

Elouan Bethuel, Magali Hennion and Olivier Kirsh

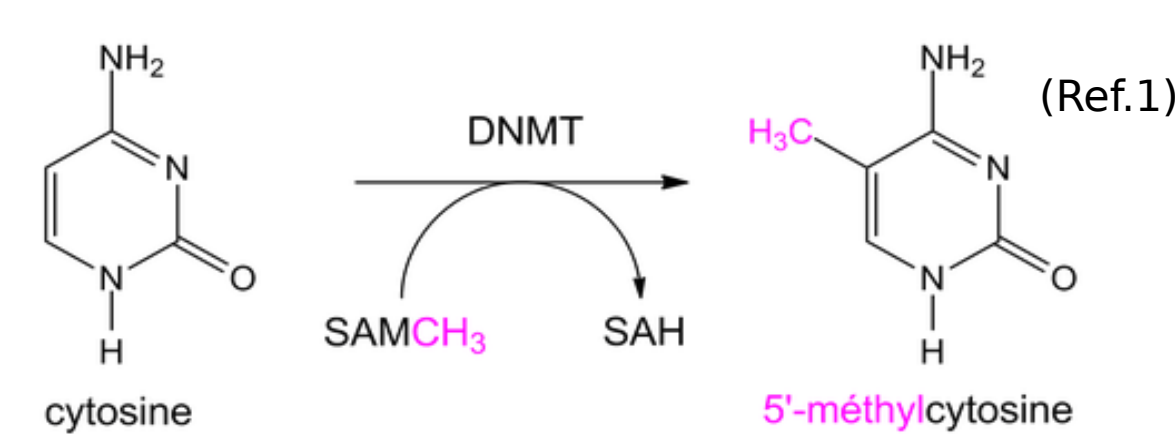
Plateforme BiBS, Université Paris Cité, CNRS, Epigenetics and Cell Fate, F-75013 Paris, France

✉ bibs@parisepigenetics.com

🌐 <https://parisepigenetics.github.io/bibs>

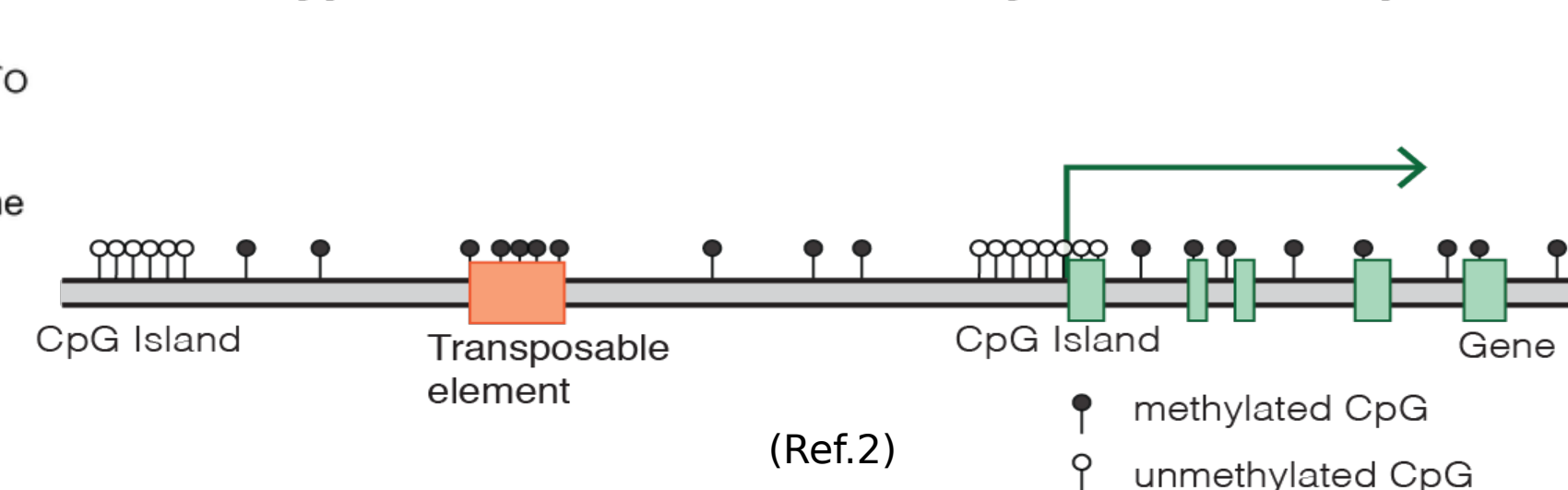
Abstract Epigenetic marks regulate gene transcription, as well as genome replication, and repair. In mammals, DNA methylation on CpGs was shown to be essential for normal development and is associated with a number of key processes including genomic imprinting, X-chromosome inactivation, repression of transposable elements, aging, and carcinogenesis. Several methods allow for genome-wide analysis of this mark, but data analysis remains complex and computationally intense. We implemented Methylator, a Snakemake workflow to analyse DNA methylation data. Methylator runs in a dedicated Apptainer image to allow for reproducibility and was optimized to compute efficiently the data on HPC clusters such as the ones from the IFB and iPOP-UP networks. We aim to make those complex analyses doable by biologists with no or little bioinformatics background.

