



# **Universidad Autónoma del Estado de Hidalgo**

*Doctorado en Ciencias Económico Administrativas*

## **Temas Selectos I: Estadística para las CEA**

Complejidad Económica | Lab 43

**CUADERNO R.MARKDOWN**

Académico:

Dra. Carla Carolina Pérez Hernández

Alumna:

Ana Griselda Sanjuan Merida

263501

Fecha de entrega: 10 de marzo de 2023.



# LAB 43 (MD)

AnaGSanjuanM

2023-03-10

—————LABORATORIO 43—————

—————PRETTYHEATMAP—————

—ACADÉMICO: DRA. CARLA CAROLINA PÉREZ HERNÁNDEZ—

————ALUMNA: ANA GRISEL SANJUAN MERIDA————

Laboratorio: Mapa de calor térmico

Datos genicos tomados de Sahir Bhatnagar

Objetivo: Realizar un heatmap con datos geneticos

En este ejercicio vamos a:

1. Cargar matriz hipotética de datos y dataframes adicionales
2. Realizar varios heatmaps

El mapa de calor es una representación gráfica de datos que utiliza un sistema de codificación de colores para representar diferentes valores

Instalar paquetería: `install.packages("pheatmap")`

Abrir la librería

```
library(pheatmap)
```

Importar datos

Como son tres archivos que se ocuparán, debo activar las tres rutas

`file.choose`

El primer archivo se convertirá en matriz

Objeto matricial llamado genes que será leída como matriz

Está en formato csv.

Está separado por comas `sep=","`

En el encabezado están los títulos `header=T`

En la primera columna están los nombres de los genes `row.names=1`

```
genes <- as.matrix(  
  read.csv("C:\\Users\\Lenovo\\Documents\\GitHub\\LAB-43\\L43 Input\\heatmap_data.csv",  
           sep=",",  
           header=T,  
           row.names=1))
```

En el environment se ve el objeto generado genes

Los otros dos archivos serán dataframes

Objeto llamado annotation\_row

Está en formato csv.

En el encabezado están los títulos header=T

En la primera columna están los nombres de los genes row.names=1

```
annotation_row <- read.csv("C:\\Users\\Lenovo\\Documents\\GitHub\\LAB-43\\L43 Input\\annotation_row.csv",  
                           row.names=1)
```

Objeto llamado annotation\_col

Está en formato csv.

En el encabezado están los títulos header=T

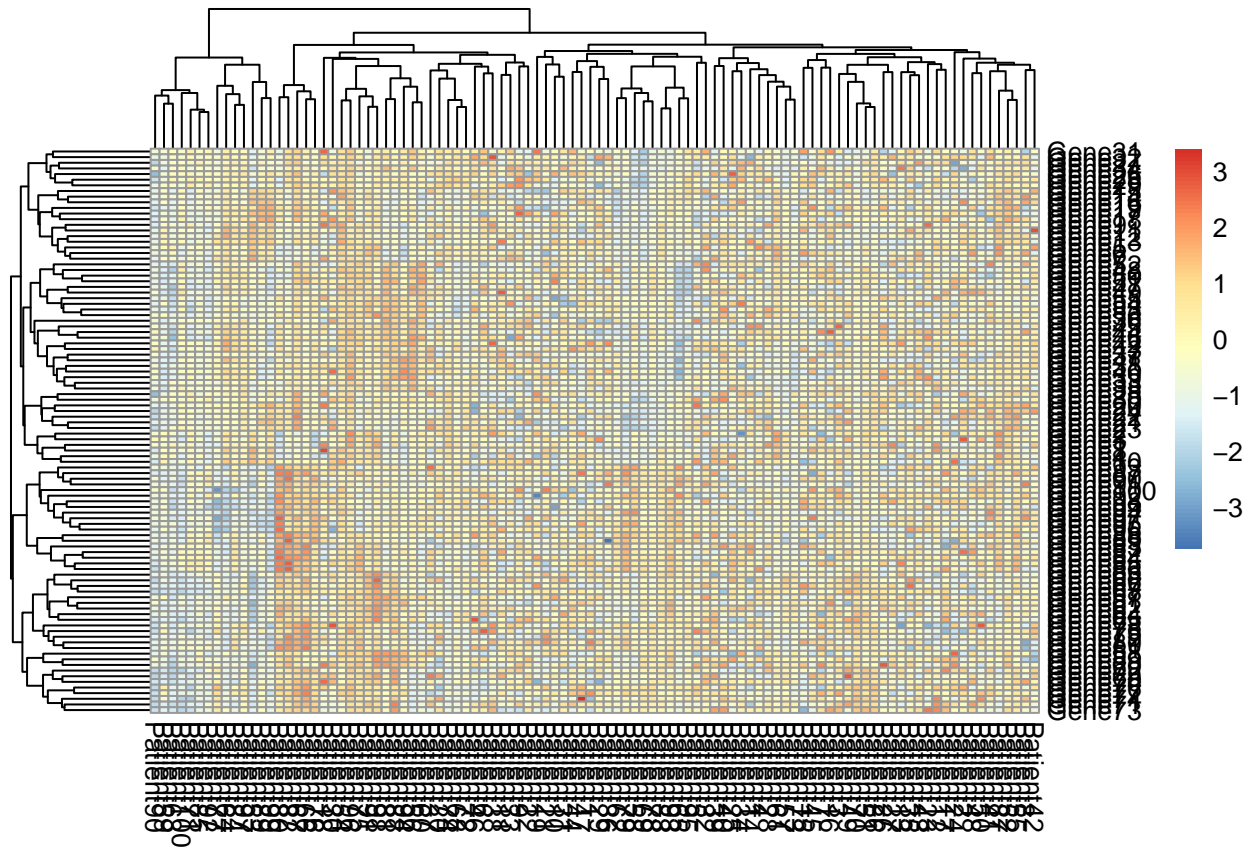
En la primera columna están los nombres de los genes row.names=1

```
annotation_col <- read.csv("C:\\Users\\Lenovo\\Documents\\GitHub\\LAB-43\\L43 Input\\annotation_col.csv",  
                           row.names=1)
```

