

Bi plaid models

```
library("biclust")
```

```
## Warning: package 'biclust' was built under R version 3.5.1
```

```
## Loading required package: MASS
```

```
## Loading required package: grid
```

```
## Loading required package: colorspace
```

```
## Loading required package: lattice
```

```
library("superbiclust")
```

```
## Warning: package 'superbiclust' was built under R version 3.5.1
```

```
## Loading required package: fabia
```

```
## Loading required package: Biobase
```

```
## Loading required package: BiocGenerics
```

```
## Loading required package: parallel
```

```
##  
## Attaching package: 'BiocGenerics'
```

```
## The following objects are masked from 'package:parallel':  
##  
##   clusterApply, clusterApplyLB, clusterCall, clusterEvalQ,  
##   clusterExport, clusterMap, parApply, parCapply, parLapply,  
##   parLapplyLB, parRapply, parSapply, parSapplyLB
```

```
## The following objects are masked from 'package:stats':  
##  
##   IQR, mad, sd, var, xtabs
```

```
## The following objects are masked from 'package:base':
##
##   anyDuplicated, append, as.data.frame, basename, cbind,
##   colMeans, colnames, colSums, dirname, do.call, duplicated,
##   eval, evalq, Filter, Find, get, grep, grepl, intersect,
##   is.unsorted, lapply, lengths, Map, mapply, match, mget, order,
##   paste, pmax, pmax.int, pmin, pmin.int, Position, rank, rbind,
##   Reduce, rowMeans, rownames, rowSums, sapply, setdiff, sort,
##   table, tapply, union, unique, unsplit, which, which.max,
##   which.min
```

```
## Welcome to Bioconductor
##
##   Vignettes contain introductory material; view with
##   'browseVignettes()'. To cite Bioconductor, see
##   'citation("Biobase")', and for packages 'citation("pkgname")'.
```

```
## +-----+
## |.....|
## |.....|
## |.....#####.....| ##### # ##### ### #
## |.....#####.....| # # # # # # #
## |.....#####.....| # # # # # # #
## |.....#####.....| ##### # ##### # # #
## |.....#####.....| # ##### # # # #####
## |.....#####.....| # # # # # # #
## |.....#####.....| # # # ##### ### # #
## |.....#####.....|
## |.....|
## +-----+
```

```
## Citation: S. Hochreiter et al.,
## FABIA: Factor Analysis for Biclust Acquisition,
## Bioinformatics 26(12):1520-1527, 2010.
## BibTex: enter 'toBibtex(citation("fabia"))'
##
## Homepage: http://www.bioinf.jku.at/software/fabia/fabia.html
##
## FABIA Package Version 2.26.0
```

```
##
## Attaching package: 'superbiclust'
```

```
## The following object is masked from 'package:Biobase':
##
##   combine
```

```
## The following object is masked from 'package:BiocGenerics':  
##  
##      combine
```

```
library("fabia")
```

```
mirnadata<-read.csv("C:\\Users\\samfero\\Documents\\Mirnadata.csv")  
mirnaplaid<-read.csv("C:\\Users\\samfero\\Documents\\Mirna_plaid.csv")
```

