

REPORTE: Análisis de expresión diferencial

Lunes 13 de Febrero del 2023

López Ángeles B. Elizabeth

El análisis de expresión diferencial que se encuentra en este repositorio fue generado con los datos obtenidos desde [recount3](#), los datos utilizados corresponden al proyecto: SRP118914 "Transcriptional and Chromatin Dynamics of Muscle Regeneration after Severe Trauma", el paper puede ser consultado en el siguiente path:

```
Proyecto_Bioinformatica/Reports/Associated_paper.pdf
```

Después de ser dañado, el tejido muscular adulto lleva a cabo a una secuencia de eventos bien coordinados de carácter fisiológico y molecular que promueve la regeneración. El propósito de la investigación es entender los mecanismos epigenéticos y transcripcionales que controlan estos eventos. La información obtenida a partir de este estudio se empleará para llevar a cabo un análisis de expresión diferencial utilizando paquetes de Bioconductor en R.

Análisis de los resultados obtenidos:

El código utilizado para obtener los resultados puede ser consultado en el siguiente path:

```
Proyecto_Bioinformatica/Analisis_expresion_diferencial.R
```

Si lo desea, también puede consultar la bitácora que se realizó con forme de avanzaba en los pasos del proyecto, sin embargo, esta fue escrita con el único propósito de llevar más orden en los pasos.

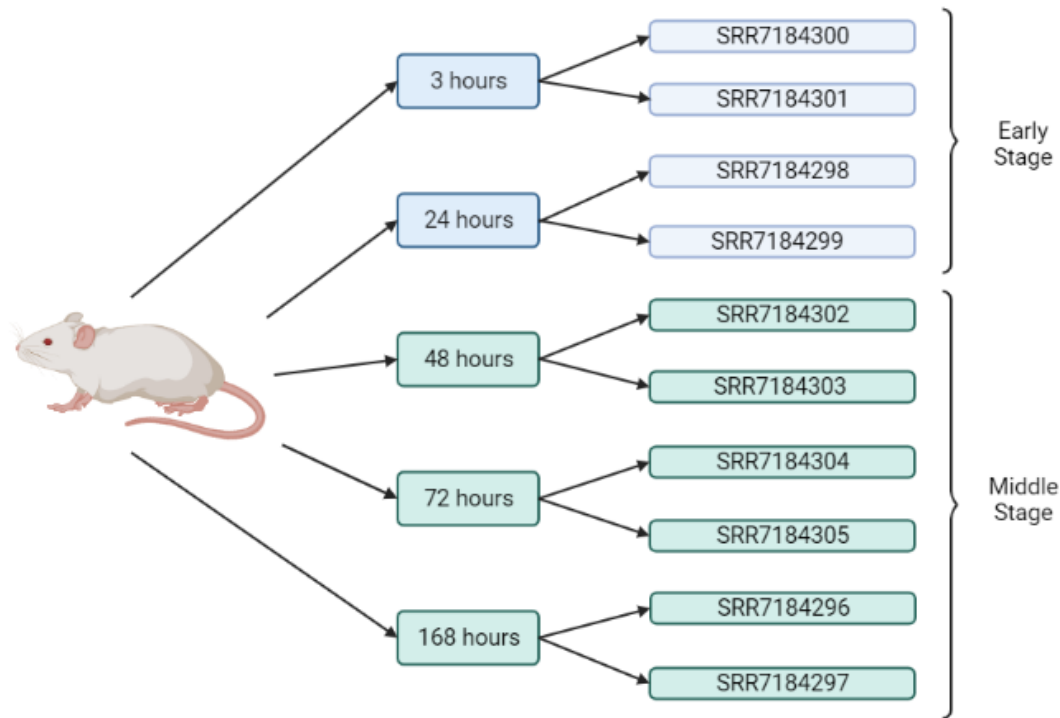
```
Proyecto_Bioinformatica/Reports/Bitacora.md
```

Data

Los ratones que utilizaron para este estudio fueron heridos en el musculo tibial anterior y posteriormente se procedió a medir los niveles de expresión génica transcurrido 3, 24, 48, 72 y 168 horas. El set de datos consta de 10 muestras diferentes y fueron clasificadas en estado early o middle según en intervalo de tiempo en el cual fueron tomadas.

- Early: 3 y 24 hrs.
- Middle: 48, 72 y 168 hrs.

La siguiente imagen ilustra la cantidad, nombres y clasificación de cada muestra.