Dibujando el heatmap

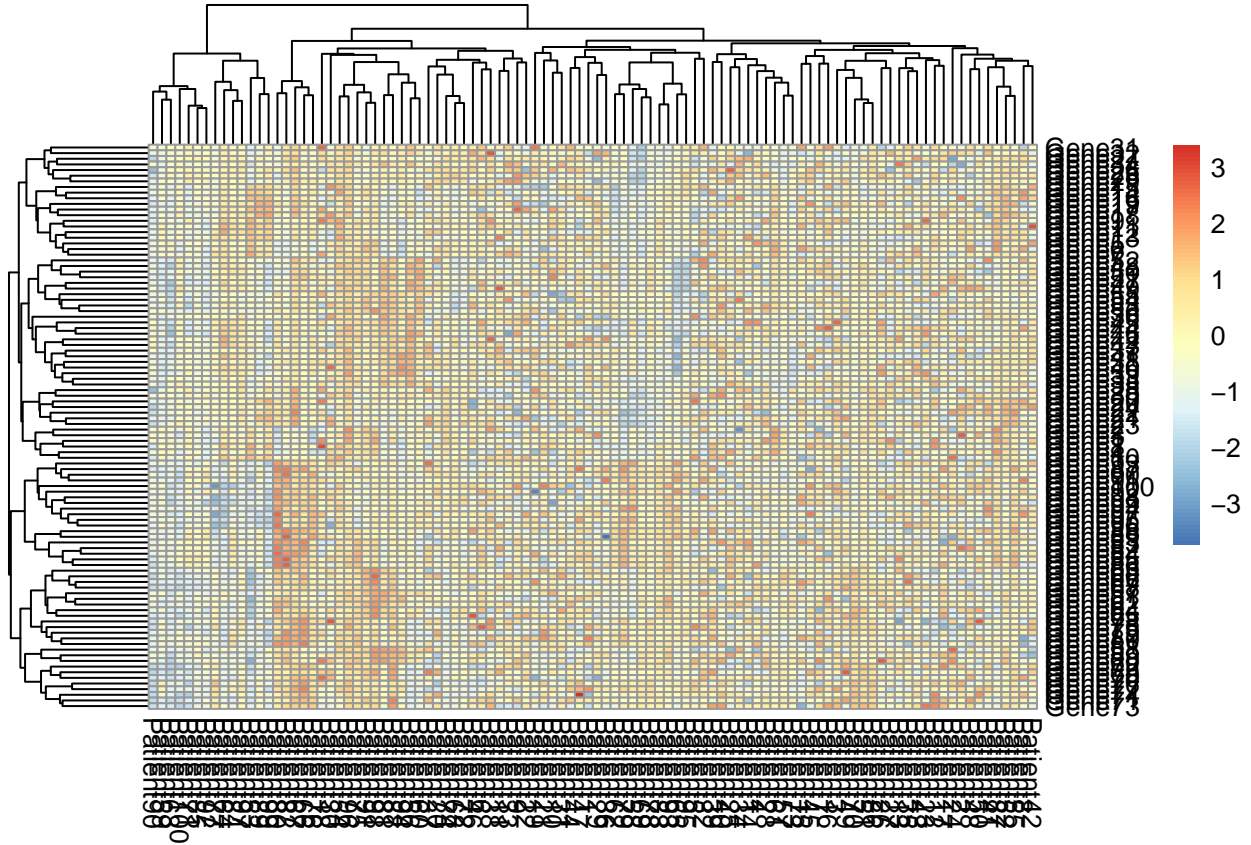
Comando pheatmap y la matriz genes

```
pheatmap(genes)
```



Cambiando el tamaño de la fuente (los cuadritos)

```
heatmap(genes, fontsize=6)
```

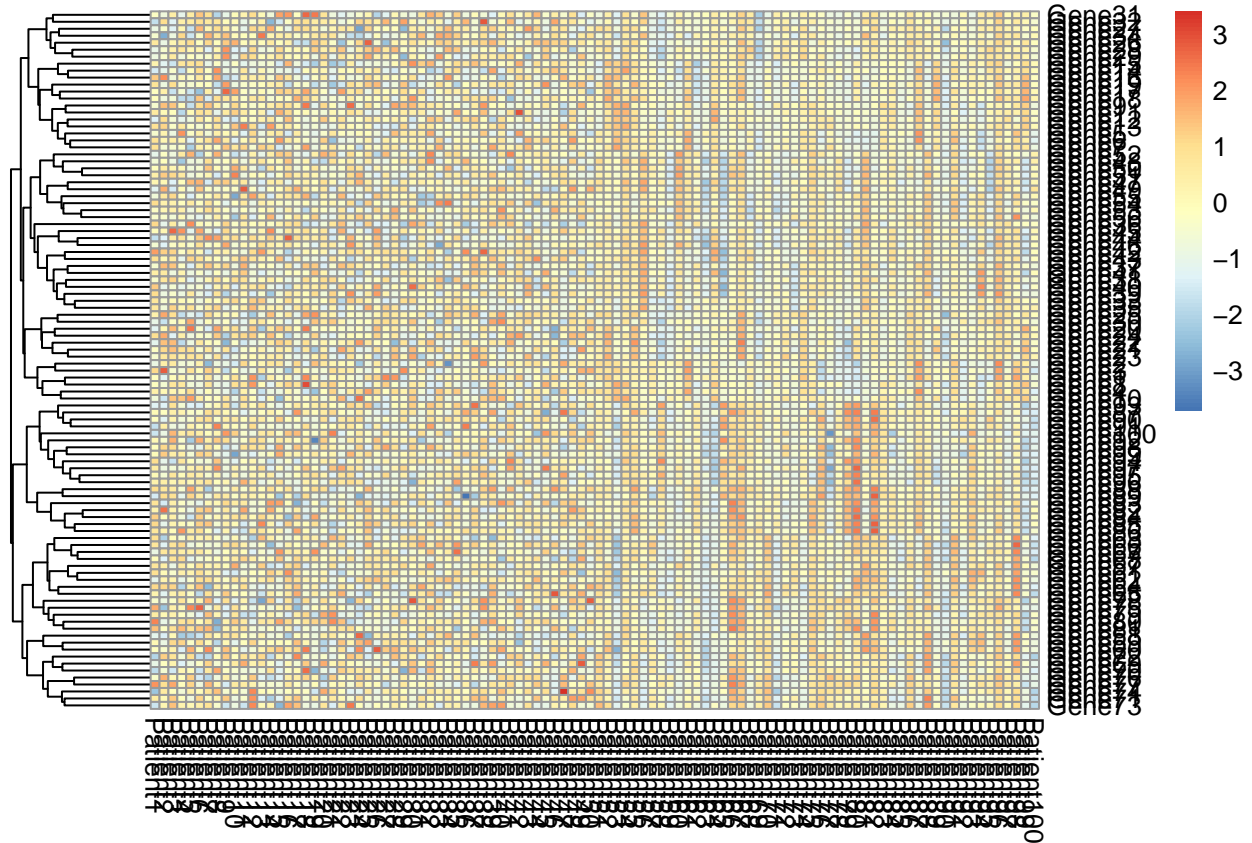


Quitando uno de los dendrogramas

Se quitará el dendrograma de las columnas, por lo que los renglones si interesan (true)

