

## Identificación de genes regulados por interacción p53-SET

**Autor:** Héctor Ulises Gaspar Eslava

### Introducción:

La acetilación es una modificación común en una gran cantidad de proteínas, sin embargo, su función no está del todo clara. Una de las acetilaciones conocidas es aquella en el dominio C-terminal (CTD) de p53. Dado que la eliminación de este dominio resulta en la activación de p53, se ha sugerido que el CTD podría actuar como sitio de unión de represores de p53.

Se ha dilucidado que, dependiendo de las modificaciones presentes en el CTD, diferentes proteínas interactuarán con p53. En particular, se ha encontrado que SET interactúa con este dominio cuando carece de desacetilación, mientras que otras modificaciones, tales como metilación o ubiquitinación, no tienen efectos con esta interacción. Igualmente, se identificó que esta interacción se encuentra principalmente en células no estresadas, mientras que, al inducir daño al DNA, los niveles de p53-SET se disminuyen profundamente, sugiriendo que SET actúa como un repesor de la actividad transcripcional de p53.

### Objetivo principal:

Realizar un análisis de expresión diferencial de células U2OS control, con inhibición de SET mediante siRNA, knockout de p53, y knockout de p53 – inhibición de SET

### Metodología:

#### 1. Descarga y procesamiento de archivos FASTQ:

Se descargaron los archivos de la base de datos GEO, usando el identificador SRR para cada grupo de muestras. SRR3711076 y SRR3711077 corresponden a dos réplicas de células U2OS control, SRR3711079 y SRR3711080 a dos réplicas donde se inhibió SET usando siRNA, SRR371108 y SRR3711083 a dos réplicas donde se eliminó p53 usando CRISPR, y finalmente SRR3711084 y SRR3711085 corresponden a muestras donde se eliminó p53 y se inhibió SET. Los archivos contenían lecturas single-end.

Se realizó la visualización de los archivos, y se determinó que, mientras que todos tienen buena calidad de lecturas, si se podía mejorar con *trim\_galore*. Posteriormente, se alinearon los archivos contra el genoma de referencia provisto en la clase usando *hisat2*. Finalmente, los archivos generados del alineamiento se convirtieron a archivos *bam*. Las líneas de comandos utilizadas se encuentran en el repositorio presente, en la carpeta *src* con el nombre de *bash*.

Para obtener el número de cuentas de cada transcrito de los archivos generados, se utilizó R. El script utilizado se encuentran en la carpeta *src* con el nombre *counts*.

#### 2. Análisis de expresión diferencial:

Se utilizó la librería *EdgeR*. Se descartaron los genes con bajo nivel de expresión para este análisis, y se generó un objeto *EdgeR* y una matriz de diseño. Se visualizaron los niveles de expresión de las muestras mediante un gráfico de PCA. No se observó ninguna disonancia entre los grupos de replicados biológicos.

Se realizó un análisis de expresión diferencial replicando el artículo replicado. En particular, se comparó el grupo de prueba contra el grupo con inhibición de SET, y este último contra un grupo con eliminación de p53 e inhibición de SET. Se consideró que si un gen estaba diferencialmente expresado, debería tener

un nivel de expresión 4 veces mayor o menor, así como una tasa de descubrimiento de falsos negativos menor a  $1e-5$ . Se podcedió a realizar un heatmap comparando todos los grupos de muestras.

El script utilizado se encuentra en la carpeta *src* con el nombre *Transcriptomica.rmd*

### 3. Análisis de ontología:

Con los genes identificados en ambas comparaciones, se realizó un análisis de ontología usando la función *enrichGO*, para identificar procesos biológicos, funciones y componentes celulares enriquecidos.

## Resultados:

### 1. Calidad de lecturas:

Se visualizó la calidad de los archivos FASTQ antes y después de realizar la limpieza de los archivos. En la imagen 1, se observa la calidad de las lecturas antes de la limpieza, y en la figura 2, previa a esta. Todos los archivos tuvieron un comportamiento similar

