

# Tercer Taller

## Estadística genómica

Juan David Henao Sánchez

22 de septiembre de 2015

**Sobre datos del GEO del NCBI de su elección (que comparen dos condiciones biológicas con al menos 5 réplicas) realice los siguientes pasos luego de normalizar:**

### 1. Realice un MA-plot

```
> library(GEOquery)
> library(vsn)
> gds <- getGEO("GDS3750")
> eset <- GDS2eSet(gds, do.log2 = TRUE)
> eset

ExpressionSet (storageMode: lockedEnvironment)
assayData: 22277 features, 8 samples
  element names: exprs
protocolData: none
phenoData
  sampleNames: GSM430339 GSM430340 ... GSM430346 (8 total)
  varLabels: sample genotype/variation description
  varMetadata: labelDescription
featureData
  featureNames: 1007_s_at 1053_at ... AFFX-TrpnX-M_at (22277 total)
  fvarLabels: ID Gene title ... GO:Component ID (21 total)
  fvarMetadata: Column labelDescription
experimentData: use 'experimentData(object)'
  pubMedIds: 20395301
Annotation:

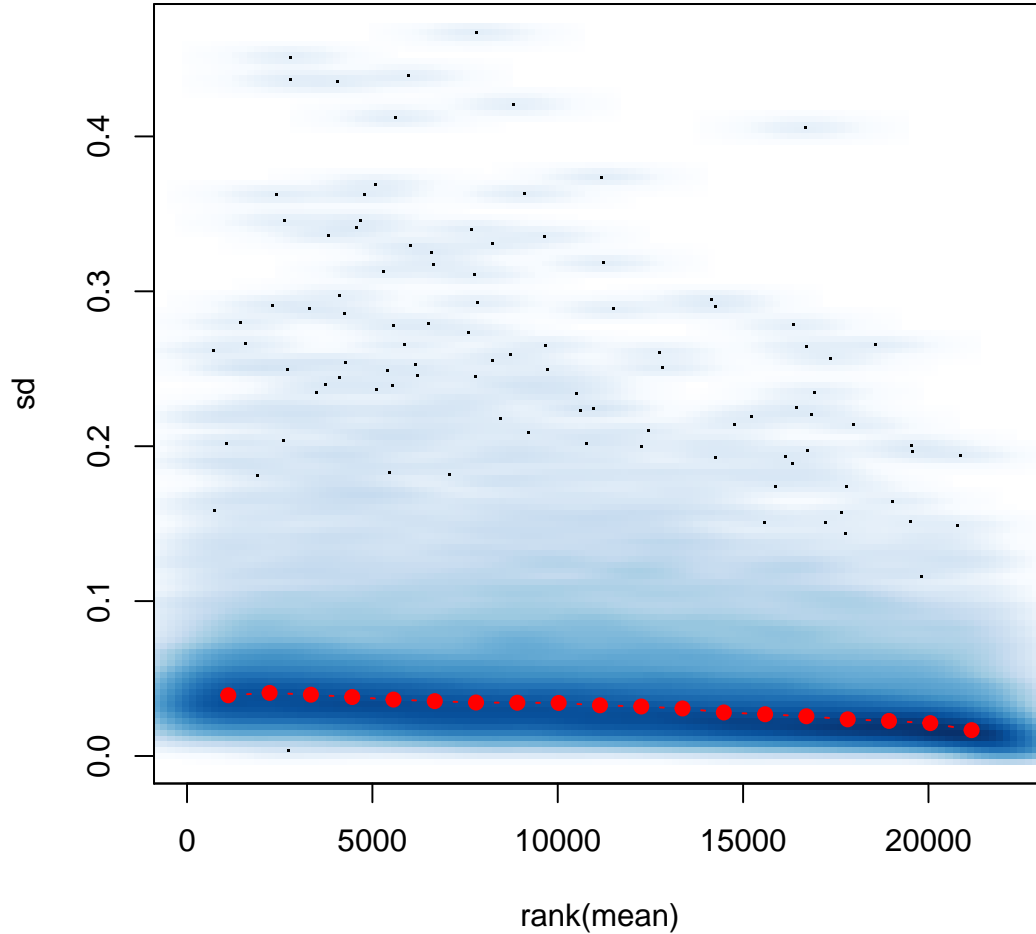
> dim(eset)

Features  Samples
  22277      8

> nml <- justvsd(eset)

> meanSdPlot(eset)
> title(main='Datos normalizados',font=2)
```