Biological context

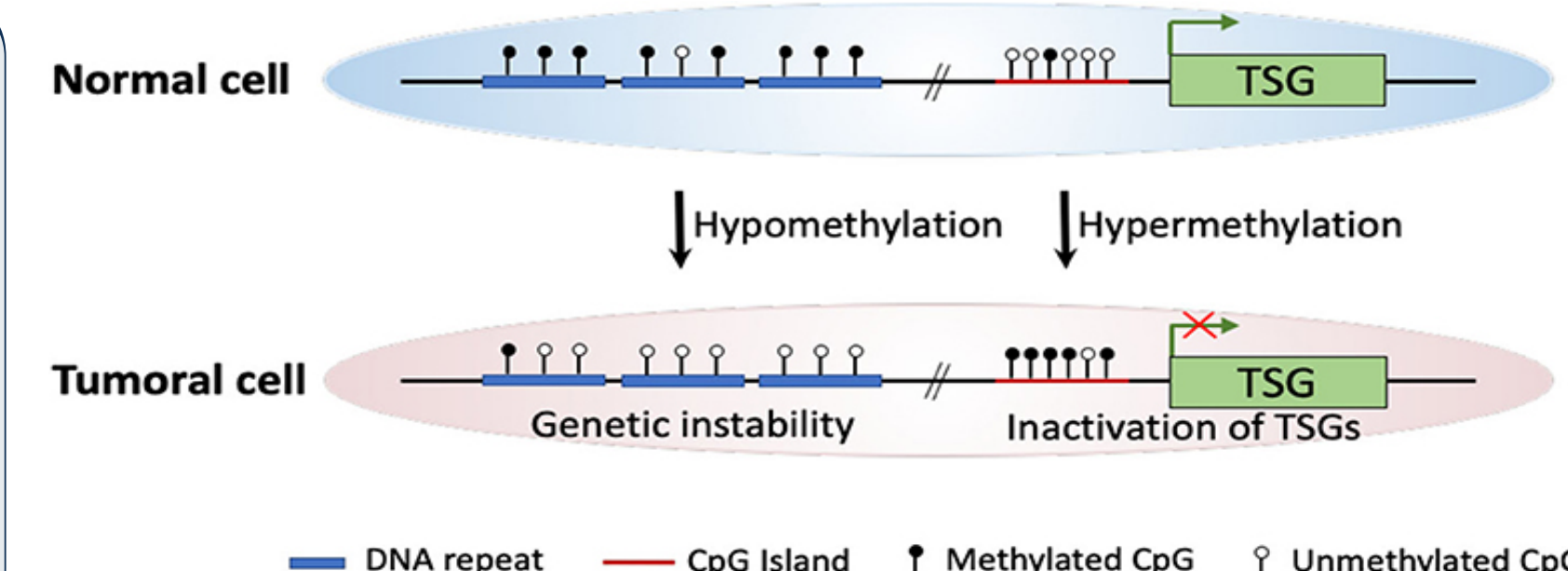


Cytosine methylation
(mainly in CpG context)

