




Bacterial Epigenomics: Coming of Age

 Pedro H. Oliveira^a

^aGénomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Université Évry, Université Paris-Saclay, Évry, France

ABSTRACT Epigenetic DNA methylation in bacteria has been traditionally studied in the context of antiparasitic defense and as part of the innate immune discrimination between self and nonself DNA. However, sequencing advances that allow genome-wide analysis of DNA methylation at the single-base resolution are nowadays expanding and have propelled a modern epigenomic revolution in our understanding of the extent, evolution, and physiological relevance of methylation. Indeed, as the number of mapped bacterial methylomes recently surpassed 4,000, increasing evidence supports roles for methylation in gene expression regulation, virulence, and host colonization, among others. In this paper, I summarize lessons taken from high-dimensional methylome data analyses and recent efforts that we and others are developing to leverage such findings into meaningful biological insights and overarching frameworks. Ultimately, I highlight anticipated research avenues and technological developments likely to unfold in the coming years.

KEYWORDS methylation, metaepigenomics, holoeigenomics, antimicrobial, single cell, epigenetics

Epigenomics refers to the systematic analysis of heritable, yet reversible, molecular modifications to both DNA and chromatin, the most extensively studied of which is DNA methylation. In bacteria, three major forms of DNA methylation have been detected: N6-methyladenine (6mA, the most abundant), N4-methylcytosine (4mC), and 5-methylcytosine (5mC) (1). These marks are mediated by DNA methyltransferases (MTases) associated with restriction-modification (R-M) systems or by orphan MTases (lacking a cognate restriction endonuclease [REase]) (2, 3). Propelled by recent progresses in third-generation sequencing technologies—single-molecule real-time sequencing (SMRT-seq) by Pacific Biosciences and nanopore sequencing by Oxford Nanopore Technologies—more than 4,000 methylomes have been mapped to date (4, 5). As a consequence, the field of bacterial epigenomics is witnessing a remarkable expansion beyond single methylome analyses to the realm of multi-omic data integration. As an example, we recently performed a large-scale DNA methylome and transcriptome analysis in the key human pathogen *Clostridioides difficile* and found a conserved orphan MTase whose inactivation impacted fundamental phenotypes involved in its transmission to a host (6). Such findings add to the growing number of studies integrating multi-omics profiling to identify putative epigenetic regulation networks. Fueled by this exciting momentum, my laboratory combines high-throughput (epi)genomic technologies and bioinformatic approaches to address outstanding questions put forward in the bacterial epigenomics field. What phenotypes are impacted by DNA MTases? Can we develop novel antimicrobial strategies by harnessing methylation systems? What is the interplay between stress adaptation and the stable inheritance of certain DNA methylation marks? In this paper, I describe my vision on how these lines of study will unfold and call out the challenges ahead.


THE UNDERRATED BACTERIAL 5mC METHYLOME

Bisulfite sequencing has traditionally been regarded as the gold standard approach enabling genome-wide 5mC mapping. This particularly holds true for eukaryotes, where 5mC is the most common DNA modification and is associated with a variety of

Citation Oliveira PH. 2021. Bacterial epigenomics: coming of age. *mSystems* 6:e00747-21. <https://doi.org/10.1128/mSystems.00747-21>.

Copyright © 2021 Oliveira. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to pcoutool@genoscope.cns.fr.

 A newly published manuscript authored by @pholive81 from @seqlab Genoscope highlights key research avenues likely to unfold in the coming years in the emergent field of bacterial epigenomics, and shares his vision on how his lab might give a helping hand.

Conflict of Interest Disclosures: P.H.O. has nothing to disclose.

The views expressed in this article do not necessarily reflect the views of the journal or of ASM.

This article is part of a special series sponsored by Floré.