BIPLAID MODEL

First approach : intercept only model

```
mirnadat.plaid<-as.matrix(mirnaplaid[-1])  
set.seed(1237)  
bics<-biclust(mirnadat.plaid, method=BCPlaid(), cluster="b", fit.model = y ~ m,  
background = TRUE, row.release = 0.7, col.release = 0.7,  
shuffle = 3, back.fit = 0, max.layers = 20, iter.startup = 5,  
iter.layer = 10, verbose = TRUE)
```

```

## layer: 0
## 521002.6
## layer: 1
## [1] 0 363 204
## [1] 1 764 189
## [1] 2 764 180
## [1] 3 764 177
## [1] 10 764 177
## [1] 11 764 177
## [1] 12 764 143
## [1] 13 764 143
## [1] 14 764 130
## [1] 15 764 130
## [1] 16 764 122
## [1] 17 764 122
## [1] 18 764 119
## [1] 19 764 119
## [1] 20 764 119
## [1] 4
## [1] 8041667 0 0 0
## layer: 2
## [1] 0 202 205
## [1] 1 764 161
## [1] 2 764 147
## [1] 3 764 144
## [1] 10 764 144
## [1] 11 745 144
## [1] 12 745 111
## [1] 13 745 111
## [1] 14 745 102
## [1] 15 745 102
## [1] 16 745 100
## [1] 17 745 100
## [1] 18 745 99
## [1] 19 745 99
## [1] 20 745 99
## [1] 4
## [1] 1518384 0 0 0
## layer: 3
## [1] 0 204 242
## [1] 1 743 156
## [1] 2 764 133
## [1] 3 764 127
## [1] 4 764 124
## [1] 5 764 123
## [1] 6 764 121
## [1] 10 764 121
## [1] 11 227 121
## [1] 12 227 76
## [1] 13 227 76
## [1] 14 227 67
## [1] 15 227 67
## [1] 16 227 62

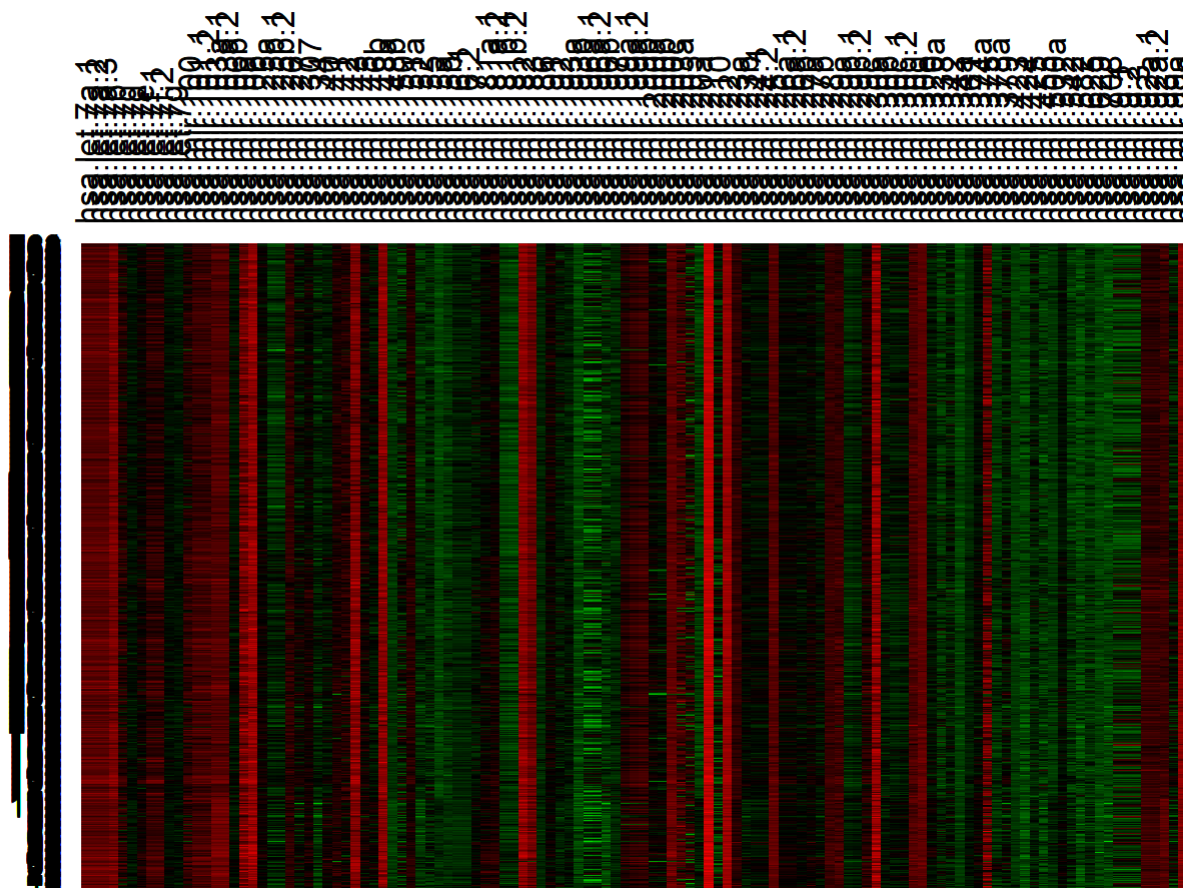
```

```
## [1] 17 227 62
## [1] 18 227 58
## [1] 19 227 58
## [1] 20 227 58
## [1] 7
## [1] 93481.58      0.00      0.00      0.00
## layer: 4
## [1] 0 227 273
## [1] 1 738 181
## [1] 2 727 154
## [1] 3 715 148
## [1] 4 714 146
## [1] 5 712 145
## [1] 6 711 144
## [1] 7 710 144
## [1] 8 710 143
## [1] 9 708 143
## [1] 10 708 143
## [1] 11 0 143
## [1] 12
## [1] 0 0 0 0
##
## Layer Rows Cols Df          SS          MS Convergence Rows Released
##    0  764 1446  1  521002.57  521002.57          NA          NA
##    1  764  119  1  8041666.66  8041666.66          1          0
##    2  745   99  1  1518384.25  1518384.25          1         19
##    3  227   58  1   93481.58   93481.58          1        537
##
## Layer Cols Released
##    0          NA
##    1          58
##    2          45
##    3          63
```

```
print(bics)
```

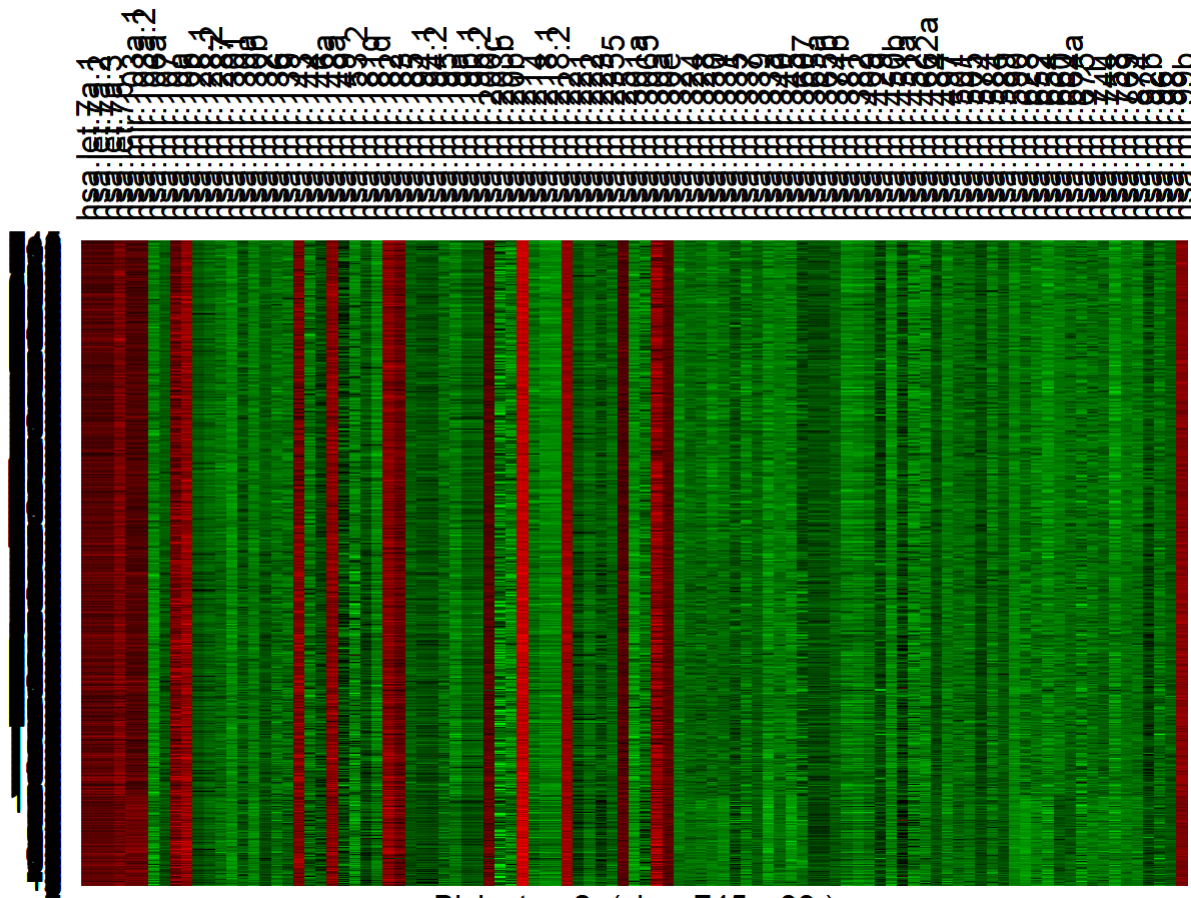
```
##
## An object of class Biclust
##
## call:
## biclust(x = mirnadat.plaid, method = BCPlaid(), cluster = "b",
##         fit.model = y ~ m, background = TRUE, row.release = 0.7,
##         col.release = 0.7, shuffle = 3, back.fit = 0, max.layers = 20,
##         iter.startup = 5, iter.layer = 10, verbose = TRUE)
##
## Number of Clusters found: 3
##
## First 3 Cluster sizes:
##              BC 1 BC 2 BC 3
## Number of Rows:   764  745  227
## Number of Columns: 119   99   58
```

```
drawHeatmap(mirnadat.plaid,bics,1)
```