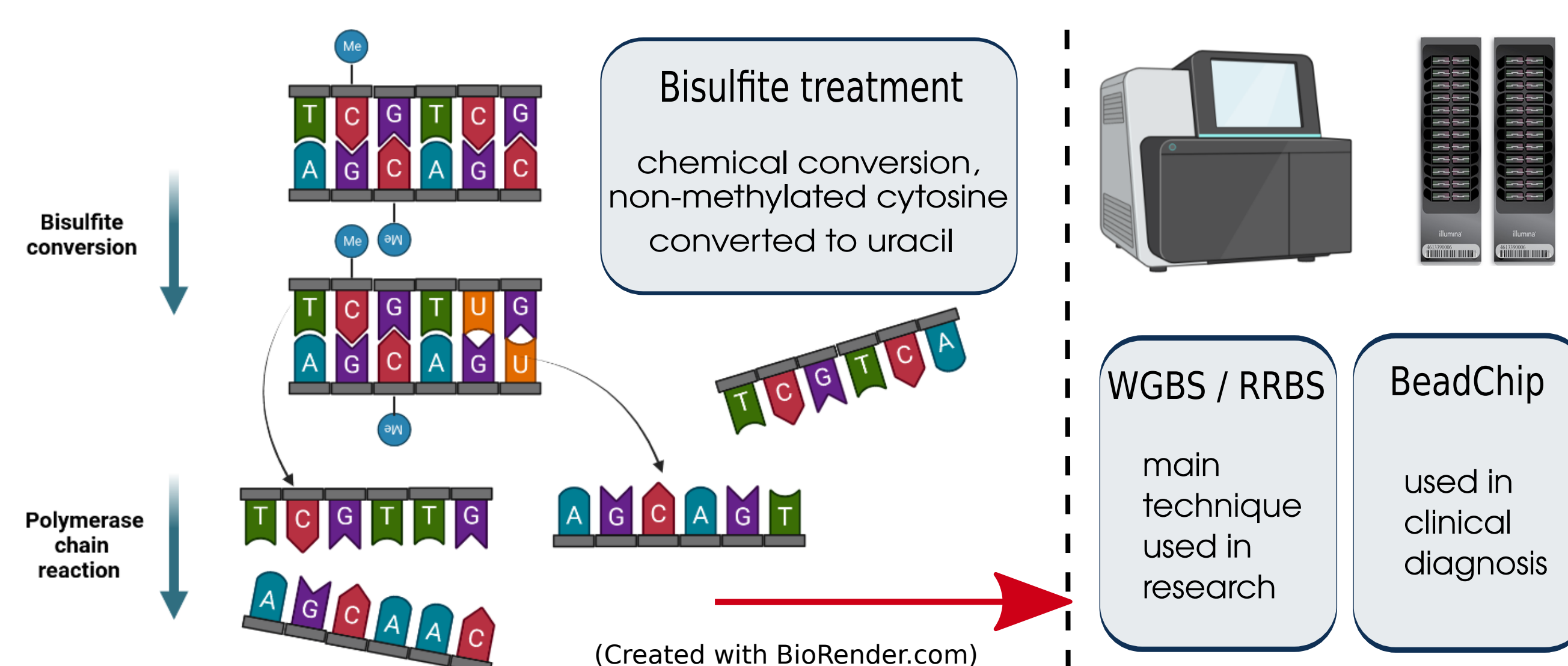
Typical mammalian DNA methylation landscape



In cancer cells, global hypomethylation at repetitive sequences might cause genetic instability, whereas site-specific CpG island promoter hypermethylation could lead to silencing of TSGs (tumor suppressor genes, ref. 3)