Como las columnas de pacientes (eje de las X) no estarán, ponemos False

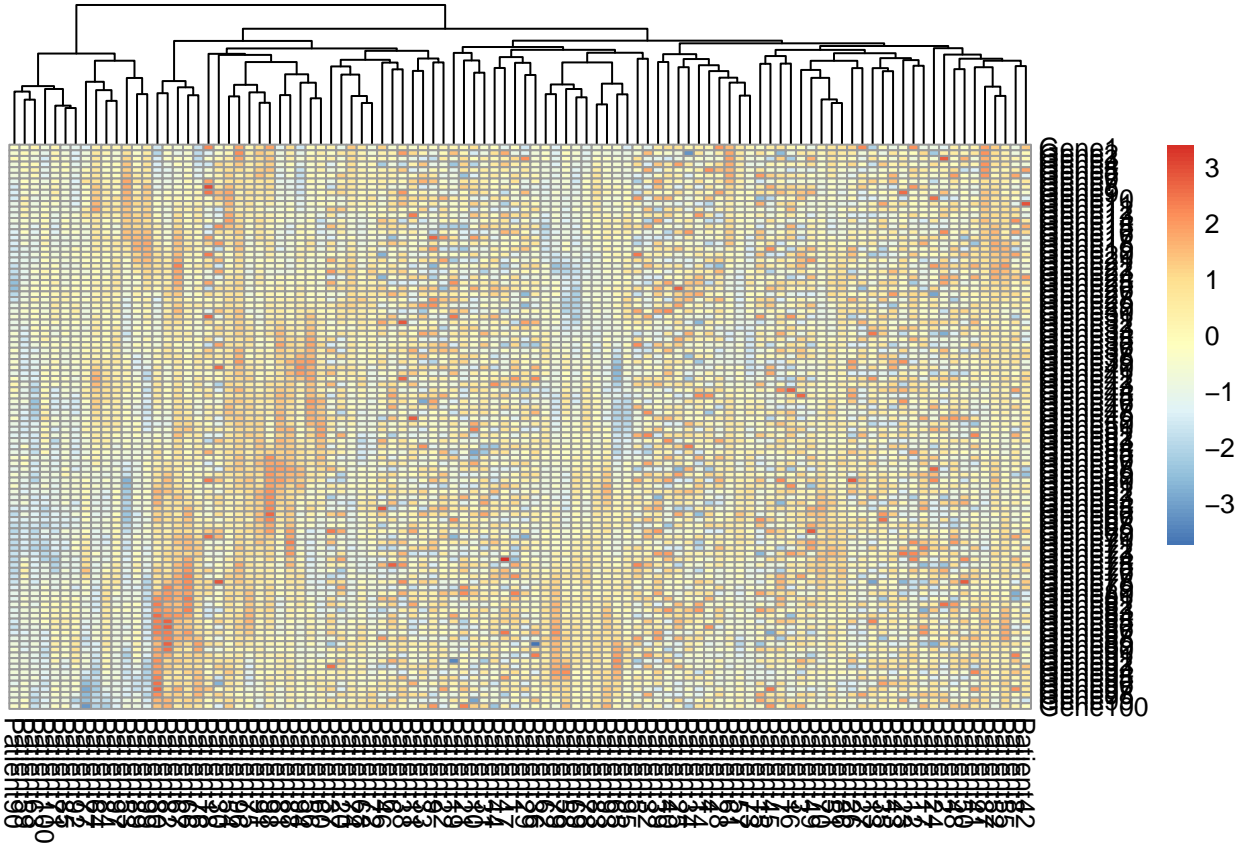
```
ph heatmap(genes, fontsize=6, cluster_rows = T, cluster_cols = F)
```



Se quitará el dendrograma de los renglones, por lo que las columnas si interesan (true)

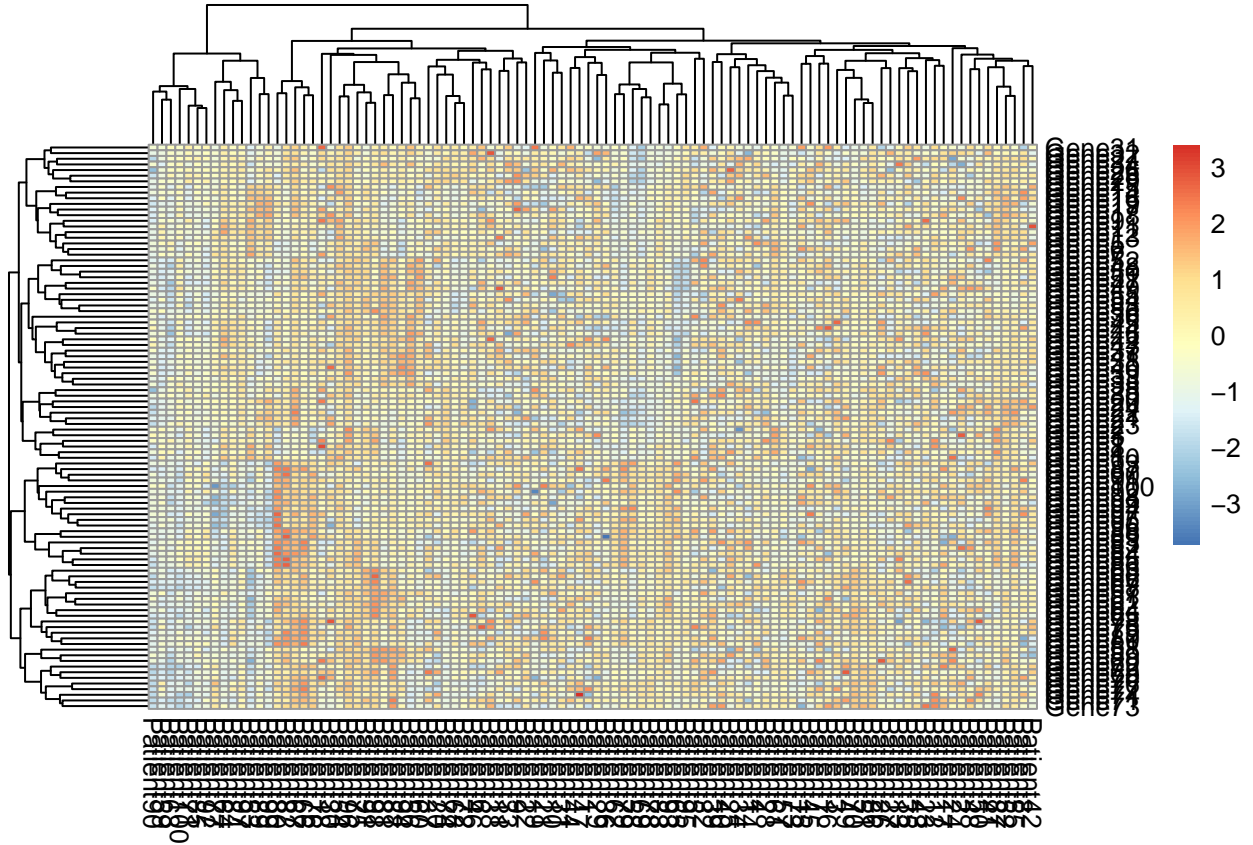
Como los renglones de genes (eje de las Y) no estarán, ponemos False

```
pheatmap(genes, fontsize=6, cluster_rows = F, cluster_cols = T)
```



