

Departamento de Ciencias de la Vida y de la Agricultura
Carrera de Ingeniería en Biotecnología

Trabajo de Titulación, Previo a la Obtención del título de Ingeniera en
Biotecnología

“Identificación de genes expresados diferencialmente en ratones con fibrosis pulmonar expuestas a dos terapias de reposición mitocondrial utilizando herramientas bioinformáticas”

Elaborado por

Jarre Moreano, Alanis Rafaela

Director

Flores Flor, Francisco Javier Ph.D

Fecha

07 de marzo del 2024



Tabla de contenido

1 Introducción

2 Objetivos e hipótesis

3 Materiales y Métodos

4 Resultados y Discusión

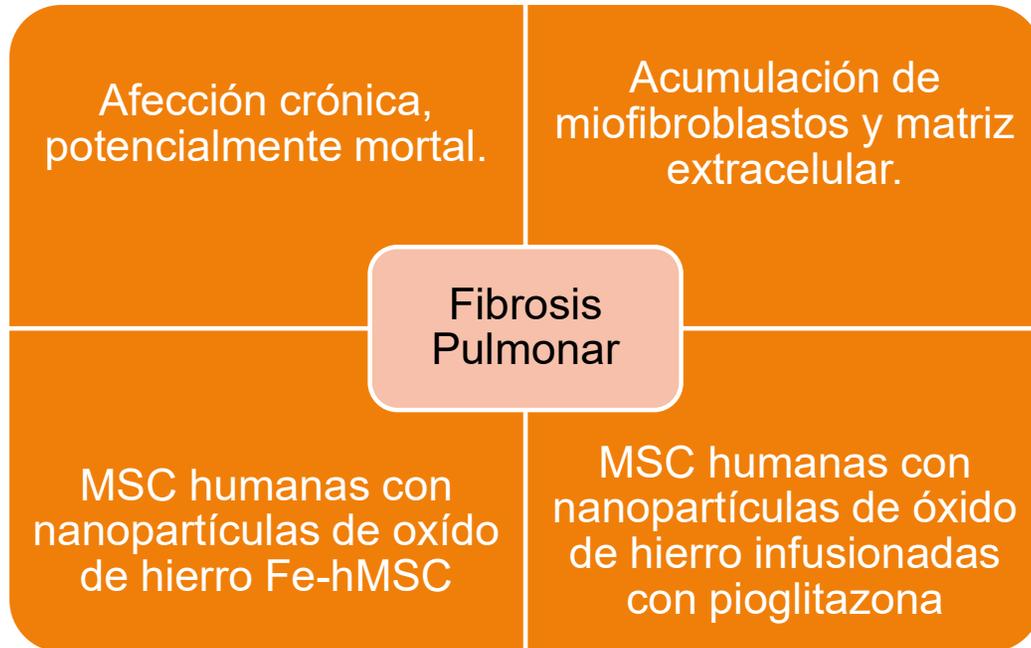
5 Conclusiones

6 Recomendaciones

7 Agradecimientos



Introducción- Fibrosis Pulmonar



4760
nuevos
casos al
año

(Huang et al., 2023; Chen et al., 2018; Li et al., 2021; Wu et al., 2020; Xin et al., 2019).

