

EM-seq Differential Methylation Analysis Report

Contents

1	Executive Summary	2
2	Output Directory Structure	2
3	Analysis Overview	3
3.1	Dataset Description	3
3.2	Analysis Pipeline Components	3
3.2.1	1. Differential Methylation Detection	3
3.2.2	2. Genomic Annotation	3
3.2.3	3. Visualization and Reporting	3
3.3	Technical Specifications	3
3.3.1	Software and Packages	3
3.3.2	Quality Control	3
3.3.3	Analysis Scripts	4
4	TE integration with DMRs (in the DMRs folder)	4
4.1	Primary Analysis Results	4
4.2	Individual Circos Plots	4
4.3	Comparative Analysis Outputs	4
4.4	Analysis Insights	4
4.4.1	Genomic Distribution Patterns	4
4.4.2	Transposable Element Integration	5
5	Homer Motif Enrichment Analysis	5
5.1	Analysis Structure	5
5.2	Key Output Files	5

6	DMRichR Genomic Enrichment Analysis	5
6.1	CpG Island Enrichment	5
6.2	Genic Enrichment Analysis	6
6.3	Key Output Files	6
7	GREAT Functional Enrichment Analysis	6
7.1	Analysis Types	6
7.2	Key Output Files	6
7.3	Analysis Structure	6
8	Key Comparative Plots	7

1 Executive Summary

This report presents a comprehensive analysis of EM-seq (Enzymatic Methyl-seq) data examining differential methylation patterns in response to ionizing radiation treatment. The analysis includes both Differentially Methylated Loci (DML) and Differentially Methylated Regions (DMR) identification in CG context, with integrated transposable element (TE) annotation and gene expression correlation.

2 Output Directory Structure

```
dss/
  [comparison_name]/
    DMRs/
      DMRs_annotated_with_TE_dge.xlsx
      GREAT/          # Functional enrichment analysis
        enrichment_in_*.tsv
        distance_to_TSS_*.pdf
        Annotated_*.bed
    HOMER/
      both/           # All DMRs motif analysis
      hyper/          # Hypermethylated DMRs
      hypo/           # Hypomethylated DMRs
    DMRichments/
      *_CpG_enrichments.pdf
      *_genic_enrichments.pdf
      multi_plot.pdf files
    DML_[comparison].tsv
    circos_DML_[comparison].pdf
    circos_DMR_[comparison].pdf
    circos_DMR_comparative.pdf
    circos_DML_comparative.pdf
```

3 Analysis Overview

3.1 Dataset Description

- **Technology:** EM-seq (Enzymatic Methyl-seq)
- **Treatment Groups:** Ionizing radiation at multiple doses and timepoints
- **Comparisons Analyzed:** 8 treatment comparisons
 - IR2Gy6d_vs_NIR, IR2Gy24h_vs_NIR
 - IR10Gy6d_vs_NIR, IR10Gy24h_vs_NIR
 - IR2Gy6d_vs_IR2Gy24h, IR10Gy6d_vs_IR10Gy24h
 - IR10Gy6d_vs_IR2Gy6d, IR10Gy24h_vs_IR2Gy24h

3.2 Analysis Pipeline Components

3.2.1 1. Differential Methylation Detection

- **DML Analysis:** Single-cytosine resolution using DSS package with smoothing
- **DMR Analysis:** Regional methylation changes using DMRichR/DSS with smoothing
- **Statistical Thresholds:** FDR < 0.05, methylation difference > 10%

3.2.2 2. Genomic Annotation

- **Transposable Element Annotation:** Integration with RepeatMasker annotations
- **Gene Expression Correlation:** DESeq2 differential expression integration
- **Genomic Context:** Promoter, gene body, and intergenic region classification

3.2.3 3. Visualization and Reporting

- **Circos Plots:** Genome-wide visualization of methylation patterns
- **Enrichment Analysis:** TE family and gene-level enrichment testing
- **Comparative Analysis:** Cross-treatment pattern identification

3.3 Technical Specifications

3.3.1 Software and Packages

- **R Environment:** Statistical analysis and visualization
- **Key Packages:** DSS, DMRichR, dmrseq, circlize, ComplexHeatmap
- **Annotation:** RepeatMasker, GENCODE gene annotations
- **Visualization:** Enhanced circos plots with custom density binning

3.3.2 Quality Control

- **Internal Controls:** Lambda phage (unmethylated) and pUC19 (methylated) controls
- **Statistical Validation:** Multiple testing correction and effect size thresholds

3.3.3 Analysis Scripts

- `create_DMR_circos_plots.R`: DMR visualization functions
- `create_DML_circos_plots.R`: DML visualization functions
- `annotate_dmrs_with_TE.R`: Main analysis pipeline

