

# CrispR\_2\_jieun.Rmd

Jieun Jeong

2025-04-19

Analysis pipeline:

```
require("CRISPRcleanR")
data("KY_Library_v1.0")
wd <- "/Users/jieun/Work/Crispr"
wdf <- function(...) file.path(wd,...)
fn <- file.path( system.file("extdata", package = "CRISPRcleanR"),
                 "HT-29_counts.tsv")
ccr.AnalysisPipeline(
  file_counts = fn,
  outdir=wdf('/HT29_pipeline/'), EXPname = 'HT29',
  library_builtin = "KY_Library_v1.0",
  run_mageck = FALSE,
  ncontrols = 1
)
```

```
## [1] "#####"
## [1] "Input parameters"
## [1] "#####"
## [1] "Parameter EXPname: HT29 (class: character-character)"
## [1] "Parameter outdir: /Users/jieun/Work/Crispr//HT29_pipeline/ (class: character-character)"
## [1] "Parameter ncontrols: 1 (class: numeric-double)"
## [1] "Parameter min_reads: 30 (class: numeric-double)"
## [1] "Parameter method: ScalingByTotalReads (class: character-character)"
## [1] "Parameter FDRth: 0.05 (class: numeric-double)"
## [1] "Parameter library_builtin: KY_Library_v1.0 (class: character-character)"
## [1] "Parameter file_counts: /Library/Frameworks/R.framework/Versions/4.4-arm64/Resources/library/CRISPRcleanR/extdata/HT-29_counts.tsv (class: character-character)"
## [1] "Parameter aligner: Rsubreads (class: character-character)"
## [1] "Parameter maxMismatches: 0 (class: numeric-double)"
## [1] "Parameter nBestLocations: 2 (class: numeric-double)"
## [1] "Parameter nTrim5: 0 (class: character-character)"
## [1] "Parameter nTrim3: 0 (class: character-character)"
## [1] "Parameter strand: F (class: character-character)"
## [1] "Parameter duplicatedSeq: keep (class: character-character)"
## [1] "Parameter nthreads: 1 (class: numeric-double)"
## [1] "Parameter indexMemory: 2000 (class: numeric-double)"
## [1] "Parameter fastqc_plots: FALSE (class: logical-logical)"
## [1] "Parameter min.ngenes: 3 (class: numeric-double)"
## [1] "Parameter alpha: 0.01 (class: numeric-double)"
## [1] "Parameter nperm: 10000 (class: numeric-double)"
## [1] "Parameter p.method: hybrid (class: character-character)"
```

```
## [1] "Parameter min.width: 2 (class: numeric-double)"
## [1] "Parameter kmax: 25 (class: numeric-double)"
## [1] "Parameter nmin: 200 (class: numeric-double)"
## [1] "Parameter eta: 0.05 (class: numeric-double)"
## [1] "Parameter trim: 0.025 (class: numeric-double)"
## [1] "Parameter undo.splits: none (class: character-character)"
## [1] "Parameter undo.prune: 0.05 (class: numeric-double)"
## [1] "Parameter undo.SD: 3 (class: numeric-double)"
## [1] "Parameter run_mageck: FALSE (class: logical-logical)"
## [1] "Parameter path_to_mageck: mageck (class: character-character)"
## [1] "Parameter nseed: 679661 (class: numeric-double)"

## NULL
```

Analysis results in outdir can be used for gene set analysis etc., but one may also check what numbers support those results.

```
dr <- wdf("HT29_pipeline/data")
show_gene <- function(gene, data_dir = dr) {
  cat("gene summary\n")
  gs <- read.delim(file.path(data_dir, "gene_summary.tsv"), sep = "\t")
  gs <- gs[gs$gene == gene, ]
  print(gs)
  cat("\nrow counts\n")
  gs <- read.delim(file.path(data_dir, "raw_counts.tsv"), sep = "\t")
  gs <- gs[gs$gene == gene, ]
  print(gs[,2:ncol(gs)], row.names=FALSE)
  gs <- read.delim(file.path(data_dir, "counts_corrected.tsv"), sep = "\t")
  gs <- gs[gs$gene == gene, ]
  cat("\nnormalized counts\n")
  print(gs[,2:ncol(gs)], row.names=FALSE)
}
show_gene("CDC27")
```

```
## gene summary
##           logFC SigDep
## CDC27 -5.487711  TRUE
##
## row counts
##   gene ERS717283.plasmid HT29_c904R1 HT29_c904R2 HT29_c904R3
## CDC27           336           0           0           14
## CDC27           460           33           0           0
## CDC27           155           0           0           18
## CDC27           323           10          12           19
## CDC27           173           10           6           9
##
## normalized counts
##   gene ERS717283.plasmid HT29_c904R1 HT29_c904R2 HT29_c904R3
## CDC27           87.48764  0.000000  0.000000  2.797876
## CDC27          119.77474  7.507912  0.000000  0.000000
## CDC27           40.35888  0.000000  0.000000  3.597269
## CDC27           84.10270  2.275125  2.716059  3.797118
## CDC27           45.04572  2.275125  1.358029  1.798635
```