Introducción- Análisis de Expresión Diferencial de Genes

Patrón de alteraciones genéticas en condiciones específicas

Variaciones de transcripciones genéticas dentro del transcriptoma

Información biológica

Herramientas probadas y poderosas

Lista de genes

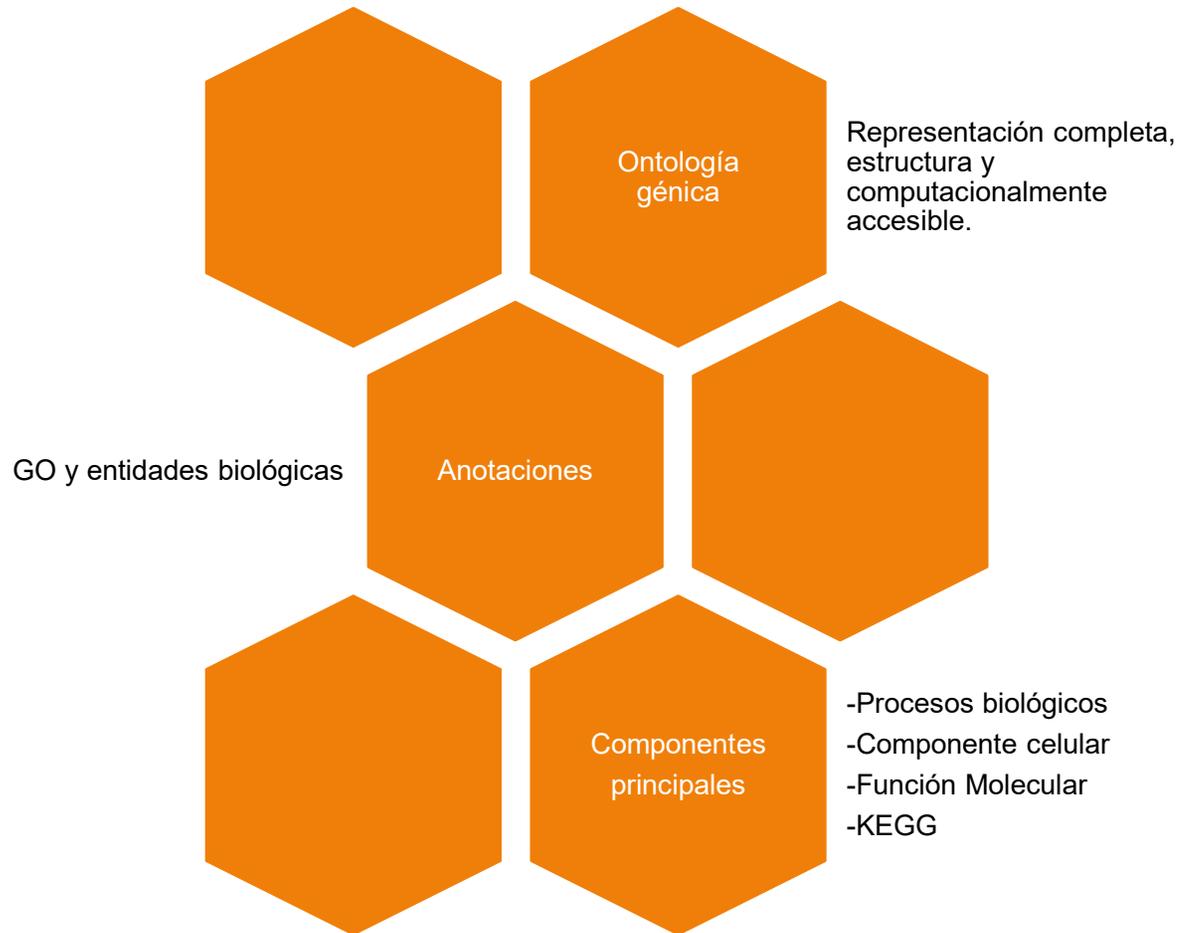
Procesos afectados



 **Galaxy**
EUROPE

(Rapaport et al., 2013; Farhang-Ghahremani et al., 2023; Chen & Wong, 2019)

Introducción- Enriquecimiento funcional



(Chen et al., 2021; Carbon et al., 2021; Wood et al., 2020; Jezernik et al., 2022)

Tabla de contenido

1 Introducción

2 Objetivos e hipótesis

3 Materiales y Métodos

4 Resultados y Discusión

5 Conclusiones

6 Recomendaciones

7 Agradecimientos

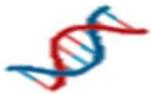


Objetivos

Objetivo General

Identificar genes expresados diferencialmente en ratones con fibrosis pulmonar expuestas a dos terapias de reposición mitocondrial utilizando herramientas bioinformáticas.

Objetivos específicos



- Obtener secuencias crudas de RNAseq de ratones con fibrosis pulmonar expuestas a dos terapias de reposición mitocondrial a partir de bases de datos públicas.
- Limpiar y normalizar las secuencias de RNAseq obtenidas
- Identificar genes con expresión diferencial entre los tratamientos.
- Realizar un análisis funcional de los genes identificados

Hipótesis

Al tratar la fibrosis pulmonar con dos tipos de células madre mesenquimales (MSC) modificadas: una con nanopartículas de óxido de hierro (Fe-hMSC) y la otra con nanopartículas de óxido de hierro infundidas con pioglitazona (Pg-Fe-hMSC), se observa una expresión génica diferencial de los mismos genes.

Tabla de contenido

1 Introducción

2 Objetivos e hipótesis

3 Materiales y Métodos

4 Resultados y Discusión

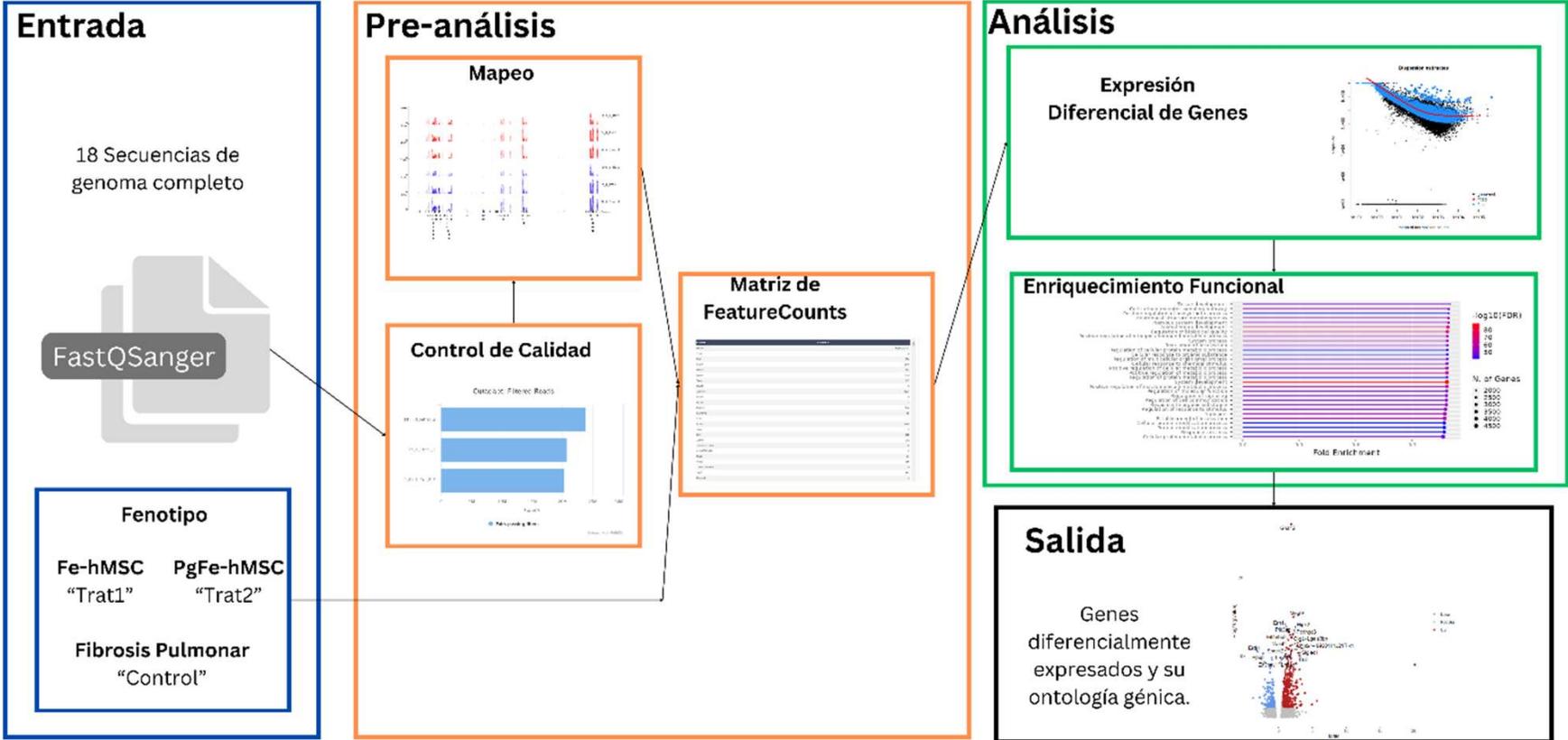
5 Conclusiones

6 Recomendaciones

7 Agradecimientos



Materiales y métodos



Materiales y métodos

Fase de Pre-análisis



Set de datos

- “Intervención eficiente para la fibrosis pulmonar mediante transferencia mitocondrial promovida por la biogénesis mitocondrial”

Descarga set de datos

- Se extrajeron ENA, donde estaban bajo el código: PRJNA948365
- 85728664428 bp.

