

Taller 3A

Estadística genómica

Juan David Henao Sánchez

23 de septiembre de 2015

Sobre unos datos del GEO (una sola medición en el tiempo) que comparen dos condiciones, identifique los genes diferencialmente expresados usando acde:

```
> library(acde)
> library("DESeq")
> #####
> data <- read.table("GSE58972_RELA_6h_processed_data.txt", h=T)
> dim(data)

[1] 23284      8

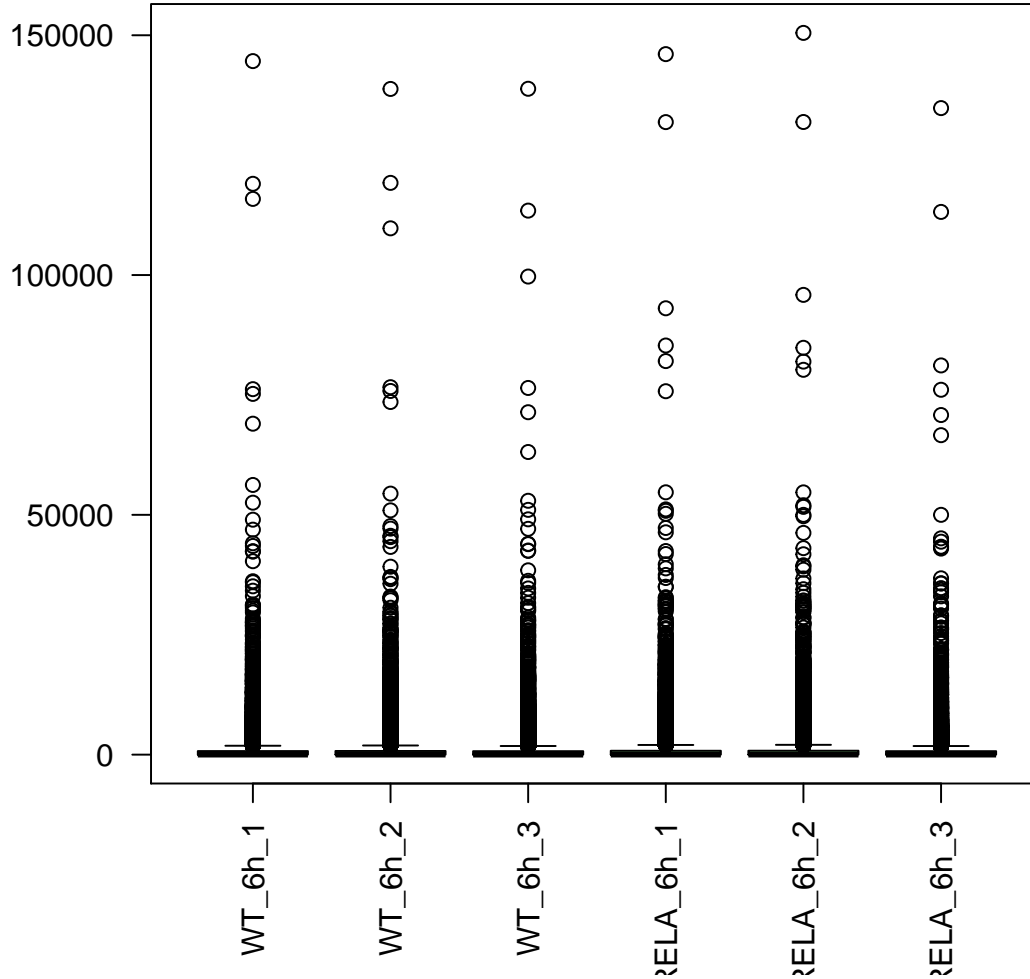
> head(data)
```

	gene description	WT_6h_1	WT_6h_2	WT_6h_3	RELA_6h_1	RELA_6h_2
1	0610005C13Rik	na	9	12	11	16
2	0610007C21Rik	na	534	581	512	537
3	0610007L01Rik	na	2124	2197	2168	2811
4	0610007N19Rik	na	17	26	21	18
5	0610007P08Rik	na	1410	1504	1543	1546
6	0610007P14Rik	na	2049	2008	2074	2192

```
  RELA_6h_3
1         10
2        501
3       2265
4         15
5       1353
6       1968

> boxplot(data[, 3:8], main="Boxplot Mus musculus", col="lightgreen", cex.names=0.2, las=2)
```

Boxplot Mus musculus



```
> rownames(data) <- data$gene
> countsTable <- data[,3:8]
> head(countsTable)
```

	WT_6h_1	WT_6h_2	WT_6h_3	RELA_6h_1	RELA_6h_2	RELA_6h_3
0610005C13Rik	9	12	12	11	16	10
0610007C21Rik	534	581	562	512	537	501
0610007L01Rik	2124	2197	2168	2811	2706	2265
0610007N19Rik	17	26	21	18	20	15
0610007P08Rik	1410	1504	1543	1546	1577	1353
0610007P14Rik	2049	2008	2074	2192	2192	1968

```
> conds <- factor( c( "S1", "S2", "S3", "M1","M2","M3"))
> newcds <- newCountDataSet( countsTable, conds )
```