4 TE integration with DMRs (in the DMRs folder)

4.1 Primary Analysis Results

- **Individual DMR Files**: `DMRs_annotated_with_TE_dge.xlsx` (per comparison)
- **TE Enrichment Results**: Family and gene-level enrichment tables

4.2 Individual Circos Plots

- `circos_DMR_[sample].pdf`
- `circos_DML_[sample].pdf`

4.3 Comparative Analysis Outputs

- **`circos_DMR_comparative.pdf`**: Multi-sample DMR density comparison
 - Genome-wide view of DMR distribution across all treatments
 - Color-coded by treatment comparison
 - Density bins show regional clustering patterns
 - Identifies treatment-specific methylation hotspots
- **`circos_DML_comparative.pdf`**: Multi-sample DML density comparison
 - Single-cytosine resolution methylation changes
 - Comparative density patterns across treatments
 - Highlights dose and time-dependent responses
 - Shows genome-wide methylation susceptibility regions

4.4 Analysis Insights

4.4.1 Genomic Distribution Patterns

The comparative circos plots reveal several key patterns:

- **Chromosome-Specific Responses**: Certain chromosomes show higher susceptibility to radiation-induced methylation changes
- **Regional Clustering**: DMRs and DMLs cluster in specific genomic regions, suggesting coordinated regulatory responses

- **Dose-Response Relationships:** Higher radiation doses (10 Gy vs 2 Gy) show distinct methylation patterns
- **Temporal Dynamics:** Time-course analysis (6d vs 24h) reveals both immediate and delayed methylation responses

4.4.2 Transposable Element Integration

- **TE Family Enrichment:** Specific TE families show preferential methylation changes
- **Genomic Context:** TE-associated methylation changes correlate with gene expression alterations
- **Regulatory Impact:** TE methylation changes may influence nearby gene regulation

5 Homer Motif Enrichment Analysis

Each comparison includes comprehensive motif enrichment analysis using Homer, examining transcription factor binding sites enriched in DMRs:

5.1 Analysis Structure

For each comparison (e.g., IR10Gy24h_vs_NIR), Homer analysis includes: - **All DMRs:** Combined hypermethylated and hypomethylated regions - **Hypermethylated DMRs:** Regions with increased methylation - **Hypomethylated DMRs:** Regions with decreased methylation

5.2 Key Output Files

- `knownResults.txt` - Detailed motif enrichment statistics
- `knownResults.html` - Interactive HTML summary
- `DMRs_annot.tab` - Genomic annotation of DMRs
- `DMRs_annot_stat.tab` - Annotation statistics

The Homer analysis did not identify any transcription factors whose binding sites are significantly enriched in radiation-induced DMRs.

6 DMRichR Genomic Enrichment Analysis

DMRichR provides comprehensive genomic context analysis for identified DMRs, examining enrichment patterns across different genomic features:

6.1 CpG Island Enrichment

- **All DMRs:** Overall CpG island context analysis
- **Hypermethylated DMRs:** CpG enrichment in regions gaining methylation
- **Hypomethylated DMRs:** CpG enrichment in regions losing methylation

6.2 Genic Enrichment Analysis

- Promoter regions
- Gene bodies
- Intergenic regions
- UTR regions

6.3 Key Output Files

For each comparison, DMRichR generates: - *_CpG_enrichments.pdf - CpG island enrichment plots - *_CpG_enrichments.xlsx - Detailed CpG enrichment statistics - *_genic_enrichments.pdf - Genomic feature enrichment plots - *_genic_enrichments.xlsx - Detailed genic enrichment data - CpG_multi_plot.pdf - Comparative CpG enrichment visualization - genic_multi_plot.pdf - Comparative genic enrichment visualization

The DMRichR analysis reveals whether radiation-induced methylation changes preferentially occur in specific genomic contexts, such as CpG islands, promoters, or gene bodies.

7 GREAT Functional Enrichment Analysis

GREAT (Genomic Regions Enrichment of Annotations Tool) performs comprehensive functional enrichment analysis to identify biological pathways and processes associated with DMRs:

7.1 Analysis Types

For each comparison, GREAT analyzes both hypermethylated and hypomethylated DMRs separately across multiple pathway databases:

- **Gene Ontology Biological Process (GOPB):** Cellular and molecular processes
- **KEGG Pathways:** Metabolic and signaling pathways
- **Reactome Pathways:** Biological reactions and pathways
- **Hallmark Gene Sets:** Well-defined biological states and processes

7.2 Key Output Files

- enrichment_in_*.tsv - Detailed enrichment statistics for each pathway database
- distance_to_TSS_*.pdf - Distance distribution plots to transcription start sites
- Annotated_*.bed - DMRs annotated with associated genes and pathways