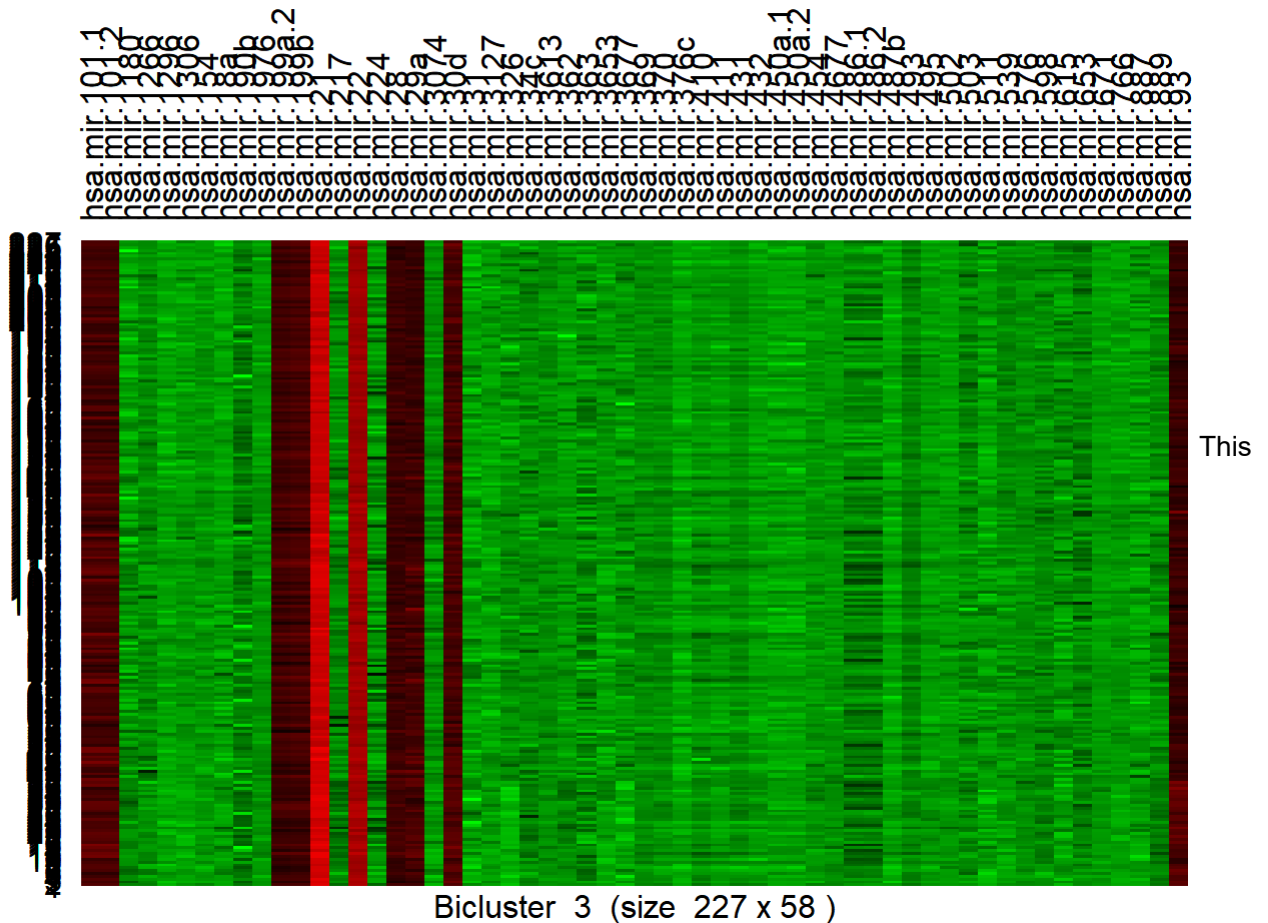
Biclust 1 (size 764 x 119)

```
drawHeatmap(mirnadat.plaid,bics,2)
```



Bicluster 2 (size 745 x 99)

```
drawHeatmap(mirnadat.plaid,bics,3)
```



model gave us 3 clusters.

