

Proyecto

Jocelyn Trujillo Gutierrez

2025-01-31

Librerías que se utilizan

```
library("recount3")
```

Datos utilizados

Abstract de los datos obtenidos:

Se recolectaron 30 muestras de 15 sujetos en dos momentos temporales (a nivel del llano y a gran altitud). La significancia se determinó comparando los perfiles de expresión a gran altitud con los del llano.

Diseño general: 15 voluntarios viajaron desde un punto de partida a 1400 m hasta una zona a 5300 m en un período de 4 días.

```
## Todos los proyectos con datos de humano en recount3
human_projects <- available_projects()

## Obtencion de datos del proyecto SRP131480
proj_info <- subset(
  human_projects,
  project == "SRP131480" & project_type == "data_sources"
)

rse_gene_SRP131480 <- create_rse(proj_info)
```

Una vez tenemos los datos, debemos explorarlos de manera general

```
# Explorar el objeto rse_SRP131480
rse_gene_SRP131480

## class: RangedSummarizedExperiment
## dim: 63856 30
## metadata(8): time_created recount3_version ... annotation recount3_url
## assays(1): raw_counts
## rownames(63856): ENSG00000278704.1 ENSG00000277400.1 ...
##   ENSG00000182484.15_PAR_Y ENSG00000227159.8_PAR_Y
## rowData names(10): source type ... havana_gene tag
## colnames(30): SRR6514110 SRR6514111 ... SRR6514138 SRR6514139
## colData names(175): rail_id external_id ...
##   recount_pred.curated.cell_line BigWigURL
```

Como podemos observar este objeto contiene 30 muestras y 63856 genes.

Con el proposito de facilitarnos la utilizacion de los datos, se modificara un poco el objeto `rse_gene_SRP131480`

```
rse_gene_SRP131480 <- expand_sra_attributes(rse_gene_SRP131480)

# Se verifica que todos los datos se vean bien
colData(rse_gene_SRP131480)[
  ,
  grepl("^sra_attribute", colnames(colData(rse_gene_SRP131480)))
]
```

```
## DataFrame with 30 rows and 4 columns
##           sra_attribute.altitude sra_attribute.lls_score
##                               <character>             <character>
## SRR6514110                Plain                0
## SRR6514111                Plain                0
## SRR6514112                Plain                0
## SRR6514113                Plain                0
## SRR6514114                Plain                0
## ...                        ...                  ...
## SRR6514135 High altitude exposure                7
## SRR6514136 High altitude exposure                6
## SRR6514137 High altitude exposure                3
## SRR6514138 High altitude exposure                1
## SRR6514139 High altitude exposure                1
##           sra_attribute.source_name sra_attribute.tissue
##                               <character>             <character>
## SRR6514110                blood                blood
## SRR6514111                blood                blood
## SRR6514112                blood                blood
## SRR6514113                blood                blood
## SRR6514114                blood                blood
## ...                        ...                  ...
## SRR6514135                blood                blood
## SRR6514136                blood                blood
## SRR6514137                blood                blood
## SRR6514138                blood                blood
## SRR6514139                blood                blood
```

Nuestros datos no muestran ningun problema aparente por lo que procederemos a analizar un poco la diferencia entre nuestros datos de acuerdo a la fraccion del total de fragmentos asignados por 'featureCounts' que se unieron a un gen en especifico

```
rse_gene_SRP131480$assigned_gene_prop <- rse_gene_SRP131480$recount_qc.gene_fc_count_all.assigned / rse_gene_SRP131480$recount_qc.gene_fc_count_all
summary(rse_gene_SRP131480$assigned_gene_prop)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.3561  0.4015  0.4119  0.4195  0.4258  0.5469
```

Ahora procederemos a visualizar estos mismos resultados pero separando las muestras de mayor altitud a las que se encuentran en plano.

```
with(colData(rse_gene_SRP131480), tapply(assigned_gene_prop, sra_attribute.altitude, summary))
```

```
## $'High altitude exposure'  
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.  
## 0.3561  0.3993  0.4110  0.4068  0.4154  0.4286  
##  
## $Plain  
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.  
## 0.3582  0.4101  0.4219  0.4322  0.4455  0.5469
```