit decreasing trends with depth (Table S4;

283   Figure 3; Figure 5).

284   Only 35 CpG sites were mapped to contigs receiving KEGG Orthology annotations. Of

285   these 35 CpG sites, chitinase gene annotations were recovered for 14 comparable sites that could

286   be traced back to six contigs with higher-confidence PhymmBL classifications (Figure 6). A total

287   of 12 quantifiable sites exhibiting differential methylation states within the same gene were

288   identified for Actinomycetales and Thermoanaerobacterales.

289   Quantifiable states of 73 CpG sites were recovered for transposase genes identified by

290   BLASTX alignments (Figure 7). Transposase CpG sites that were methylated in surface samples

291   tended to remain in methylated states across depths, although several methylated sites undergo

292   shifts into hemimethylated or non-methylated sites in deeper samples. CpG sites existing within

293   the same contig tended to shift to similar methylation states from 12-15 cm to 24-27 cm.

294   **Discussion**

295   DNA methylation has an established association with RM systems and gene regulation

296   within cultured prokaryotes (Casadesús and Low 2006; Blow *et al.* 2016). However, its presence

297   and function within uncultured environmental microbes is not well understood, especially

298   considering the extent of undiscovered taxa (Locey and Lennon, 2016). Current methylation

299   detection protocols for NGS sequencing are improbable for sediment samples due to community

300   complexity, high required concentration and molecular weight of DNA samples, and error rate.

301   We provide evidence of differential m5C methylation at CpG sites within estuarine sediment

302   communities using a restriction enzyme-based Illumina sequencing method capable of

303   reconstructing methylation profiles from environmental DNA samples. This ability for

304   continuous CpG methylation scoring within metagenomic DNA opens future possibilities for

305   better understanding how microbes in the subsurface can respond to various conditions and

306   stressors.

307   Sediment communities observed through both 16S rRNA gene and metagenomic gDNA

308   sequencing do not appear to be governed highly by factors such as porosity and sediment

309   porewater geochemistry, and this notion is supported by previous research of Broadkill River

310   sediments (Cheng 2013) and other estuarine microbe communities (Koretsky *et al.* , 2005).

311   However, shifts in community composition appear to be more closely related to the drastic

312   change in sediment age suggested by radionuclide constraints.

313   Analysis of 16S rRNA gene amplicons suggests that the 3-6 cm sample communities

314   exhibit greater diversity than the deeper samples (Figure S2). It is likely that surface sediments

315   are more aerobic than deeper sediments due to regular cycling and deposition, and these higher

316   oxygen levels could be a factor in this higher diversity. Obligate anaerobes and facultative

317   aerobes were observed in the shallow sample as well. The 3-6 cm sample also encompasses the

318   transition zone from young, fresh sediment to older, established sediment at 4-5 and 5-6 cm, so

319   overlap in communities was expected. Our results clearly show increases in age from 3-6 to 12-

320   15 cm, and the deeper depth of 24-27 cm is certainly older although our tests could not measure

321   an exact age between 15-24 cm. Results support the presence of a drastically shifting downcore

322   age gradient with higher anaerobic community potential at depth. Methane and porewater ion

323   profiles are more varied within surface sediments, suggesting a bioturbated or tidally mixed

324   region of fresh sediment in line with ^{7}Be and ^{210}Pb activity constraints. CpG methylation profiles

325   recovered from these sediments were mapped to taxa and genes, and exhibit dynamic shifts in

326   methylation state.

327   Of the CpG site populations assigned to taxa, 22% had < 5% methylation gains or losses

328   between depths. Two possible explanations are presented here. First, MM can propagate

329   methylated CpGs and result in a higher conservation of methylated states. Second, MM does not

330   act upon non-methylated CpG sites, and as such these sites tend to remain non-methylated

331   (Casadesús and Low, 2006). While these sites remained within generally equivalent states,

332   9.22% of recovered CpGs were observed to have differential methylation shifts >= 50% between

333   depths. This is representative of the standard binary response associated with the concept of an

334   epigenetic on/off switch (Marsh and Pasqualone 2014; Hernday *et al.* 2004). Shifts between

335   highly methylated and fractionally methylated states suggest the presence of dynamic CpG sites

336   that contribute to a mixed population. However, we cannot rule out the potential effect of gene or

337   whole genome duplication on methylation scoring, as newly replicated DNA would contain

338   fewer methylated bases and is highly dependent upon maintenance methylation (Casadesús and

339   Low, 2006).

340   CpG methylation states were shown to vary for specific genes including chitinases and

341   transposases. Chitinolytic bacteria are widely distributed in sediment environments

342   (Bhattacharya *et al.* 2007; Souza *et al.* 2011), and are responsible for converting this insoluble

343   source of carbon and nitrogen into a widely used form. It has been previously noted that chitin is

344   rapidly removed from an estuary within the first 10 cm of sediment (Gooday 1990). The most

345   abundant shifts from the 3-6 cm to 12-15 cm samples are from methylated to unmethylated states

346   within lineages of Actinobacteria and Thermoanaerobacter, both of which can degrade chitin

347   (Bhattacharya *et al.* 2007; Spanevello and Patel 2015). We interpret the variations in these

348   signals to suggest regulation, and not just a signature of cellular replication, considering that

349   methylation losses are greater upon transition to the anaerobic 12-15 cm depth. These same CpG

350   sites are again methylated within the 24-27 cm depth as they leave the assumed zone of available

351   chitin. While chitin was not concurrently measured, the often noted correlation between

352   cultivable chitinolytic bacteria and chitin abundances suggests that this process is one that would

353   not be maintained if chitin were not present (Gooday 1990). The evidence for anaerobic

354   organisms only reducing methylation from chitinase CpG sites within the anaerobic sedimentary

355   horizons suggests that this is methylation-based regulation of an metabolically energetically

356   costly process. We postulate that this is an initial glimpse into how marine sediment microbes

357   potentially utilize DNA methylation to regulate biogeochemical processes that are vital for

358   nutrient cycling.

359   DNA methylation's potential to regulate gene transcription also applies to the expression

360   of genes involved in the transport of transposable elements, i.e. transposases. We provide

361   evidence for differential methylation of multiple CpG sites within sediment bacterial transposase

362   genes (Figure 7), as a lack of methylation within CpG sites or shifts into hemimethylated states

363   could hint at the potential activity of transposases. The regulation of transposases and transposon

364   mobility could play a role in rapid acclimation responses by influencing transcriptional activity

365   and the ability for mobile genes to be inserted into a genome. Transposase regulation has been

366   observed to take place via adenine methylation at GATC sites in *E. coli* (Dodson and Berg 1989;

367   Reznikoff 1993; Roberts *et al.* 1985; Yin *et al.* 1988; Spilemann-Ryser *et al.* 1991)*,* yet the

368   regulatory mechanisms of one model organism do not necessarily apply to the entire bacterial

369   domain. Horizontal gene transfer is speculated to occur in estuarine sediment microbes

370   (Angermeyer *et al.* 2016), and extracellular transposases could play a potential role in this

371   process as substantial numbers of horizontally-transferred genes in several bacteria species are

372   attributed to foreign DNA such as transposons (Ochman *et al.* 2000). Due to the known influence

373   of DNA methylation within bacterial transposons and the results of this study, we speculate that

374   DNA methylation could act as a regulator of transposition within the subsurface.

375   As epigenetic research shifts from model systems towards potentially novel organisms

376   within natural environments, there is a pressing need for the development of assays capable of

377   detecting epigenetic signatures within environmental samples. This study provides a community-

378   level insight into the dynamic behavior of a well-known and conserved methylation site within

379   estuarine sediments. A benefit of this Illumina assay is that it requires less DNA than single-

380   molecule approaches, and allows for CpG site mapping to specific taxa and genes. Future

381   modifications tailored for metagenomic samples could pave the way for the reconstruction of

382   dynamic methylation profiles within genomes obtained from the environment.