Fitting $Y \sim M + A + B$ (Ordinary Least squares model with biclustering)

```
set.seed(1200)
bics.linear<-biclust(mirnadat.plaid, method=BCPlaid(), cluster="b", fit.model = y ~ m + a + b,
background = TRUE,
shuffle = 3, back.fit = 2, max.layers = 20, iter.startup = 5,
iter.layer = 10, verbose = TRUE)
```



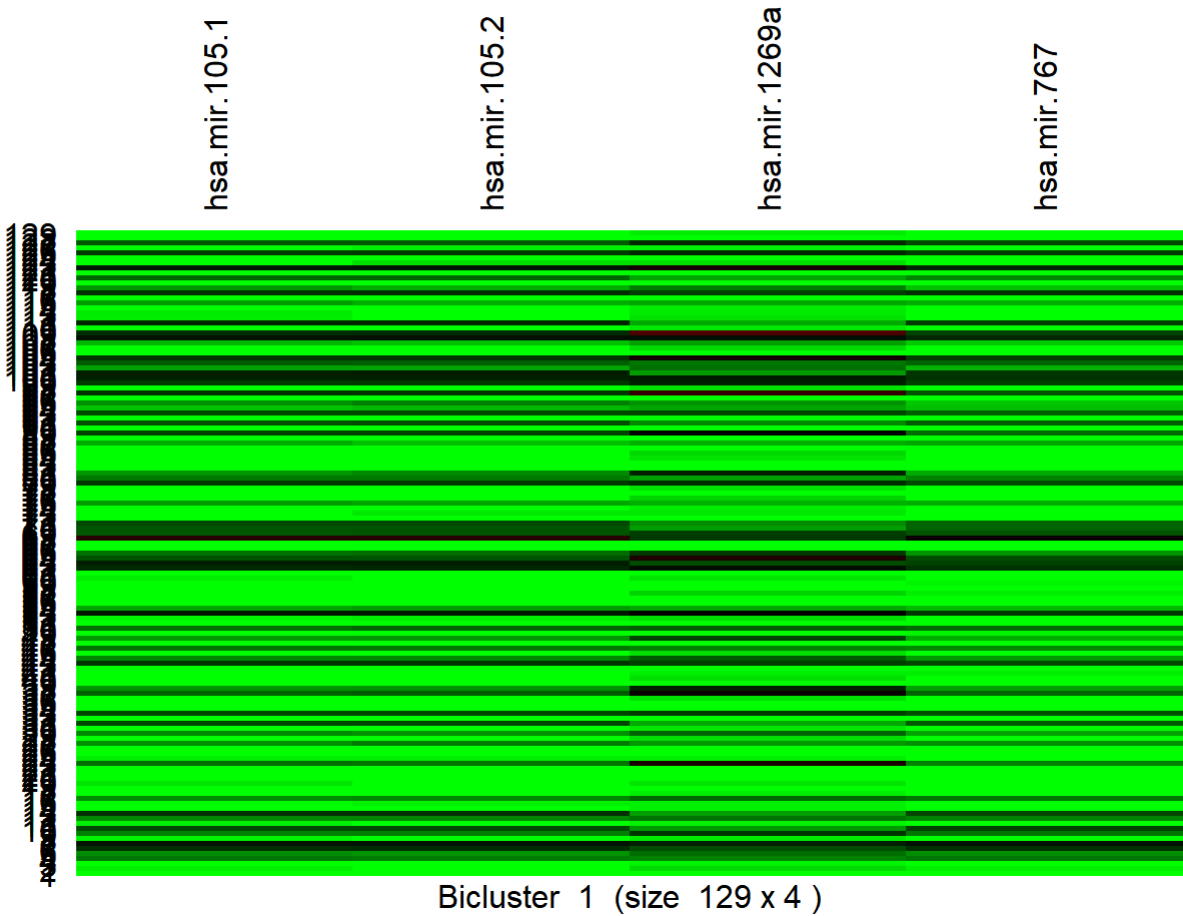
```

## layer: 0
## 12776297
## layer: 1
## [1] 0 344 4
## [1] 1 423 4
## [1] 2 414 4
## [1] 3 408 4
## [1] 4 407 4
## [1] 5 404 4
## [1] 6 403 4
## [1] 7 404 4
## [1] 8 402 4
## [1] 9 401 4
## [1] 10 402 4
## [1] 11 158 4
## [1] 12 158 4
## [1] 13 139 4
## [1] 14 139 4
## [1] 15 130 4
## [1] 16 130 4
## [1] 17 129 4
## [1] 18 129 4
## [1] 19 129 4
## [1] 20 129 4
## [1] 20
## [1] 7576.685 0.000 0.000 0.000
## back fitting 2 times
## layer: 2
## [1] 0 343 118
## [1] 1 294 117
## [1] 2 283 117
## [1] 3 278 117
## [1] 4 276 117
## [1] 10 276 117
## [1] 11 3 117
## [1] 12 3 84
## [1] 13 3 84
## [1] 14 3 81
## [1] 15 3 81
## [1] 16 3 80
## [1] 17 3 80
## [1] 20 3 80
## [1] 5
## [1] 1124.414 0.000 0.000 0.000
## back fitting 2 times
## layer: 3
## [1] 0 350 164
## [1] 1 329 151
## [1] 2 300 143
## [1] 3 268 137
## [1] 4 240 133
## [1] 5 225 132
## [1] 6 215 132

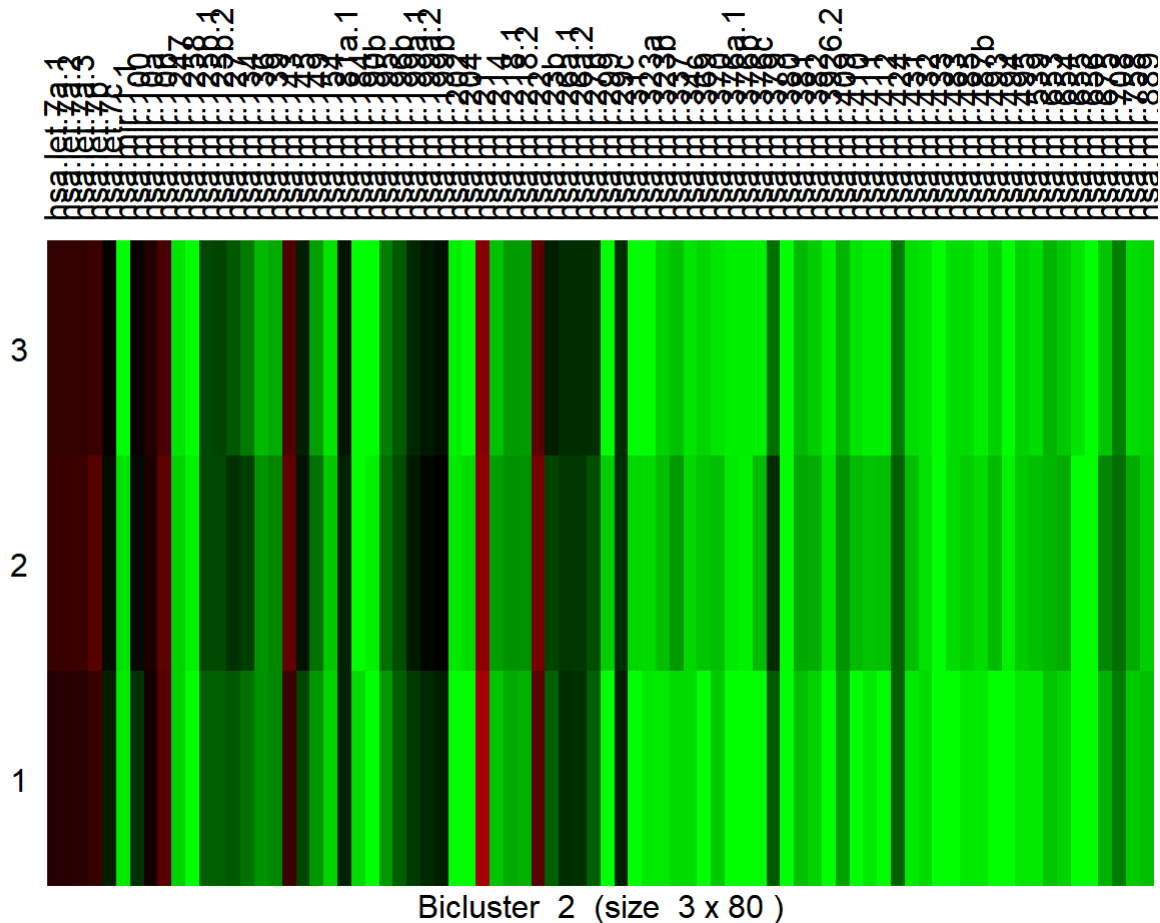
```

```
## [1] 7 206 132
## [1] 8 200 132
## [1] 9 196 131
## [1] 10 195 131
## [1] 11 0 131
## [1] 12
## [1] 0 0 0 0
##
## Layer Rows Cols Df SS MS Convergence Rows Released
## 0 764 1446 2209 12781543.23 5786.12 NA NA
## 1 129 4 132 8107.08 61.42 0 273
## 2 3 80 82 1254.50 15.30 1 273
##
## Layer Cols Released
## 0 NA
## 1 0
## 2 37
```

```
drawHeatmap(mirnadat.plaid,bics.linear,1)
```



```
drawHeatmap(mirnadat.plaid,bics.linear,2)
```



```
print(bics.linear)
```

```
##
## An object of class Biclust
##
## call:
## biclust(x = mirnadat.plaid, method = BCPlaid(), cluster = "b",
##         fit.model = y ~ m + a + b, background = TRUE, shuffle = 3,
##         back.fit = 2, max.layers = 20, iter.startup = 5, iter.layer = 10,
##         verbose = TRUE)
##
## Number of Clusters found: 2
##
## First 2 Cluster sizes:
##               BC 1 BC 2
## Number of Rows:   129   3
## Number of Columns:  4  80
```

