

Generando un informe

Guillermo Ayala Gallego

Generando un informe

Guillermo Ayala Gallego

5/16/23

Justificación

- Hemos visto cómo obtener subconjuntos de genes con expresión diferencial.
- Ya tenemos la lista.
- Queremos explorar este grupo de genes.
- En la práctica estadística clásica se generaba un informe y el investigador interpreta los resultados y redacta un informe.
- En este contexto el problema es mayor.
- Básicamente se explora en la gran cantidad de genes y por ello los resultados han de ser validados y, posiblemente, nuevas hipótesis generadas.
- El modo más útil posiblemente sea generar un archivo en html.
- Nos proponemos el siguiente objetivo:
 - Generar un fichero en que cada fila contenga distintos identificadores de aquellos genes que declaramos con expresión diferencial significativa así como los p-valores originales y los ajustados.

Datos y análisis previo

- Leemos los datos.

```
1 pacman::p_load(Biobase)
2 data(gse20986,package="tamidata")
3
4   • Hacemos análisis de expresión diferencial.
5
6 1 eset = gse20986[, c(2, 3, 5, 10:12)]
7 2 tissue = factor(rep(1:2, each = 3), levels = 1:2,
8 3               labels = c("retina", "huvec"))
9 4 tt = genefilter::rowttests(eset, tissue)
10 5 padj = p.adjust(tt[, "p.value"], method="BH")
11 6 sig = which(padj < 0.05)
```

- Consideramos los p-valores y los p-valores ajustados.

```
1 pvalor = tt[sig,"p.value"]
2 pajustado = padj[sig]
```

Obteniendo identificadores con annotate

- Utilizamos `annotate::lookUp()`

```
1 pacman::p_load(annotate)
```

- ¿Qué paquete de anotación hemos de usar?

```
1 annotation(eset)
```

```
[1] "hgu133plus2"
```

- Cargamos el paquete correspondiente.

```
1 pacman::p_load(hgu133plus2.db)
```

- Obtenemos los identificadores de nuestros genes.

```
1 ID = featureNames(eset)[sig]
2 head(ID)
```

```
[1] "1552487_a_at" "1552626_a_at" "1552701_a_at" "1552703_s_at" "1552730_at"
[6] "1552760_at"
```

- Podemos determinar los nombres abreviados de nuestros genes.

```
1 lookUp(ID, "hgu133plus2.db", "SYMBOL")
2 getSYMBOL(ID, "hgu133plus2.db")
```

- Sus nombres con

```
1 lookUp(ID, "hgu133plus2.db", "GENENAME")
```

- Los identificadores ENSEMBL

```
1 lookUp(ID, "hgu133plus2.db", "ENSEMBL")
```

- O en Gene Ontology,

```
1 lookUp(ID, "hgu133plus2.db", "GO")
```

- O los identificadores ENTREZ.

```
1 lookUp(ID, "hgu133plus2.db", "ENTREZID")
```

Generando informe con ReportingTools

```
1 pacman::p_load(ReportingTools)
```

- Obtenemos distintos identificadores.

```

1 ID = featureNames(eset)[sig]
2 Name = as.character(lookup(ID,"hgu133plus2.db", "GENENAME"))
3 entrezid = as.character(lookup(ID, "hgu133plus2.db", "ENTREZID"))
4 ID[ID == "NA"] = NA
5 Name[Name == "NA"] = NA
6 entrezid = ifelse(entrezid == "NA",NA,
7   paste0("<a href='http://www.ncbi.nlm.nih.gov/gene/?term=",
8     entrezid,">",entrezid,"</a>"))

```

- Posiblemente queramos añadir a estos descriptores los p-valores originales así como los ajustados.

```

1 df = data.frame(ID = ID,Name = Name,entrezid = entrezid,
2   pvalor = pvalor,pajustado = pajustado,stringsAsFactors=FALSE)

```

- Utilizando `ReportingTools::HTMLReport()` fijamos el nombre del fichero en que guardamos la información así como el directorio en donde estará dicho fichero.

```

1 foutput = "gse20986.DE"
2 htmlRep1 = HTMLReport(shortName = foutput,title = foutput,
3   reportDirectory = "./reports")

```

- Guardamos la información y cerramos el fichero con las funciones `ReportingTools::publish()` y `ReportingTools::finish()`.

```

1 publish(df,htmlRep1)
2 finish(htmlRep1)

[1] "./reports/gse20986.DE.html"

1 browseURL("./reports/gse20986.DE.html")

```

Utilizando `AnnotationDbi::select`

```

1 a = AnnotationDbi::select(hgu133plus2.db,keys=ID,
2   columns=c("ENTREZID","ENSEMBL"),
3   keytype="PROBEID")
4 b = BiocGenerics::match(ID,a[, "PROBEID"])
5 df0 = a[b,]
6 names(df0)

[1] "PROBEID" "ENTREZID" "ENSEMBL"

1 df = data.frame(ID = ID,Name = Name,entrezid = df0$ENTREZID,
2   pvalor = pvalor,pajustado = pajustado,stringsAsFactors=F)

```

Y lo demás igual que en la opción previa.