Published 17 August 2021

biological phenomena such as gene silencing, genomic imprinting, X chromosome inactivation, RNA splicing, and silencing of transposable elements (7). Perhaps due to the lower genomic predominance and anticipated minor role of 5mC in bacterial gene regulation, the use of bisulfite sequencing has been rather limited in bacteria (8–10). Moreover, 5mC detection by third-generation sequencing technologies has been hampered either by the need for very high sequence coverage/ten-eleven translocation dioxygenase hypermodification in SMRT-seq (11) or by the limited availability of methods to perform *de novo* fine mapping of methylation type and recognition motif for nanopore sequencing (12). Notwithstanding, 5mC methylation is emerging as an important mechanism in bacterial epigenetics, as recent studies have thrown light on previously underappreciated roles in virulence and host adaptation. Some notorious examples include the control of bacterial cell shape, adherence to host cells, natural competence for DNA uptake, and envelope formation (9, 13). Perhaps more surprising was the recent finding of a substantial number of highly conserved 5mC bacterial MTases for which little is known regarding the underlying epigenetic mechanisms regulating cellular phenotypes (14). Another outstanding question concerns the origin and evolution of DNA methylation across the tree of life. In the case of 5mC, despite the different target sequence contexts in which it takes place, methylation is established and maintained by a family of DNA MTases that share a catalytic domain containing 10 conserved small motifs, suggesting a common origin (15). Also, the impact of 5mC on the intrinsic structure and mechanical properties of DNA (reduced flexibility and widening of the major groove) is expected to be consistent across different organisms, which may nevertheless take advantage of such conformational changes under very distinct genetic conditions. The extent to which such DNA structural changes are a function of sequence context and how they impact recognition by DNA-binding proteins are still unclear. Hence, understanding the full significance of 5mC methylation, its functional consequences, and evolution remains an exciting challenge for the future.

HARNESSING METHYLATION SYSTEMS AS AN ANTIMICROBIAL STRATEGY

Orthodox Type IIP R-M systems are composed of one homodimeric or homotetrameric REase and one monomeric MTase and are able to operate separately and independently from each other. Such a feature allows these systems to behave as toxin-antitoxin addiction modules and facilitate programmed cell death by postsegregational killing (16). Given these observations, it is reasonable to ask whether bacterial R-M systems could be exploited for clinical purposes. In particular, could we envisage an antimicrobial chemotherapeutic strategy based on molecules that selectively interfere with the R-M balance through binding to the MTase or by enhancing its rate of proteolysis? Such a strategy would result in the loss of protection provided by epigenetic methylation, followed by cleavage of chromosomal DNA by the cognate REase and ultimately cell death. One downside of this approach is that R-M genes are frequently exchanged between bacteria by horizontal gene transfer and evolve very quickly, making it more likely to be used as a narrow-spectrum therapy against a particular species or emergent strain where the targeted R-M system would be significantly conserved and expressed. An alternative strategy would be to target orphan MTases that are conserved at a given taxonomic rank (e.g., species level). The former are typically encoded by core/quasicore genes and frequently found to be conditionally essential. One interesting possibility for MTase inhibition would be to use analogs of the methyl donor *S*-adenosyl-L-methionine (SAM) or bisubstrate inhibitors that simultaneously target SAM- and substrate-binding sites (17). Such a scenario was recently proposed for the core MTase CamA of *C. difficile* (18). Another example is that of the well-characterized *Escherichia coli* Dam enzyme, whose inhibition reportedly weakens bacterial pathogenicity *in vivo* (19, 20). Dam methylation was also found to play a role in drug potentiation, by curbing the therapeutic activity of the β -lactam and quinolone classes of antibiotics (21). In this view, Dam represents an attractive

target for epigenetic inhibition of the multiple biological processes that it regulates (e.g., virulence), as it lacks mammalian homologs while being conserved in several enteric pathogens. While some selective inhibitors of Dam were previously proposed (22), there have been no further advances over the past decade. Since *camA* and *dam* are part of a much larger list of 145 genes recently reported in a study investigating highly conserved MTases in bacteria (14), I foresee a renewed interest in the exploitation of such targets for the development of next-generation epigenetic drugs (23–25).

