

Analisis de datos de microarrays

Estudio de datos

Preparación de los datos para el analisis

En la carpeta GSE134178_CEL se encuentran 8 archivos binarios en formato .CEL con los datos en crudo del estudio. 4 De ellos correspondientes a muestras tratadas para inhibir los factores de necrosis tumoral (TNF_KO) y 4 muestras de control.

También existe un archivo metadata con una asociación entre los nombres de los archivos y el grupo al que pertenecen.

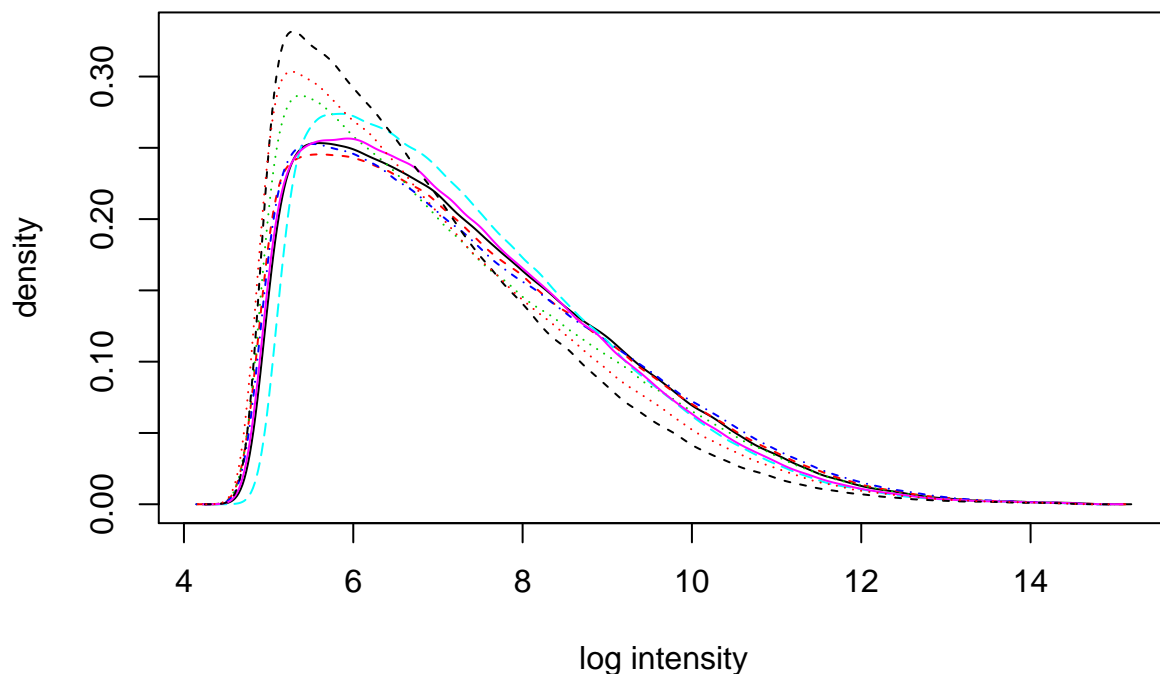
```
setwd('./data');  
metadata <- read_csv("/Users/iagolast/Dropbox/master/ADO/PEC_1/data/metadata.csv", col_types = cols(Grou  
metadata$sampleNames <- metadata$FileName  
rawData <- ReadAffy(phenoData = AnnotatedDataFrame(metadata))
```

Vamos a analizar los datos en crudo para comprobar la presencia de posibles errores.

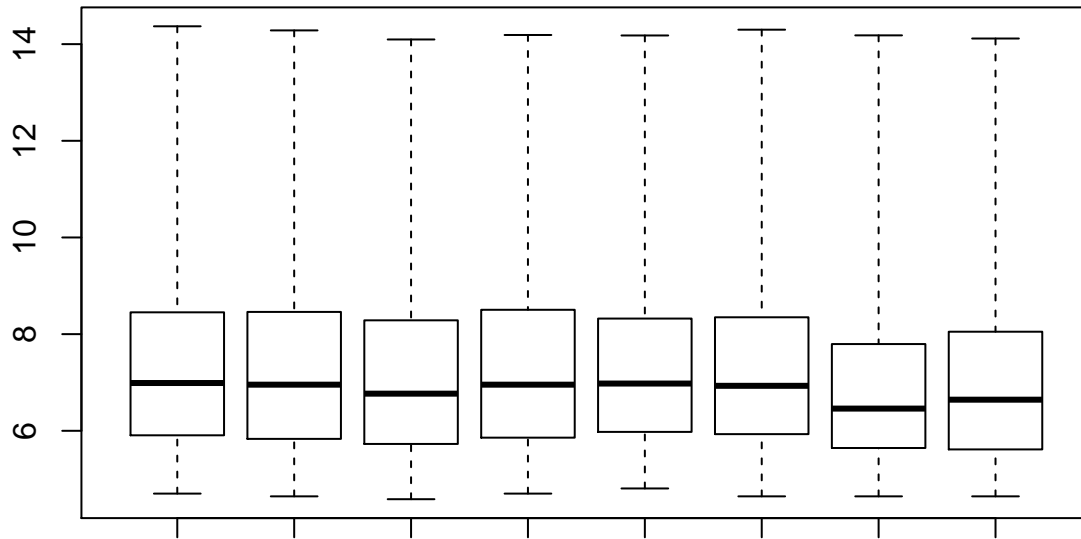
```
setwd('./data');  
rawData <- ReadAffy()
```

```
hist(rawData);
```

```
##
```



```
boxplot(rawData);
```



GSM3938443_TNF_KO_CEL.CEL

GSM3938448_TNF_OK.CEL

Gráficamente podemos apreciar que las muestras presentan una distribución similar de intensidad por lo que a priori no sospechamos de muestras corruptas. Vamos a utilizar una librería específica para asegurarnos.

A continuación pasamos a normalizar los datos

```
eset_rma <- affy::rma(rawData)
```

```
## Background correcting  
## Normalizing  
## Calculating Expression
```