UHASSELT

KNOWLEDGE IN ACTION

Bisulfite Sequencing Analysis Workflow

Mai Nguyen

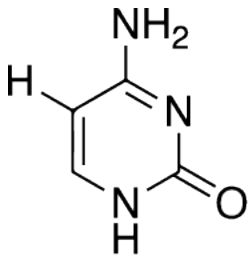
Faculty of Science
Master of Biostatistics
Specialization in Bioinformatics

April 22, 2019

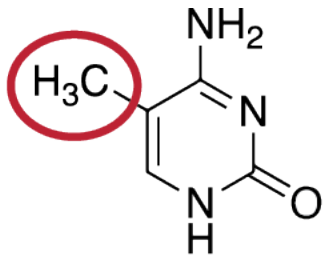
DNA Methylation

From Wikipedia

- A process by which methyl groups are added to the DNA molecule
- Methylation can change the activity of a DNA segment without changing the sequence



Cytosine

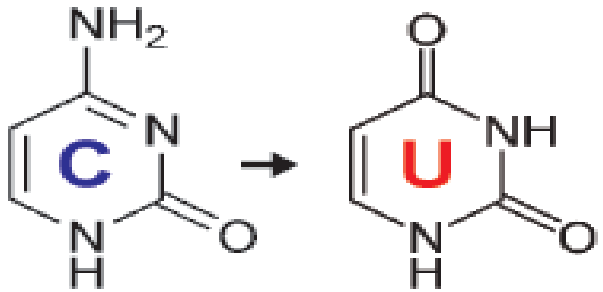


methylated Cytosine

DNA Methylation

Unmethylated Cytosine

C in single-stranded DNA converts into U and recognized as T in subsequent PCR amplification and sequencing



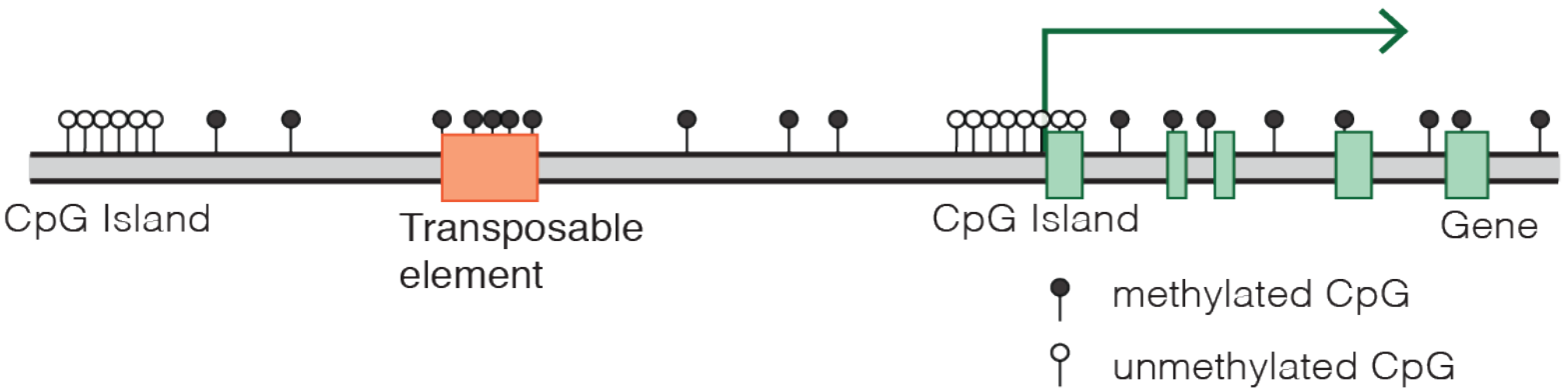
5 Methylated Cytosine (5mCs)

Immune to the conversion to Uracil and remain as C, 5mCs are distinguished from unmethylated cytosines

DNA Methylation

- Most commonly occurs at the C5 position of cytosines within CpG dinucleotides
- Role: gene expression, embryonic development, cellular proliferation, differentiation and chromosome stability