## Datos normalizados



### MA PLOT

```
> iref = seq(1, 7, by=2)
> ismp = seq(2, 8, by=2)
> M= exprs(nml)[,ismp]-exprs(nml)[,iref] #log ratios
> A=(exprs(nml)[,ismp]+exprs(nml)[,iref])/2 #promedio
> dim(M)

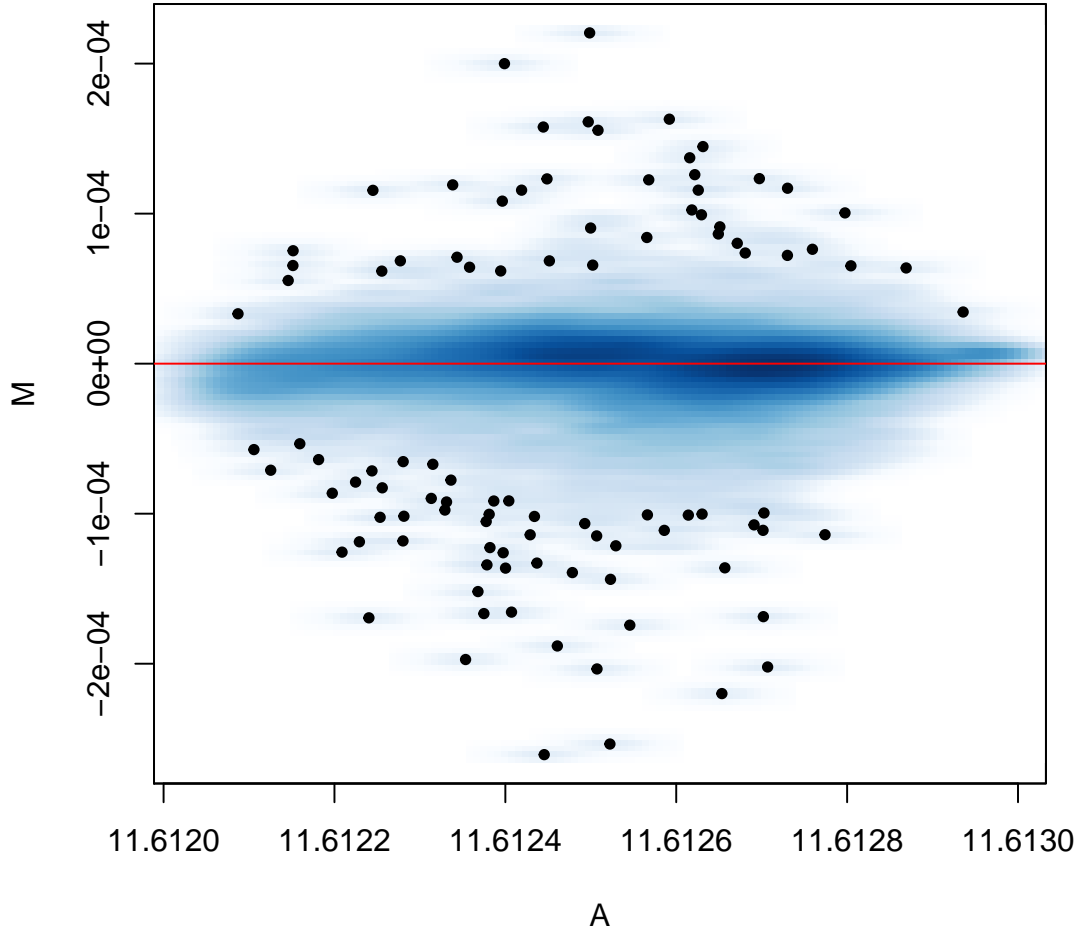
[1] 22277      4

> dim(A)

[1] 22277      4

> smoothScatter(rowMeans(A), rowMeans(M), main=" ", xlab="A", ylab="M", pch=20)
> title(main='MA PLOT',font=2)
> abline(h=0, col="red")
```

## MA PLOT



2. Identifique los genes diferencialmente expresados con pruebas T múltiples (rowttest) y resáltelos en el MAplot