Fitting a linear model just gave u two clusters. Tried to map it ER+ AND ER- subtypes with not good results.

Fitting Fabia model

Hardcoded 5 clusters to match the subtypes. But difficult to interpret the clusters. Need help for this.

```
mirnadat<-as.matrix(mirnadata[-1])  
res <- fabia(mirnadat,5,0.01,500)
```

```
## Running FABIA on a 1446x764 matrix with
```

```
##   Number of biclusters ----- p: 5
```

```
##   Sparseness factor ----- alpha: 0.01
```

```
##   Number of iterations ----- cyc: 500
```

```
##   Loading prior parameter ----- spl: 0
```

```
##   Factor prior parameter ----- spz: 0.5
```

```
##   Initialization loadings----- random: 1 = interval
```

```
##   Nonnegative Loadings and Factors -----: 0 = No
```

```
##   Centering ----- center: 2 = median
```

```
##   Quantile scaling (0.75-0.25): ---- norm: 1 = Yes
```

```
##   Scaling loadings per iteration -- scale: 0 = No
```

```
##   Constraint variational parameter -- lap: 1
```

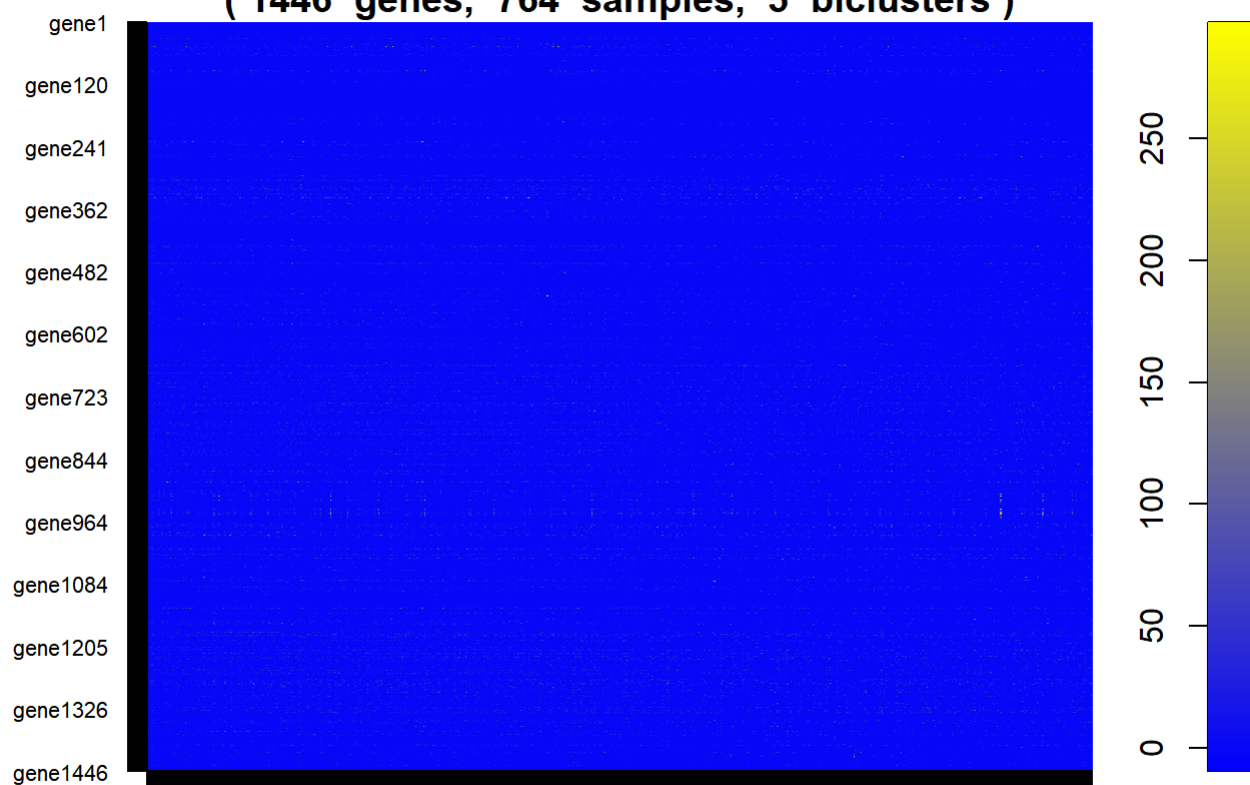
```
##   Max. number of biclusters per row -- nL: 0 = no limit
```

```
##   Max. number of row elements / biclu. lL: 0 = no limit
```

```
## Cycle: 0  
Cycle: 20  
Cycle: 40  
Cycle: 60  
Cycle: 80  
Cycle: 100  
Cycle: 120  
Cycle: 140  
Cycle: 160  
Cycle: 180  
Cycle: 200  
Cycle: 220  
Cycle: 240  
Cycle: 260  
Cycle: 280  
Cycle: 300  
Cycle: 320  
Cycle: 340  
Cycle: 360  
Cycle: 380  
Cycle: 400  
Cycle: 420  
Cycle: 440  
Cycle: 460  
Cycle: 480  
Cycle: 500
```