Typical mammalian DNA methylation landscape

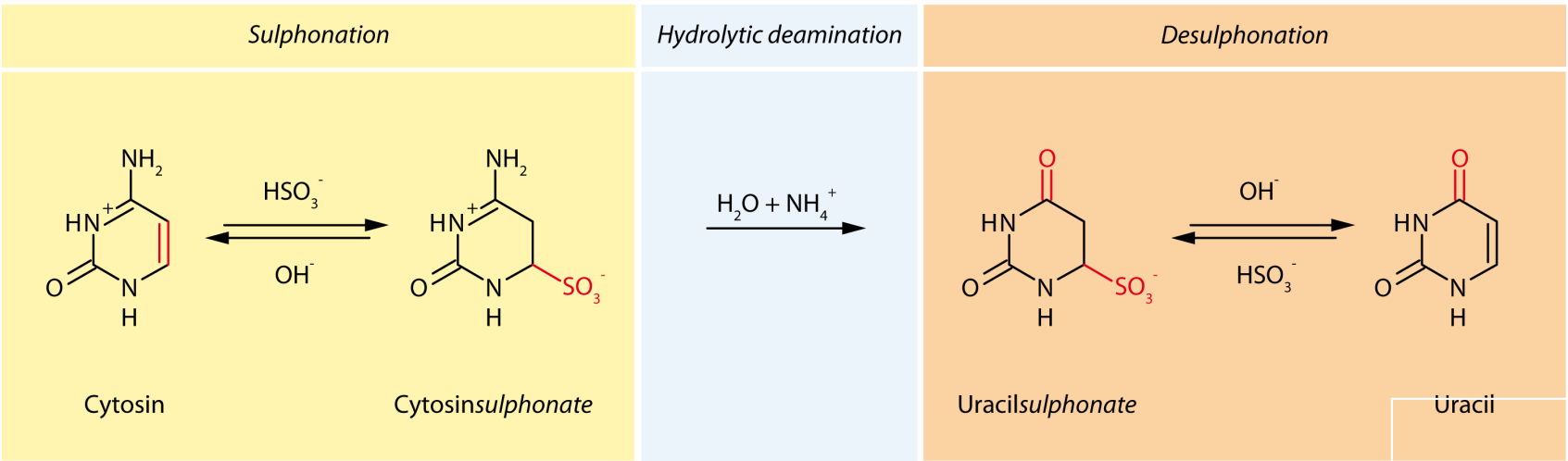


Bisulfite Genomic Sequencing

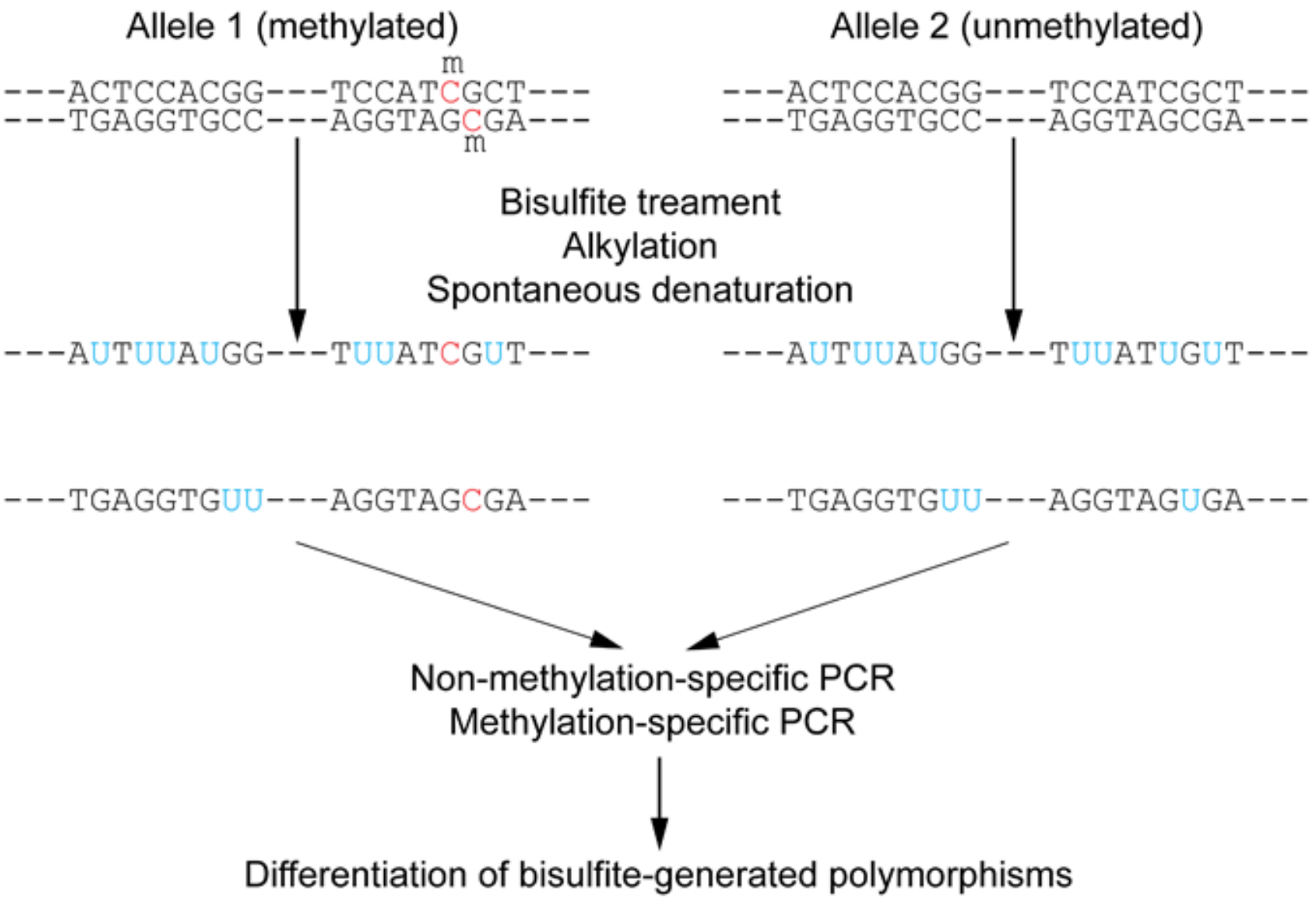
From Wikipedia

The use of bisulfite treatment of DNA before routine sequencing to determine the pattern of methylation

Bisulfite-mediated conversion of cytosine to uracil



Tollefsbol T (ed.): Handbook of Epigenetics: The New Molecular and Medical Genetics. 1st edition. London, San Diego: Academic Press, 2011.


Allele 1 (methylated)
Allele 2 (unmethylated)
---ACTCCACGG---TCCATCGCT---
---TGAGGTGCC---AGGTAGCGA---
m
m
---ACTCCACGG---TCCATCGCT---
---TGAGGTGCC---AGGTAGCGA---
Bisulfite treament
Alkylation
Spontaneous denaturation
---AUTUUAUGG---TUUATCGUT---
---AUTUUAUGG---TUUATUGUT---
---TGAGGTGUU---AGGTAGCGA---
---TGAGGTGUU---AGGTAGUGA---
Non-methylation-specific PCR
Methylation-specific PCR
Differentiation of bisulfite-generated polymorphisms

Measuring Approaches

Bisulfite Sequencing Analysis

Bisulfite treatment has been coupled with next generation sequencing (NGS)

- Whole genome bisulfite sequencing (WGBS)
- Reduced representation bisulfite sequencing (RRBS)
- Targeted bisulfite sequencing (TBS)

Reduced representation bisulfite sequencing

- Focusing on reduced, representative sample of the whole genome: enrich for CpGs rich regions
- RRBS targets CpG-regions
- RRBS approach can capture approximately 70% of gene promoters and 85% of CpG islands [3]

