

geod71810

Cristian Pérez and Guillermo Ayala

2024-04-02

Table of contents

Packages	1
Downloading and preprocessing the data	1
Normalization	2
Making the ExpressionSet	2
Adding identifiers	3

Packages

The data can be found at

```
library(Biobase)
library(GEOquery)
library(affy)
library(hgu133plus2.db)
```

Downloading and preprocessing the data

```
GEOquery::getGEOSuppFiles("GSE71810")
system("tar xvf GSE71810/GSE71810_RAW.tar")
# Reading CEL files to a affybatch
affybatch = ReadAffy()
system("rm -fr GSE71810")
system("rm *CEL.gz")
```

Normalization

- Normalization using RMA

```
eset = rma(affybatch)
```

- Obtaining expression matrix.

```
ExprMatrix = exprs(eset)
```

Making the ExpressionSet

- Experimental data. The file ‘

```
infoData = new('MIAME',
  name='Johann PD et al.',
  lab='Not Specified',
  contact='Marcel Kool <m.kool@dkfz.de>',
  title = 'Atypical Teratoid/Rhabdoid Tumors Are Comprised of Three
  Epigenetic Subgroups with Distinct Enhancer Landscapes.',
  abstract = 'Summary: B lymphocytes are important mediators of
  adaptive immunity. They drive naïve T cells to different T-helper profiles.
  Th2 responses are implicated in allergic reactions.
  The goal of this study was to characterize the gene expression profile of B
  cells under Th2-skewing conditions.
```

Overall design B lymphocytes were isolated from human tonsils and cultured for 24 hours under Th2-skewing stimulation (anti-CD40 and IL4 treatment) or with no stimulation.

RNA was isolated from 4 biological replicates for each condition. 8 samples in total were hybridized to Affymetrix gene expression microarrays.

variable: treatment: antiCD40-IL4: antiCD40-IL4_1, antiCD40-IL4_2, antiCD40-IL4_3, antiCD40-IL4_4

variable: treatment: unstimulated: Control_1, Control_2, Control3, Control_4

repeat: biological replicate: antiCD40-IL4_1, antiCD40-IL4_2, antiCD40-IL4_3, antiCD40-IL4_4
Control_4',

```
url = 'http://www.ebi.ac.uk/arrayexpress/experiments/E-GEOD-71810/')
```

```
pd0 = data.frame(type = factor(c(0,1,0,1,0,1,0,1),
  levels = 0:1,labels=c("control","treatment")))
```

```

rownames(pd0) = colnames(ExprMatrix)
metadatos = data.frame(labelDescription = colnames(pd0), row.names=colnames(pd0))
datosfenotipo = new("AnnotatedDataFrame", data = pd0,
  varMetadata = metadatos)
geod71810 = new("ExpressionSet",
  exprs=ExprMatrix,
  phenoData = datosfenotipo,
  experimentData = infoData,
  annotation='hgu133plus2')

```

Adding identifiers

```

a = AnnotationDbi::select(hgu133plus2.db::hgu133plus2.db,
  keys=featureNames(geod71810), column=c(
b = match(featureNames(geod71810), a[, "PROBEID"])
fData(geod71810) = a[b,]
dim(geod71810)
dim(fData(geod71810))
save(geod71810, file="geod71810.rda")

```