

# Chromcall Workflow Example

Bo Wang

2025-08-20

## Package & Data Preparation

The ‘chromcall’ package provides a workflow for integrative analysis of chromatin profiling data, quantifying and comparing chromatin mark enrichment and transcriptional activity across predefined genomic regions.

### 1. Load Input Files

All examples utilize bundled data included with the chromcall package.

- genome\_file: Chromosome metadata (name, length, circularity, genome build)
- region\_file: Genomic regions of interest (e.g., promoters)
- blacklist\_file (optional): Regions to exclude from analysis
- expr\_files (optional): Expression profiles per region
- bam\_files: BAM files for H3K4me3, H3K27me3, and matched controls

```
library(chromcall)

# File paths
genome_file    <- system.file("extdata", "genome.txt", package = "chromcall")
region_file    <- system.file("extdata", "example.bed", package = "chromcall")
blacklist_file <- system.file("extdata", "blacklist.bed", package = "chromcall")
expr_file1     <- system.file("extdata", "expression_sampleA.bed", package = "chromcall")
expr_file2     <- system.file("extdata", "expression_sampleB.bed", package = "chromcall")
bam_dir        <- system.file("extdata", package = "chromcall")
```

### 2. Preview Input Content

```
# Genome file
head(read.table(genome_file, header = FALSE))
```

```
##      V1      V2    V3   V4
## 1 chr1 248956422 FALSE hg38
## 2 chr2 242193529 FALSE hg38
## 3 chr3 198295559 FALSE hg38
## 4 chr4 190214555 FALSE hg38
## 5 chr5 181538259 FALSE hg38
## 6 chr6 170805979 FALSE hg38
```

```
# BED regions
head(read.table(region_file, header = FALSE))
```

```
##      V1      V2      V3
## 1 chr20 297560 297660
## 2 chr20 324560 324660
## 3 chr20 381290 381390
```

```
# Blacklist
head(read.table(blacklist_file, header = FALSE))
```

```
##      V1      V2      V3
## 1 chr1  628087 635923
## 2 chr1  789516 792613
## 3 chr1 2651095 2657690
## 4 chr1 2658436 2659484
## 5 chr1 2660040 2662057
## 6 chr1 2663196 2666377
```

```
# Expression data
head(read.table(expr_file1, header = FALSE))
```

```
##      V1      V2      V3
## 1 chr20 297610 12.5
## 2 chr20 324610  8.7
## 3 chr20 381340 16.2
```

```
head(read.table(expr_file2, header = FALSE))
```

```
##      V1      V2      V3
## 1 chr20 297610 12.5
## 2 chr20 324610  8.7
## 3 chr20 381340 16.2
```

```
# BAM files available
list.files(bam_dir, pattern = "sample[AB].*\\.bam$")
```

```
## [1] "control_sampleA.bam" "control_sampleB.bam" "h3k27me3_sampleA.bam"
## [4] "h3k27me3_sampleB.bam" "h3k4me3_sampleA.bam" "h3k4me3_sampleB.bam"
```

## Full Workflow

### Step 1 Construct Sample Objects

```
# Sample A
sampleA <- build_chromcall_sample(
  sample_name      = "SampleA",
  paired = F,
  experiments      = list(
    H3K27me3 = file.path(bam_dir, "h3k27me3_sampleA.bam"),
    H3K4me3  = file.path(bam_dir, "h3k4me3_sampleA.bam"),
    Control  = file.path(bam_dir, "control_sampleA.bam")
  ),
  control_name     = "Control",
  genome_file      = genome_file,
  region_file      = region_file,
  window_size      = 10000,
  blacklist_file    = blacklist_file,
  expression_file  = expr_file1
)

# Sample B
sampleB <- build_chromcall_sample(
  sample_name      = "SampleB",
  paired = F,
  experiments      = list(
    H3K27me3 = file.path(bam_dir, "h3k27me3_sampleB.bam"),
    H3K4me3  = file.path(bam_dir, "h3k4me3_sampleB.bam"),
    Control  = file.path(bam_dir, "control_sampleB.bam")
  ),
  control_name     = "Control",
  genome_file      = genome_file,
  region_file      = region_file,
  window_size      = 10000,
  blacklist_file    = blacklist_file,
  expression_file  = expr_file2
)
```