Muestras

- Control: *Mus musculus* con Fibrosis Pulmonar.
- Tratamiento 1: Fe-hMSC
- Tratamiento 2: Pg-Fe-hMSC.

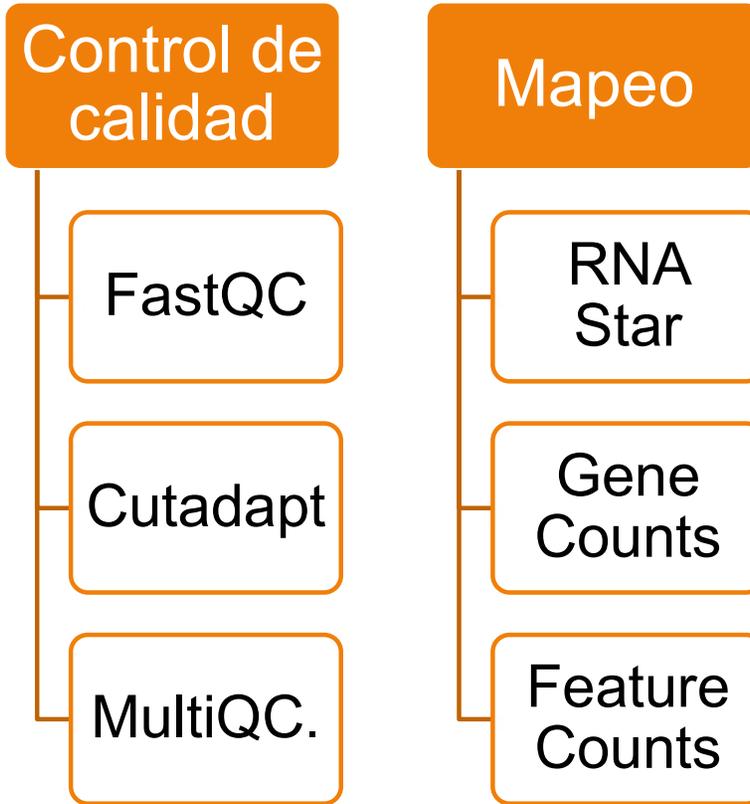


(Huang et al., 2023)



Materiales y métodos

Fase de Pre-Análisis

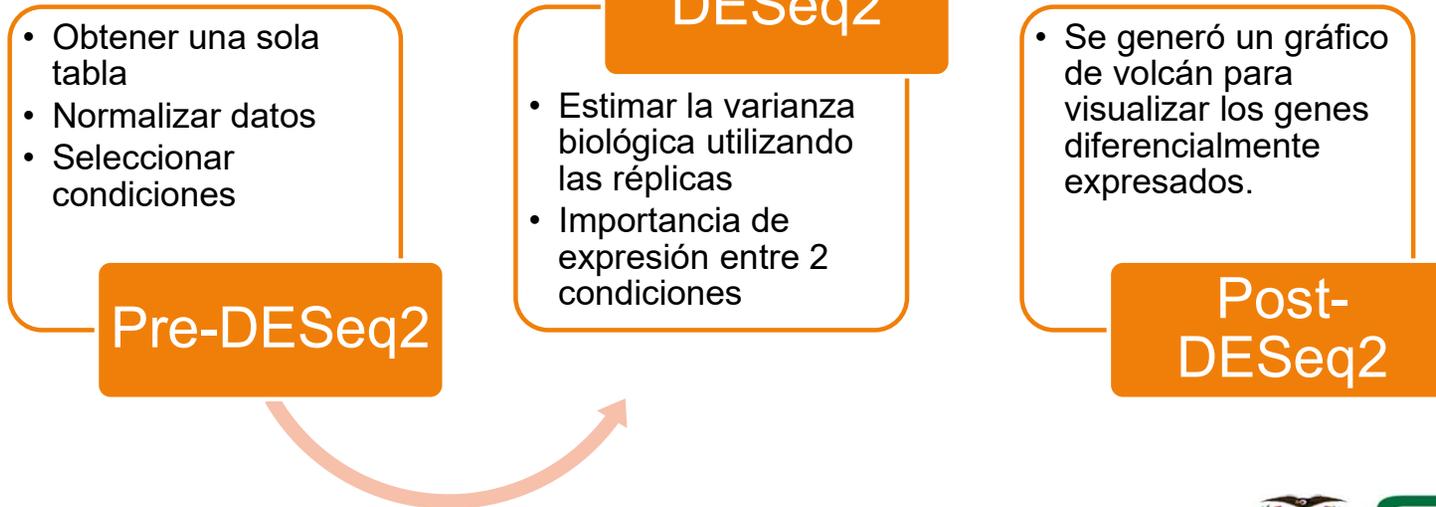


(Tekman et al., 2021; Subbotin et al., 2019)

Materiales y métodos

Fase de Análisis

Análisis de expresión diferencial de genes



(Ritchie et al., 2015; Damiani et al., 2020; Afgan et al., 2018)

Material es y métodos

Fase de Anál isis

Anál isis de enriquecimiento funcional



(Ge & Jung, 2021; Ge et al., 2020)