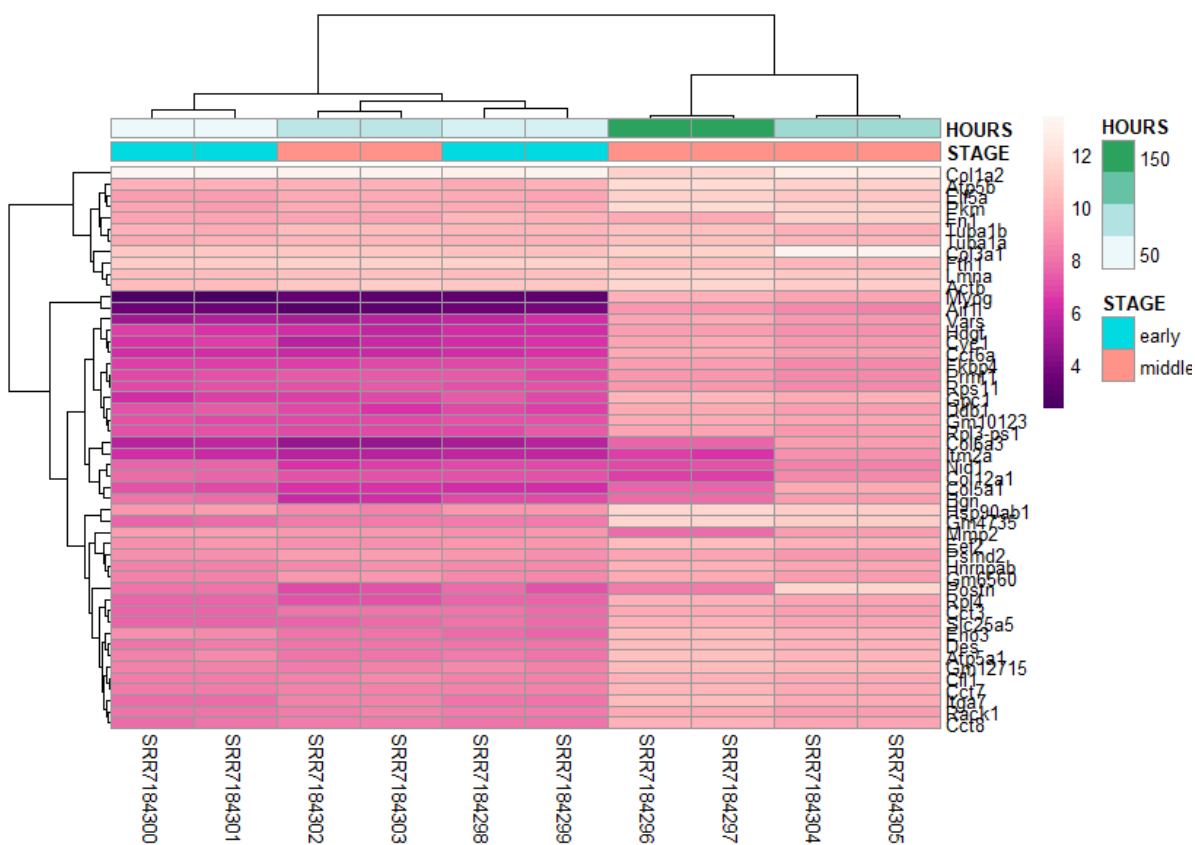
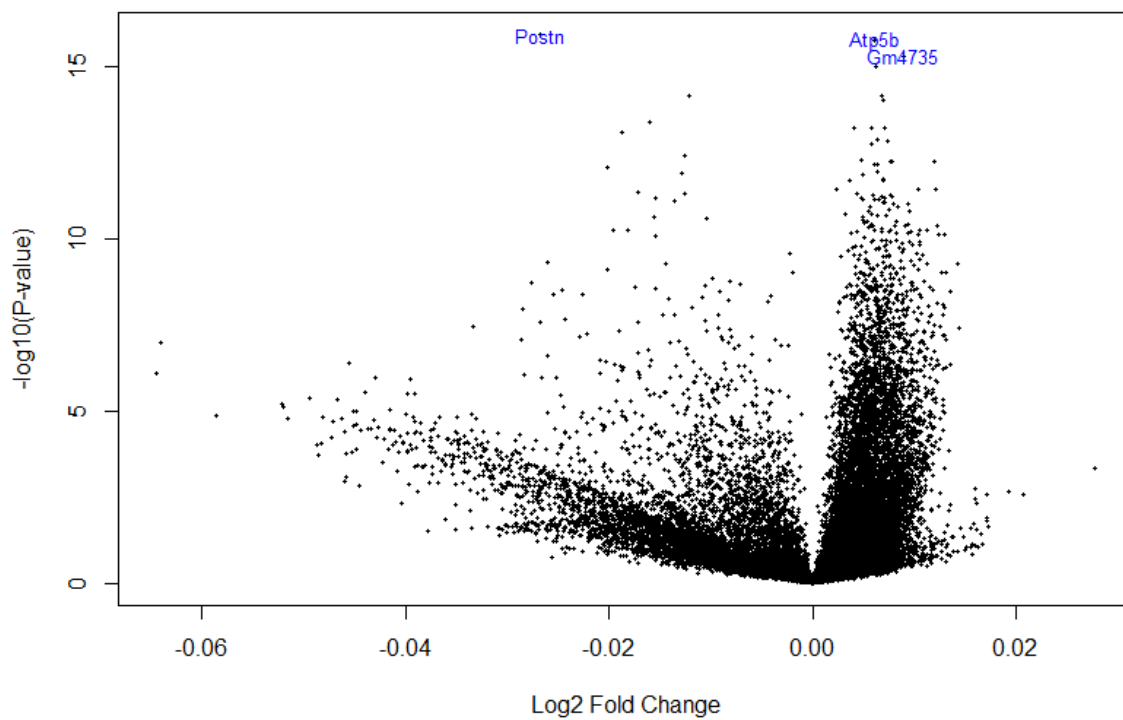
Created in BioRender.com bio

