

Unlocking DNA methylation with BSXplorer

Yuditskiy K.I., Bezdvornyykh I.V., Samsonova A.A., Kanapin A.A.*

Institute of Translational Biomedicine, Saint Petersburg State University, St. Petersburg, Russia

* akanapin@spbu.ru

Key words: epigenomics; bisulphite sequencing; exploratory analysis; transcription regulation; methylation density

Motivation and Aim: Cytosine methylation is a common epigenetic mark that is generally found in eukaryotes, including vertebrates, insects, fungi, and plants, exhibiting distinct patterns and pathways of DNA methylation [1, 2]. Bisulfite sequencing detects and quantifies DNA methylation patterns, contributing to our understanding of gene and transposon expression regulation, genome stability maintenance, conservation of epigenetic mechanisms. Despite being a technology of choice to profile DNA methylation for many years now there are surprisingly few lightweight and robust standalone tools available for efficient graphical analysis of data in non-model systems. Graphical representation of methylation data is crucial in exploring epigenetic regulation on a genome-wide scale however this feature is missing in many of existing analytical packages [3, 4]. This is especially relevant for non-model organisms with poorly annotated genomes and/or organisms where genome sequences are not yet assembled on chromosome level. BSXplorer is a tool specifically developed to assist researchers in explorative data analysis and in visualising and interpreting bisulfite sequencing data more easily.

Methods and Algorithms: BSXplorer is implemented in Python (version 3.9 or higher). The package runs on most modern systems and its functions are available through both Python API and command-line interface. The data processing speed is mostly limited by I/O capacity of the storage. BSXplorer is publicly available at <https://github.com/shitohana/BSXplorer> or <https://pypi.org/project/bsxplorer/> where a comprehensive user manual is provided. The tool processes methylation data quickly and offers API and CLI capabilities, along with the ability to create high-quality figures suitable for publication.

Results: BSXplorer provides in-depth graphical and exploratory analysis of bisulphite sequencing data and includes the following features: profiling of methylation levels in metagenes or in user-defined regions, generation of summary statistics charts, comparative analyses of methylation patterns across experimental samples, methylation contexts and species, and, finally, identification of modules sharing similar methylation signatures in various functional genomic elements.

Conclusions: Overall, BSXplorer is a lightweight and flexible instrument that facilitates explorative analyses of DNA methylation patterns, including BS-seq data mining, contrasting and visualization of methylation patterns in genomes of model and non-model organisms. The tool can be used on its own or seamlessly incorporated into epigenomic data processing workflows.

Funding: The study is supported by Russian Scientific Foundation (Grant No. 23-14-00134).

References

1. Greenberg M.V.C., Bourc'his D. The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol.* 2019;20(10):590-607. doi 10.1038/s41580-019-0159-6
2. Zhang H., Lang Z., Zhu J. K. Dynamics and function of DNA methylation in plants. *Nat Rev Mol Cell Boil.* 2018;19(8):489-506. doi 10.1038/s41580-018-0016-z
3. Müller F., Scherer M., Assenov Y., Lutsik P., Walter J., Lengauer T., Bock C. RnBeads 2.0: comprehensive analysis of DNA methylation data. *Genome Biol.* 2019;20(1):55. doi 10.1186/s13059-019-1664-9
4. Ki H., Si M., Par N., Kwo K., Ki J., Ki J. msPIPE: a pipeline for the analysis and visualization of whole-genome bisulfite sequencing data. *BMC Bioinformatics.* 2022;23(1):383. doi 10.1186/s12859-022-04925-2