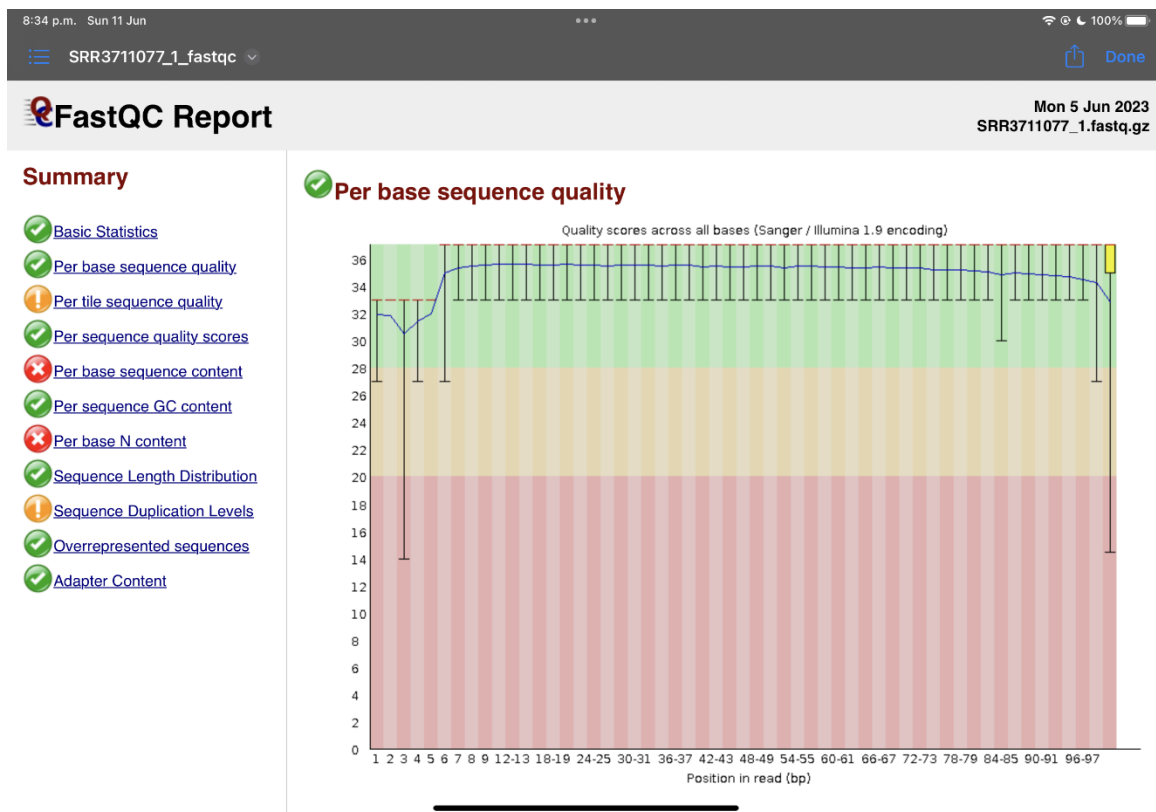


Figura 1. Análisis de calidad de archivo FASTQ antes de la limpieza

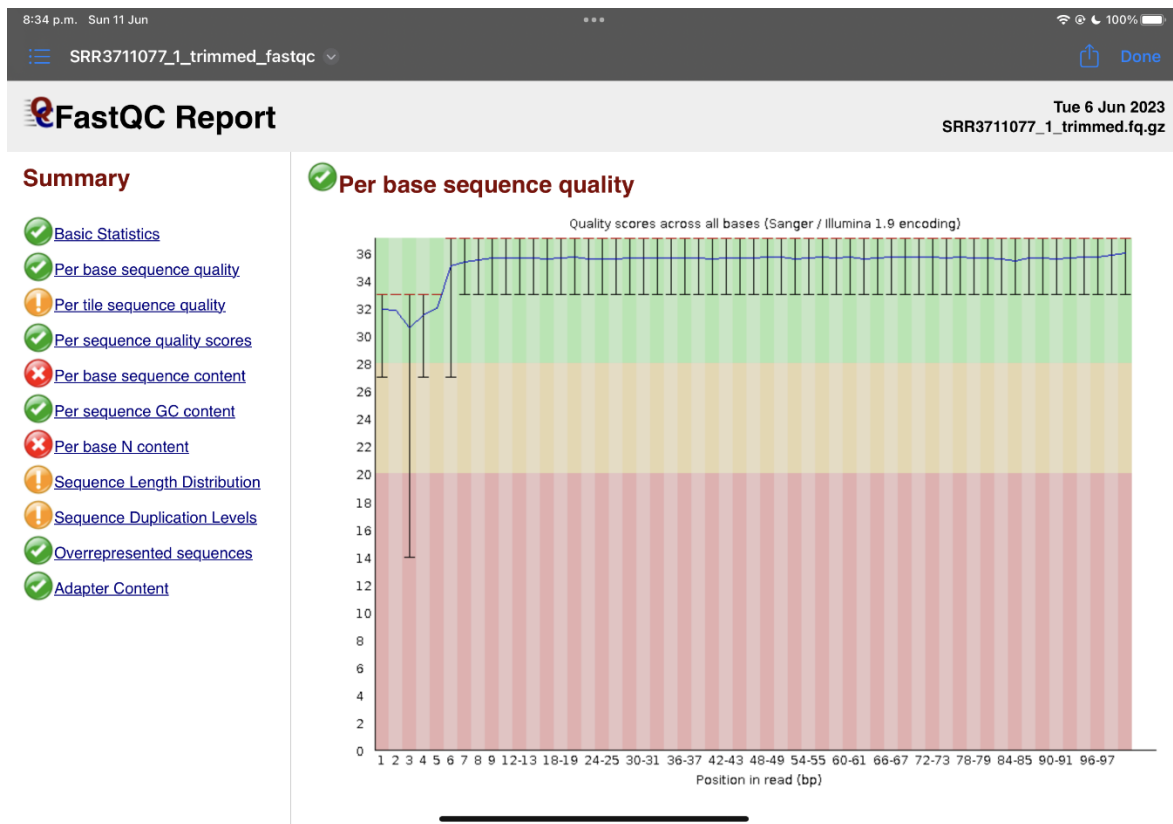


Figura 2. Análisis de calidad de archivo FASTQ después de la limpieza

## 2. Análisis de expresión diferencial:

P previo al análisis, se generó una gráfica PCA para visualizar como las muestras se relacionan entre si de manera general (figura 3). Se observa que las réplicas control (WT en el PCA) se agrupan entre ellas en menor medida. No se ve disonancia entre los niveles de expresión de los diferentes grupos que se tienen.