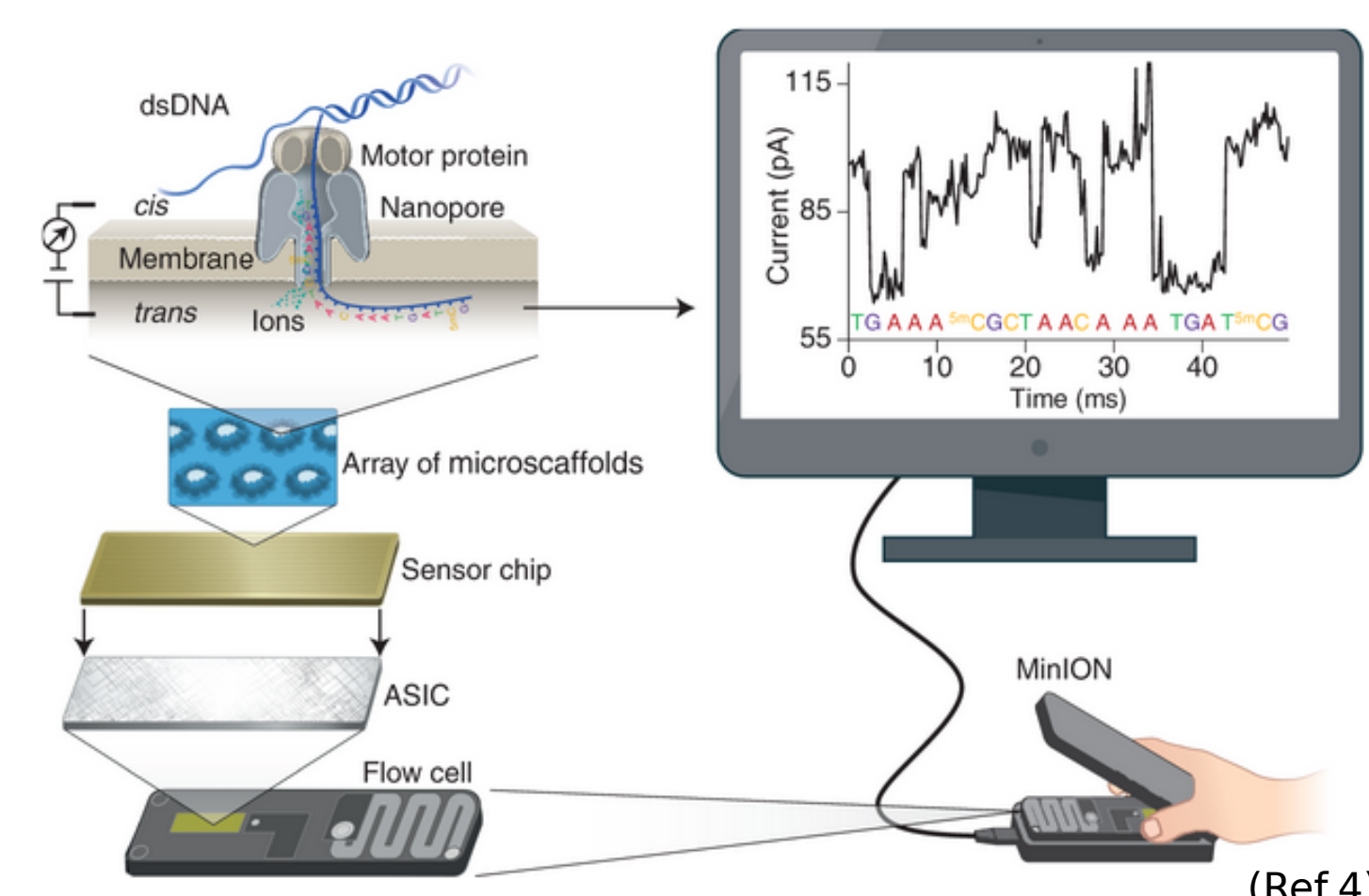


Profiling genome-wide DNA methylation



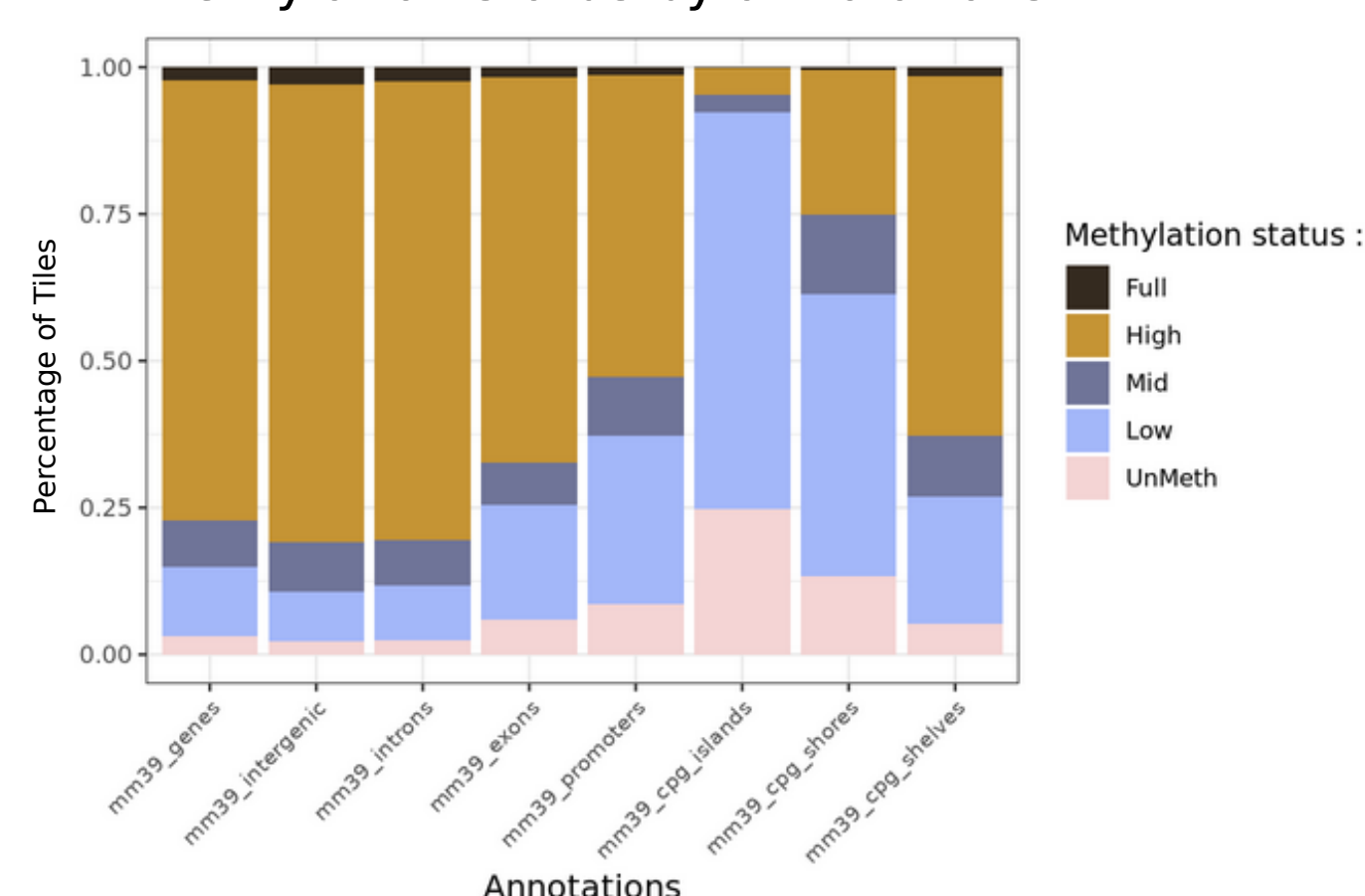
Nanopore sequencing

- long reads
- no DNA damage
- detecte 5-methylcytosine
- 5-hydroxy-methylcytosine