383

Conflict of Interest

    The software platform designed for processing DNA methylation profiles from NGS sequence data is licensed by the University of Delaware to Genome Profiling LLC, a company co-founded by Adam G. Marsh and developed with support from an Innovation Corps Grant from the National Science Foundation to AGM. Ian Rambo and Jennifer Biddle declare no financial involvement with any commercial entity.

Datasets for this project are publicly available from BCO-DMO at http://www.bco-dmo.org/dataset/628253.

17

References:

Albertsen, M., Hugenholtz, P., Skarshewski, A., Nielsen, K. L., Tyson, G. W. and Nielsen, P. H. (2013) 'Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes.', *Nature biotechnology*, 31(6), pp. 533–8. doi: 10.1038/nbt.2579.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) 'Gapped BLAST and PSI-BLAST: A new generation of protein database search programs', *Nucleic Acids Research*, 25(17), pp. 3389–3402. doi: 10.1093/nar/25.17.3389.

Angermeyer, A., Crosby, S. C. and Huber, J. A. (2016) 'Decoupled distance-decay patterns between dsrA and 16S rRNA genes among salt marsh sulfate-reducing bacteria', *Environmental Microbiology*, 18(1), pp. 75–86. doi: 10.1111/1462-2920.12821.

Bhattacharya, D., Nagpure, A. and Gupta, R. K. (2007) 'Bacterial chitinases: Properties and Potential.', *Critical reviews in biotechnology*, 27(1), pp. 21–28. doi: 10.1080/07388550601168223.

Bird, A. (2002) 'DNA methylation patterns and epigenetic memory.', *Genes & development*, 16(1), pp. 6–21. doi: 10.1101/gad.947102.

Blow, M. J., Clark, T. A., Daum, C. G., Deutschbauer, A. M., Fomenkov, A., Fries, R., Froula, J., Kang, D. D., Malmstrom, R. R., Morgan, R. D., Posfai, J., Singh, K., Visel, A., Wetmore, K., Zhao, Z., Rubin, E. M., Korlach, J., Pennacchio, L. A. and Roberts, R. J. (2016) 'The Epigenomic Landscape of Prokaryotes', *PLoS Genetics*, 12(2), pp. 1–28. doi: 10.1371/journal.pgen.1005854.

Brady, A. and Salzberg, S. (2011) 'PhymmBL expanded: confidence scores, custom databases, parallelization and more', *Nature methods*, 8(5), pp. 2011–2013. doi: 10.1038/nmeth0511-367.

Brunet, Y. R., Bernard, C. S., Gavioli, M., Lloubès, R. and Cascales, E. (2011) 'An epigenetic switch involving overlapping fur and DNA methylation optimizes expression of a type VI secretion gene cluster.', *PLoS genetics*, 7(7), p. e1002205. doi: 10.1371/journal.pgen.1002205.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. a, Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. a, Mcdonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. a, Widmann, J., Yatsunenko, T., Zaneveld, J. and Knight, R. (2010) 'correspondence QIIME allows analysis of high- throughput community sequencing data Intensity normalization improves color calling in SOLiD sequencing', *Nature Publishing Group*. Nature Publishing Group, 7(5), pp. 335–336. doi: 10.1038/nmeth0510-335.

Casadesús, J. and Low, D. (2006) 'Epigenetic gene regulation in the bacterial world.', *Microbiology and molecular biology reviews : MMBR*, 70(3), pp. 830–56. doi: 10.1128/MMBR.00016-06.

Cheng, B. (2013) *Variations of Archaeal Communities in Sediments of Coastal Delaware*. University of Delaware, Thesis.

Chernov, A. V., Reyes, L., Peterson, S. and Strongin, A. Y. (2015) 'Depletion of CG-specific methylation in Mycoplasma hyorhinis genomic DNA after host cell invasion', *PLoS ONE*, 10(11), pp. 1–14. doi: 10.1371/journal.pone.0142529.

Cole, J. R., Wang, Q., Fish, J. a., Chai, B., McGarrell, D. M., Sun, Y., Brown, C. T., Porras-Alfaro, A., Kuske, C. R. and Tiedje, J. M. (2014) 'Ribosomal Database Project: Data and tools for high throughput rRNA analysis', *Nucleic Acids Research*, 42(D1), pp. 633–642. doi: 10.1093/nar/gkt1244.

Collier, J. (2009) 'Epigenetic regulation of the bacterial cell cycle.', *Current opinion in microbiology*, 12(6), pp. 722–9. doi: 10.1016/j.mib.2009.08.005.

Darling, A. E., Jospin, G., Lowe, E., Matsen, F. a, Bik, H. M. and Eisen, J. a (2014) 'PhyloSift: phylogenetic analysis of genomes and metagenomes.', *PeerJ*, 2, p. e243. doi: 10.7717/peerj.243.

Eddy, S. R. (2011) 'Accelerated profile HMM searches', *PLoS Computational Biology*, 7(10). doi: 10.1371/journal.pcbi.1002195.

Edgar, R. C. (2010) 'Search and clustering orders of magnitude faster than BLAST', *Bioinformatics*, 26(19), pp. 2460–2461. doi: 10.1093/bioinformatics/btq461.

Edgar, R. C. (2013) 'UPARSE: highly accurate OTU sequences from microbial amplicon reads.', *Nature Methods*, 10(10), pp. 996–8. doi: 10.1038/nmeth.2604.

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. and Knight, R. (2011) 'UCHIME improves sensitivity and speed of chimera detection', *Bioinformatics*, 27(16), pp. 2194–2200. doi: 10.1093/bioinformatics/btr381.

Fuhrman, J. a (2009) 'Microbial community structure and its functional implications.', *Nature*, 459(7244), pp. 193–9. doi: 10.1038/nature08058.

Gonzalez, D., Kozdon, J. B., McAdams, H. H., Shapiro, L. and Collier, J. (2014) 'The functions of DNA methylation by CcrM in Caulobacter crescentus: a global approach.', *Nucleic acids research*, 42(6), pp. 3720–35. doi: 10.1093/nar/gkt1352.

Gooday, G. W. (1990) 'The ecology of chitin degradation', *Advances in Microbial Ecology*, 11, pp. 387–430.

