Choice of DNA methylation analyses in invertebrates requires consideration of genome characteristics and methylation landscape

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

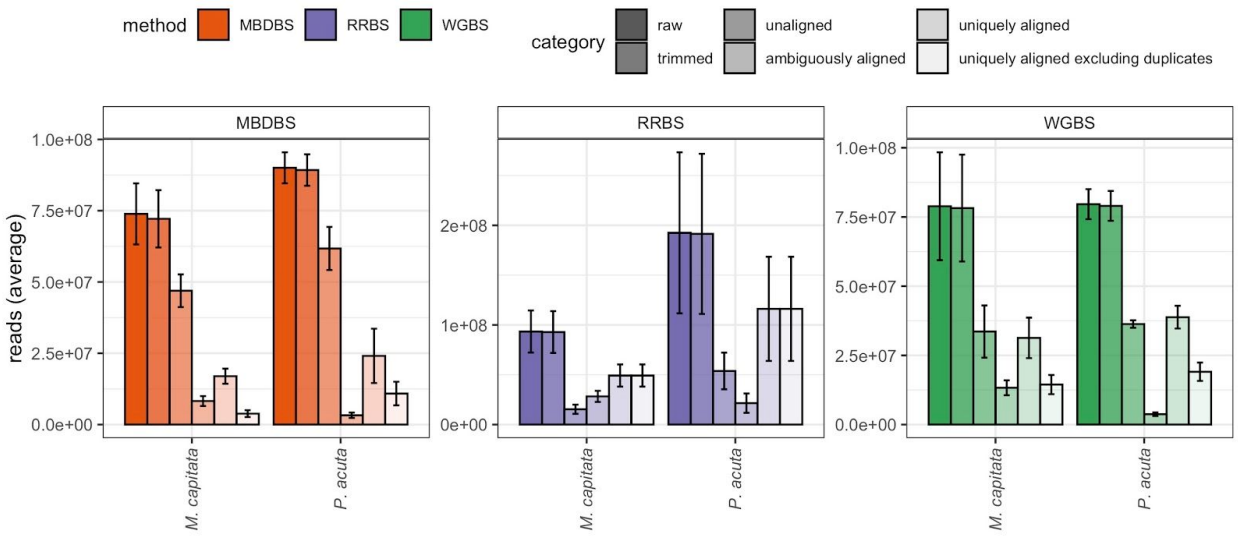
[1] University of Washington, School of Aquatic and Fishery Sciences 1122 NE Boat St. Seattle, WA, 98195, USA
[2] NOAA Northwest Fisheries Science Center Montlake 2725 Montlake Blvd E, Seattle, WA, 98112, USA
[3] Rutgers
[4] Department of Marine and Environmental Sciences, Northeastern University, 430 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]Department of Biological Sciences, University of Rhode Island, Kingston, RI 02881, USA

[+]corresponding author: hputnam@uri.edu
* equal contribution

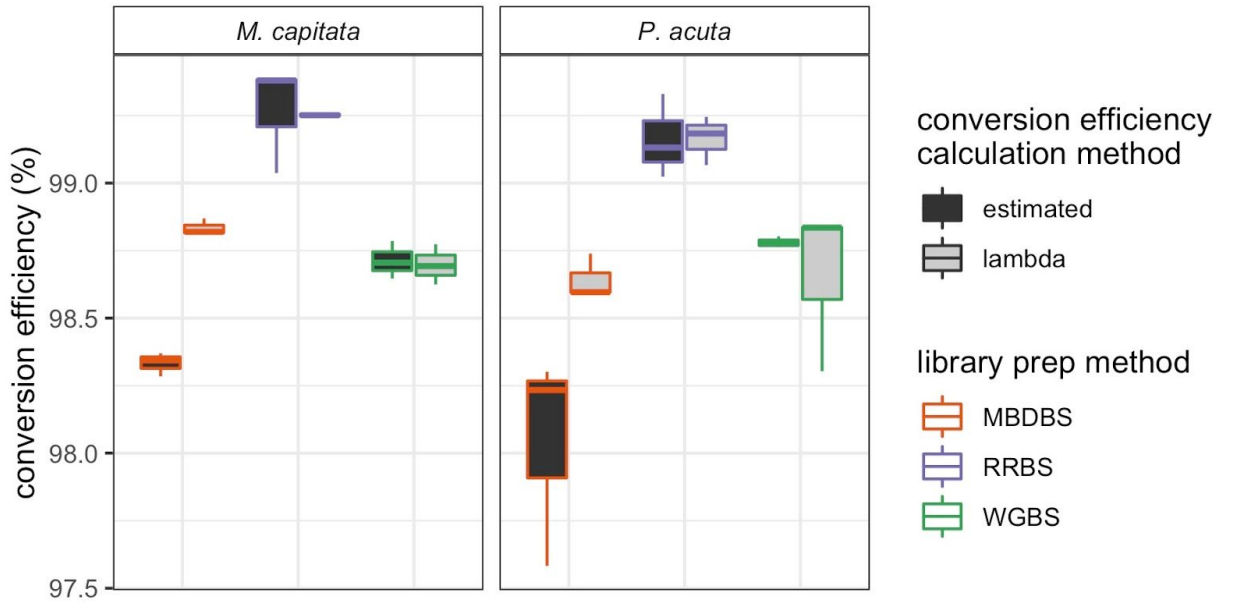
Table of Contents:

perMANOVA tests were used to investigate differences between two sequencing methods. Beta-dispersion models were only used with pairwise perMANOVA tests if the global beta-dispersion model indicated a significant effect of group variance (ANOVA *P*-value < 0.05) on the global perMANOVA result.

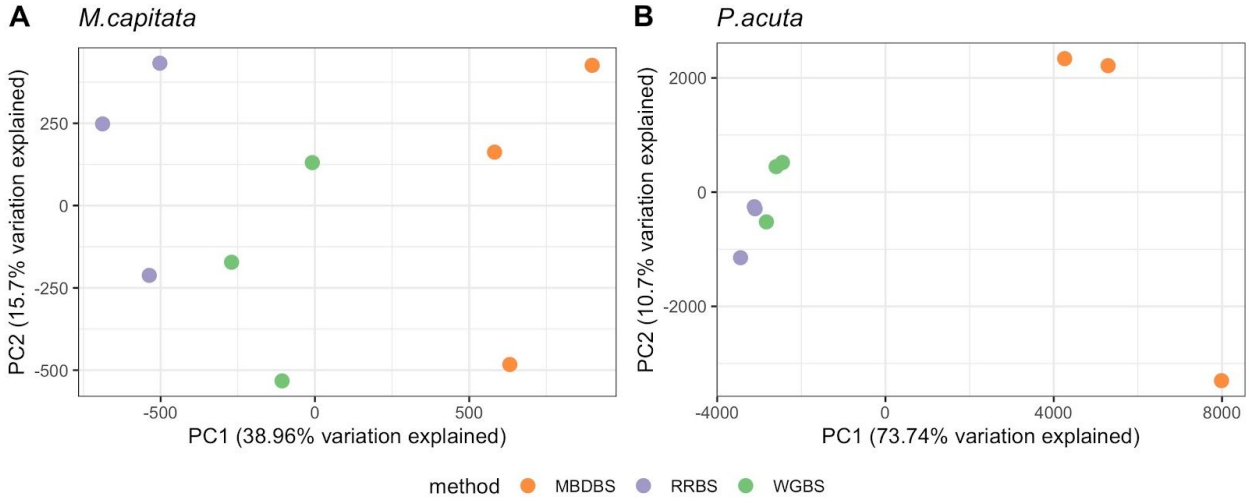
Supplementary Table 7: Contingency test results for *M. capitata* and *P. acuta*