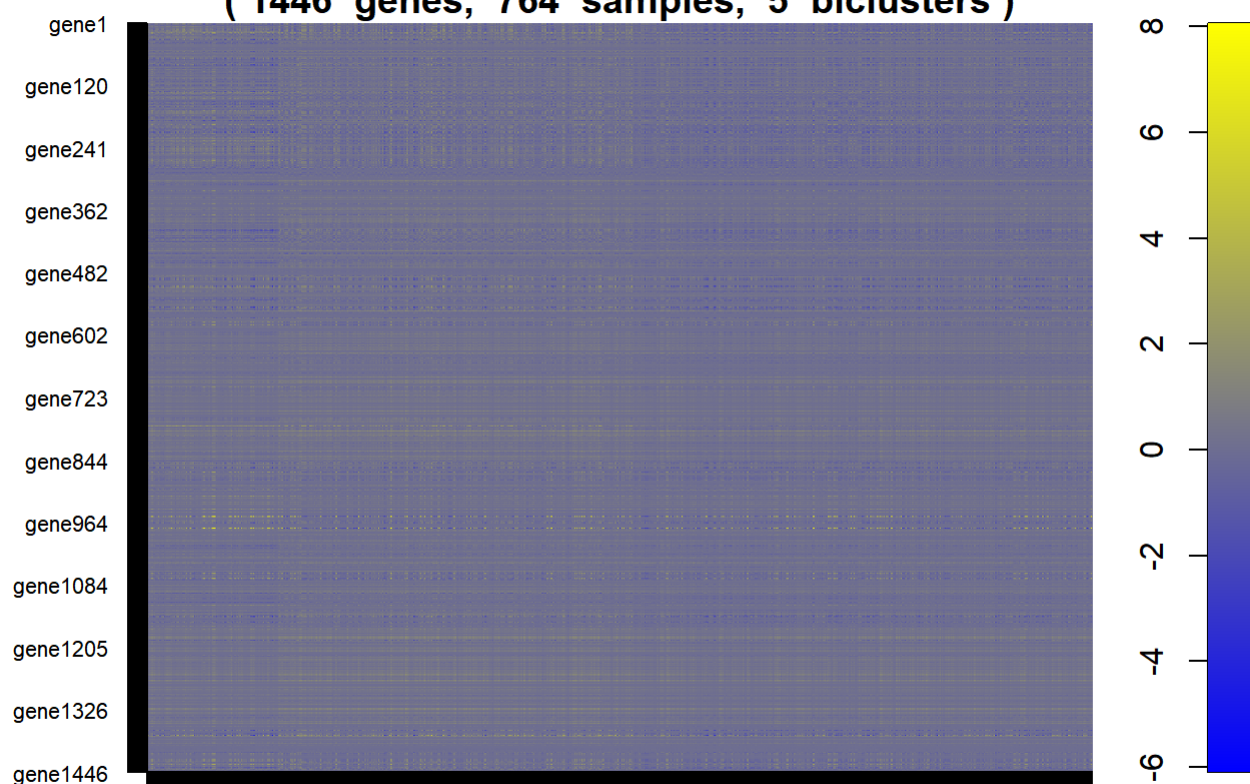
```
extractPlot(res,ti="FABIA")
```


FABIA: data **(1446 genes, 764 samples, 5 biclusters)**



TCGA.A7.A0DC.01A.11R.A010.13 TCGA.A2.A0CR.01A.11R.A22I.13 