Henikoff, S. and Henikoff, J. G. (1992) 'Amino acid substitution matrices from protein blocks', *Proceedings of the National Academy of Sciences of the United States of America*, 89(November), pp. 10915–10919.

Hernday, A., Braaten, B. A. and Low, D. A. (2004) 'The intricate workings of a bacterial

494   epigenetic switch', *Advances in experimental medicine and biology*, 547.
495
496   Hoehler, T. M. and Jørgensen, B. B. (2013) 'Microbial life under extreme energy limitation.',
497   *Nature reviews. Microbiology*. Nature Publishing Group, 11(2), pp. 83–94. doi:
498   10.1038/nrmicro2939.
499
500   Hua, Z.-S., Han, Y.-J., Chen, L.-X., Liu, J., Hu, M., Li, S.-J., Kuang, J.-L., Chain, P. S., Huang,
501   L.-N. and Shu, W.-S. (2014) 'Ecological roles of dominant and rare prokaryotes in acid mine
502   drainage revealed by metagenomics and metatranscriptomics.', *The ISME journal*. Nature
503   Publishing Group, pp. 1–15. doi: 10.1038/ismej.2014.212.
504
505   Ichida, H., Matsuyama, T., Abe, T. and Koba, T. (2007) 'DNA adenine methylation changes
506   dramatically during establishment of symbiosis', *FEBS Journal*, 274(4), pp. 951–962. doi:
507   10.1111/j.1742-4658.2007.05643.x.
508
509   Inagaki, F., Kubo, Y., Bowles, M. W., Heuer, V. B., Ijiri, A., Imachi, H., Ito, M., Kaneko, M.,
510   Lever, M. A., Morita, S., Morono, Y., Tanikawa, W., Bihan, M., Bowden, S. A., Elvert, M.,
511   Glombitza, C., Gross, D., Harrington, G. J., Hori, T., Li, K., Limmer, D., Murayama, M.,
512   Ohkouchi, N., Ono, S., Purkey, M., Sanada, Y., Sauvage, J., Snyder, G., Takano, Y., Tasumi, E.,
513   Terada, T., Tomaru, H., Wang, D. T. and Yamada, Y. (2015) 'Exploring deep microbial life in
514   coal-bearing sediment down to ~2.5 km below the ocean floor', *Science*, 349(6246), pp. 420–
515   424.
516
517   Kahramanoglou, C., Prieto, A. I., Khedkar, S., Haase, B., Gupta, A., Benes, V., Fraser, G. M.,
518   Luscombe, N. M. and Seshasayee, A. S. N. (2012) 'Genomics of DNA cytosine methylation in
519   Escherichia coli reveals its role in stationary phase transcription', *Nature Communications*.
520   Nature Publishing Group, 3, p. 886. doi: 10.1038/ncomms1878.
521   Kallmeyer, J., Pockalny, R., Adhikari, R. R., Smith, D. C. and D'Hondt, S. (2012) 'Global
522   distribution of microbial abundance and biomass in subseafloor sediment.', *Proceedings of the
523   National Academy of Sciences of the United States of America*, 109(40), pp. 16213–6. doi:
524   10.1073/pnas.1203849109.
525
526   Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. and Tanabe, M. (2016) 'KEGG as a
527   reference resource for gene and protein annotation', *Nucleic Acids Research*, 44(D1), pp. D457–
528   D462. doi: 10.1093/nar/gkv1070.
529
530   Kawai, M., Uchiyama, I., Takami, H. and Inagaki, F. (2015) 'Low frequency of endospore-
531   specific genes in subseafloor sedimentary metagenomes', *Environmental Microbiology Reports*,
532   7(2), pp. 341–350. doi: 10.1111/1758-2229.12254.
533
534   Koretsky, C. M., Van Cappellen, P., DiChristina, T. J., Kostka, J. E., Lowe, K. L., Moore, C. M.,
535   Roychoudhury, A. N. and Viollier, E. (2005) 'Salt marsh pore water geochemistry does not
536   correlate with microbial community structure', *Estuarine, Coastal and Shelf Science*, 62(1–2),
537   pp. 233–251. doi: 10.1016/j.ecss.2004.09.001.
538
539   Kuhlmann, M., Borisova, B. E., Kaller, M., Larsson, P., Stach, D., Na, J., Eichinger, L., Lyko,

540 F., Ambros, V., Söderbom, F., Hammann, C. and Nellen, W. (2005) 'Silencing of
541 retrotransposons in Dictyostelium by DNA methylation and RNAi', *Nucleic Acids Research*,
542 33(19), pp. 6405–6417. doi: 10.1093/nar/gki952.

544 Kumar, R. and Rao, D. N. (2012) 'Role of DNA Methyltransferases in Epigenetic Regulation in
545 Bacteria', in *Epigenetics: Development and Disease*, pp. 81–102. doi: 10.1007/978-94-007-
546 4525-4.

548 Løbner-Olesen, A., Skovgaard, O. and Marinus, M. G. (2005) 'Dam methylation: Coordinating
549 cellular processes', *Current Opinion in Microbiology*, 8(2), pp. 154–160. doi:
550 10.1016/j.mib.2005.02.009.

552 Locey, K. J. and Lennon, J. T. (2016) 'Scaling laws predict global microbial diversity',
553 *Proceedings of the National Academy of Sciences*, Early Edit, pp. 1–6. doi:
554 10.7287/peerj.preprints.1451v1.

556 Lomstein, B. A., Langerhuus, A. T., D'Hondt, S., Jørgensen, B. B. and Spivack, A. J. (2012)
557 'Endospore abundance, microbial growth and necromass turnover in deep sub-seafloor
558 sediment.', *Nature*. Nature Publishing Group, 484(7392), pp. 101–4. doi: 10.1038/nature10905.

560 Low, D. A., Weyand, N. J. and Mahan, M. J. (2001) 'Roles of DNA Adenine Methylation in
561 Regulating Bacterial Gene Expression and Virulence', *Infection and Immunity*, 69(12), pp.
562 7197–7204. doi: 10.1128/IAI.69.12.7197.

564 Low, D. and Casadesús, J. (2008) 'Clocks and switches: bacterial gene regulation by DNA
565 adenine methylation.', *Current opinion in microbiology*, 11(2), pp. 106–12. doi:
566 10.1016/j.mib.2008.02.012.

568 Marinus, M. G. and Casadesus, J. (2009) 'Roles of DNA adenine methylation in host-pathogen
569 interactions: Mismatch repair, transcriptional regulation, and more', *FEMS Microbiology
570 Reviews*, 33(3), pp. 488–503. doi: 10.1111/j.1574-6976.2008.00159.x.

572 Marsh, A. G. and Pasqualone, A. A. (2014) 'DNA methylation and temperature stress in an
573 Antarctic polychaete , Spiophanes tcherniai', *Frontiers in Physiology*, 5(May), pp. 1–9. doi:
574 10.3389/fphys.2014.00173.