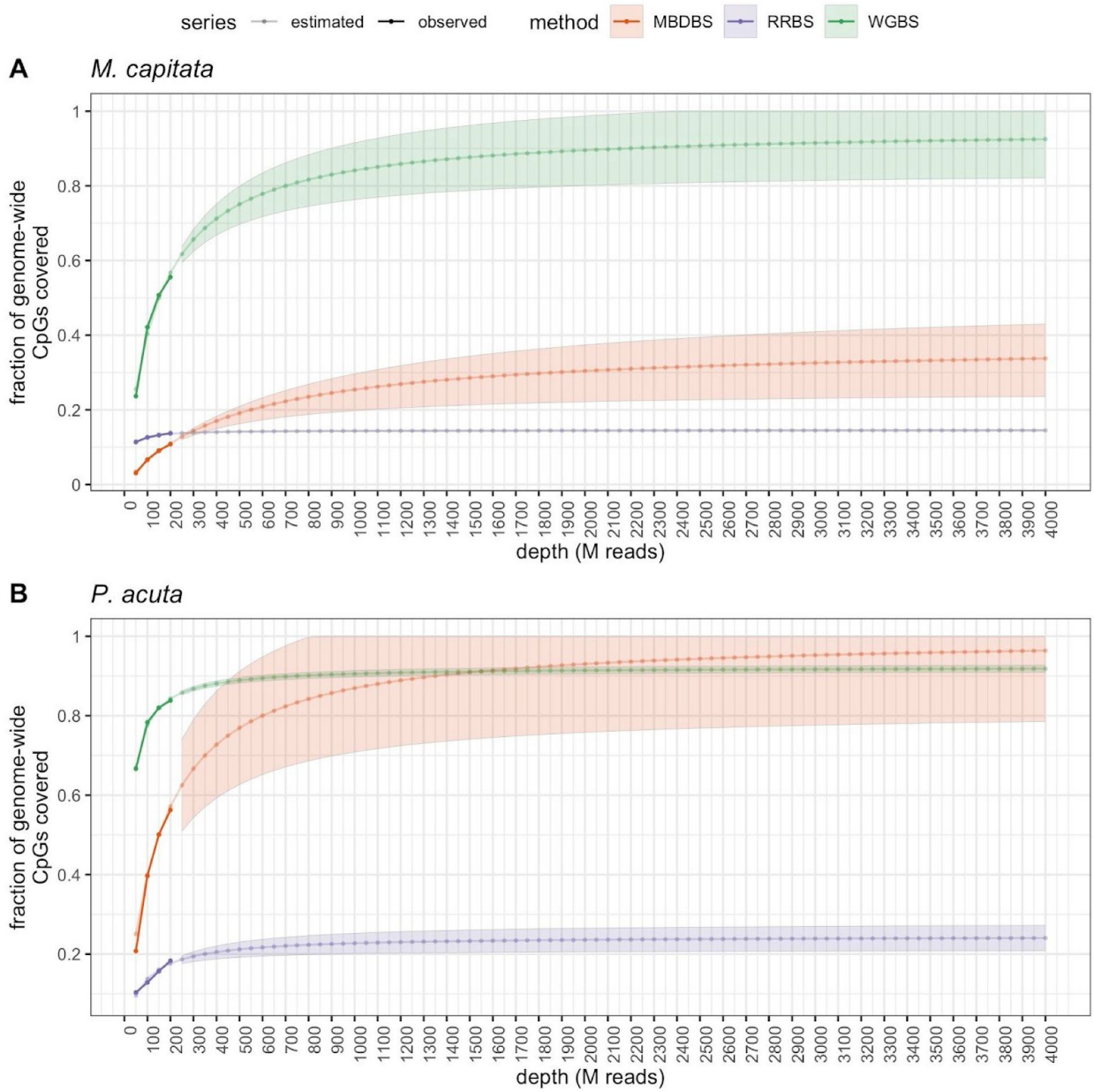
Supplementary Figure 1. Summary of sequencing depth and alignments for all libraries. Bars show average number of reads for each method and species and error bars show standard deviation.