7.3 Analysis Structure

GREAT/

```
enrichment_in_gopb_hyper_DMR.tsv      # GO enrichment for hypermethylated
enrichment_in_gopb_hypo_DMR.tsv       # GO enrichment for hypomethylated
```

enrichment_in_kegg_hyper_DMR.tsv	# KEGG pathways for hypermethylated
enrichment_in_kegg_hypo_DMR.tsv	# KEGG pathways for hypomethylated
enrichment_in_reactome_*.tsv	# Reactome pathway enrichments
enrichment_in_hallmark_*.tsv	# Hallmark gene set enrichments
distance_to_TSS_hyper_DMR.pdf	# TSS distance plots
distance_to_TSS_hypo_DMR.pdf	

8 Key Comparative Plots

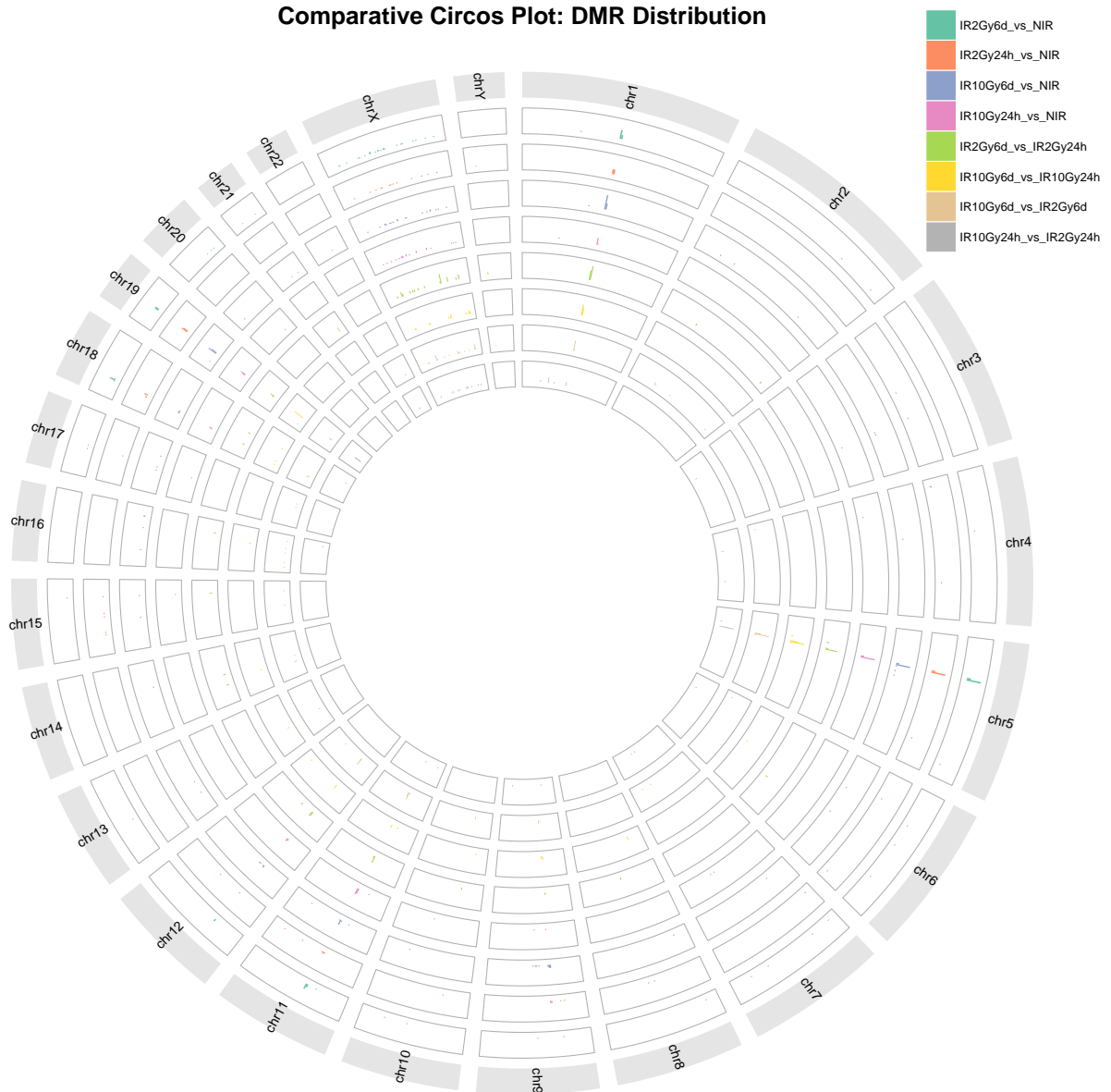


Figure 1: DMR Comparative Circos Plot - Genome-wide visualization of differentially methylated regions (DMRs) across all radiation treatment comparisons. The circular plot displays DMR density and methylation changes across chromosomes, with color-coded tracks representing

different treatment conditions. This comprehensive view enables identification of chromosomal regions with consistent radiation-induced methylation alterations.