Tabla de contenido

1 Introducción

2 Objetivos e hipótesis

3 Materiales y Métodos

4 Resultados y Discusión

5 Conclusiones

6 Recomendaciones

7 Agradecimientos



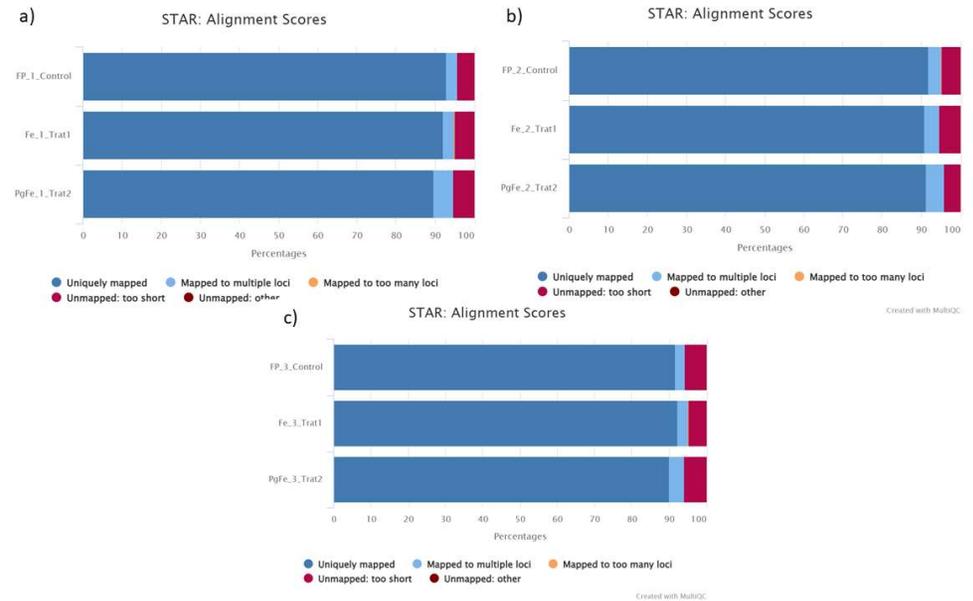
ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

Resultados y discusión

Control de calidad y mapeo

Nombre de la muestra	% Duplicados	% GC	Seq en Millones
FP_1_Control_forward	52.2%	51%	23.9
FP_1_Control_reverse	51.6%	51%	23.9
Fe_1_Trat1_forward	51.5%	51%	20.8
Fe_1_Trat1_reverse	50.3%	51%	20.8
PgFe_1_Trat2_forward	60.6%	49%	20.3
PgFe_1_Trat2_reverse	59.5%	49%	20.3
FP_2_Control_forward	51.4%	51%	21.5
FP_2_Control_reverse	50.7%	51%	21.5
Fe_2_Trat1_forward	52.1%	50%	22.3
Fe_2_Trat1_reverse	49.5%	50%	22.3
PgFe_2_Trat2_forward	57.3%	48%	21.9
PgFe_2_Trat2_reverse	53.4%	48%	21.9
FP_3_Control_forward	50.3%	51%	22.2
FP_3_Control_reverse	48.8%	51%	22.2
Fe_3_Trat1_forward	49.7%	50%	22.3
Fe_3_Trat1_reverse	47.9%	50%	22.3
PgFe_3_Trat2_forward	54.9%	50%	23.6
PgFe_3_Trat2_reverse	52.6%	50%	23.6

STAR: Scores alineados, donde: a) primera repetición, b) segunda repetición y, c) tercera repetición.



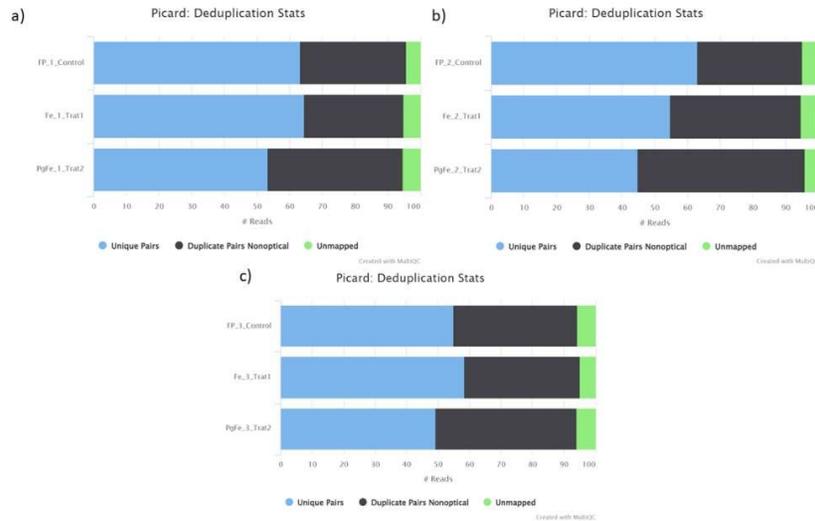
Datos estadísticos generales obtenidos de Galaxy utilizando MultiQC

(Dobin et al., 2013; Wang et al., 2012)

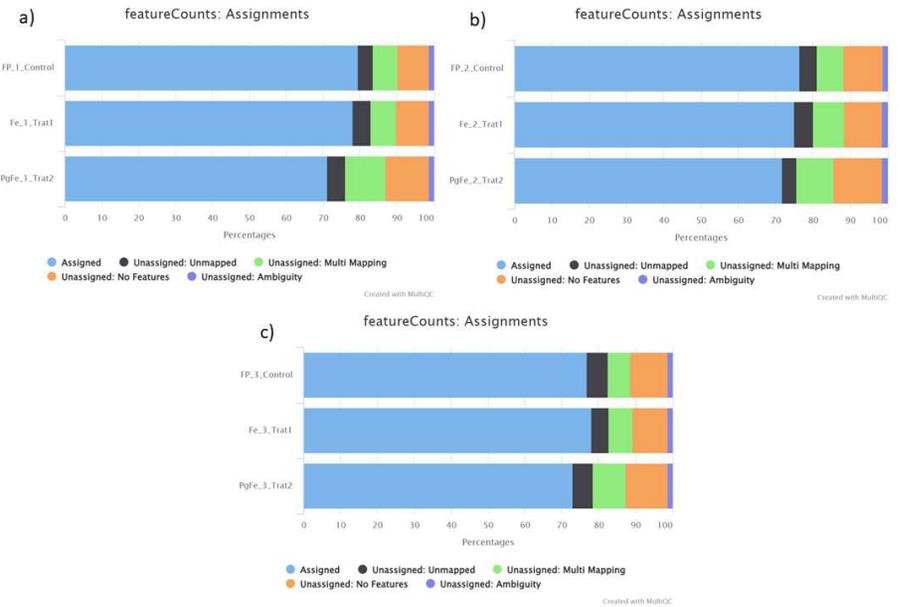
Resultados y discusión

Control de calidad y mapeo

Estadísticas de duplicación, donde: a) primera repetición, b) segunda repetición y c) tercera repetición.



