Methods Code:

1. Data Acquisition and Preprocessing

Downloading Sequencing Data: ChIP-seq raw reads were downloaded from the SRA repository for both wild-type and mutant-type cells using the SRA Toolkit.
Command used for wild-type:

fasterq-dump SRR18355704 --split-files --outdir ./fastq_files/wild_type

Command used for mutant-type:

fasterq-dump SRR18355705 --split-files --outdir ./fastq_files/mutant_type

Quality Control: FASTQ files were assessed for quality using FastQC for both conditions:

fastqc ./fastq_files/wild_type/SRR18355704_1.fastq
./fastq_files/wild_type/SRR18355704_2.fastq -o ./qc_reports/wild_type

fastqc ./fastq_files/mutant_type/SRR18355705_1.fastq
./fastq_files/mutant_type/SRR18355705_2.fastq -o ./qc_reports/mutant_type

Reference Genome Preparation: The hg38 human genome reference sequence was downloaded and indexed using Bowtie2.

Commands used:

wget ftp://ftp.ensembl.org/pub/release-109/fasta/homo_sapiens/dna/Homo_sapiens.GRCh38.dna.primary_assembly.fa.gz

gunzip Homo_sapiens.GRCh38.dna.primary_assembly.fa.gz

bowtie2-build Homo_sapiens.GRCh38.dna.primary_assembly.fa hg38_index

2. Read Alignment

Aligning Reads: Reads were aligned to the hg38 genome using Bowtie2 for both wild-type and mutant-type samples.

Wild-type:

bowtie2 -x hg38_index -1 ./fastq_files/wild_type/SRR18355704_1.fastq -2
./fastq_files/wild_type/SRR18355704_2.fastq -S ./alignment/wild_type.sam

Mutant-type:

bowtie2 -x hg38_index -1 ./fastq_files/mutant_type/SRR18355705_1.fastq -2 ./fastq_files/mutant_type/SRR18355705_2.fastq -S ./alignment/mutant_type.sam

Conversion and Sorting: SAM files were converted to BAM format, sorted, and indexed for both wild-type and mutant-type:

Wild-type:

samtools view -bS ./alignment/wild_type.sam > ./alignment/wild_type.bam

samtools sort ./alignment/wild_type.bam -o ./alignment/wild_type_sorted.bam

samtools index ./alignment/wild_type_sorted.bam

Mutant-type:

samtools view -bS ./alignment/mutant_type.sam > ./alignment/mutant_type.bam

samtools sort ./alignment/mutant_type.bam -o ./alignment/mutant_type_sorted.bam

samtools index ./alignment/mutant_type_sorted.bam

3. Peak Calling

MACS2 Peak Calling: Peaks were identified for both conditions using MACS2.
Wild-type:

macs2 callpeak -t ./alignment/wild_type_sorted.bam -f BAM -g hs -n TP53_WT_ChIP --outdir ./macs2_output/wild_type

Mutant-type:

macs2 callpeak -t ./alignment/mutant_type_sorted.bam -f BAM -g hs -n TP53_MT_ChIP --outdir ./macs2_output/mutant_type

4. Visualization

Preparation for Visualization: Peaks and alignment data were formatted for visualization using bedtools and DeepTools.
Sorting peaks:

bedtools sort -i ./macs2_output/wild_type/TP53_WT_ChIP_peaks.narrowPeak > ./visualization/sorted_peaks_wild_type.bed

bedtools sort -i ./macs2_output/mutant_type/TP53_MT_ChIP_peaks.narrowPeak > ./visualization/sorted_peaks_mutant_type.bed

Genome Browser Visualization: Aligned reads and peaks for both wild-type and mutant-type were visualized using IGV:

- **Wild-type:**
  - Aligned reads: wild_type_sorted.bam
  - Peaks: sorted_peaks_wild_type.bed
- **Mutant-type:**
  - Aligned reads: mutant_type_sorted.bam
  - Peaks: sorted_peaks_mutant_type.bed

The TP53 locus (chr17:7668402-7687550) was examined for peak coverage and alignment differences between conditions.

5. Data Analysis

Peak Annotation: Peaks were annotated to nearby genes using HOMER and ChIPseeker for both conditions.

Wild-type:

annotatePeaks.pl ./visualization/sorted_peaks_wild_type.bed hg38 > ./annotation/wild_type_annotated.txt

Mutant type:

annotatePeaks.pl ./visualization/sorted_peaks_mutant_type.bed hg38 > ./annotation/mutant_type_annotated.txt

Quality and Summary Metrics: Alignment metrics and mapping rates were computed using samtools flagstat and idxstats for both conditions:

Wild-type:

samtools flagstat ./alignment/wild_type_sorted.bam > ./metrics/wild_type_alignment_metrics.txt

Mutant type:

samtools flagstat ./alignment/mutant_type_sorted.bam > ./metrics/mutant_type_alignment_metrics.txt

Comparative Analysis: Peak distributions and enrichment metrics were compared between wild-type and mutant-type using DeepTools:

computeMatrix scale-regions -S wild_type_sorted.bam mutant_type_sorted.bam -R ./regions/TP53.bed -o ./comparative/matrix.gz

plotHeatmap -m ./comparative/matrix.gz -out ./comparative/heatmap.png