Advantages

- Reducing the reads for sequencing
- Low cost methylation analysis

Pipeline for Bilsulfite sequencing analysis

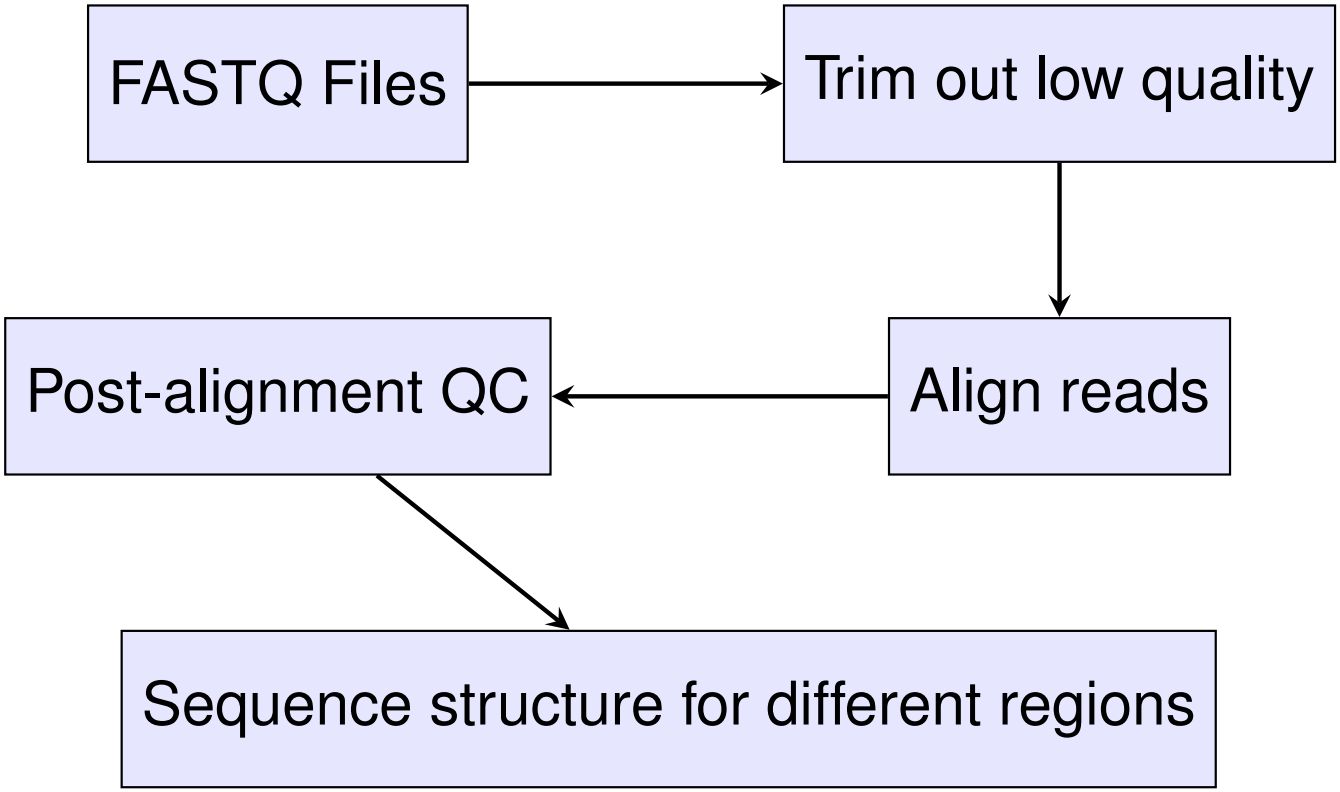


Figure: Analysis Pipeline for Bisulfite sequencing analysis for different methylation regions (DMR) within a sample

