

Service name : Whole genome bisulfite sequencing (WGBS)

- methylation-seq analysis of delivered samples of 64 gDNA samples.
- Quantitative and qualitative parameters of DNA samples:

Sample Type	Remarks	Amount	Volume	Concentration	Purity
Genomic DNA	Strongly Recommended	≥5 µg	≥20 µL	≥50ng/µL	no degradation , no contamination
	Required	≥2.5 µg			

Requirements for methodological procedure:

- Quality control of the delivered DNA samples
- 200~400 bp insert bisulfite treat DNA library
- Use of the most up-to-date platforms such as illumina *Novaseq 6000* or *Hi-seq* platforms.
- 150 bp pair-end reads
- In case any sample does not meet the specification above, the company is responsible to inform the customer immediately (no later than the term specified in article III of contract) about it. The customer will deliver new sample(s) without any purposeless delay.
- Minimum **40 million reads** (reading frames) per sample, **12Gb** raw data.
- Data quality requirement: Phred value at least Q30 for at least 80% of the reads (reading frames) and at the same time Phred value at least Q20 for at least 90% of the reads (reading frames)

- After the quality check, the following data analyses need to be done:

Data quality control (filtering reads containing adapter or with low quality; Q20, Q30, error rate distribution, GC distribution, total bases)
Mapping onto reference genome (mapping rate, duplication rate, sequencing depth, reads coverage)
mCs detection, methylation level calculation.
(1) Methylation level and frequency distribution in different sequence context (CG, CHG, CHH) (2) Methylation level and frequency distribution in different chromosomes (3) Methylation level and frequency distribution in different functional elements (promoter, 5'UTR, exon, intron, 3'UTR)
Differentially Methylated Site (DMS) detection
Differentially methylated regions (DMRs), Differentially Methylated Promoters (DMPs) detection and annotation
Function enrichment (Gene Ontology and KEGG Pathway, GO enrichment) of DMR-associated genes and DMP-associated genes
Visualization of BS seq data – i.e., Circos plots

- Consultation on sequencing and data analysis service need to be there to explain the entire pipeline to the researchers as per request and requirement.
- Updated version of the softwares needs to be used for performing the data analysis to meet high quality publication requirement.
- The analysis pipeline can be modified for the sake of high quality publication output if required.