Inspect structure:

```
sampleA

## class: RangedSummarizedExperiment
## dim: 3 3
## metadata(3): sample_name control_name genome_tiles
## assays(1): counts
## rownames: NULL
## rowData names(2): blacklist expression
## colnames(3): H3K27me3 H3K4me3 Control
## colData names(3): name control lambda_g
```

```
SummarizedExperiment::assayNames(sampleA)
```

```
## [1] "counts"
```

```
SummarizedExperiment::colData(sampleA)
```

```
## DataFrame with 3 rows and 3 columns
##           name      control  lambda_g
##      <character> <logical>  <numeric>
## H3K27me3      H3K27me3      FALSE 0.00201308
## H3K4me3       H3K4me3       FALSE 0.00138967
## Control       Control       TRUE  0.00114291
```

```
head(SummarizedExperiment::assay(sampleA, "counts"))
```

```
##      H3K27me3 H3K4me3 Control
## [1,]        0        2        0
## [2,]        0        1        0
## [3,]        0        1        0
```

## Step 2 Run Region-Level Testing

```
resultA <- test_region_counts(sampleA)
resultB <- test_region_counts(sampleB)
```

Preview assay outputs:

```
SummarizedExperiment::assayNames(resultA)
```

```
## [1] "counts"  "logFC"   "lambda_t" "p_value" "p_adj"   "score"
```

```
SummarizedExperiment::colData(resultA)
```

```
## DataFrame with 3 rows and 3 columns
##           name      control  lambda_g
##      <character> <logical>  <numeric>
## H3K27me3      H3K27me3      FALSE 0.00201308
## H3K4me3       H3K4me3       FALSE 0.00138967
## Control       Control       TRUE  0.00114291
```

```
head(SummarizedExperiment::assay(resultA, "counts"))
```

```
##      H3K27me3 H3K4me3 Control
## [1,]        0        2        0
## [2,]        0        1        0
## [3,]        0        1        0
```

```
head(SummarizedExperiment::assay(resultA, "logFC"))
```

```
##      H3K27me3  H3K4me3 Control
## [1,]      0 7.651052      0
## [2,]      0 6.658211      0
## [3,]      0 6.658211      0
```

```
head(SummarizedExperiment::assay(resultA, "lambda_t"))
```

```
##      H3K27me3      H3K4me3      Control
## [1,] 0.002013079 0.001389674 0.001142909
## [2,] 0.002013079 0.001389674 0.001142909
## [3,] 0.002013079 0.001389674 0.001142909
```

```
head(SummarizedExperiment::assay(resultA, "p_value"))
```

```
##      H3K27me3      H3K4me3 Control
## [1,]      1 9.647022e-07      1
## [2,]      1 1.388708e-03      1
## [3,]      1 1.388708e-03      1
```

```
head(SummarizedExperiment::assay(resultA, "p_adj"))
```

```
##      H3K27me3      H3K4me3 Control
## [1,]      1 2.894107e-06      1
## [2,]      1 1.388708e-03      1
## [3,]      1 1.388708e-03      1
```

### Step 3 Export Single-Sample Results

```
write_experiment_results(resultA, experiment = "H3K4me3", file = "SampleA_H3K4me3.tsv")
head(read.delim("SampleA_H3K4me3.tsv"))
```

```
##  seqnames  start    end width counts_control counts_experiment  logFC
## 1   chr20 297560 297660   101          0              2 7.651052
## 2   chr20 324560 324660   101          0              1 6.658211
## 3   chr20 381290 381390   101          0              1 6.658211
##      score modification_factor test_lambda      pvalue      padj expression
## 1 7.463325                  1 0.001389674 9.647022e-07 2.894107e-06      12.5
## 2 6.470485                  1 0.001389674 1.388708e-03 1.388708e-03       8.7
## 3 6.470485                  1 0.001389674 1.388708e-03 1.388708e-03      16.2
##  significant
## 1      sig
## 2      ns
## 3      ns
```