Porcentaje de genes asignados con Feature Counts

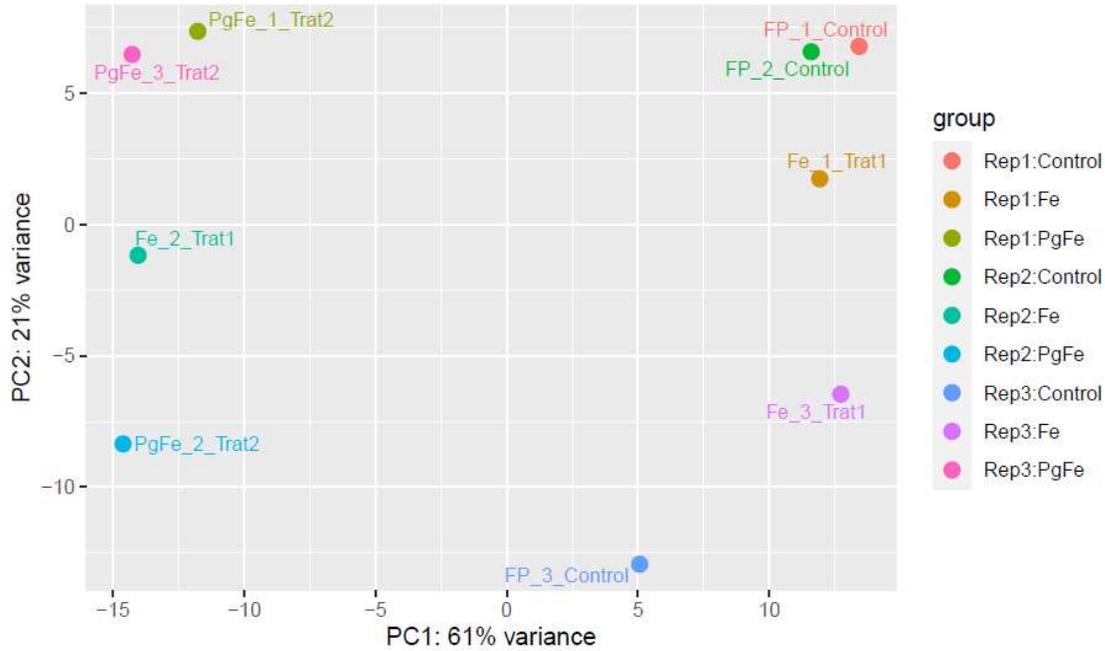


(Liao et al., 2014; Koch et al., 2018)

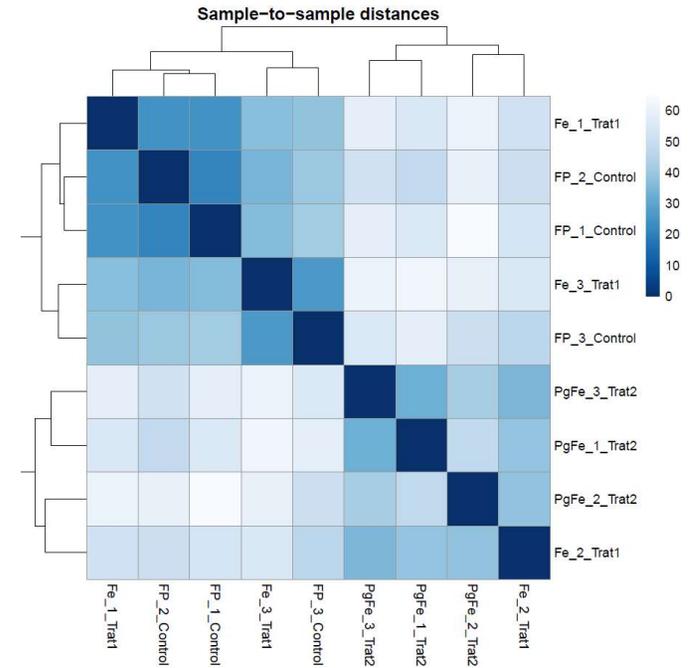
Resultados y discusión

Expresión diferencial de genes

Gráfico PCA de las dos primeras dimensiones de un análisis de componentes principales, realizado con los recuentos normalizados de las muestras.



Mapa de calor de la matriz de distancias de muestra a muestra (con agrupamiento) basado en los recuentos normalizados.

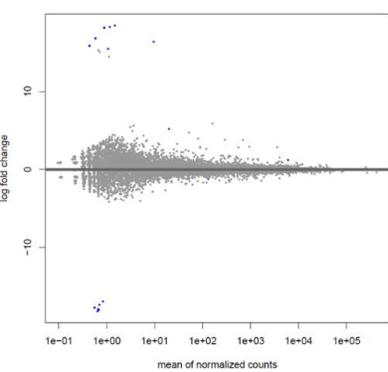


(Camargo, 2022; Love et al., 2014)