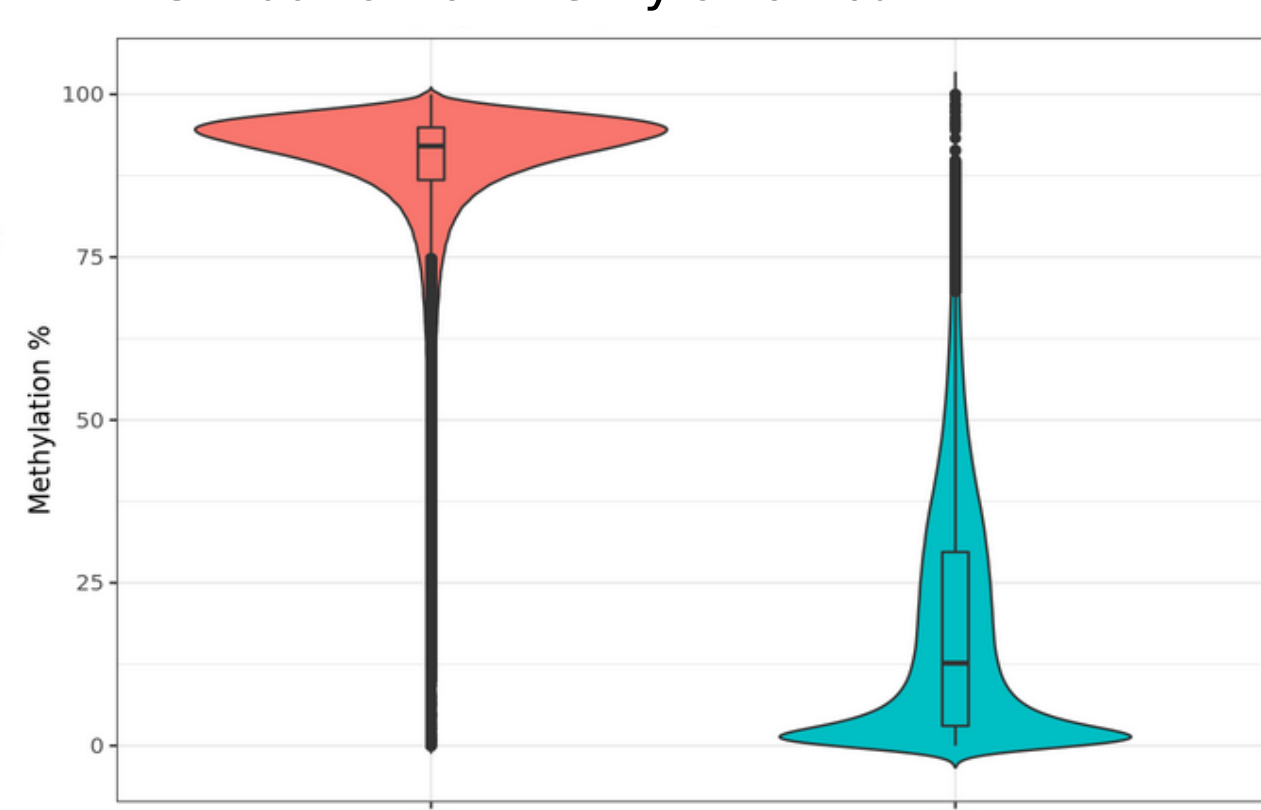


Results

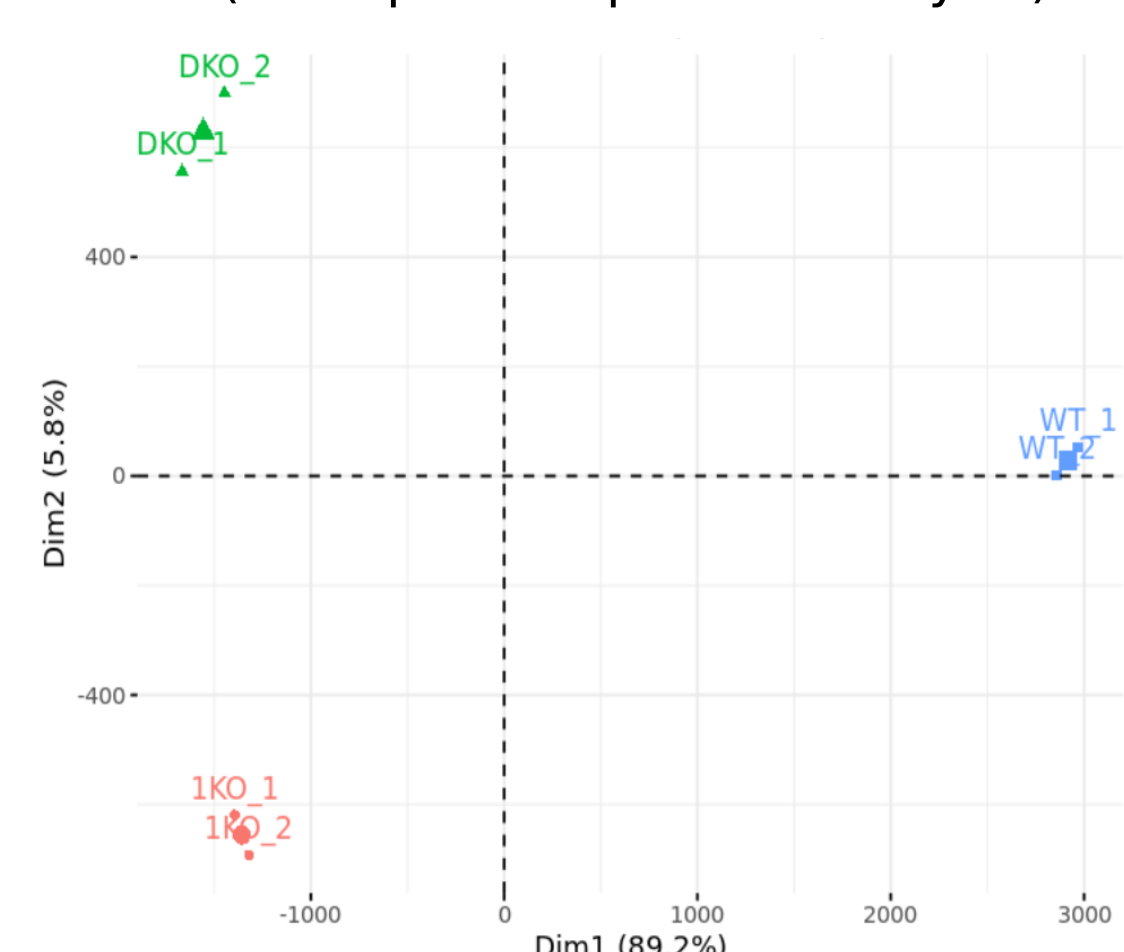
Methylation status by annotations