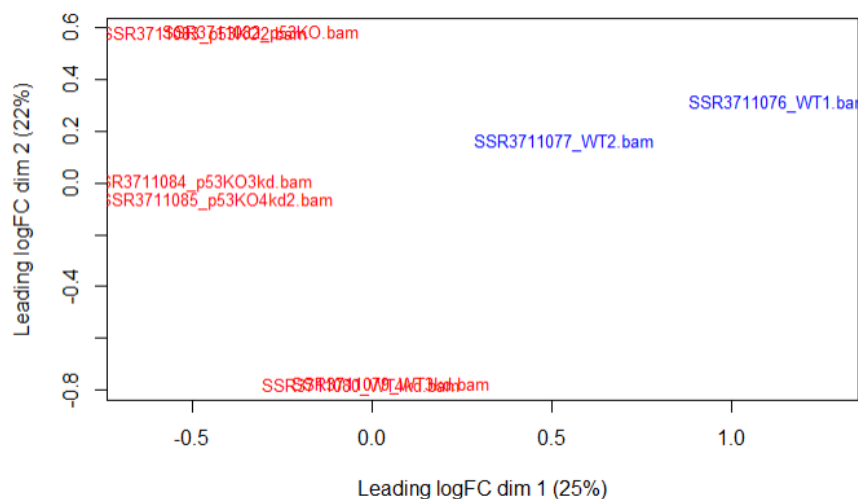


Figura 3. PCA de niveles de expresión de los diferentes grupos de muestras.

Posterior al análisis de expresión diferencial, se visualizaron los genes regulados al alta o a la baja en los dos contrastes realizados en un smear plot. Los puntos entre las líneas azules representan genes que no tuvieron cambios significantes en su expresión. La imagen 4 corresponde al contraste ente las muestras control contra las muestras con inhibición de SET. La segunda imagen corresponde al contraste entre muestras con p53 eliminado e inhibición de SET contra muestras con solo inhibición de SET.

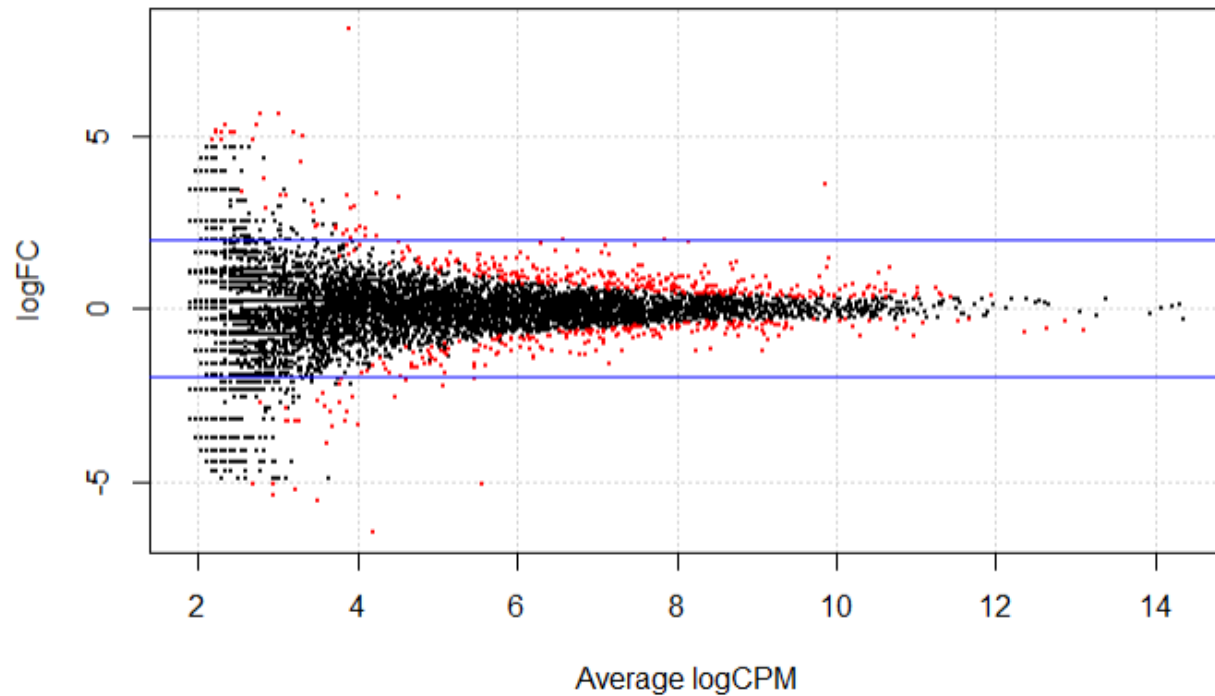


Figura 4. Smear plot de genes expresados diferencialmente en el contraste WT-vs-WT-SETkd