Para que se muestren nuevamente los dendrogramas en ambos ejes

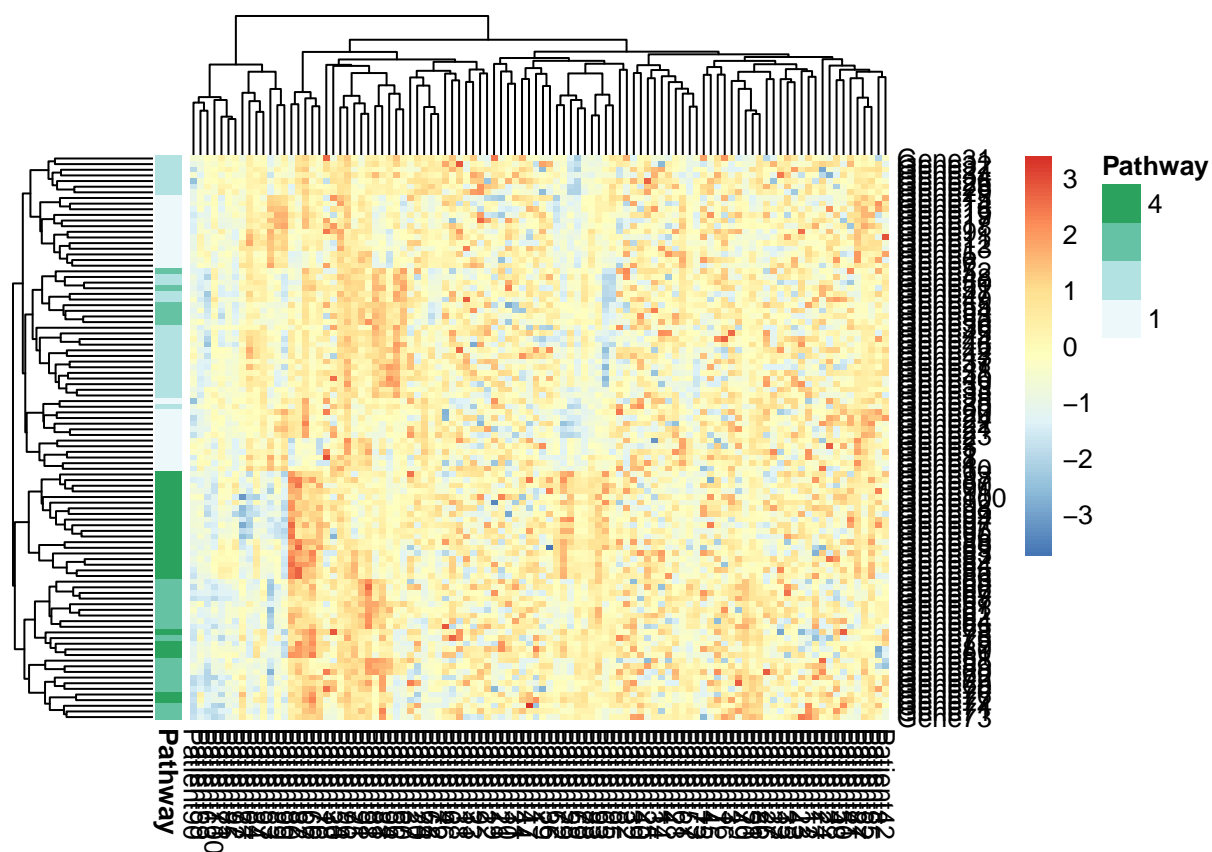
```
heatmap(genes, fontsize=6, cluster_rows = T, cluster_cols = T)
```



Identificar si hay patrones sibyacentes a las anotaciones de columnas y renglones

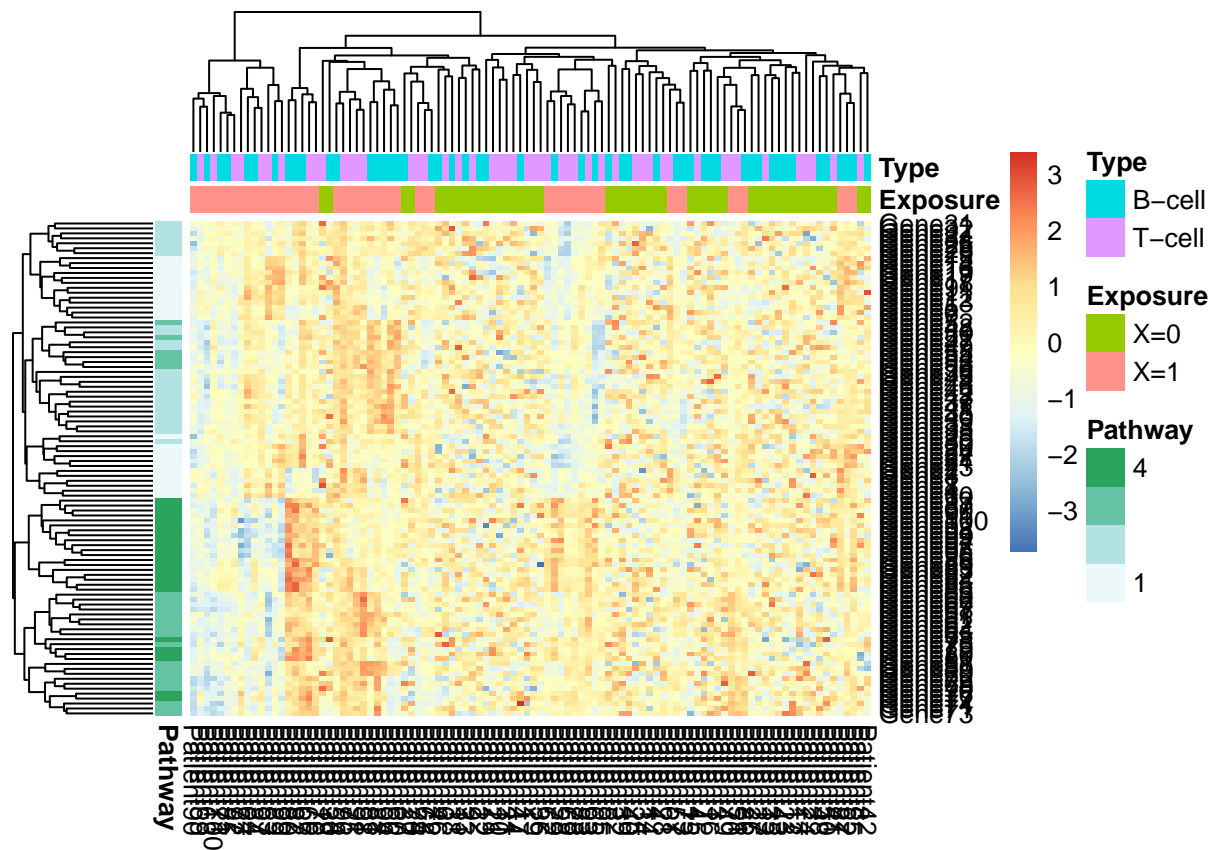
```
heatmap(genes, fontsize=6, cluster_rows = T, cluster_cols = T, annotation_row = annotation_row)
```