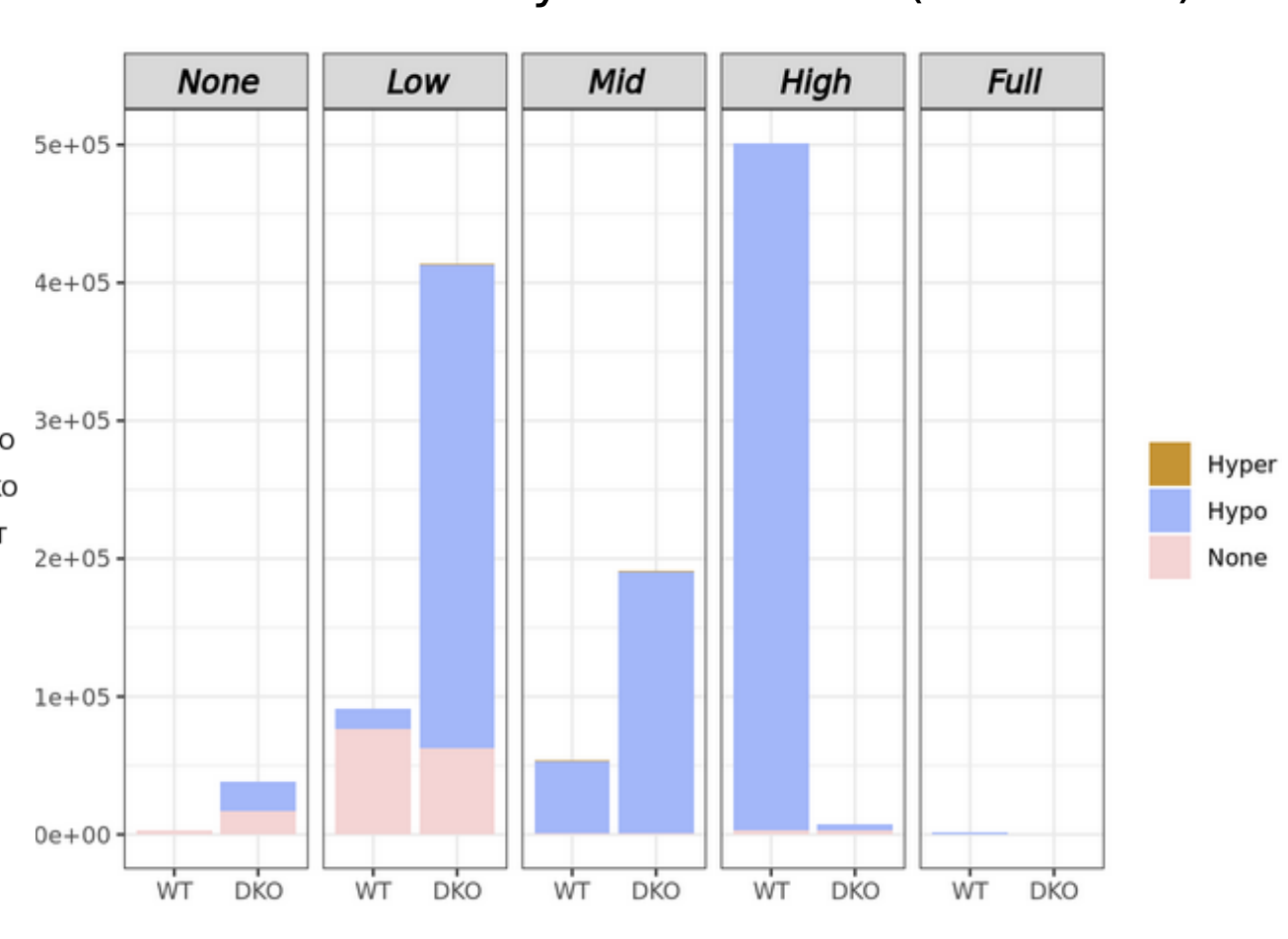
Distribution of methylation %



PCA (Principal Component Analysis)