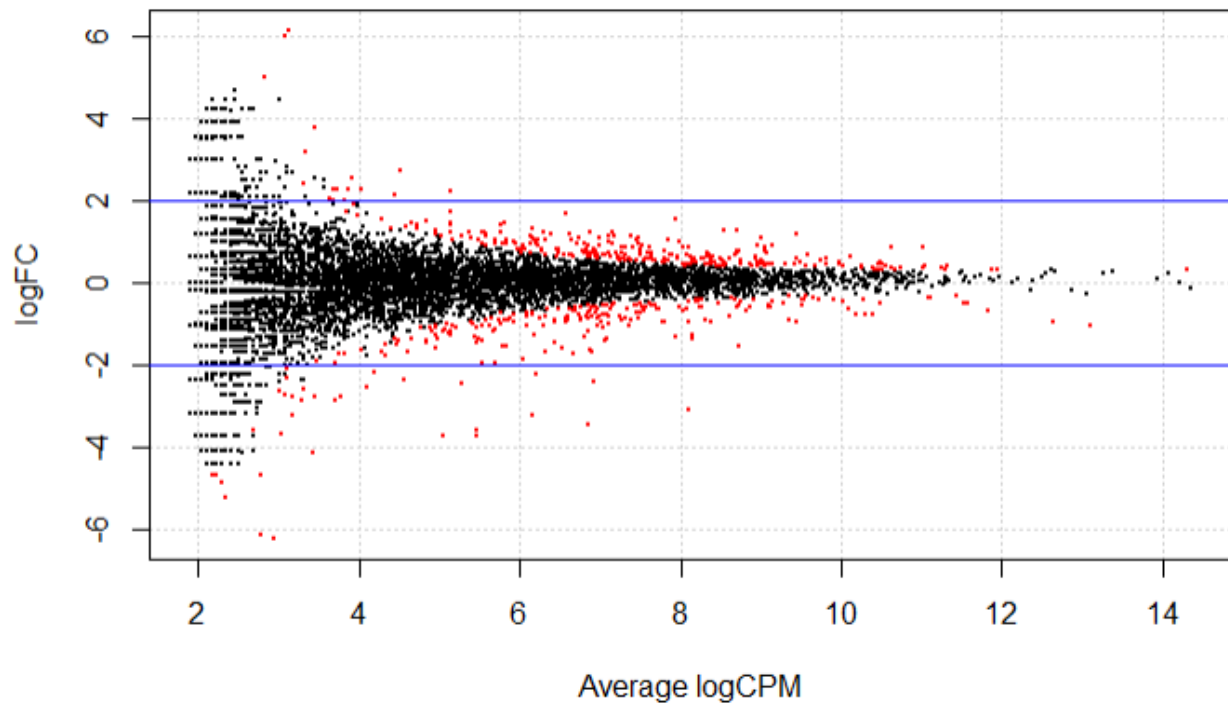


Figura 5. Smear plot de genes expresados diferencialmente en el contraste p53KO-SETkd-vs-WT-SETkd