Protocol

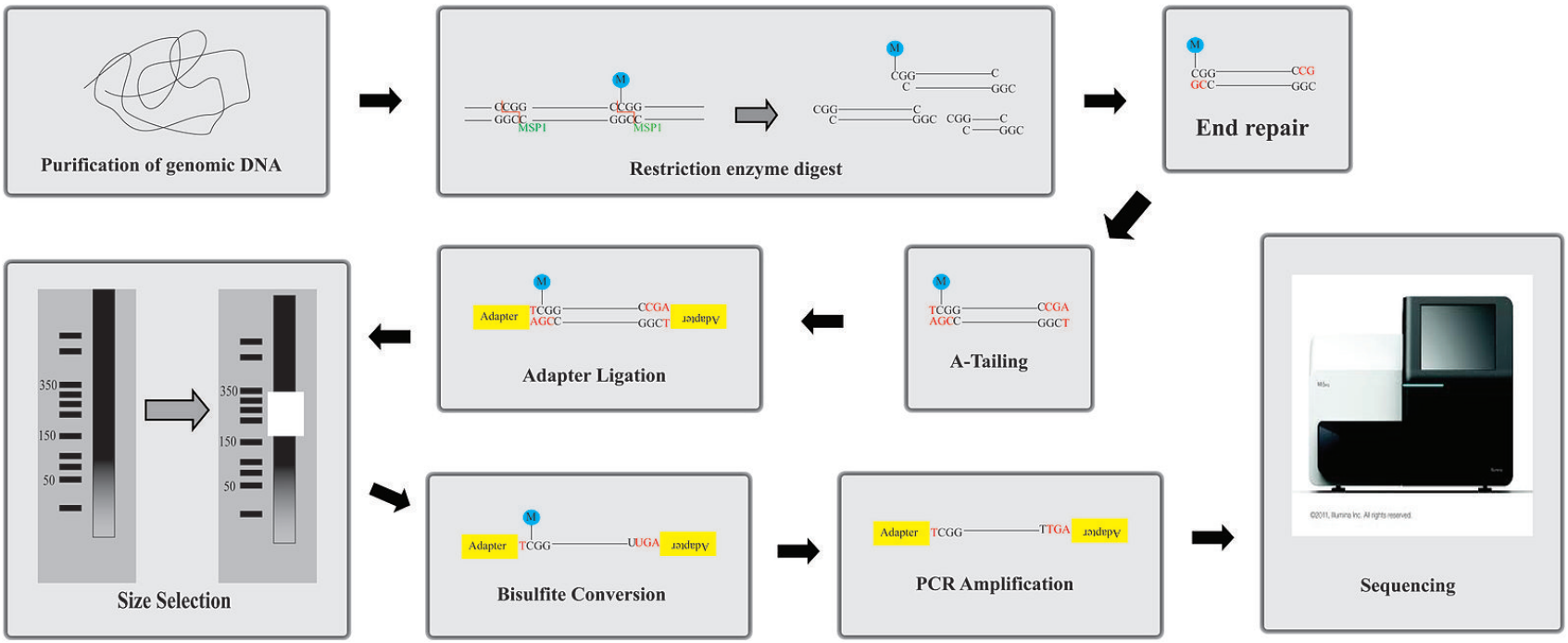


Figure: Caption

Statistical analysis

- Normalisation the count data after alignment
- Region finding by analysing which region have higher expression of methylation between converted sample and non-converted sample.
- Methylated regions within sample can be analyse by comparing across regions