Para añadir anotaciones en las columnas

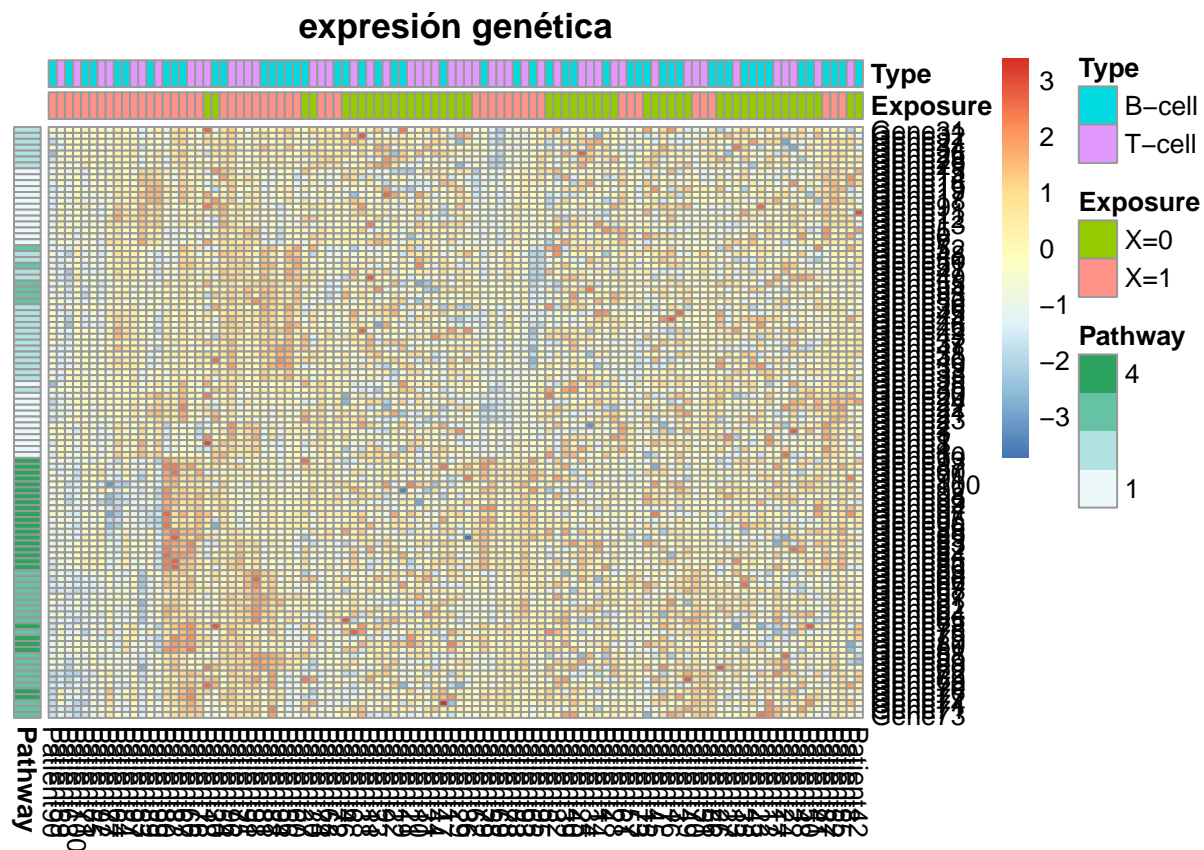
```
heatmap(genes, fontsize=6, cluster_rows = T, cluster_cols = T, annotation_row = annotation_row, annot
```



Generar gráfico de manera completa

Se quita dendrograma de renglones

```
heatmap(genes, fontsize=6, cluster_rows = T, cluster_cols = T, annotation_row = annotation_row, annot
```



Tomando una pequeña muestra de la base general para crear un subset

El subconjunto de datos proviene de la base matricial llamada genes

Extraerá ciertos datos de un vector (del gen 1 al gen 5)

Interesándonos por los pacientes del 55 al 60

También, del mismo vector, extraer del gen 1 al 5 y pacientes del 20 al 35

También, del mismo vector, extraer pacientes del 55 al 60

```
sub <- genes [c(1:5, 55:60), c(1:5, 20:35, 55:60)]
```

Generar mapa de calor del subconjunto llamado sub

```
pheatmap(sub, fontsize=6, cluster_rows = T, cluster_cols = T, annotation_row = annotation_row, annotat
```