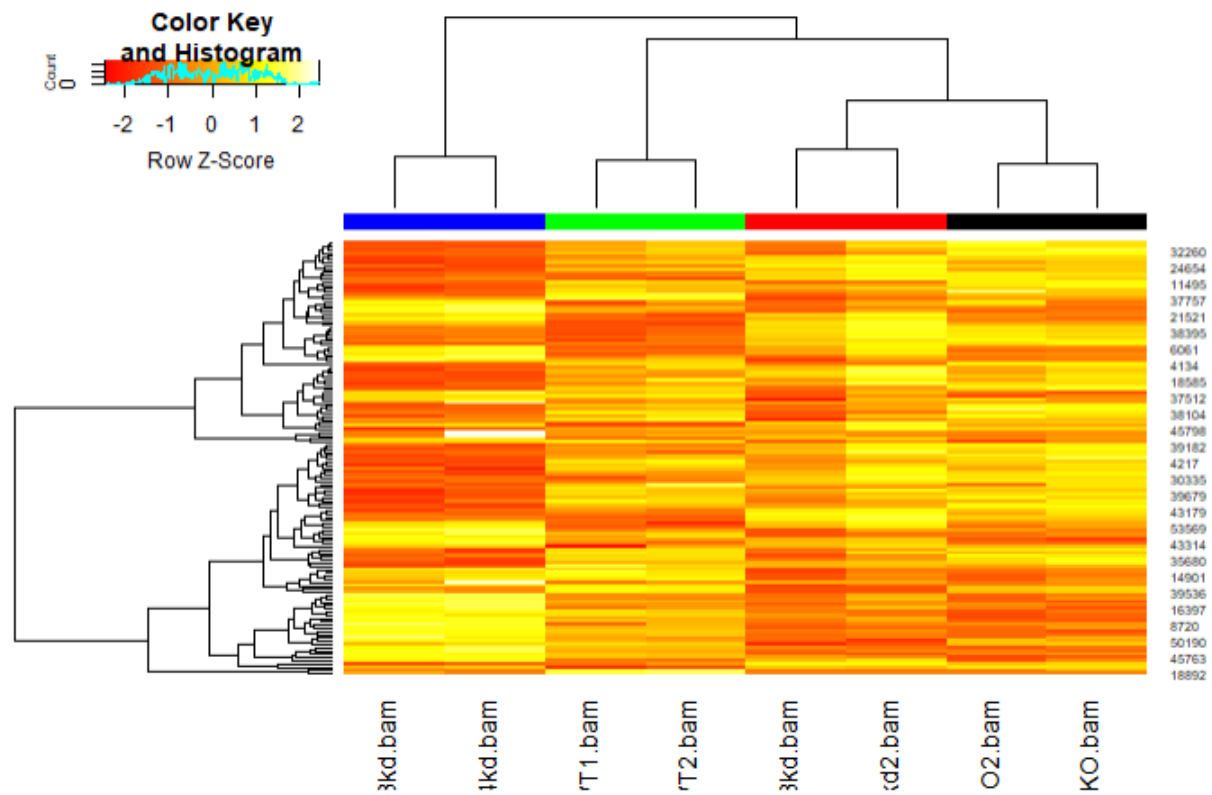
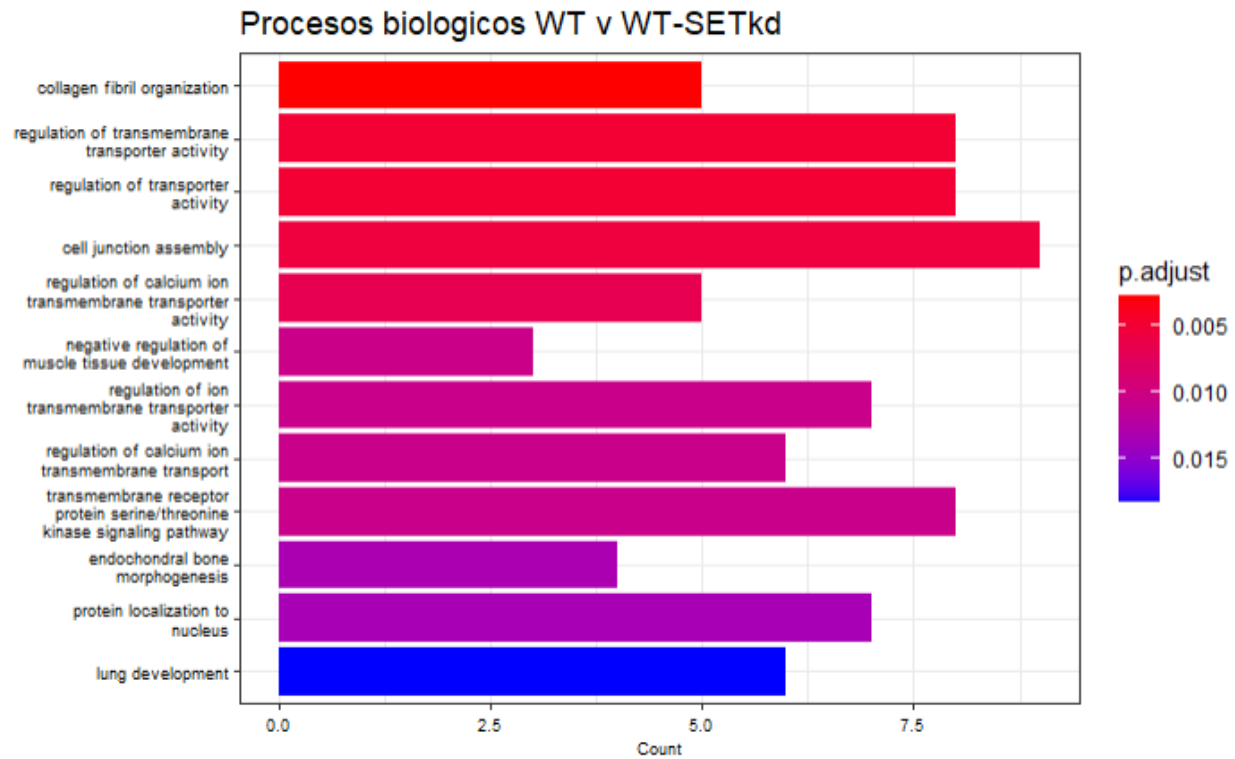