META- AND HOLOEPIGENOMICS

It is estimated that more than 99% of the potentially 10^{11} to 10^{12} species that make up all microbial diversity on Earth remain unexplored to date and that only a small fraction can be culturable under standard laboratory conditions (26). Culture-independent techniques such as metagenomic sequencing have provided a greater depth of understanding of the biodiversity and functional capabilities of microbial communities. The introduction of third-generation sequencing technologies has substantially improved metagenome assemblies and holds the potential to change our understanding of the hidden diversity of methylomes across different ecological niches (Fig. 1). In two recent metaepigenomic studies performed in aquatic ecosystems, important advances were made both in the finding of previously undescribed target methylation sites and in the understanding of the coevolutionary history of methylation systems and host genome (27, 28). These findings add to recent bioinformatic developments exploring endogenous epigenetic barcodes as complements to coverage and composition features in order to improve strain-level resolution of metagenomes and link mobile genetic elements to their host genomes in microbial samples (29). It is expected that in the next few years, metaepigenomic analyses of bacteria from different ecological niches will significantly deepen our understanding on the evolution of methylation systems and on the impacts of DNA methylation in shaping the composition of such niches. More broadly, the systematic search for antiphage defense hot spots in metagenomic data sets is expected to uncover novel immune systems, with the potential to be adapted into useful molecular tools. While the metaepigenome encompasses the ensemble of epigenetic changes in a community within a nonliving environment, there is an increasing interest in studying the holoeptigenome, which by definition implies an epigenetic interaction between the host and its symbionts (the holobiont) (Fig. 1). Such interactions can affect key biological processes of both hosts and microorganisms and have the power to shape their coevolution. For example, dysbiosis and reduction of microbial diversity can change the proportion of metabolites acting as regulators of DNA and histone modifications in the host. Alternatively, the secretion or injection/translocation of nucleus-targeted effectors—termed nucleomodulins—from a bacterial pathogen into the host cytosol can subvert the host epigenome through interference with histone and DNA modifications, regulation of transcription, interference on the cell cycle, and regulation of cell signaling pathways. For example, the nucleomodulins Mhy from *Mycoplasma hyorhinis* and Rv2966c from *Mycobacterium tuberculosis* are capable of acting as mammalian DNA MTases and regulate proliferation-specific pathways (30). Hence, it is foreseeable that the next years will bring additional research on nucleomodulin diversity in bacterial pathogens and a better understanding of the mechanisms used for nuclear trafficking and modulation of the host genome.

VIEW FROM A SINGLE-CELL PERSPECTIVE

Exposure of clonal bacterial populations to environmental changes, stress, and other stimuli results in methylome alterations that modulate global gene expression patterns. Such nongenetic diversification can in turn lead to the emergence of phenotypically heterogeneous subpopulations, in which some persister cells have a better ability to withstand the change. In recent years, important technical advancements in single-cell isolation, whole-genome/transcriptome amplification, and high-throughput sequencing are paving the way for resolving cell-to-cell multi-omic heterogeneity at unprecedented resolution. However, the development of efficient high-throughput