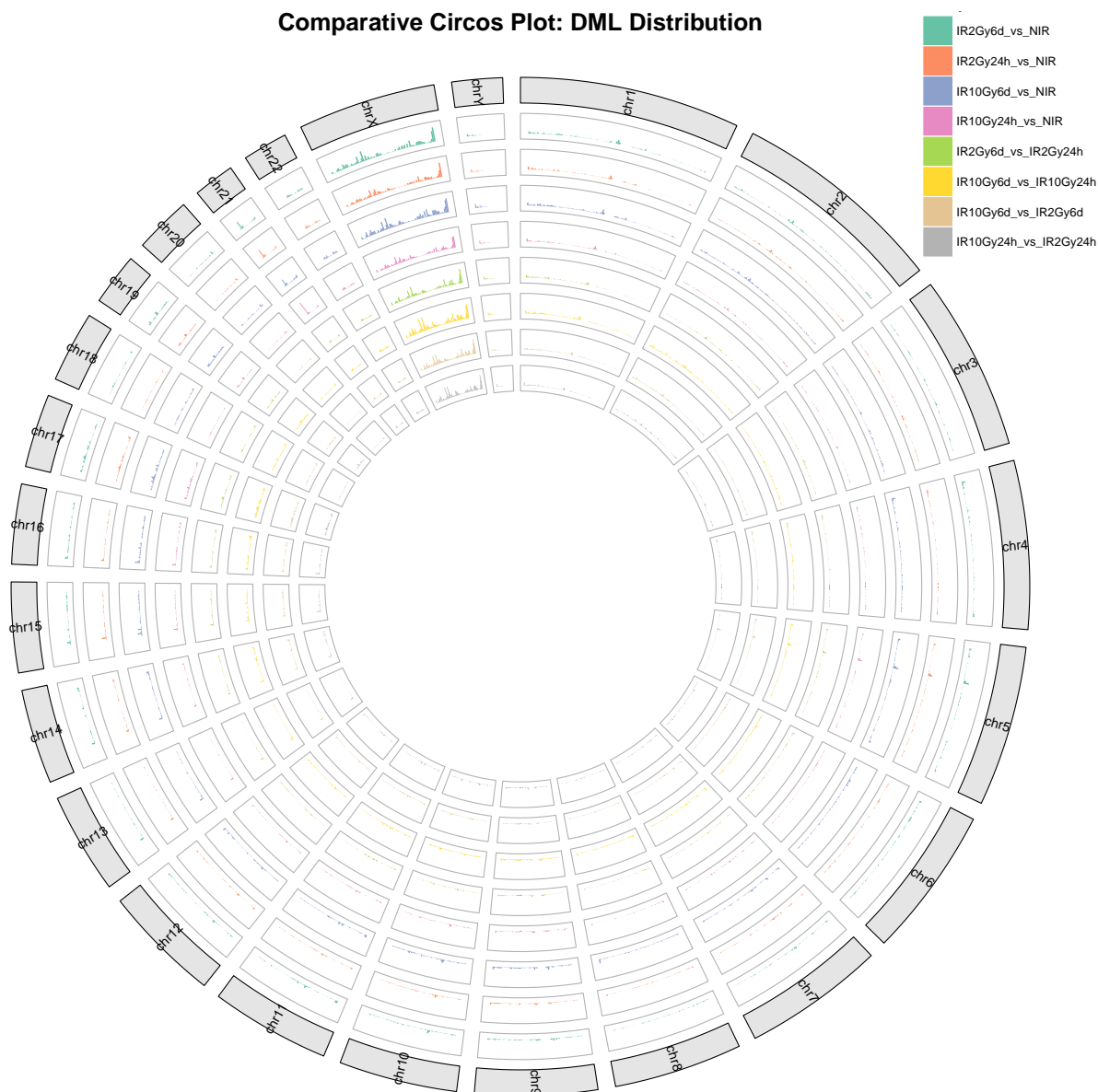


Figure 2: DML Comparative Circos Plot - Genome-wide visualization of differentially methylated loci (DMLs) at single-cytosine resolution across all radiation treatment comparisons. The circular genomic plot shows DML density patterns and methylation magnitude changes, highlighting hotspots of radiation-induced methylation alterations at the finest resolution. This single-base resolution view complements the regional DMR analysis.

These comprehensive circular genomic visualizations provide genome-wide perspectives on radiation-induced methylation changes, enabling identification of treatment-specific patterns and genomic susceptibility regions.

Report generated: August 2025

Analysis pipeline: EM-seq differential methylation with enhanced circos visualization

Contact: [Shuye Pu]