576 Noguchi, H., Park, J. and Takagi, T. (2006) 'MetaGene: Prokaryotic gene finding from
577 environmental genome shotgun sequences', *Nucleic Acids Research*, 34(19), pp. 5623–5630. doi:
578 10.1093/nar/gkl723.

580 Ochman, H., Lawrence, J. G. and Groisman, E. A. (2000) 'Lateral gene transfer and the nature of
581 bacterial innovation', *Nature*, 405, pp. 299–304.

583 Oremland, R. S. and Polcin, S. (1982) 'Methanogenesis and sulfate reduction: competitive and
584 noncompetitive substrates in estuarine sediments.', *Applied and Environmental Microbiology*,
585 44(6), pp. 1270–1276. doi: 10.1016/0198-0254(83)90262-5.

586

587 Orsi, W. D., Edgcomb, V. P., Christman, G. D. and Biddle, J. F. (2013) 'Gene expression in the
588 deep biosphere.', *Nature*. Nature Publishing Group, 499(7457), pp. 205–8. doi:
589 10.1038/nature12230.

590

591 Peng, Y., Leung, H. C. M., Yiu, S. M. and Chin, F. Y. L. (2010) 'IDBA - A practical iterative De
592 Bruijn graph De Novo assembler', *Lecture Notes in Computer Science (including subseries*
593 *Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6044 LNBI, pp.
594 426–440. doi: 10.1007/978-3-642-12683-3_28.

595

596 Prestat, E., David, M. M., Hultman, J., Taş, N., Lamendella, R., Dvornik, J., Mackelprang, R.,
597 Myrold, D. D., Jumpponen, A., Tringe, S. G., Holman, E., Mavromatis, K. and Jansson, J. K.
598 (2014) 'FOAM (Functional Ontology Assignments for Metagenomes): a Hidden Markov Model
599 (HMM) database with environmental focus.', *Nucleic acids research*, 42(19), pp. 1–7. doi:
600 10.1093/nar/gku702.

601

602 Ratel, D., Ravanat, J.-L., Berger, F. and Wion, D. (2006) 'N6-methyladenine: the other
603 methylated base of DNA.', *BioEssays : news and reviews in molecular, cellular and*
604 *developmental biology*, 28(3), pp. 309–15. doi: 10.1002/bies.20342.

605

606 Reisenauer, A. and Shapiro, L. (2002) 'DNA methylation affects the cell cycle transcription of
607 the CtrA global regulator in [i]Caulobacter[/i].', *The EMBO Journal*, 21(18), pp. 4969–4977.
608 doi: 10.1093/emboj/cdf490.

609

610 Russell, J. A., Zayas-Leon, R., Wrighton, K. and Biddle, J. F. (2016) 'Deep subsurface life from
611 North Pond: enrichment, isolation, characterization and genomes of hetertrophic bacteria',
612 *Frontiers in Microbiology*, 7(May), pp. 1–13. doi: 10.3389/fmicb.2016.00678.

613

614 Seitz, K. W., Lazar, C. S., Hinrichs, K., Teske, A. P. and Baker, B. J. (2016) 'Genomic
615 reconstruction of a novel , deeply branched sediment archaeal phylum with pathways for
616 acetogenesis and sulfur reduction', *The ISME Journal*. Nature Publishing Group, 1(10), pp. 1–
617 10. doi: 10.1038/ismej.2015.233.

618

619 Souza, C. P., Almeida, B. C., Colwell, R. R. and Rivera, I. N. G. (2011) 'The Importance of
620 Chitin in the Marine Environment', *Marine Biotechnology*, 13(5), pp. 823–830. doi:
621 10.1007/s10126-011-9388-1.

622

623 Spanevello, M. D. and Patel, B. K. C. (2015) 'Thermoanaerobacter', in *Bergey's Manual of*
624 *Systematics of Archaea and Bacteria*. John Wiley & Sons, pp. 1–6.

625

626 Srikhanta, Y., Dowideit, S., Edwards, J., Falsetta, M., Wu, H.-J., Harrison, O., Fox, K., Seib, K.,
627 Maguire, T., Wang, A.-J., Maiden, M., Grimmond, S., Apicella, M. and Jennings, M. (2009)
628 'Phasevarions mediate random switching of gene expression in pathogenic Neisseria', *PLOS*
629 *Pathology*, 5.

630

631 Srikhanta, Y. N., Maguire, T. L., Stacey, K. J. and Grimmond, S. M. (2005) 'The Phasevarion :

632   A Genetic System Controlling Coordinated , Random Switching of Expression of Multiple
633   Genes', *Proceedings of the National Academy of Sciences*, 102(15), pp. 5547–5551.
634
635   Tolba, S. T. M., Nagwa, A. A. A. and Hatem, D. (2013) 'Molecular characterization of rare
636   actinomycetes using 16S rRNA-RFLP', *African Journal of Biological Sciences*, 9(March), pp.
637   185–197.
638
639   Whitman, W. B., Coleman, D. C., Wiebe, W. J. and Colemant, D. C. (1998) 'Prokaryotes: The
640   Unseen Majority', *Proceedings of the National Academy of Sciences of the United States of*
641   *America*, 95(12), pp. 6578–6583.
642
643   Wion, D. and Casadesus, J. (2006) 'N 6 -methyl-adenine : an epigenetic signal for DNA –
644   protein interactions', *Nature Reviews Microbiology*, 4(March), pp. 183–193. doi:
645   10.1038/nrmicro1350.
646
647   Wood, D. E. and Salzberg, S. L. (2014) 'Kraken: ultrafast metagenomic sequence classification
648   using exact alignments.', *Genome biology*, 15(3), p. R46. doi: 10.1186/gb-2014-15-3-r46.
649
650

651 **Figure Legends:**

652 **Figure 1**: Taxonomic relative abundances >= 0.05% of metagenome contigs with both Kraken

653 and higher-confidence PhymmBL class assignments present in two or more samples. PhymmBL

654 annotations were paired with Kraken annotations to select against potential false positives

655 annotated by PhymmBL. These annotations suggest a presence of Actinomycetales at all depths,

656 and an increased abundance of anaerobic classes (Clostridia, Dehalococcoidia).

657 Methanogenic archaea (Methanosarcinales) are present in the 12-15 cm and 24-27 cm

658 samples.

659

660 **Figure 2**: Abundance of KEGG Orthology functional assignments across sample depths. KEGG

661 Orthology annotation results show that Oyster Rocks sediment communities at 12-15 cm and 24-

662 27 cm have higher enzymatic potential for anaerobic metabolism.

663

664 **Figure 3:** Community methylation load distributions. Histograms and kernel density overlays

665 represent overall methylation levels of CpG sites that are shared across all three samples. A

666 greater number of recovered CpG sites were highly methylated at 3-6 cm (**A**), resulting in a

667 greater methylation load at this depth. However, a trend of decreasing overall methylation was

668 seen at 12-15 cm (**B**) and 24-27 cm (**C**), with more sites experiencing transitions from highly

669 methylated to non-methylated or hemimethylated states, or persisting in non-methylated states.