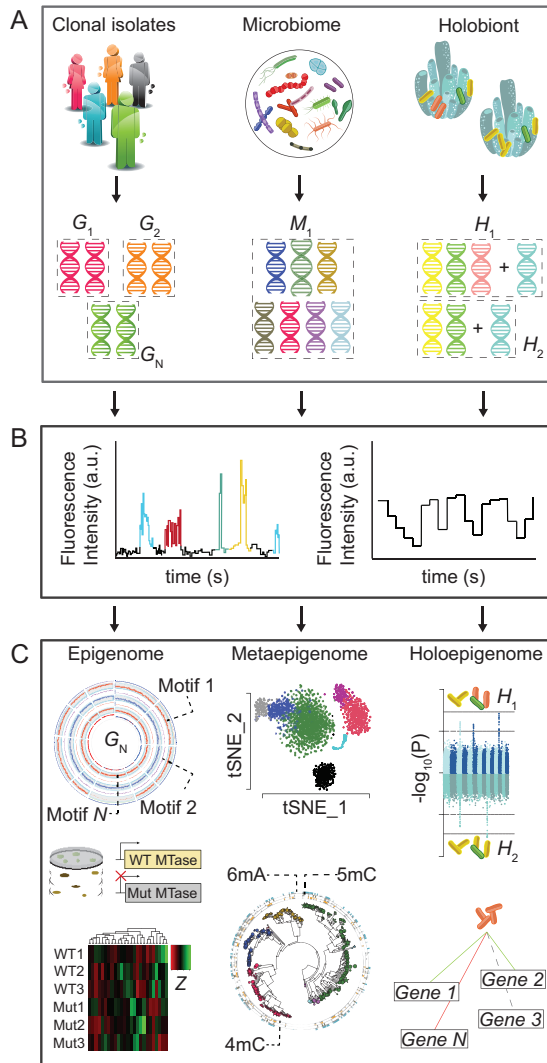


FIG 1 Approaches to study bacterial methylomes from clonal isolates, microbiomes, and holobionts. (A) Although a large abundance of methylomes profiled to date belongs to genome (*G*) isolates, there is a growing interest in the analysis of microbiome (*M*) and holobiont (*H*) methylomes. (B) Recent progress in third-generation sequencing technologies (e.g., SMRT-seq and nanopore sequencing) has enabled direct genome-wide detection of methylated positions and target motifs. (C) Relevant functional information on the epigenome can be obtained by targeted mutagenesis of DNA MTases. A comprehensive global transcriptome and functional profiling by RNA-seq offers the opportunity to further dissect the range of differentially expressed genes in a methylation-free strain. For metaepigenomes, nonlinear dimensionality reduction algorithms such as the t-distributed stochastic neighbor embedding (t-SNE) are a possible option to visualize and interpret methylation features across multiple metagenomic contigs. A phylogenetic representation of methylation systems' density across several metagenome-assembled genomes may also provide clues into the interplay between DNA methylation and factors unique to the environment of each community. In holoepigenomes, genome-wide analysis of CpG site methylation differences between multiple hosts (as shown in the Manhattan plots of *P* values) may provide insight into the network of host genes whose expression is being significantly modulated by the presence of certain symbionts.

single-cell solutions for microbial systems has lagged behind those for eukaryotes (31), mainly due to their low DNA/mRNA content, difficult lysis/permeabilization of cell walls and membranes, and lack of polyadenylation of bacterial mRNA, which limits its separation from rRNA. Since long-read sequencing technologies require a relatively large amount of starting genomic DNA for library preparation, it is conceptually challenging to perform single-cell analysis. Moreover, such technologies rely on a consensus sequence obtained from a cell population and lack the resolution required to perceive epigenetic heterogeneity. In this sense, a recently proposed bioinformatics tool allows performing single-molecule characterization of epigenetic heterogeneity in bacterial

methylomes using SMRT-seq (32). One interesting research avenue that would greatly benefit from an in-depth bacterial methylome tracking at the single-cell level is the one dealing with genetic assimilation. The latter essentially assumes that a stress-induced nongenetic change in phenotype can, during the course of selection and over multiple generations, become genetically encoded. This necessarily raises a few outstanding questions: (i) is this genetic assimilation aimed at maintaining stress-related epigenetic landscapes? (ii) are the observable changes in gene expression directly modulated by the acquisition of a particular subset of DNA methylation marks? Although recent studies have begun to provide insight into this topic (33), we will need to wait for further advances in long-read technologies applied to single-cell sequencing, in order to identify the missing pieces of what appears to be not only a complex puzzle of epigenetic-mediated persistence but also a promising gateway for the development of novel antibacterial drugs.

ACKNOWLEDGMENTS

I am grateful to my former mentors Eduardo P. C. Rocha (Pasteur Institute) and Gang Fang (Mount Sinai School of Medicine) for their support and encouragement and for many fruitful discussions that helped shape my research vision. I thank Patrick Wincker (Genoscope) for helping nurture the research activity in my lab, my students for all their dedication, and Jack Gilbert for the opportunity to participate in this special series.