TCGA.A2.A0EO.01A.11R.A035.13

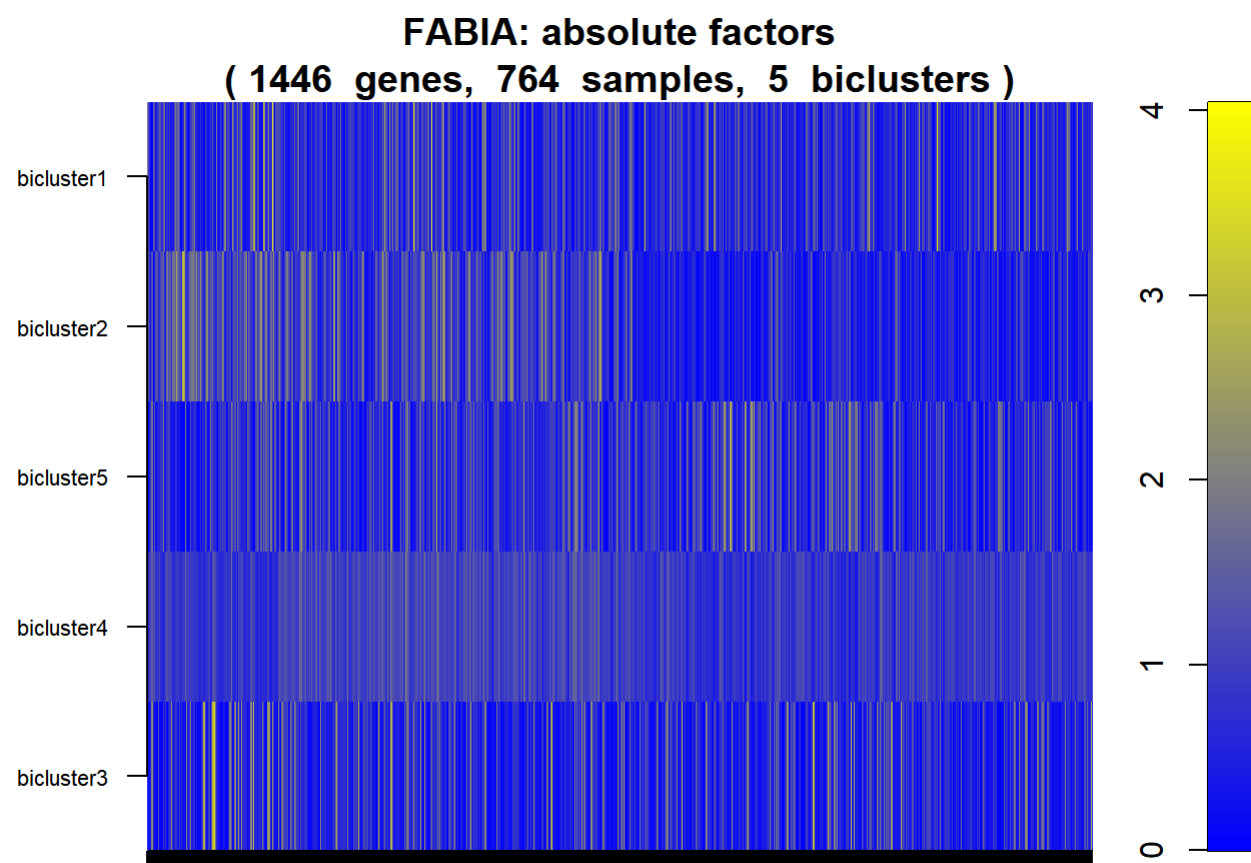
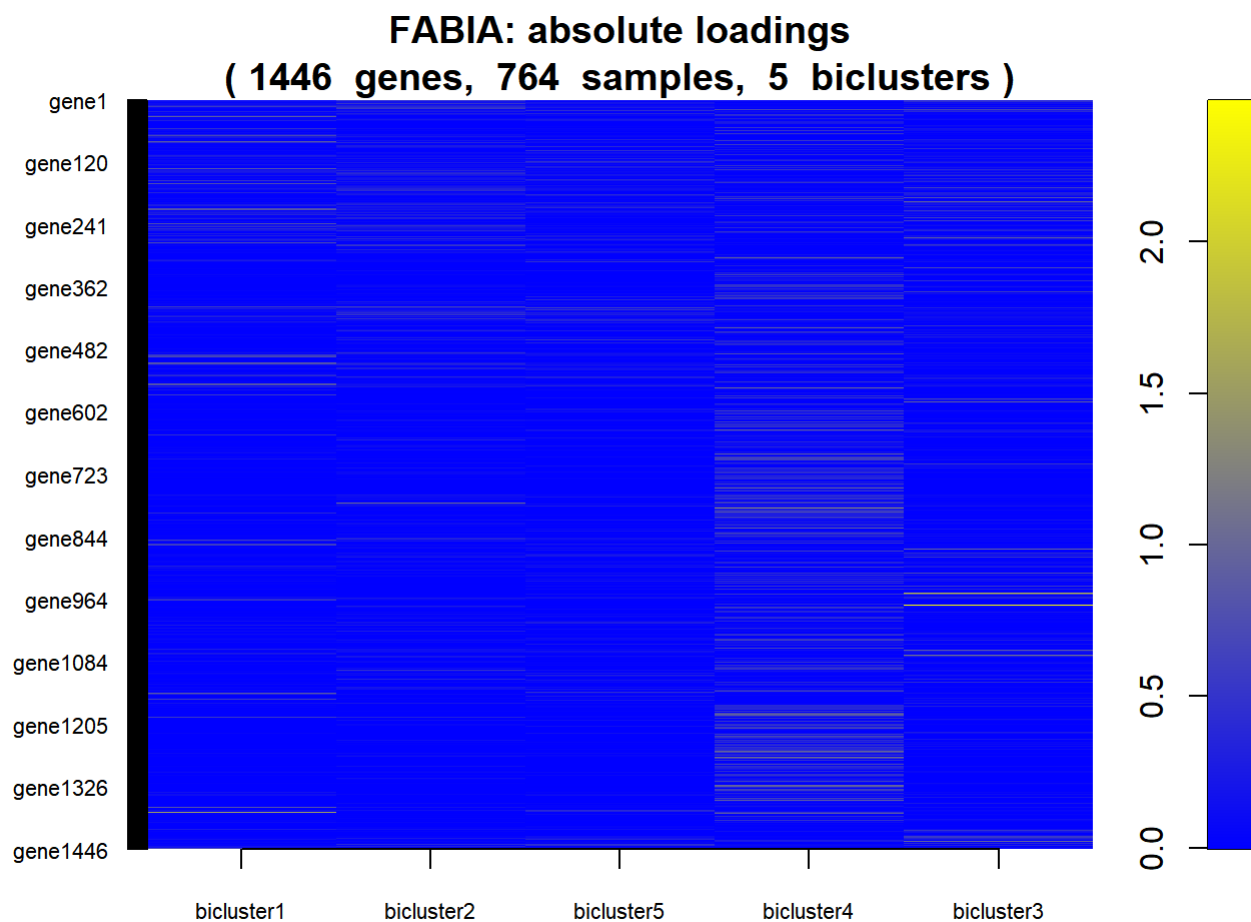
FABIA: reconstructed data **(1446 genes, 764 samples, 5 biclusters)**