670

671 **Figure 4**: CpG methylation gains and losses from (**A**) 3-6 cm to 12-15 cm and (**B**) 12-15 cm to

672 24-27 cm. Histograms and kernel density overlays are representative of the densities of

673 methylation shifts for individual CpG sites. Sites represented in (**A**) are the same sites
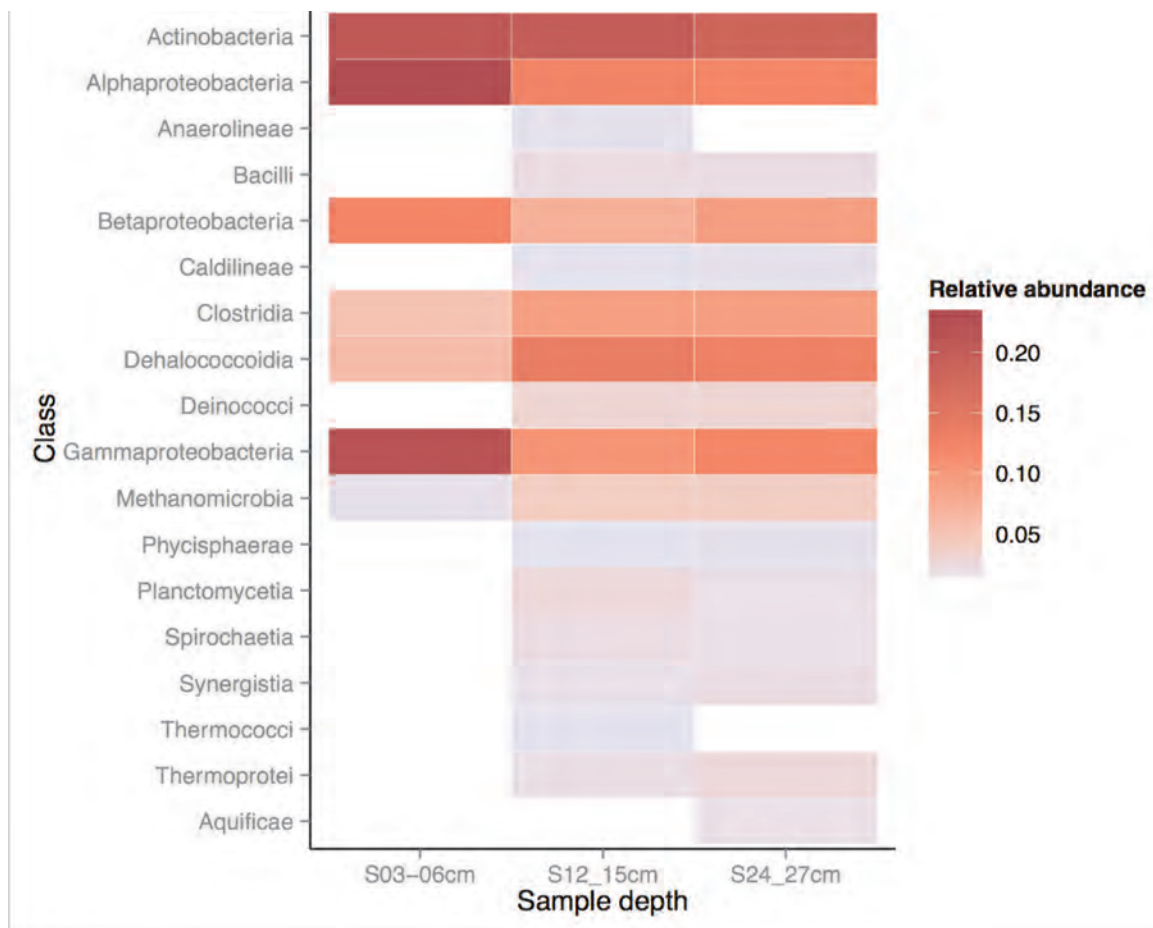
674   represented in (**B**). Shifts range from -100 (total methylation loss) to +100 (total methylation

675   gain). A significant number of CpG sites remained at equivalent methylation states from 3-6 cm

676   to 12-15 cm, yet there is an apparent increase in methylation losses ranging from ~25% to ~50%

677   from 12-15 cm to 24-27 cm.

678

679   **Figure 5:** Total CpG methylation shifts from 3-6 cm to 12-15 cm (**A**) and 12-15 cm to 24-27 cm

680   (**B**). Plots are representative of metagenome-wide methylation loads. Each point is a recovered

681   CpG site whose quantified methylation states are comparable across all three samples. The CpG

682   sites represented in (**A**) are the same as those represented in (**B**). Changes in a CpG site's

683   methylation profile can be traced from a shallower sample (x-axis) to a deeper sample (y-axis). A

684   greater number of sites remain in highly methylated states from 3-6 cm to 12-15 cm. These same

685   CpG sites experience a general trend of methylation loss from the mid sample to the deepest

686   sample. An increased number of CpG sites with methylation loads ~80% at 12-15 cm undergo

687   methylation losses ranging from 20-60% upon transitioning to 24-27 cm.

688

689   **Figure 6**: Chitinase CpG methylation dynamics. Quantifiable CpG sites were recovered for six

690   contigs mapped to four orders. Connected points are representative of changes in the methylation

691   state of a single CpG site for each sample. Numbers denote contigs to which these CpG sites are

692   mapped. Multiple CpG sites exhibiting differential methylation within the same gene were

693   recovered for the Actinobacteria and Thermoanaerobacterales. The recovered CpG site mapped

694   to Xanthomonadales was shown to persist in a non-methylated state from 3-6 cm to 12-15 cm,

695   but exhibited an apparent *de novo* methylation event from 12-15 cm to 24-27 cm.

696

697   **Figure 7**: CpG site methylation shifts for transposases mapped to Actinobacteria (**A**),

698   Alphaproteobacteria (**B**), Bacilli (**C**), and Gammaproteobacteria (**D**). Interpretation guidelines

699   are the same as those in Figure 6. CpG sites were shown to exist in differential methylation states

700   across depths. The methylation states of multiple CpG sites located within the same contig were

701   shown to shift to highly similar states in several instances.