## Step 4 Compare Two Samples

```
comparison <- compare_samples(resultA, resultB)
comparison
```

```
## GRanges object with 3 ranges and 17 metadata columns:
##      seqnames      ranges strand | SampleA_expression SampleA_H3K27me3_padj
##      <Rle>      <IRanges> <Rle> |      <numeric>      <numeric>
## [1]   chr20 297560-297660      * |          12.5          1
## [2]   chr20 324560-324660      * |           8.7          1
## [3]   chr20 381290-381390      * |          16.2          1
##      SampleA_H3K27me3_class SampleA_H3K27me3_score SampleA_H3K4me3_padj
##      <numeric>      <numeric>      <numeric>
## [1]              0          -0.264606          2.89411e-06
## [2]              0          -0.264606          1.38871e-03
## [3]              0          -0.264606          1.38871e-03
##      SampleA_H3K4me3_class SampleA_H3K4me3_score SampleB_expression
##      <numeric>      <numeric>      <numeric>
## [1]              1           7.46333          12.5
## [2]              1           6.47049           8.7
## [3]              1           6.47049          16.2
##      SampleB_H3K27me3_padj SampleB_H3K27me3_class SampleB_H3K27me3_score
##      <numeric>      <numeric>      <numeric>
## [1]              1              0          -0.19838
## [2]              1              0          -0.19838
## [3]              1              0          -0.19838
##      SampleB_H3K4me3_padj SampleB_H3K4me3_class SampleB_H3K4me3_score
##      <numeric>      <numeric>      <numeric>
## [1]              1              0          -0.172013
## [2]              1              0          -0.172013
## [3]              1              0          -0.172013
##      log2FC_expression H3K27me3_DeltaScore H3K4me3_DeltaScore
##      <numeric>      <numeric>      <numeric>
## [1]              0          0.0662258          -7.63534
## [2]              0          0.0662258          -6.64250
## [3]              0          0.0662258          -6.64250
## -----
## seqinfo: 25 sequences (1 circular) from hg38 genome
```

## Step 5 Export Comparison Results

```
write_comparison_results(comparison, "comparison_SampleA_vs_SampleB.tsv")
head(read.delim("comparison_SampleA_vs_SampleB.tsv"))
```

```
##      chr  start    end width SampleA_expression SampleA_H3K27me3_padj
## 1 chr20 297560 297660   101             12.5             1
## 2 chr20 324560 324660   101              8.7             1
## 3 chr20 381290 381390   101             16.2             1
##      SampleA_H3K27me3_class SampleA_H3K27me3_score SampleA_H3K4me3_padj
## 1                      0          -0.2646059          2.894107e-06
## 2                      0          -0.2646059          1.388708e-03
## 3                      0          -0.2646059          1.388708e-03
##      SampleA_H3K4me3_class SampleA_H3K4me3_score SampleB_expression
## 1                      1              7.463325             12.5
## 2                      1              6.470485              8.7
## 3                      1              6.470485             16.2
##      SampleB_H3K27me3_padj SampleB_H3K27me3_class SampleB_H3K27me3_score
## 1                      1                      0          -0.1983801
## 2                      1                      0          -0.1983801
## 3                      1                      0          -0.1983801
##      SampleB_H3K4me3_padj SampleB_H3K4me3_class SampleB_H3K4me3_score
## 1                      1                      0          -0.1720127
## 2                      1                      0          -0.1720127
## 3                      1                      0          -0.1720127
##      log2FC_expression H3K27me3_DeltaScore H3K4me3_DeltaScore
## 1                      0              0.0662258          -7.635338
## 2                      0              0.0662258          -6.642498
## 3                      0              0.0662258          -6.642498
```