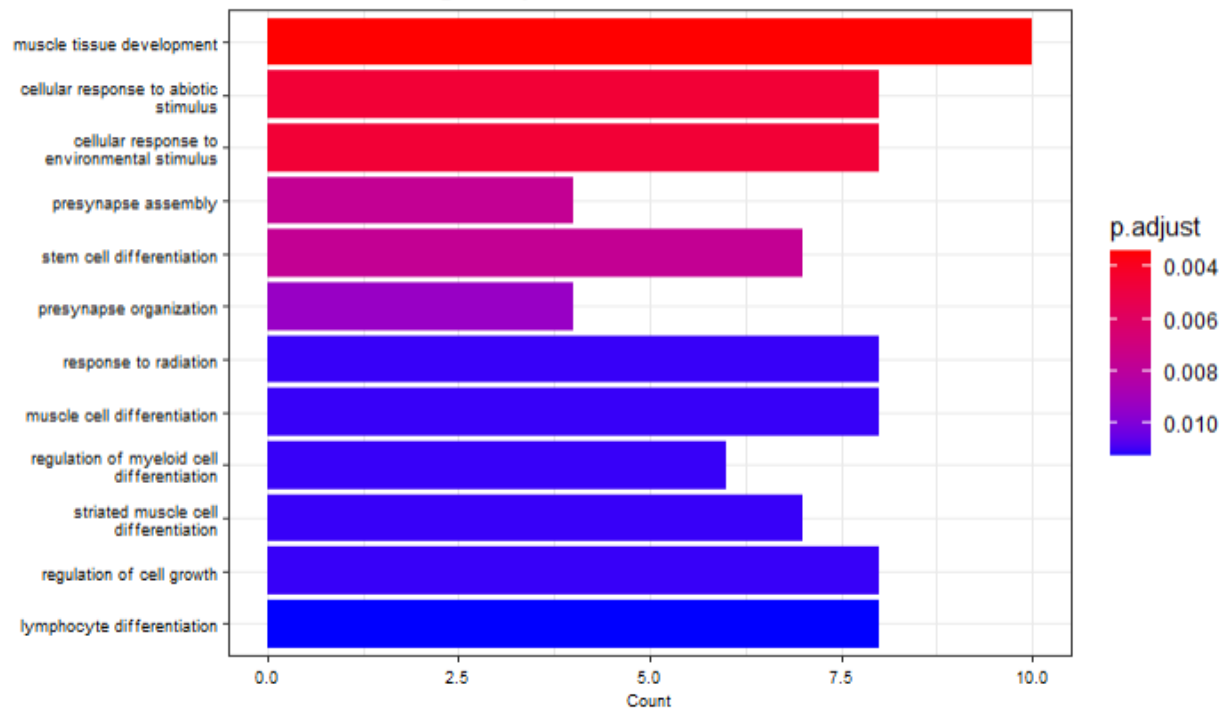
Figura 6. Heatmap de todas las muestras.

### 3. Análisis de ontología:

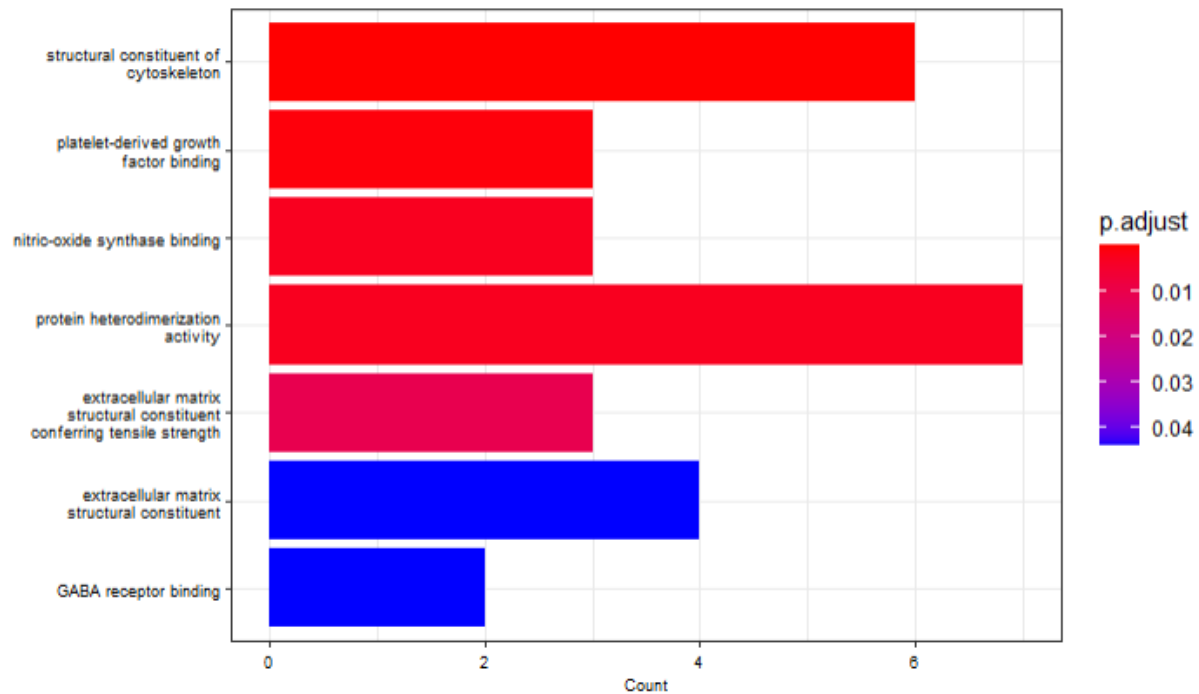
Finalmente, se realizaron análisis de ontología para identificar qué procesos biológicos están siendo afectados por la ausencia inhibición de SET, y la eliminación de p53 e inhibición de SET.