702

703   Supplemental Information includes figures and tables.

Figure 1: Taxonomic relative abundances >= 0.05% of metagenome contigs with both Kraken and higher-confidence PhymmBL class assignments present in two or more samples. PhymmBL annotations were paired with Kraken annotations to select against potential false positives annotated by PhymmBL. These annotations suggest a presence of Actinomycetales at all depths, and an increased abundance of anaerobic classes (Clostridia, Dehalococcoidia). Methanogenic archaea (Methanosarcinales) are present in the 12-15 cm and 24-27 cm samples.
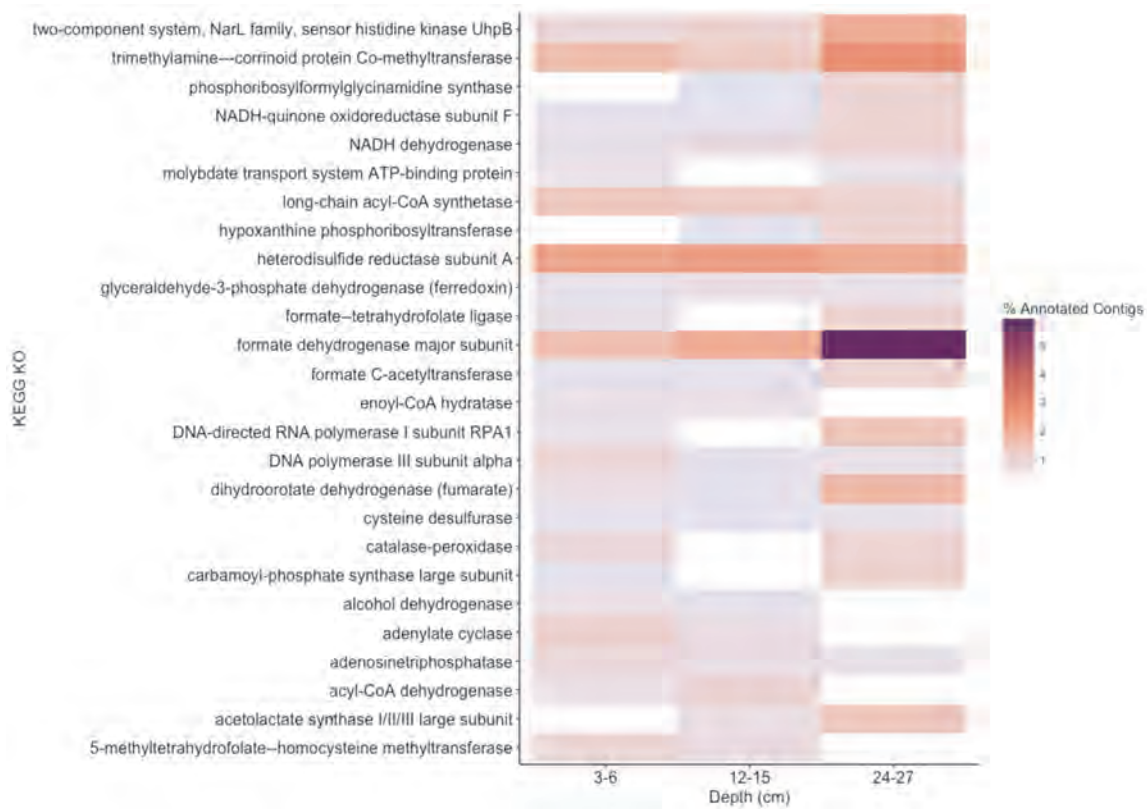
Figure 2: Abundance of KEGG Orthology functional assignments across sample depths. KEGG Orthology annotation results show that Oyster Rocks sediment communities at 12-15 cm and 24-27 cm have higher enzymatic potential for anaerobic metabolism.
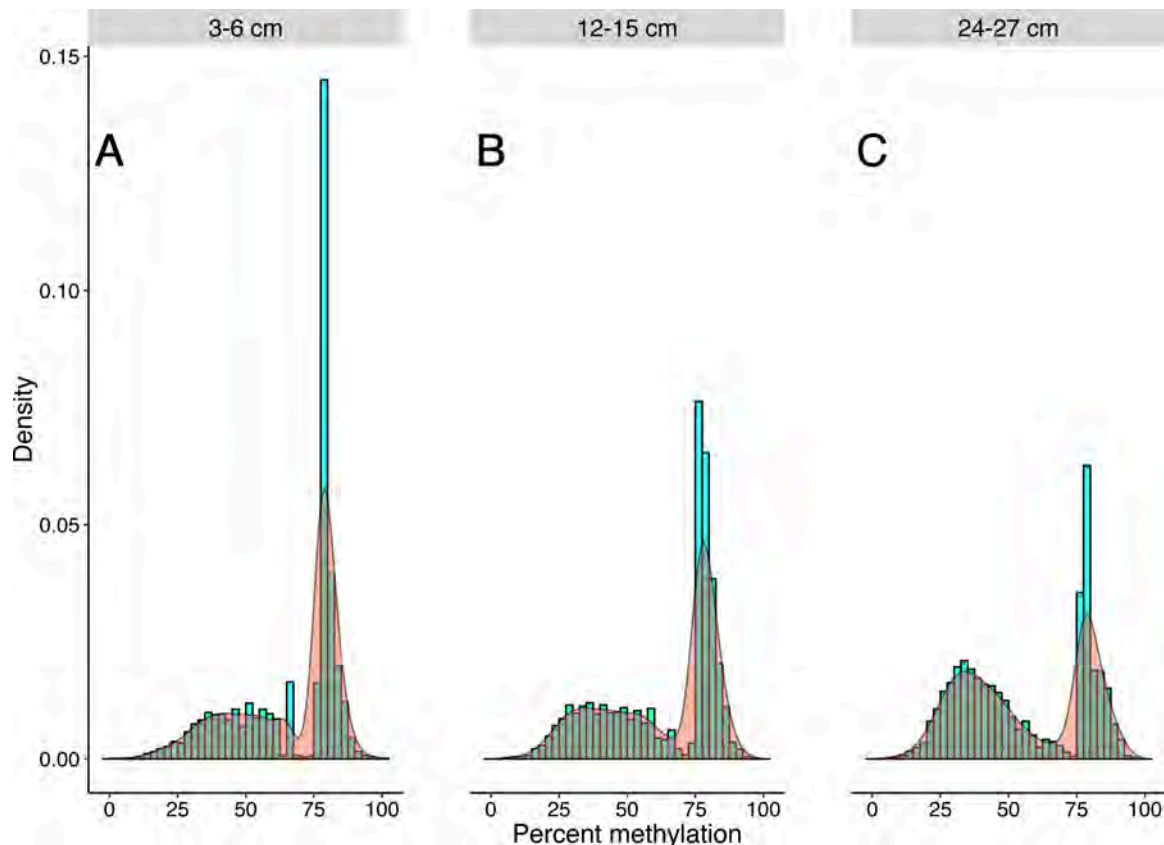
Figure 3: Community methylation load distributions. Histograms and kernel density overlays represent overall methylation levels of CpG sites that are shared across all three samples. A greater number of