Desplegar valores del gráfico recién obtenido

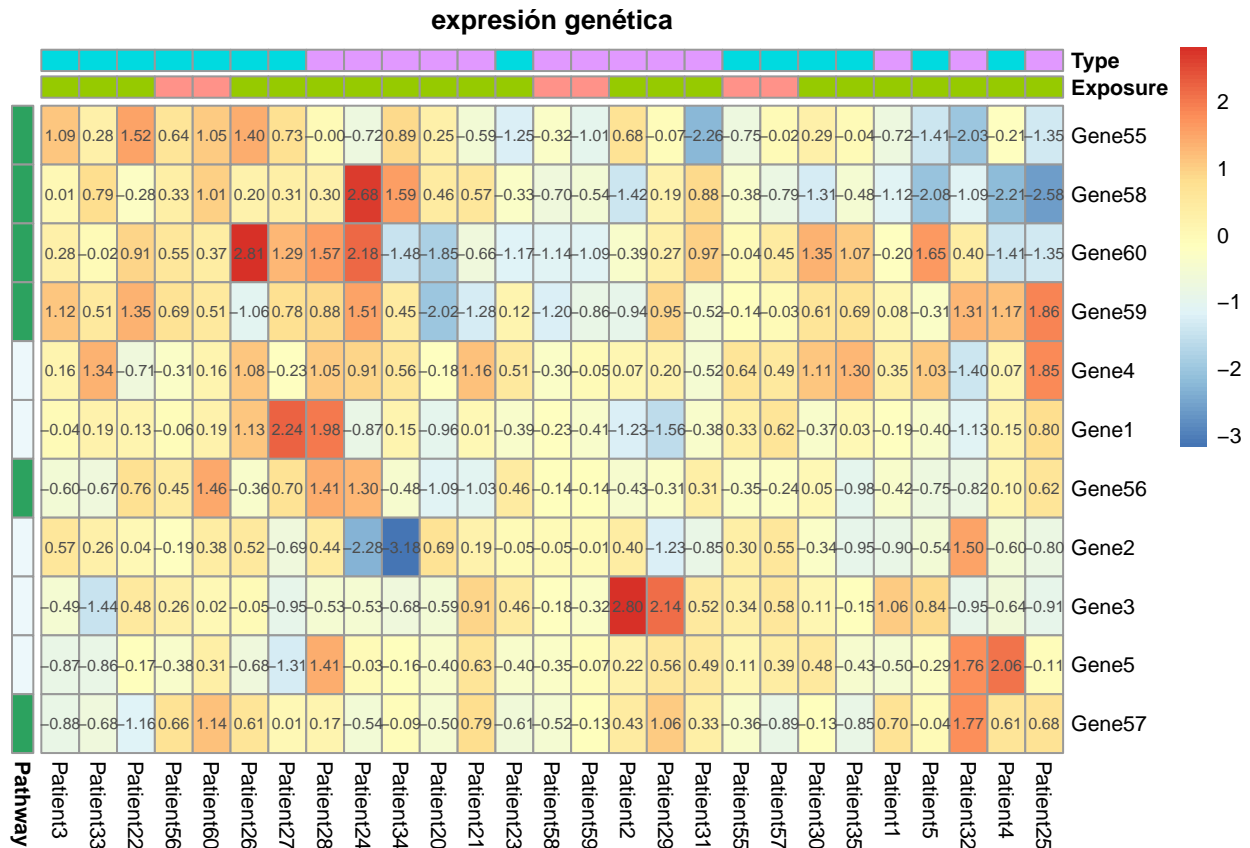
Del mapa de calor del subconjunto llamado sub el tamaño será de 8

Las anotaciones de la leyenda será falsa

Desplegar los números (True)

El tamaño de los números será de 6

```
ph heatmap(sub, fontsize=6, cluster_rows = T, cluster_cols = T, annotation_row = annotation_row, annotat
```



Para añadir elemento estético adicional

Cargamos paquetería para cambiar color:

```
install.packages("viridis")
```

```
install.packages("viridisLite")
```

Abrimos la librería

```
library(viridis)
```

```
## Loading required package: viridisLite
```