En cuanto a los genes, se comenzó trabajando con 55421, pero tras la limpieza y descarte de aquellos genes con expresiones muy bajas, se terminó conservando únicamente el 48.46% de estos, es decir 26857 genes.

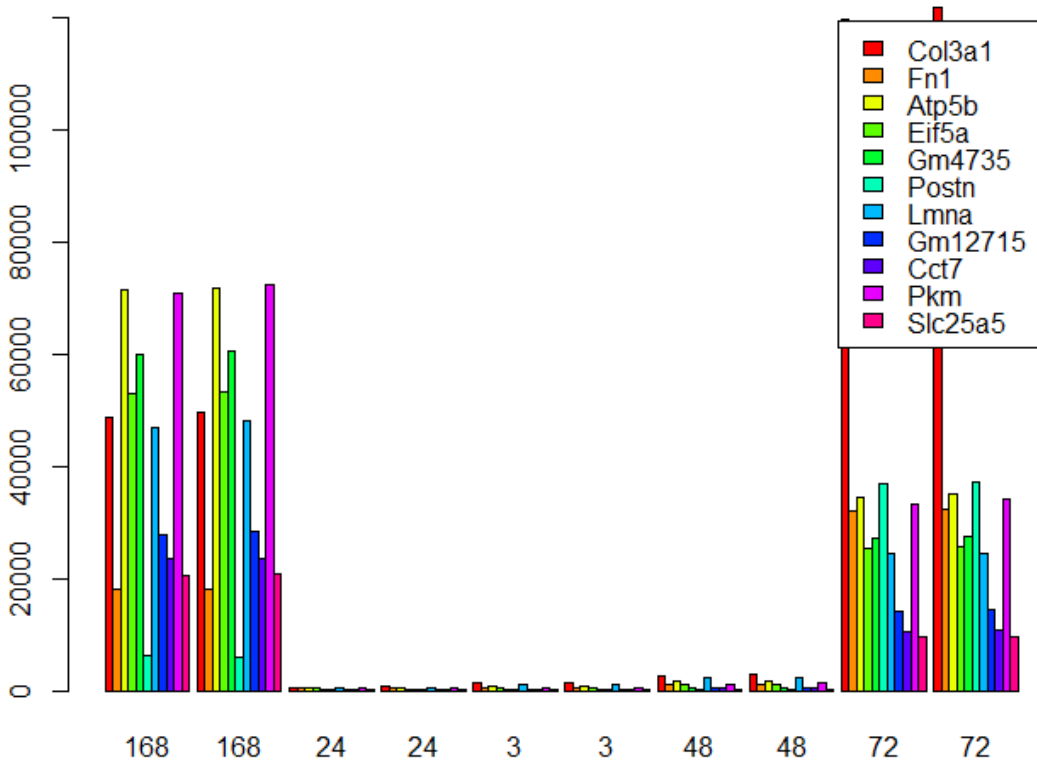
Para ver más a detalle los pasos que se siguieron y la estrategia de limpieza y normalización de datos, se puede consultar el código cuyo path se encuentra al inicio de esta sección.