recovered CpG sites were highly methylated at 3-6 cm (**A**), resulting in a greater methylation load at this depth. However, a trend of decreasing overall methylation was seen at 12-15 cm (**B**) and 24-27 cm (**C**), with more sites experiencing transitions from highly methylated to non-methylated or hemimethylated states, or persisting in non-methylated states.
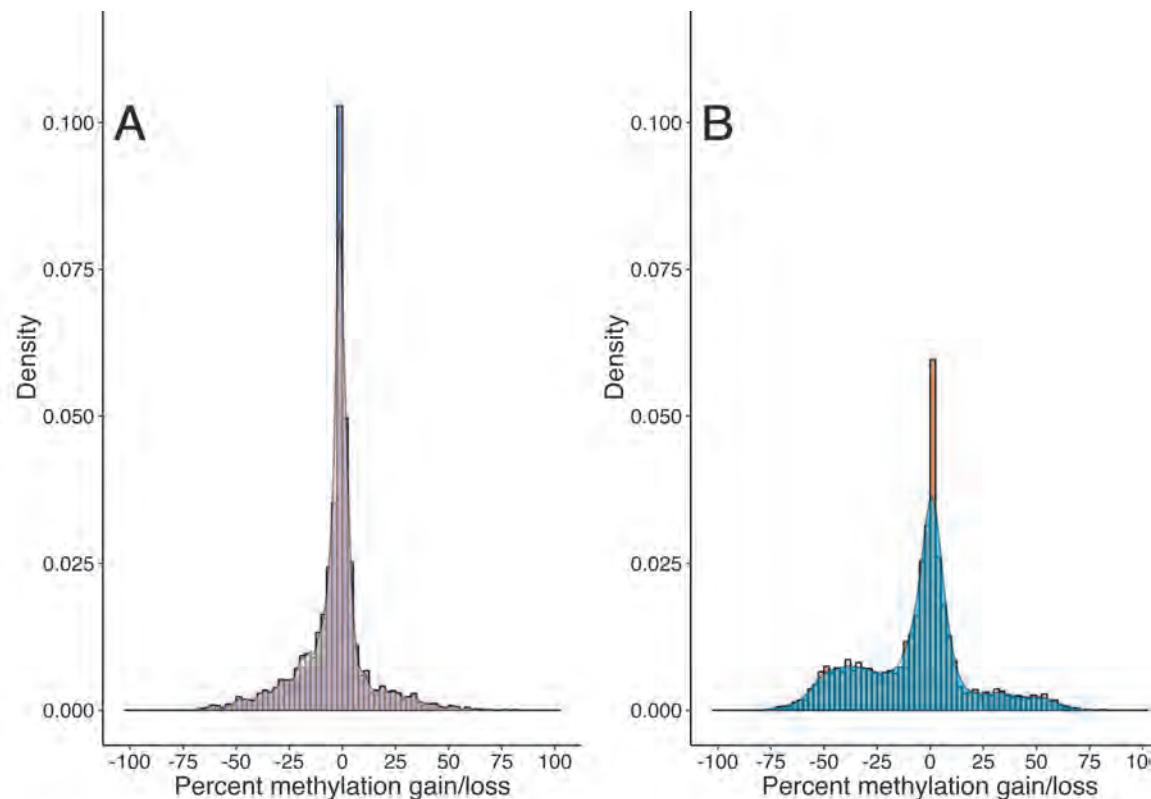
Figure 4: CpG methylation gains and losses from (**A**) 3-6 cm to 12-15 cm and (**B**) 12-15 cm to 24-27 cm. Histograms and kernel density overlays are representative of the densities of methylation shifts for individual CpG sites. Sites represented in (**A**) are the same sites represented in (**B**). Shifts range from -100 (total methylation loss) to +100 (total methylation gain). A significant number of CpG sites remained at equivalent methylation states from 3-6 cm to 12-15 cm, yet there is an apparent increase in methylation losses ranging from ~25% to ~50% from 12-15 cm to 24-27 cm.
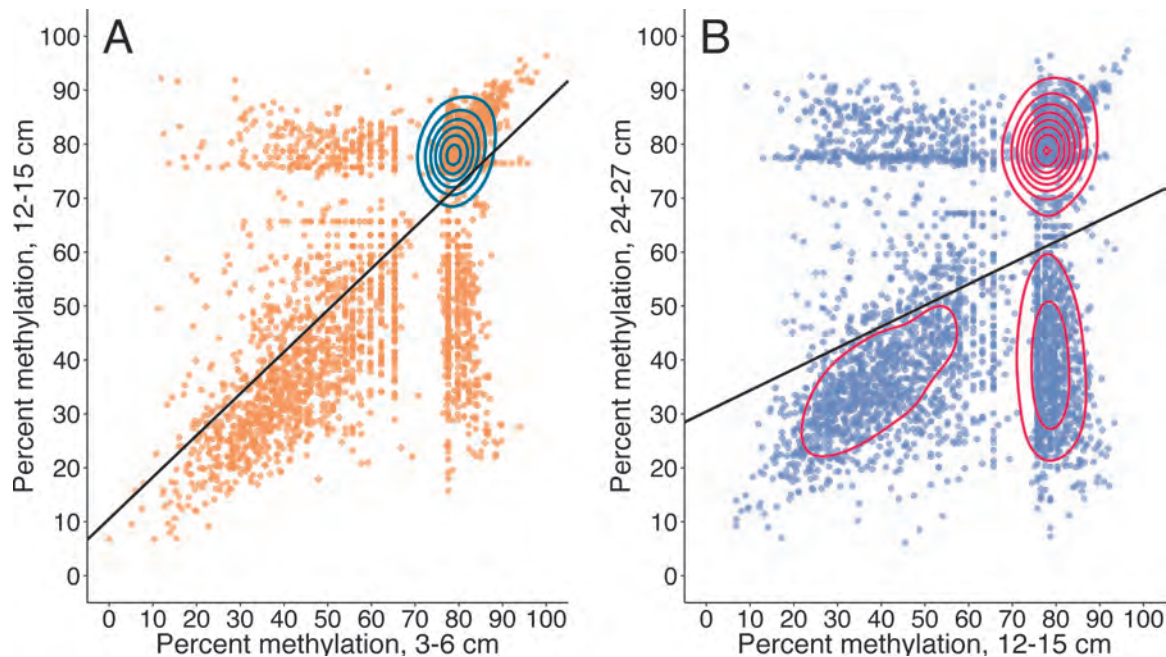
Figure 5: Total CpG methylation shifts from 3-6 cm to 12-15 cm (**A**) and 12-15 cm