Hay cuatro paletas de colores: magma, plasma, cividis e inferno

Volviendo a cargar pheatmap

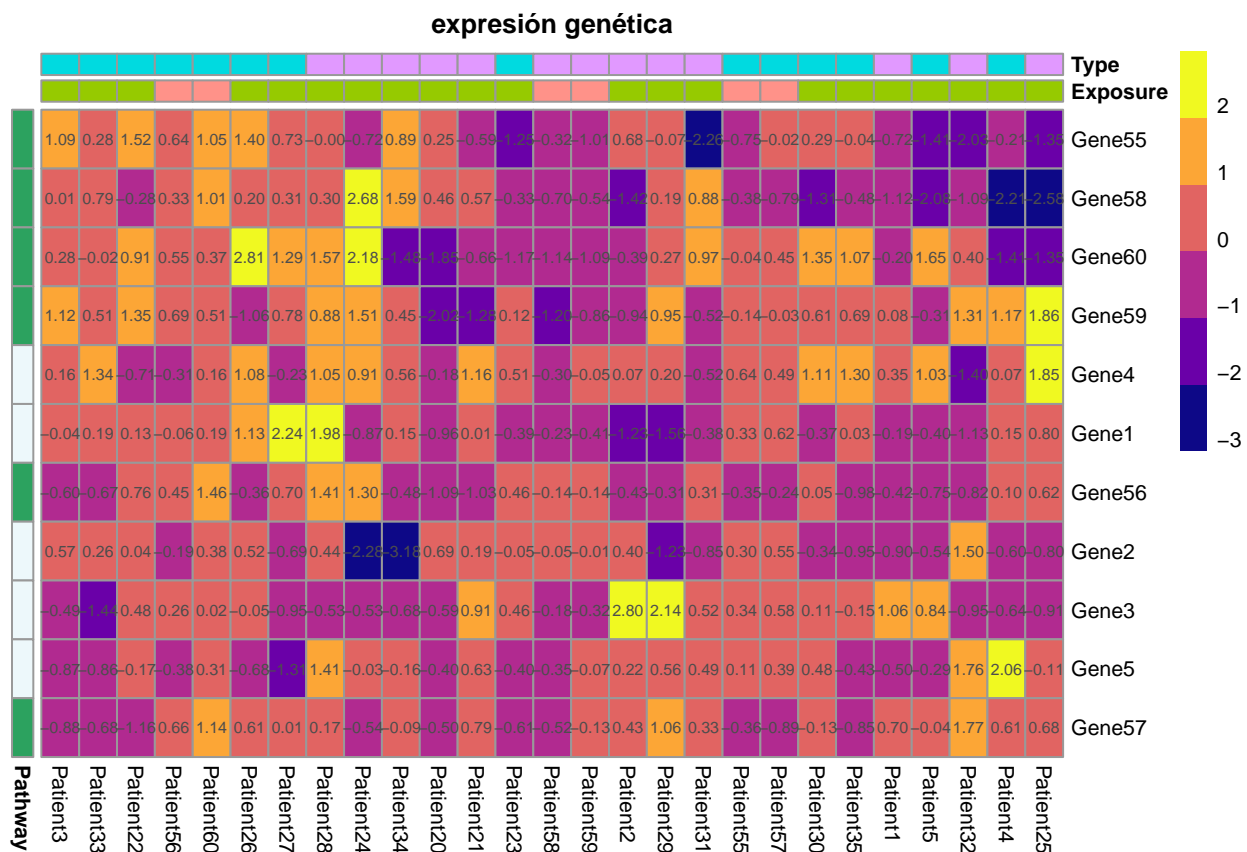
```
library(pheatmap)
```

Uso de paletas de colores para una mejor estética

Se usará paleta viridis y la opción plasma

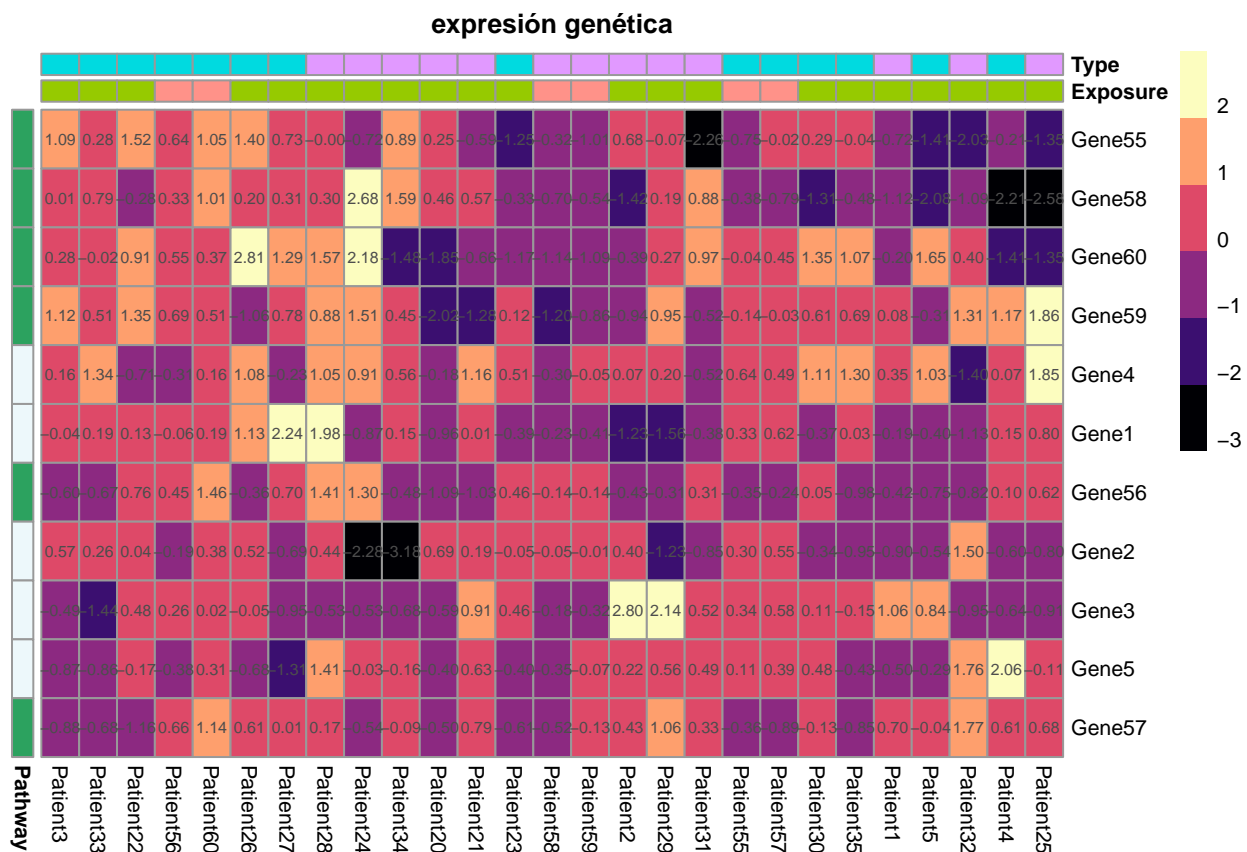
Y seis facetas diferentes

```
pheatmap(sub, fontsize=6, cluster_rows = T, cluster_cols = T, annotation_row = annotation_row,
          annotation_col = annotation_col, treeheight_row = 0, treeheight_col = 0,
          main = "expresión genética", fontsize = 8, annotation_legend = FALSE, display_numbers = TRUE,
          fontsize_number = 6, col=viridis_pal(option="plasma") (6))
```



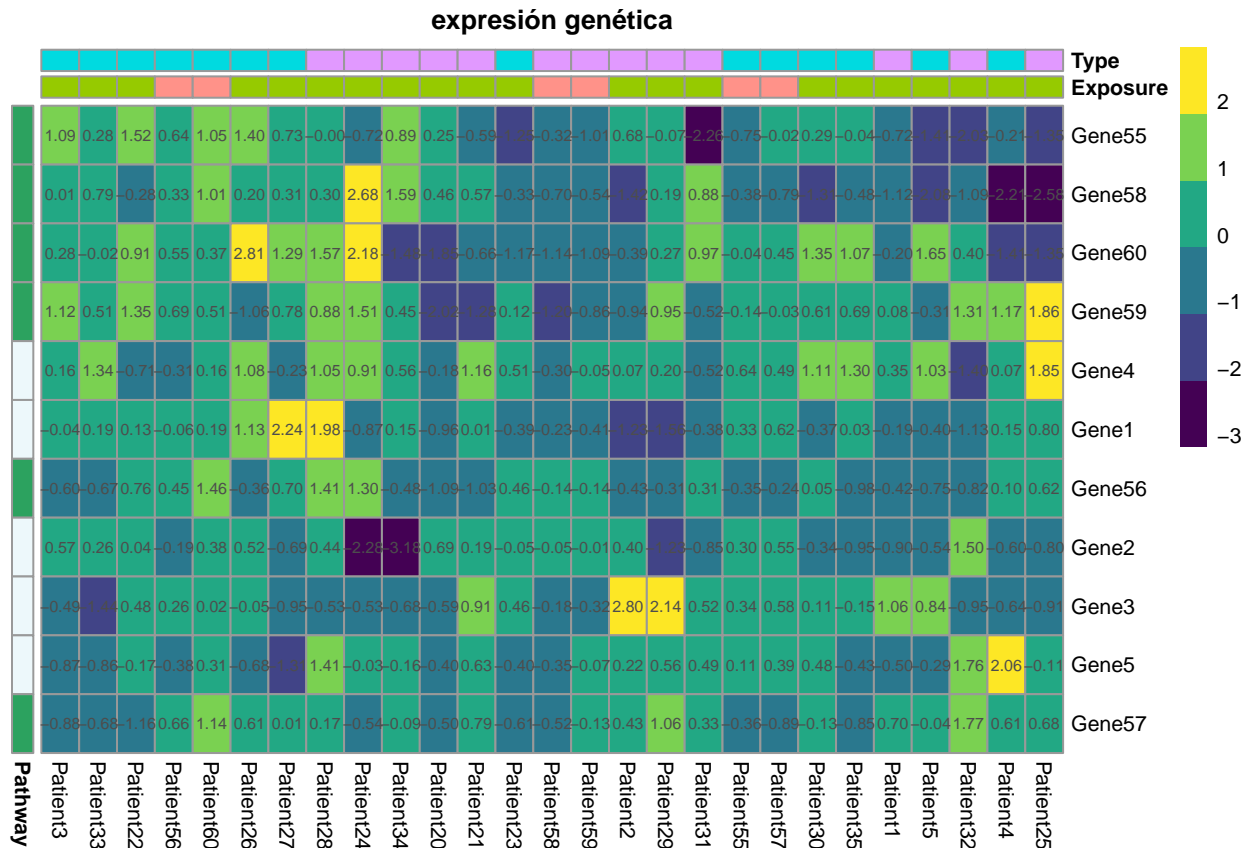
Para usar otro tono de paleta de colores: magma

```
pheatmap(sub, fontsize=6, cluster_rows = T, cluster_cols = T, annotation_row = annotation_row,
          annotation_col = annotation_col, treeheight_row = 0, treeheight_col = 0,
          main = "expresión genética", fontsize = 8, annotation_legend = FALSE, display_numbers = TRUE,
          fontsize_number = 6, col=viridis_pal(option="magma") (6))
```