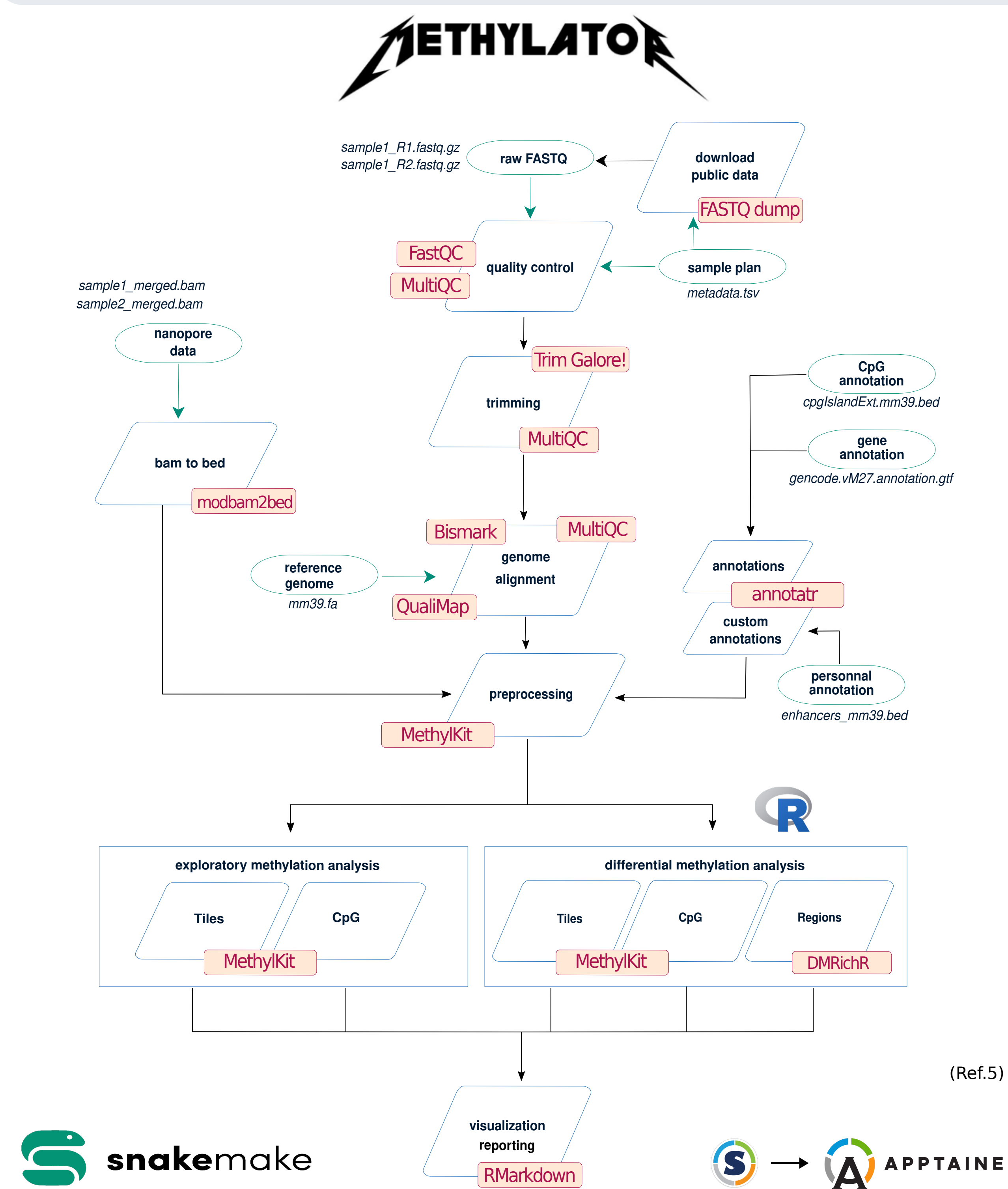


Differential methylation status (WT - DKO)

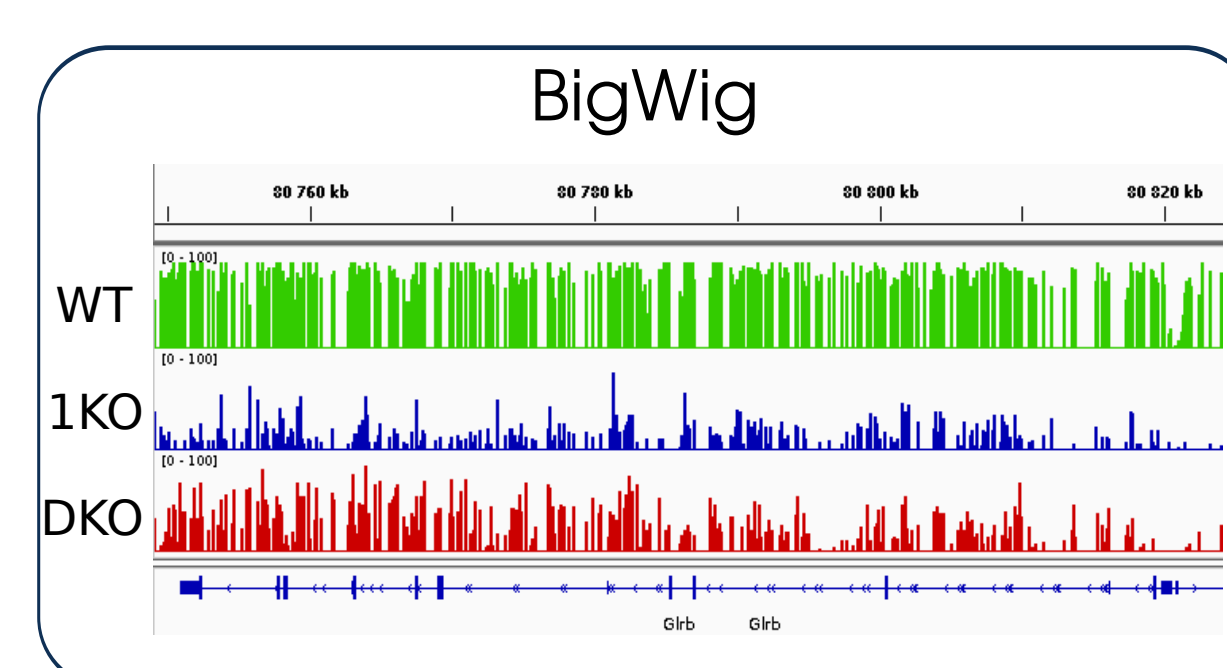


(Data from ref. 6 reanalysed by Methylator)

Workflow structure



Outputs



HTML Report



Conclusion



- Easy configuration
- Multiple data types
- Efficient computation (HPC clusters)
- Statistical analysis : exploratory and differential methylation analysis
- Reproducible analysis
- Results summarized in a user-friendly HTML report

Prospects :

- Integration of methylation BeadChip data
- Differential methylated region identification

Many thanks to

- Magali Hennion and Olivier Kirsh for their supervision and patience
- Guillaume Velasco and Laure Ferry for helpfull discussions and data
- Maya Zygadlo for the first steps of the pipeline
- BiBS steering committee
- Julien Rey and all iPOP-UP technical committee for HPC computing
- The SFBI for financial support to attend ISMB/ECCB 2023

References

- <https://planet-vie.ens.fr/thematiques/sante/pathologies/epigenetique-et-cancer>
- https://www.wikiwand.com/en/DNA_methylation
- Lin Lehang, Cheng Xu, Yin Dong. Aberrant DNA Methylation in Esophageal Squamous Cell Carcinoma: Biological and Clinical Implications. Frontiers in Oncology (2020)
- <https://nanoporetech.com>
- <https://github.com/parisepigenetics/Methylator>
- Dahlet, T., Argüeso Lleida, A., Al Adhami, H. et al. Genome-wide analysis in the mouse embryo reveals the importance of DNA methylation for transcription integrity. Nat Commun 11, 3153 (2020)