REFERENCES

- Sánchez-Romero MA, Casadesús J. 2020. The bacterial epigenome. *Nat Rev Microbiol* 18:7–20. <https://doi.org/10.1038/s41579-019-0286-2>.
- Oliveira PH, Touchon M, Rocha EP. 2014. The interplay of restriction-modification systems with mobile genetic elements and their prokaryotic hosts. *Nucleic Acids Res* 42:10618–10631. <https://doi.org/10.1093/nar/gku734>.
- Mouammine A, Collier J. 2018. The impact of DNA methylation in Alphaproteobacteria. *Mol Microbiol* 110:1–10. <https://doi.org/10.1111/mmi.14079>.
- Roberts RJ, Vincze T, Janos P, Macelis D. 2015. REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res* 43:D298–D299. <https://doi.org/10.1093/nar/gku1046>.
- Beaulaurier J, Schadt EE, Fang G. 2019. Deciphering bacterial epigenomes using modern sequencing technologies. *Nat Rev Genet* 20:157–172. <https://doi.org/10.1038/s41576-018-0081-3>.
- Oliveira PH, Ribis JW, Garrett EM, Trzilova D, Kim A, Sekulovic O, Mead EA, Pak T, Zhu S, Deikus G, Touchon M, Lewis-Sandari M, Beckford C, Zeitouni NE, Altman DR, Webster E, Oussenko I, Bunyavanich S, Aggarwal AK, Bashir A, Patel G, Wallach F, Hamula C, Huprikar S, Schadt EE, Sebra R, van Bakel H, Kasarskis A, Tamayo R, Shen A, Fang G. 2020. Epigenomic characterization of *Clostridioides difficile* finds a conserved DNA methyltransferase that mediates sporulation and pathogenesis. *Nat Microbiol* 5:166–180. <https://doi.org/10.1038/s41564-019-0613-4>.
- Breiling A, Lyko F. 2015. Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond. *Epigenetics Chromatin* 8:24. <https://doi.org/10.1186/s13072-015-0016-6>.
- Vandenbussche I, Sass A, Van Nieuwerburgh F, Pinto-Carbó M, Mannweiler O, Eber L, Coenye T. 2021. Detection of cytosine methylation in *Burkholderia cenocepacia* by single-molecule real-time sequencing and whole-genome bisulfite sequencing. *Microbiology* 167(3). <https://doi.org/10.1099/mic.0.001027>.
- Chao MC, Zhu S, Kimura S, Davis BM, Schadt EE, Fang G, Waldor MK. 2015. A cytosine methyltransferase modulates the cell envelope stress response in the cholera pathogen. *PLoS Genet* 11:e1005739. <https://doi.org/10.1371/journal.pgen.1005739>.
- Kahramanoglou C, Prieto A, Khedkar S, Haase B, Gupta A, Benes V, Fraser GM, Luscombe NM, Seshasayee ASN. 2012. Genomics of DNA cytosine methylation in *Escherichia coli* reveals its role in stationary phase transcription. *Nat Commun* 3:886. <https://doi.org/10.1038/ncomms1878>.
- Blow MJ, Clark TA, Daum CG, Deutschbauer AM, Fomenkov A, Fries R, Froula J, Kang DD, Malmstrom RR, Morgan RD, Posfai J, Singh K, Visel A, Wetmore K, Zhao Z, Rubin EM, Korchak J, Pennacchio LA, Roberts RJ. 2016. The epigenomic landscape of prokaryotes. *PLoS Genet* 12:e1006064. <https://doi.org/10.1371/journal.pgen.1005854>.
- Tourancheau A, Mead EA, Zhang XS, Fang G. 2021. Discovering multiple types of DNA methylation from bacteria and microbiome using nanopore sequencing. *Nat Methods* 18:491–498. <https://doi.org/10.1038/s41592-021-01109-3>.
- Estibariz I, Overmann A, Ailloud F, Krebs J, Josenhans C, Suerbaum S. 2019. The core genome m5C methyltransferase JHP1050 (M.Hpy99III) plays an important role in orchestrating gene expression in *Helicobacter pylori*. *Nucleic Acids Res* 47:2336–2348. <https://doi.org/10.1093/nar/gky1307>.
- Oliveira PH, Fang G. 2021. Conserved DNA methyltransferases: a window into fundamental mechanisms of epigenetic regulation in Bacteria. *Trends Microbiol* 29:28–40. <https://doi.org/10.1016/j.tim.2020.04.007>.
- Bhattacharyya M, De S, Chakrabarti S. 2021. Origin and evolution of DNA methyltransferases (DNMT) along the tree of life: a multi-genome survey. *bioRxiv* <https://doi.org/10.1101/2020.04.09.033167>.
- Mruk I, Kobayashi I. 2014. To be or not to be: regulation of restriction-modification systems and other toxin-antitoxin systems. *Nucleic Acids Res* 42:70–86. <https://doi.org/10.1093/nar/gkt711>.
- Zhang J, Zheng YG. 2016. SAM/SAH analogs as versatile tools for SAM-dependent methyltransferases. *ACS Chem Biol* 11:583–597. <https://doi.org/10.1021/acscchembio.5b00812>.
- Zhou J, Horton JR, Blumenthal RM, Zhang X, Cheng X. 2021. *Clostridioides difficile* specific DNA adenine methyltransferase CamA squeezes and flips adenine out of DNA helix. *Nat Commun* 12:3436. <https://doi.org/10.1038/s41467-021-23693-w>.
- García-Del Portillo F, Pucciarelli MG, Casadesús J. 1999. DNA adenine methylase mutants of *Salmonella typhimurium* show defects in protein secretion, cell invasion, and M cell cytotoxicity. *Proc Natl Acad Sci U S A* 96:11578–11583. <https://doi.org/10.1073/pnas.96.20.11578>.
- Watson ME, Jr, Jarisch J, Smith AL. 2004. Inactivation of deoxyadenosine methyltransferase (dam) attenuates *Haemophilus influenzae* virulence. *Mol Microbiol* 53:651–664. <https://doi.org/10.1111/j.1365-2958.2004.04140.x>.
- Cohen NR, Ross CA, Jain S, Shapiro RS, Gutierrez A, Belenky P, Li H, Collins JJ. 2016. A role for the bacterial GATC methylome in antibiotic stress survival. *Nat Genet* 48:581–586. <https://doi.org/10.1038/ng.3530>.
- Mashhoon N, Pruss C, Carroll M, Johnson PH, Reich NO. 2006. Selective inhibitors of bacterial DNA adenine methyltransferases. *J Biomol Screen* 11:497–510. <https://doi.org/10.1177/1087057106287933>.
- Ganesan A, Arimondo PB, Rots MG, Jeronimo C, Berdasco M. 2019. The timeline of epigenetic drug discovery: from reality to dreams. *Clin Epigenetics* 11:174. <https://doi.org/10.1186/s13148-019-0776-0>.
- Hupkes M, Azevedo R, Jansen H, van Zoelen EJ, Decherer KJ. 2013. Identification of novel bacterial M.SssI DNA methyltransferase inhibitors. *J Biomol Screen* 18:348–355. <https://doi.org/10.1177/1087057112465009>.
- Erdmann A, Halby L, Fahy J, Arimondo PB. 2015. Targeting DNA methylation with small molecules: what's next? *J Med Chem* 58:2569–2583. <https://doi.org/10.1021/jm500843d>.