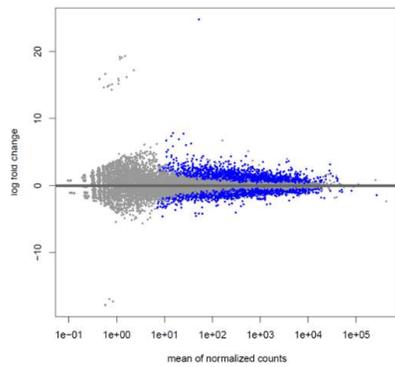
Resultados y discusión

Expresión diferencial de genes

MA-plot for Tratamientos: Control vs TratFe



MA-plot for Tratamientos: Control vs TratPgFe



No muestra significancia

MA-plot for Tratamientos: Fe vs PgFe

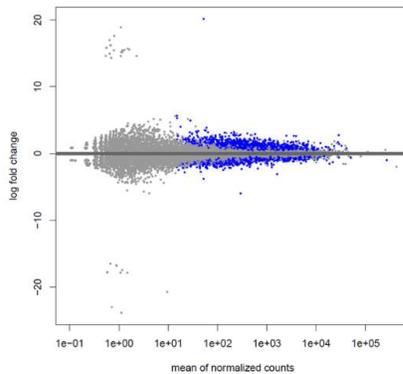
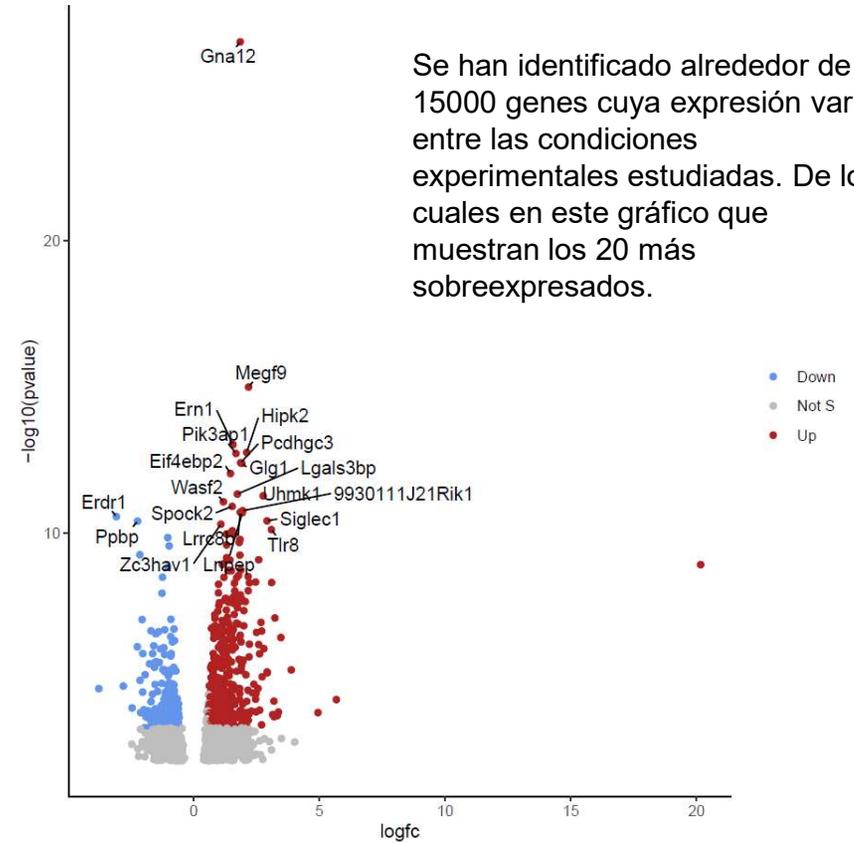


Gráfico MA de cada una de las muestras comparadas entre sí. Los genes que superaron el umbral de significancia (valor p ajustado $< 0,1$) están coloreados en azul.

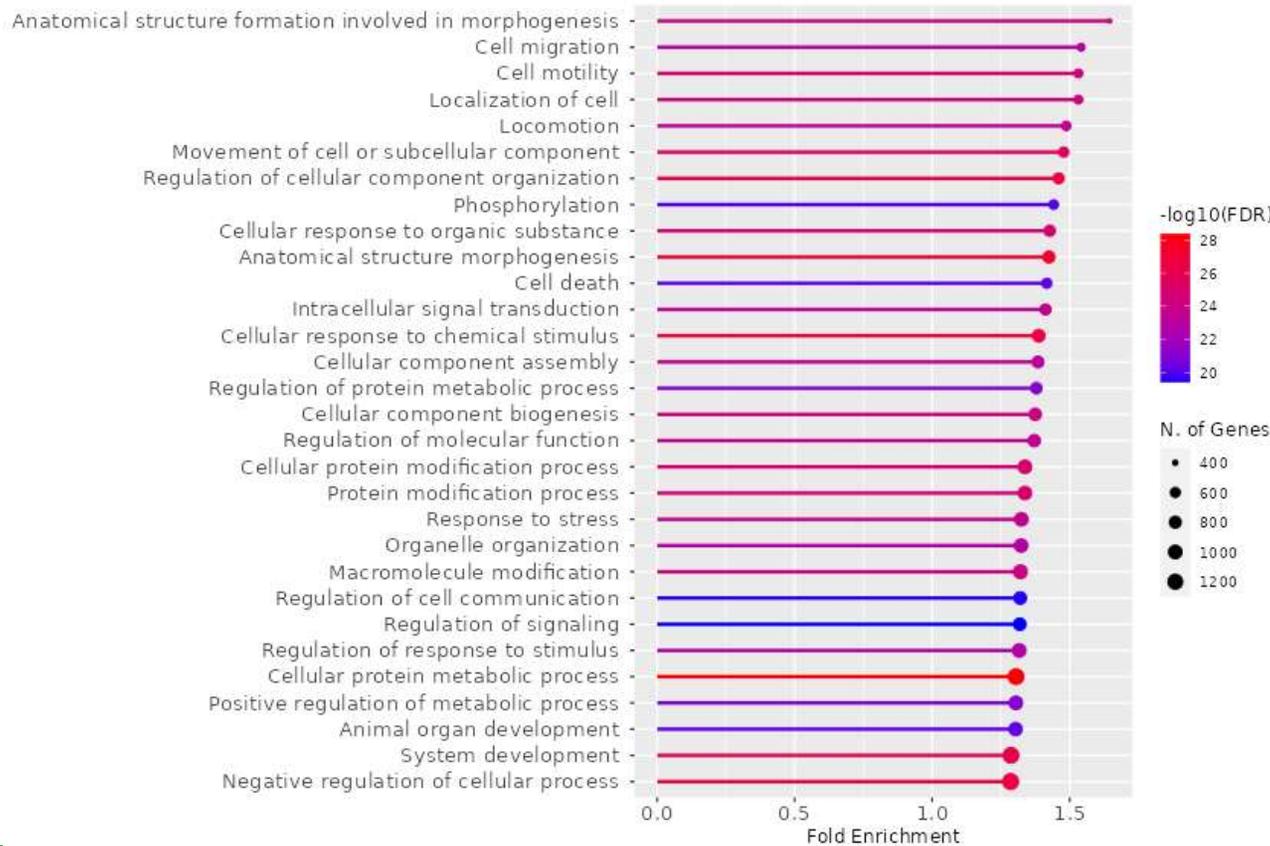


Se han identificado alrededor de 15000 genes cuya expresión varía entre las condiciones experimentales estudiadas. De los cuales en este gráfico que muestran los 20 más sobreexpresados.

