

RNA-seq analysis with DESeq2: : CHEAT SHEET



Background

The **DESeq2** package provides tools that enable you to:

1. Perform **differential expression analysis** of count data produced by RNA sequencing
2. Effectively **visualize** these count data

Installation

Run the following to install DESeq2 from **Bioconductor**

```
if (!require("BiocManager", quietly = TRUE))
  install.packages("BiocManager")

BiocManager::install("DESeq2")
```

Load DESeq2

```
library(DESeq2)
```

Create dds object

Create a **DESeqDataSet** (dds) object from counts data stored in a **counts matrix**

```
dds <- DESeqDataSetFromMatrix(
  countData = counts,
  colData = coldata,
  Design = ~ Condition
)
```

Counts matrix

Sample information

Design formula

Inspect dds object

counts(dds)

View counts matrix (only first three rows and columns are shown here)

| | sample1 | sample2 | sample3 |
|-----------------|---------|---------|---------|
| ENSG00000223972 | 0 | 0 | 0 |
| ENSG00000227232 | 14 | 28 | 17 |
| ENSG00000278267 | 8 | 4 | 3 |

colData(dds)

View sample information

| | Cell | Type | Condition |
|---------|-------|------------|-----------|
| sample1 | Cell1 | Epithelial | mock |
| sample2 | Cell1 | Epithelial | mock |
| sample3 | Cell1 | Epithelial | mock |
| sample4 | Cell1 | Epithelial | KD |
| sample5 | Cell1 | Epithelial | KD |
| sample6 | Cell1 | Epithelial | KD |

Subset dds object

dds[1:5,]

Choose first 5 genes

dds[, 1:3]

Choose first 3 samples

dds[, dds\$Condition == "mock"]

Choose samples that belong to "mock" condition

Sample PCA plots

Apply variance stabilizing transformation

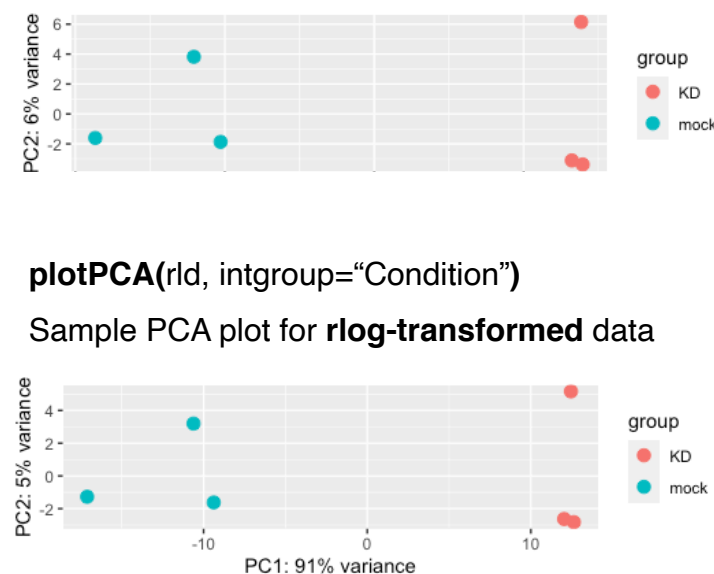
```
vsd <- vst(dds, blind=FALSE)
```

Apply regularized log transformation

```
rld <- rlog(dds, blind=FALSE)
```

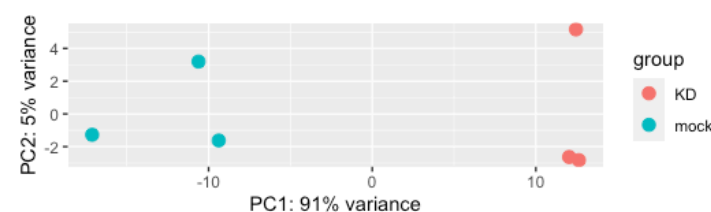
plotPCA(vsd, intgroup="Condition")

Sample PCA plot for **vst-transformed** data



plotPCA(rld, intgroup="Condition")

Sample PCA plot for **rlog-transformed** data



Differential Expression

Perform estimation of size factors, estimation of dispersion, Negative Binomial GLM fitting and Wald statistics

```
dds <- DESeq(dds)
```

results(dds)

Extract a results table

Extract a results table, specifying the numerator and denominator for fold-changes

```
res <- results(
  dds,
  contrast = c("Condition", "KD",
               "mock")
)
```

mcols(res)\$description

Print descriptions of results table columns

Visualize Results

plotMA(res)

Plot log fold changes versus mean expression (MA plot)

Log fold change shrinkage

Performs log fold change shrinkage

```
res_shrunk <- lfcShrink(
  dds,
  coef = 2,
  type = "apeglm"
)
```

Coefficient of interest

dds object

Type of shrinkage

MA plot