26. Bodor A, Bounedjoum N, Vincze GE, Erdeiné Kis Á, Laczi K, Bende G, Szilágyi Á, Kovács T, Perei K, Rákhely G. 2020. Challenges of unculturable bacteria: environmental perspectives. *Rev Environ Sci Biotechnol* 19:1–22. <https://doi.org/10.1007/s11157-020-09522-4>.
27. Hiraoka S, Okazaki Y, Anda M, Toyoda A, Nakano S, Iwasaki W. 2019. Meta-epigenomic analysis reveals the unexplored diversity of DNA methylation in an environmental prokaryotic community. *Nat Commun* 10:159. <https://doi.org/10.1038/s41467-018-08103-y>.
28. Hiraoka S, Sumida T, Hirai M, Toyoda A, Kawagucci S, Yokokawa T, Nunoura T. 2021. Diverse DNA modification in marine prokaryotic and viral communities. *bioRxiv* <https://doi.org/10.1101/2021.05.08.442635>.
29. Beaulaurier J, Zhu S, Deikus G, Mogno I, Zhang XS, Davis- Richardson A, Canepa R, Triplett EW, Faith JJ, Sebra R, Schadt EE, Fang G. 2018. Metagenomic binning and association of plasmids with bacterial host genomes using DNA methylation. *Nat Biotechnol* 36:61–69. <https://doi.org/10.1038/nbt.4037>.
30. Hanford HE, Dwingelo JV, Kwaik YA. 2021. Bacterial nucleomodulins: a coevolutionary adaptation to the eukaryotic command center. *PLoS Pathog* 17:e1009184. <https://doi.org/10.1371/journal.ppat.1009184>.
31. Clark SJ, Smallwood SA, Lee HJ, Krueger F, Reik W, Kelsey G. 2017. Genome-wide base-resolution mapping of DNA methylation in single cells using single-cell bisulfite sequencing (scBS-seq). *Nat Protoc* 12:534–547. <https://doi.org/10.1038/nprot.2016.187>.
32. Beaulaurier J, Zhang XS, Zhu S, Sebra R, Rosenbluh C, Deikus G, Shen N, Munera D, Waldor MK, Chess A, Blaser MJ, Schadt EE, Fang G. 2015. Single molecule-level detection and long read-based phasing of epigenetic variations in bacterial methylomes. *Nat Commun* 6:7438. <https://doi.org/10.1038/ncomms8438>.
33. Walworth NG, Lee MD, Dolzhenko E, Fu FX, Smith AD, Webb EA, Hutchins DA. 2021. Long-term m5C methylome dynamics parallel phenotypic adaptation in the cyanobacterium *Trichodesmium*. *Mol Biol Evol* 38:927–939. <https://doi.org/10.1093/molbev/msaa256>.