Resultados y discusión

Enriquecimiento funcional

Distribución de enriquecimiento funcional basado en procesos biológicos de los diferentes genes



La Tasa de Descubrimiento Falso (FDR) representa una corrección aplicada a los p-valores derivados del análisis de enriquecimiento funcional

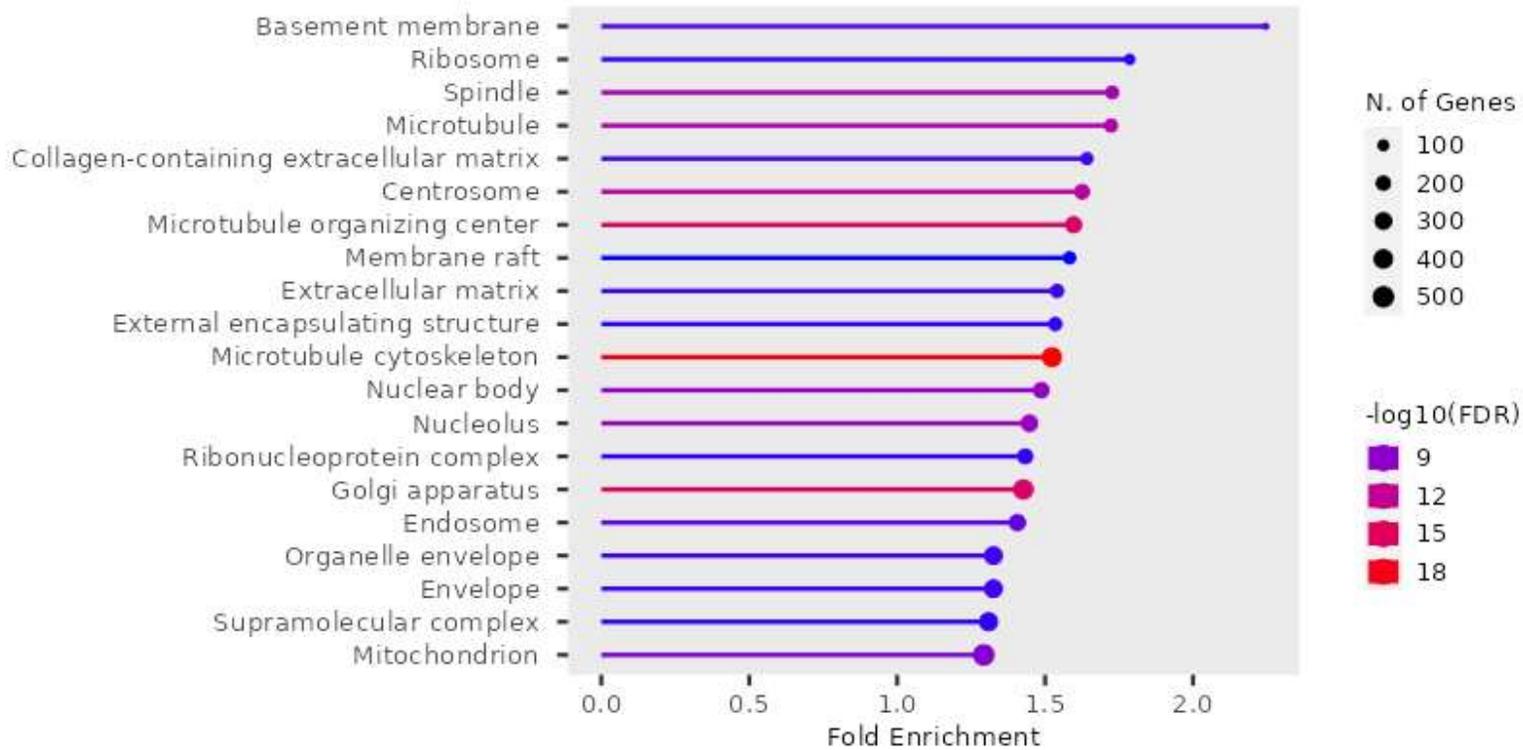
Si se observa el FDR, cuando el color es azul, es bajo o significativo, lo cual indica que la asociación que existe entre los genes y la vía, es altamente significativa. El morado indicaría asociación media y el rojo una asociación que no es estadísticamente significativa con la vía.

(Ge et al., 2020; Carbon et al., 2021)

Resultados y discusión

Enriquecimiento funcional

Distribución de enriquecimiento funcional basado en los componentes celulares de los diferentes genes