Resultados:

Hablaremos más a detalle de los resultados obtenidos de tres graficas en concreto: Heatmap, barplot y volcano plot:



Expression diferencial: Top 10



Gene	Dominant Expression Stage	Function
Atp5b	middle	Adenyl ribonucleotide binding activity; angiotensin binding activity; and proton-transporting ATPase activity, rotational mechanism. Predicted to contribute to ATP hydrolysis activity and proton-transporting ATP synthase activity, rotational mechanism.
Gm4735	middle	Processed pseudogene/Predicted gene
Postn	middle	This gene encodes a secreted extracellular matrix protein that functions in tissue development and regeneration, including wound healing and ventricular remodeling following myocardial infarction. The encoded protein binds to integrins to support adhesion and migration of epithelial cells. This protein plays a role in cancer stem cell maintenance and metastasis. Mice lacking this gene exhibit cardiac valve disease, and skeletal and dental defects.
Col3a1	middle	This gene encodes the alpha-1 subunit of the fibril-forming type III collagen found in bone, cartilage, dentin, tendon, bone marrow stroma and other connective tissue. The encoded protein forms homotrimeric type III procollagen that undergoes proteolytic processing during fibril formation. A majority of mice lacking the encoded protein die within two days of birth but about 5% of the animals survive to adulthood. The surviving mice exhibit severe cortical malformation and experience significantly shorter lifespan.
Fn1	middle	Enables peptidase activator activity. Acts upstream of or within several processes, including calcium-independent cell-matrix adhesion; cell-substrate junction assembly; and positive regulation of axon extension. Located in apical plasma membrane and basement membrane. Is expressed in several structures, including alimentary system; cardiovascular system; egg cylinder; embryo mesenchyme; and genitourinary system.
Eif5a	middle	This gene encodes an elongation initiation factor, which participates in protein synthesis. The encoded protein also plays roles in mRNA metabolism, cell proliferation, and cell cycle control. This protein contains a modified lysine residue called hypusine, which appears to be necessary for its function.

Gene	Dominant Expression Stage	Function
Lmna	middle	This gene encodes a protein that is a member of the lamin family. Nuclear lamins, intermediate filament-like proteins, are the major components of the nuclear lamina, a protein meshwork associated with the inner nuclear membrane. This meshwork is thought to maintain the integrity of the nuclear envelope, participate in chromatin organization, and regulate gene transcription. Vertebrate lamins consist of two types, A and B. This protein is an A-type and is proposed to be developmentally regulated. In mouse deficiency of this gene is associated with muscular dystrophy. Mouse lines with different mutations in this gene serve as pathophysiological models for several human laminopathies
Gm12715	middle	not found
Cct7	middle	Enables identical protein binding activity. Acts upstream of or within binding activity of sperm to zona pellucida and toxin transport. Located in cell body. Part of chaperonin-containing T-complex and zona pellucida receptor complex.
Pkm	middle	Enables mRNA binding activity and pyruvate kinase activity. Involved in positive regulation of cytoplasmic translation. Acts upstream of or within glycolytic process. Located in cilium; photoreceptor inner segment; and rough endoplasmic reticulum. Is expressed in several structures, including alimentary system; genitourinary system; nervous system; respiratory system; and sensory organ.
Slc25a5	middle	This gene encodes a transmembrane domain-containing protein that localizes to the mitochondrial inner membrane. The encoded protein facilitates the exchange of ADP from the cytoplasm with ATP from the mitochondria.
Col1a2	early	This gene encodes the pro-alpha2 chain of type I collagen whose triple helix comprises two alpha1 chains and one alpha2 chain. Type I is a fibril-forming collagen found in most connective tissues and is abundant in bone, cornea, dermis and tendon. Mutations in this gene are associated with osteogenesis imperfecta types I-IV, Ehlers-Danlos syndrome type VIIB, recessive Ehlers-Danlos syndrome Classical type, idiopathic osteoporosis, and atypical Marfan syndrome. Symptoms associated with mutations in this gene, however, tend to be less severe than mutations in the gene for the alpha1 chain of type I collagen (COL1A1) reflecting the different role of alpha2 chains in matrix integrity.

Dentro de las 10 posiciones más significativas de expresión diferencial existen algunos genes que capturan nuestra atención debido a su función y expresión. Encontramos que los genes *Postn* y *Col3a1* tienen un rol en el desarrollo y regeneración de tejidos conectivos, la ausencia de ellos usualmente causa enfermedades cardíacas, defectos dentales o de esqueleto para *Postn*. En el caso de *Col3a1*, su ausencia o deficiencia causa la muerte. Los resultados obtenidos muestran que el proceso de regeneración comienza por los menos 3 horas luego de que se produjo el estímulo en los ratones, pero va en aumento entre las horas siguientes, llegando a su máxima expresión entre las 72 y 168 horas.