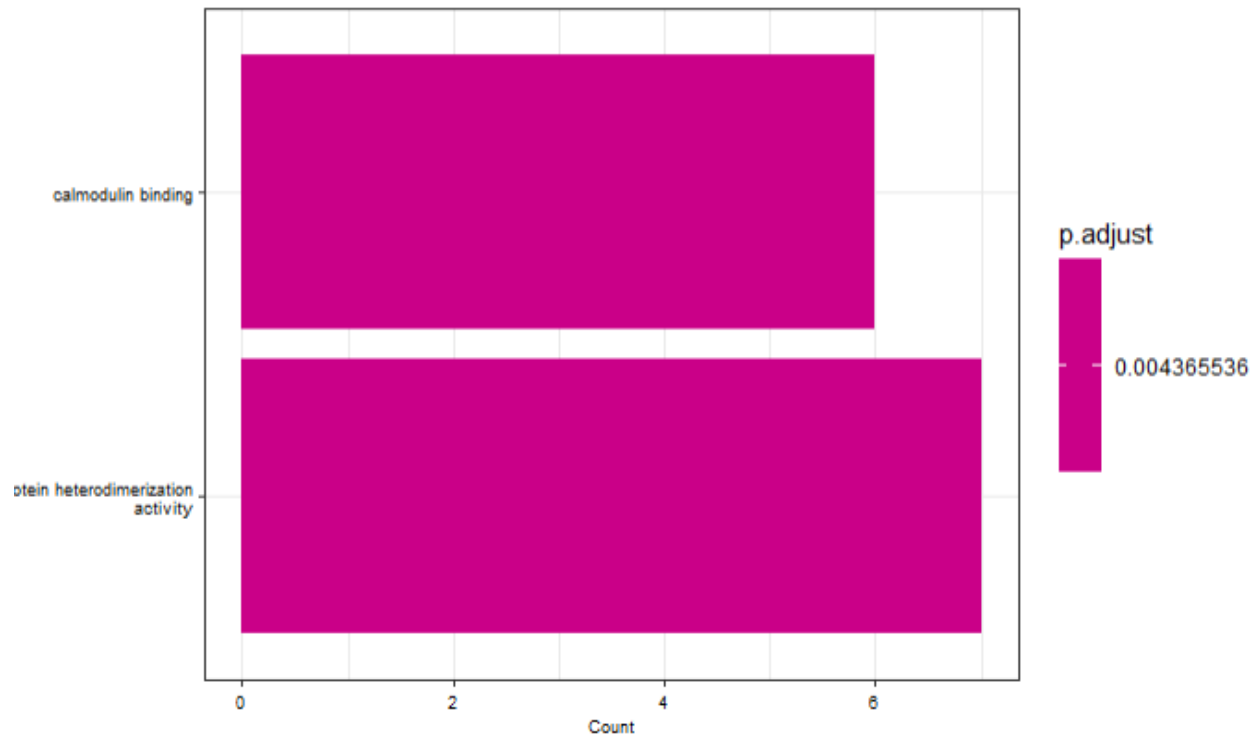
### Procesos biológicos p53KO-SETkd vs WT-SETkd