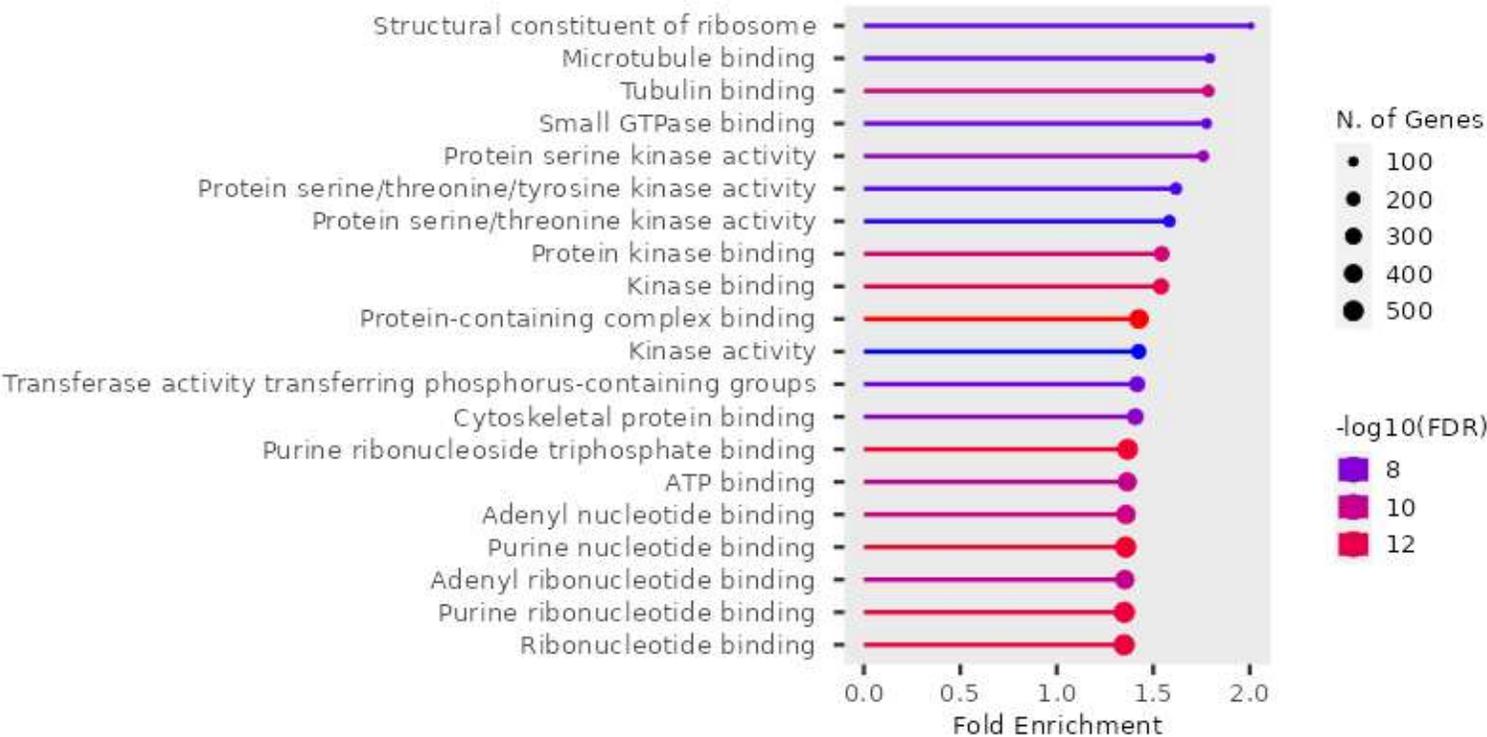
9 componentes celulares que se encuentran altamente asociados con los distintos genes.

(Ge et al., 2020, Carbon et al., 2021)

Resultados y discusión

Enriquecimiento funcional

Distribución de enriquecimiento funcional basado en la función molecular de los diferentes genes.



Observando la relación del FDR basado en la función molecular a la cual se asocian los genes, se pueden destacar 3 que son altamente significativos,

Tabla de contenido

1 Introducción

2 Objetivos e hipótesis

3 Materiales y Métodos

4 Resultados y Discusión

5 Conclusiones

6 Recomendaciones

7 Agradecimientos



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

Conclusiones

- Se obtuvieron tres réplicas de cada muestra de: control, tratamiento 1 con Fe-hMSC y tratamiento 2 con PgFe-hMSC, de bases de datos públicas, las cuales se utilizaron para realizar un análisis de expresión diferencial de genes y se obtuvieron alrededor de 15000 genes diferencialmente expresados.
- El mapeo permitió constatar que el procesamiento de las lecturas obtenidas de la base de datos generó secuencias de alta calidad y sin signos de contaminación.
- Dentro de los tratamientos, se observa una clara diferencia entre el tratamiento 1 (Fe-hMSC) y el tratamiento 2 (PgFe-hMSC). Mientras que el tratamiento 1 no mostró resultados significativos en comparación con el grupo de control, en el grupo de tratamiento 2 se observó una diferencia en la expresión de genes y a su vez demostró una alta capacidad de transferencia mitocondrial y excelentes resultados terapéuticos entre los dos tratamientos. Lo cual se puede utilizar de punto de partida para más investigaciones involucrando las células madre mesenquimales con nanopartículas de óxido de hierro y pioglitazona como terapia para la fibrosis pulmonar.
- En el análisis del enriquecimiento funcional se puede concluir que la mayoría de los genes se encuentran estrechamente relacionados a los componentes celulares, donde tuvieron más de 9 componentes celulares con mayor significancia. Por el lado de las rutas KEGG, la mayoría de genes forman parte de rutas cancerígenas, sin embargo, tomando en cuenta los probables falsos positivos se encontraron que muchos de los genes pertenecen también a la ruta metabólica de la enfermedad del COVID-19, lo cual como son dos enfermedades que afectan las vías respiratorias, se puede usar de punto de partida para posteriores investigaciones.

Tabla de contenido

1 Introducción

2 Objetivos e hipótesis

3 Materiales y Métodos

4 Resultados y Discusión

5 Conclusiones

6 Recomendaciones

7 Agradecimientos



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

Recomendaciones

- Es necesario hacer varias pruebas que permitan identificar si las muestras son de buena calidad o si se contaminaron, sobre todo si no las estamos tomando nosotros, sino que son tomadas de una base de datos. Es por esto que es indispensable hacer un control de calidad y mapeo.
- Al comparar dos tratamientos para realizar un análisis de expresión diferencial de genes, es importante observar cómo se comporta cada tratamiento en comparación con el control y entre sí. Por lo tanto, es necesario obtener gráficos separados para cada comparación. Aunque DESeq2 proporciona un único gráfico, se puede ejecutar el análisis dos veces más para obtener las comparaciones deseadas. Además, es aconsejable generar un mapa de volcán para visualizar el análisis de expresión diferencial de genes y identificar los genes con mayor grado de expresión.
- Cuando se hace un análisis de enriquecimiento funcional, es necesario observar cada uno de los campos y la clasificación de los genes con sus respectivos valor-p antes de tomar la decisión de cuál va a ser la vía final de los genes.
- Cuando se hace un análisis de enriquecimiento funcional, es necesario observar cada uno de los campos y la clasificación de los genes con sus respectivos valor-p antes de tomar la decisión de cuál va a ser la vía final de los genes. ShinyGO es una herramienta sencilla para poder observar toda la estadística y sus gráficos de manera eficiente y con una interfaz amigable.

Tabla de contenido

1 Introducción

2 Objetivos e hipótesis

3 Materiales y Métodos

4 Resultados y Discusión

5 Conclusiones

6 Recomendaciones

7 Agradecimientos



Agradecimientos



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

Francisco Flores P.h.D.
Director del Proyecto

**FAMILIA Y
AMIGOS**