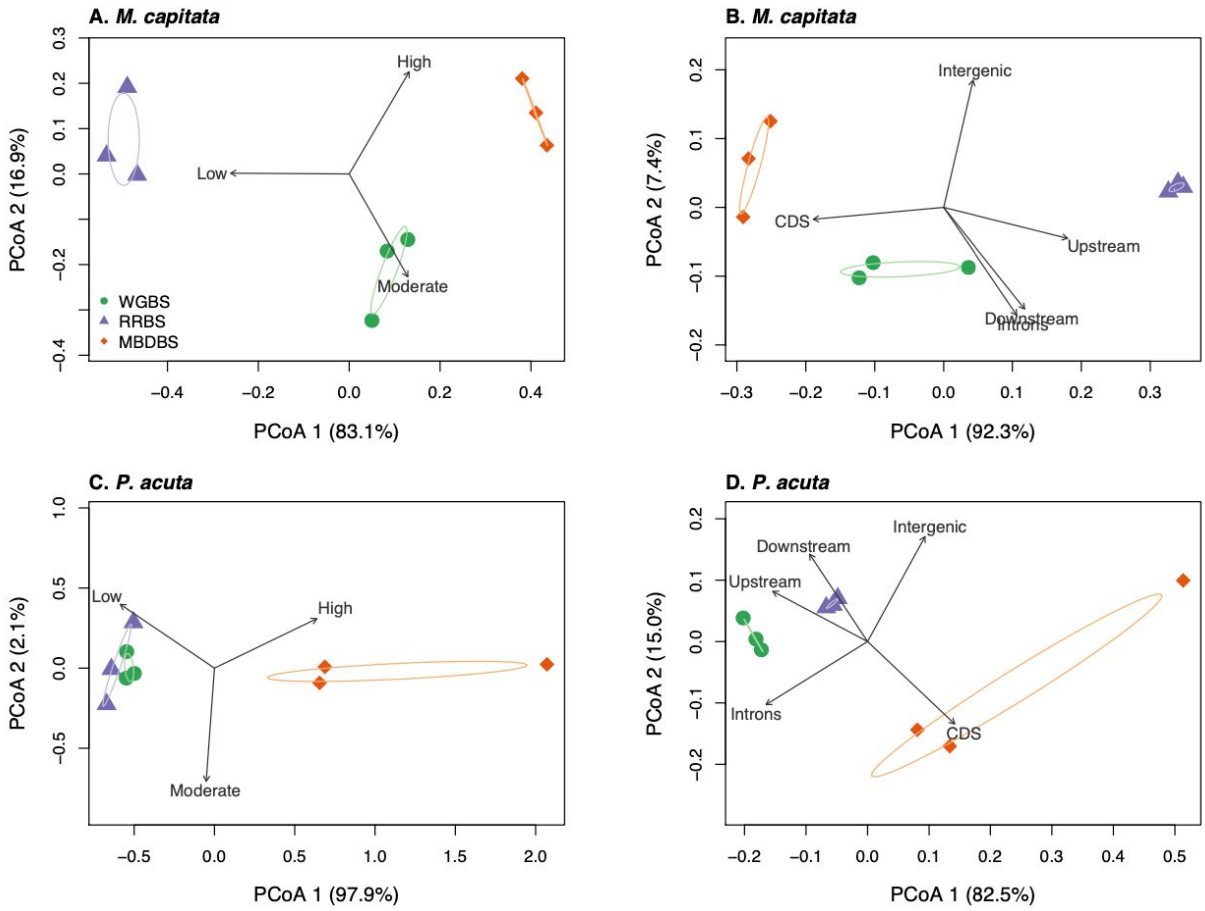
Supplementary Figure 2. Bisulfite conversion efficiency assessment. Bisulfite conversion efficiency is different across library preparation methods whether calculated from lambda alignments or estimated from non-CpG methylation from coral alignments for *M. capitata* libraries and *P. acuta* libraries. Generally, bisulfite conversion efficiency calculation methods are not different.