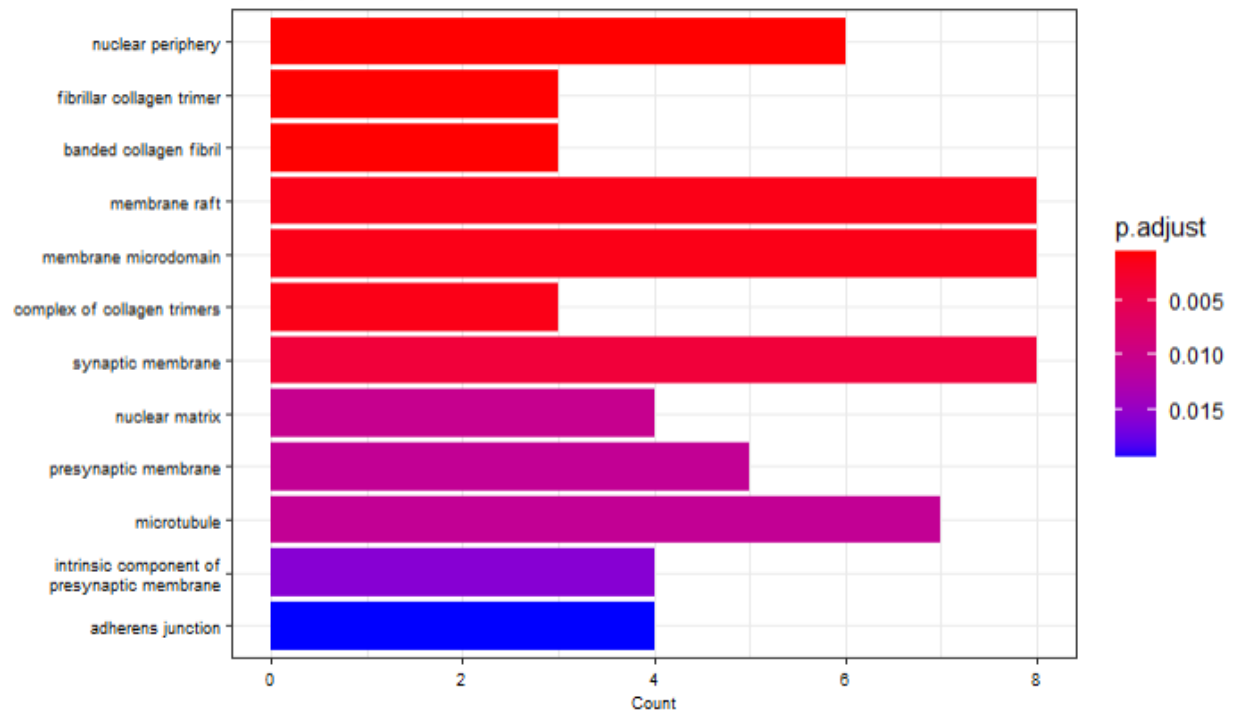
### Función molecular WT vs WT-SETkd



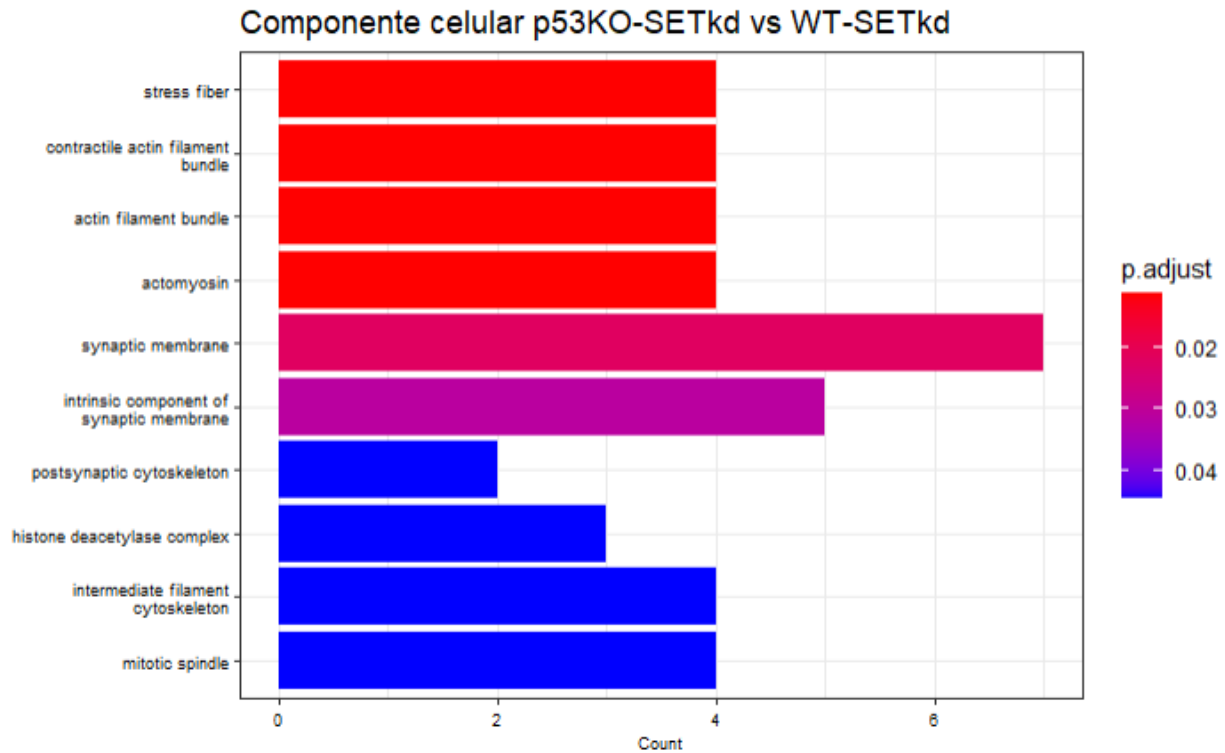
### Función molecular p53KO-SETkd vs WT-SETkd



### Componente celular WT vs WT-SETkd







### Discusión y conclusión:

En el presente reporte se presentan los resultados de un análisis de expresión diferencial, comparando muestras de células U2OS control contra células U2OS con inhibición del represor de p53 SET. Para comprender el rol de SET en el contexto de su interacción con p53, se comparan también muestras con eliminación genética de p53 e inhibición de SET contra células con solo inhibición de SET.