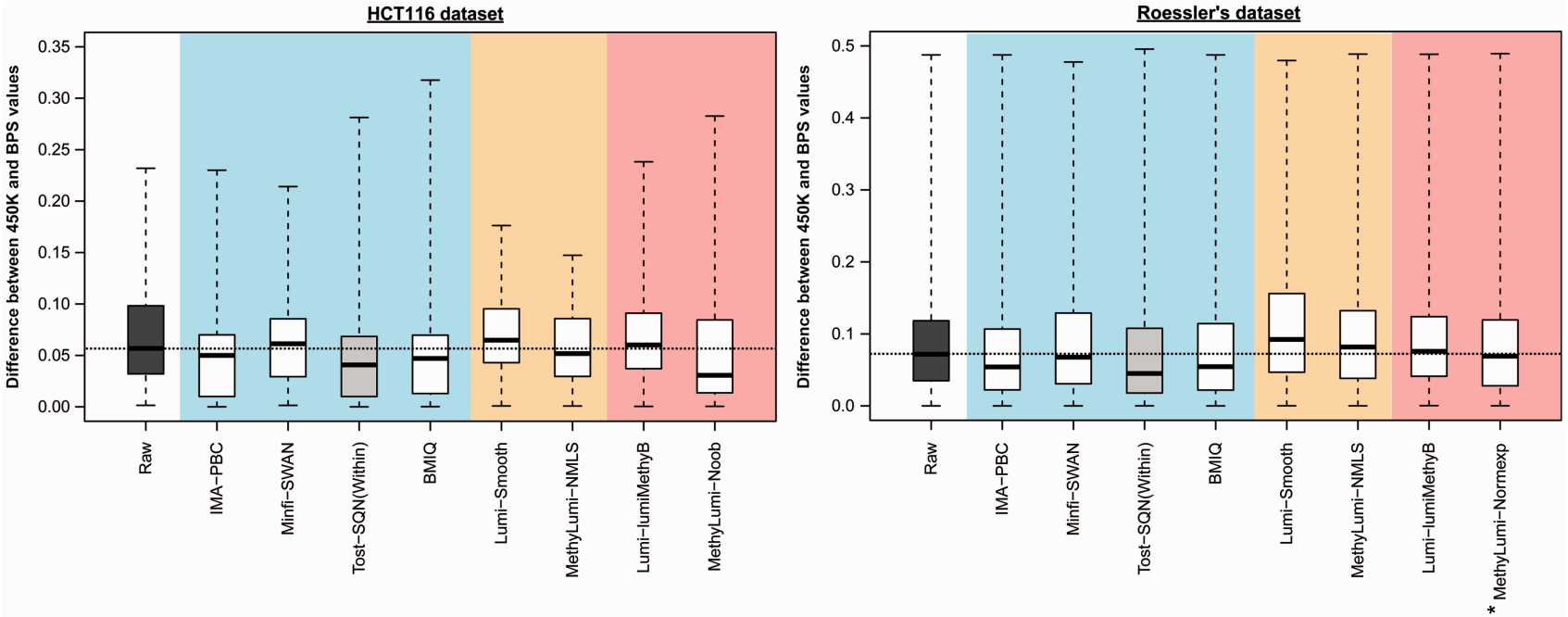


Figure: Methylated region within sample [5]

Statistical Analysis

Bisulfite Sequencing Analysis within sample[4]

- Methylated CpG (DMC) sites
- Methylated regions (DMR)

- local auto-correlation patterns between CpGs
- patterns of both methylation and auto-correlation
- CpGs are unevenly distributed across the genome

Bisulfite Sequencing Analysis within sample[4]

Hidden Markov Model for within sample Bisulfite sequence analysis

Reference Sources I

Li, Y., Tollefsbol, T.O. (2011)
DNA methylation detection: Bisulfite genomic sequencing analysis.
Methods in molecular biology (Clifton, N.J.),791,11-21, doi:10.1007/978-1-61779-316-52

Frommer M. et al. (1992)
A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands..
Proceedings of the National Academy of Sciences of the United States of America, 89(5), 18271831

Reference Sources II

Chen, Y. et al. (2018)
Differential methylation analysis of reduced representation bisulfite sequencing experiments using edgeR.
F1000Research, 6, 2055. doi:10.12688/f1000research.13196.2

Shokoohi F. et. al. (2018
A Hidden Markov Model for Identifying Differentially Methylated Sites in Bisulfite Sequencing Data
Biometrics. doi:10.111/biom.12965

Dedeurwaerder S. et. al(2014)
A comprehensive overview of Infinium HumanMethylation450 data processing, Briefings in Bioinformatics
Volume 15, Issue 6, November 2014, Pages 929941,
https://doi.org/10.1093/bib/bbt054