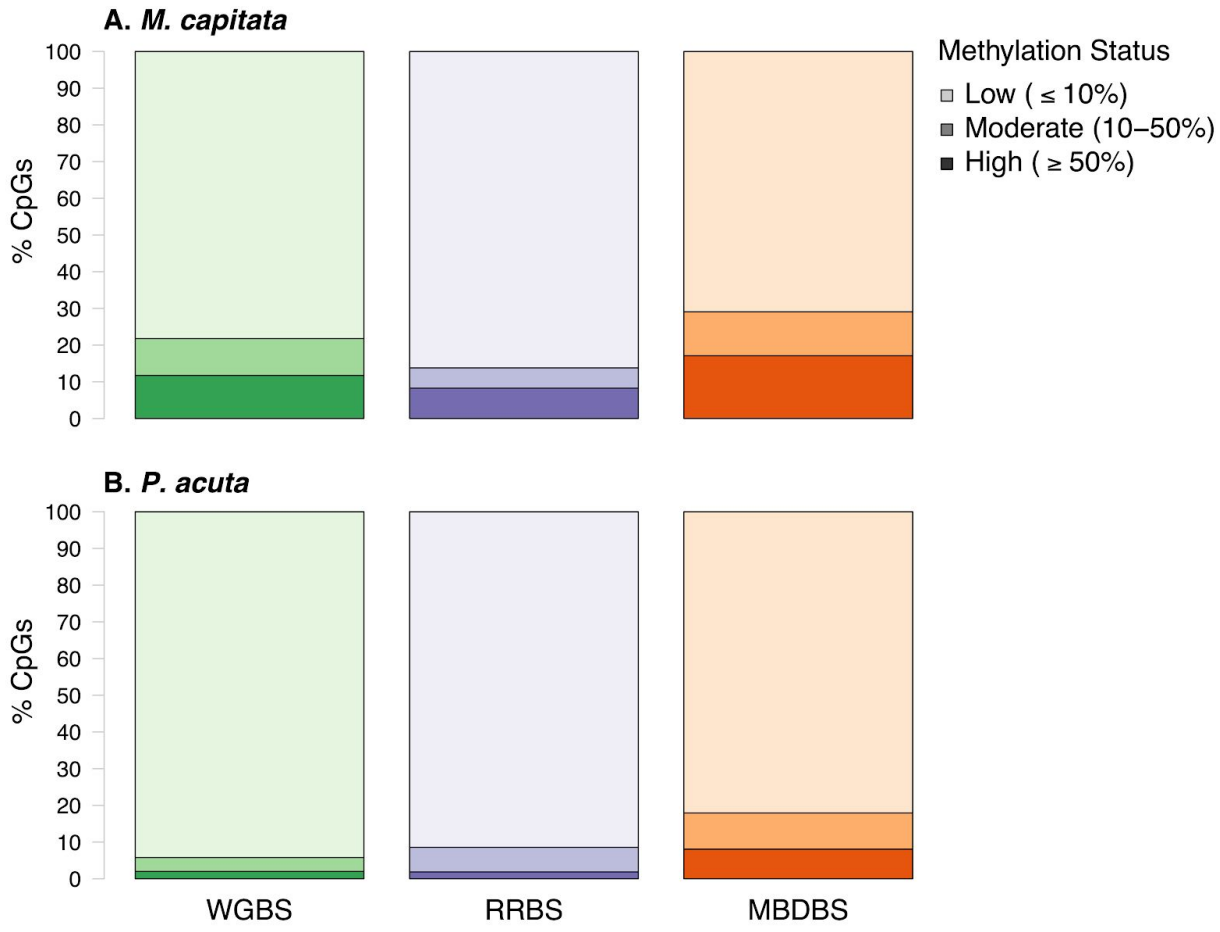
Supplementary Figure 3. PCA of CpG methylation for loci covered at 5x read depth in all samples for (**A**) *M. capitata* and (**B**) *P. acuta*.



Supplementary Figure 4. Estimated fraction of CpG sites in the genome covered by at least 5 reads at different sequencing depths (number of M read pairs) for (**A**) *M. capitata* and (**B**) *P. acuta* samples for each bisulfite sequencing method. 'Observed' (opaque line and dots) denotes the fraction of genome-wide CpG loci covered by at least 5 reads determined from pooled data that was subsampled at 50M, 100M, 150M, and 200M reads. 'Estimated' (translucent line and dots) denotes the fraction of genome-wide CpG loci covered by at least 5 reads estimated by michaelis-menten modelling of the 'observed' data with standard error shown by shaded areas. All samples within a bisulfite sequencing method were pooled for the downsampling analyses.



Supplementary Figure 5. Principal Coordinate Analyses associated with perMANOVA and beta-dispersion tests (see **Supplementary Table 6**). Differences in proportion of CpGs that are highly (≥ 50%), moderately (10-50%), or lowly (≤ 10%) methylated in **A**) *M. capitata* and **C**) *P. acuta*. Differences in proportion of CpGs in various genomic locations (CDS, introns, upstream flanks, downstream flanks, and intergenic regions) for **B**) *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 α = 0.05 level.



Supplementary Figure 6. Percent of highly methylated (≥ 50%; darkest shade), moderately methylated (10-50%; medium shade), and lowly methylated CpGs (< 10%; lightest shade) detected by each method A) for *M. capitata* and B) *P. acuta*.