to 24-27 cm (**B**). Plots are representative of metagenome-wide methylation loads.

Each point is a recovered CpG site whose quantified methylation states are

comparable across all three samples. The CpG sites represented in (**A**) are the same

as those represented in (**B**). Changes in a CpG site's methylation profile can be

traced from a shallower sample (x-axis) to a deeper sample (y-axis). A greater

number of sites remain in highly methylated states from 3-6 cm to 12-15 cm. These

same CpG sites experience a general trend of methylation loss from the mid sample

to the deepest sample. An increased number of CpG sites with methylation loads

~80% at 12-15 cm undergo methylation losses ranging from 20-60% upon
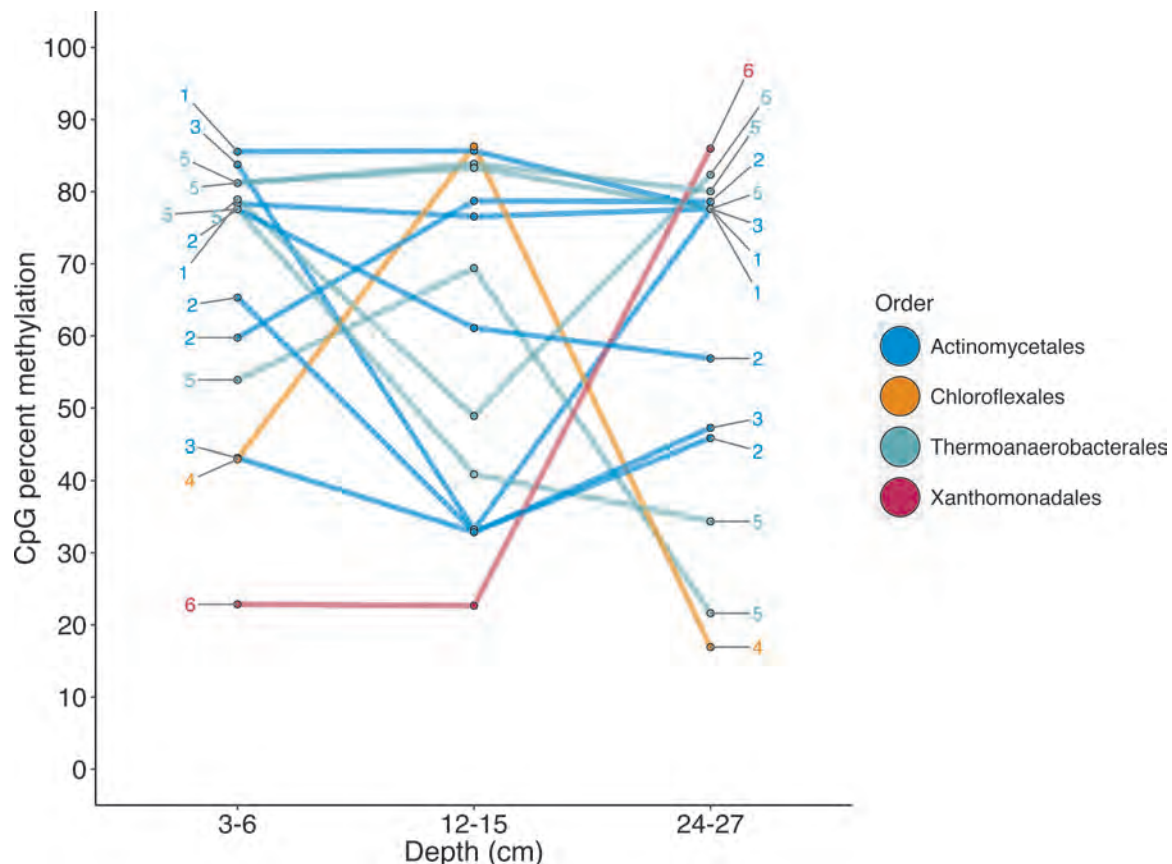
transitioning to 24-27 cm.

Figure 6: Chitinase CpG methylation dynamics. Quantifiable CpG sites were recovered for six contigs mapped to four orders. Connected points are representative of changes in the methylation state of a single CpG site for each sample. Numbers denote contigs to which these CpG sites are mapped. Multiple CpG sites exhibiting differential methylation within the same gene were recovered for the Actinobacteria and Thermoanaerobacterales. The recovered CpG site mapped to Xanthomonadales was shown to persist in a non-methylated state from 3-6 cm to 12-15 cm, but exhibited an apparent *de novo* methylation event from 12-15 cm to 24-27 cm.
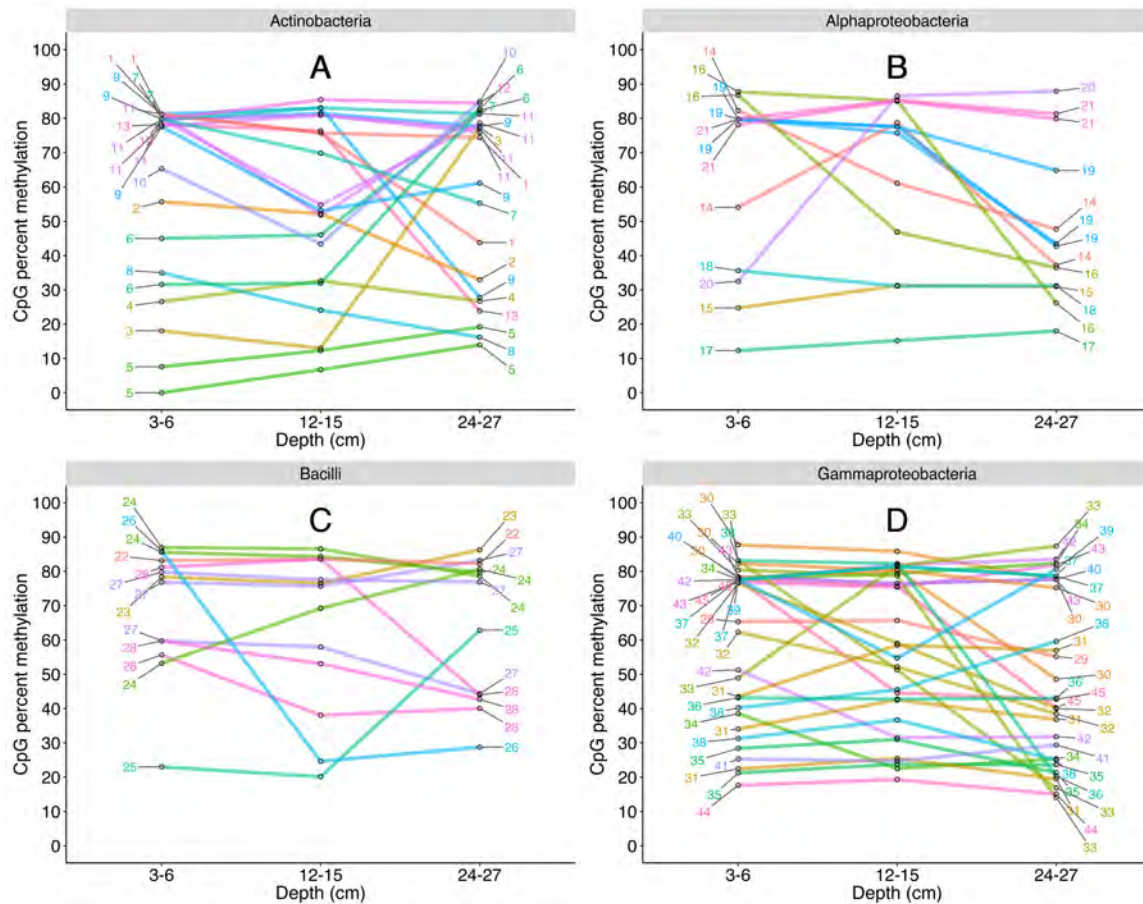
Figure 7: CpG site methylation shifts for transposases mapped to Actinobacteria

(**A**), Alphaproteobacteria (**B**), Bacilli (**C**), and Gammaproteobacteria (**D**).

Interpretation guidelines are the same as those in Figure 6. CpG sites were shown to

exist in differential methylation states across depths. The methylation states of

multiple CpG sites located within the same contig were shown to shift to highly

similar states in several instances.