Para usar otro tono de paleta de colores: viridis

```
pheatmap(sub, fontsize=6, cluster_rows = T, cluster_cols = T, annotation_row = annotation_row,
          annotation_col = annotation_col, treeheight_row = 0, treeheight_col = 0,
          main = "expresión genética", fontsize = 8, annotation_legend = FALSE, display_numbers = TRUE,
          fontsize_number = 6, col=viridis_pal(option="viridis") (6))
```



Algunos elementos adicionales

Identificar la distribución de las distancias

```
dist(sub)
```

```
##          Gene1      Gene2      Gene3      Gene4      Gene5      Gene55      Gene56      Gene57
## Gene2    6.506125
## Gene3    7.823569  7.021725
## Gene4    5.253565  7.649124  6.516104
## Gene5    6.411847  5.977640  5.967513  6.184570
## Gene55   5.703940  6.969997  7.096321  6.837653  7.534618
## Gene56   4.544832  6.723925  6.542745  5.805165  5.150859  6.028094
## Gene57   6.124657  6.069362  5.550487  6.004035  3.881691  7.122986  5.209746
## Gene58   7.417422  8.796956  8.462521  7.874145  8.030439  6.777444  6.292359  7.669524
## Gene59   6.189649  8.293720  7.977707  6.115718  5.821355  7.317126  4.835770  6.104449
## Gene60   6.623226  8.133474  7.665999  6.837342  7.659167  7.569942  6.373711  7.296198
##          Gene58      Gene59
## Gene2
## Gene3
## Gene4
## Gene5
## Gene55
## Gene56
## Gene57
## Gene58
```



```
## Gene59 8.312043
## Gene60 7.813793 6.992657
```

Identificar el mapa de calor de la correlación de los datos de pacientes

Visualizar la correlación existente entre los genes

Identificar matriz transpuesta

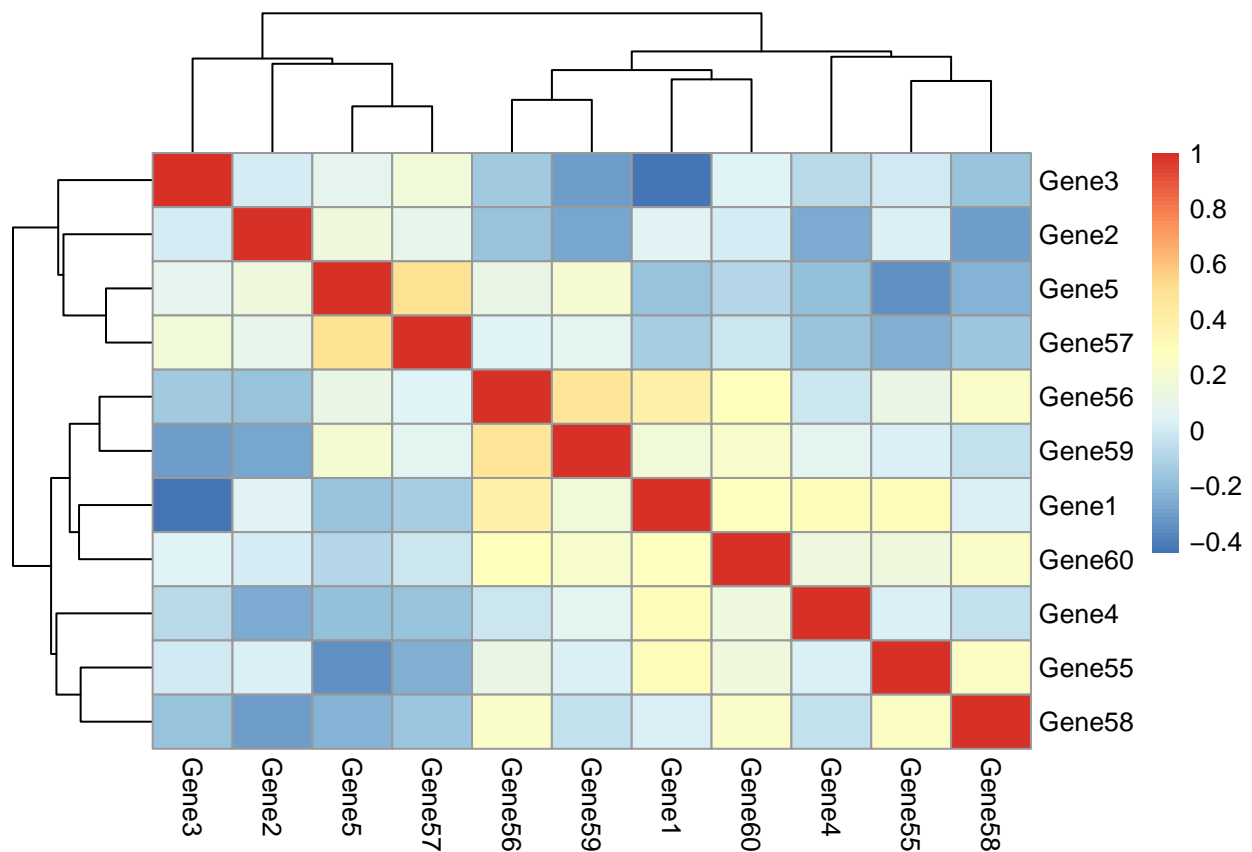
Crear objeto llamado trans

Dicho dataframe será igual a la matriz transpuesta de la sub base de datos

```
trans <- t(sub)
```

Mapa de calor de la correlación de los genes con la matriz transpuesta

```
pheatmap(cor(trans))
```



FIN DE LABORATORIO 43