TCGA.A7.A0DC.01A.11R.A010.13 TCGA.A2.A0CR.01A.11R.A22I.13 TCGA.A2.A0EO.01A.11R.A035.13

FABIA: error





TCGA.A7.A0DC.01A.11R.A010.13 TCGA.A2.A0CR.01A.11R.A22I.13 TCGA.A2.A0EO.01A.11R.A035.13

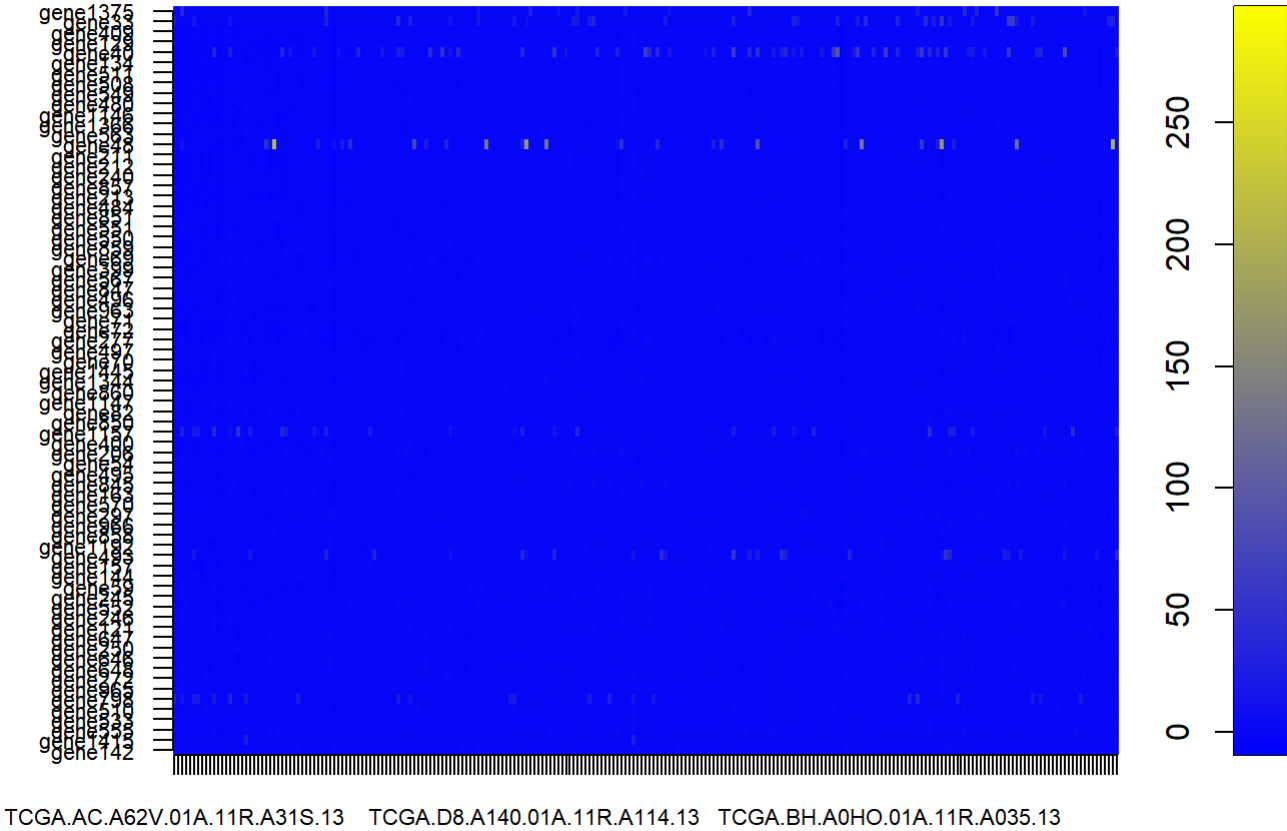
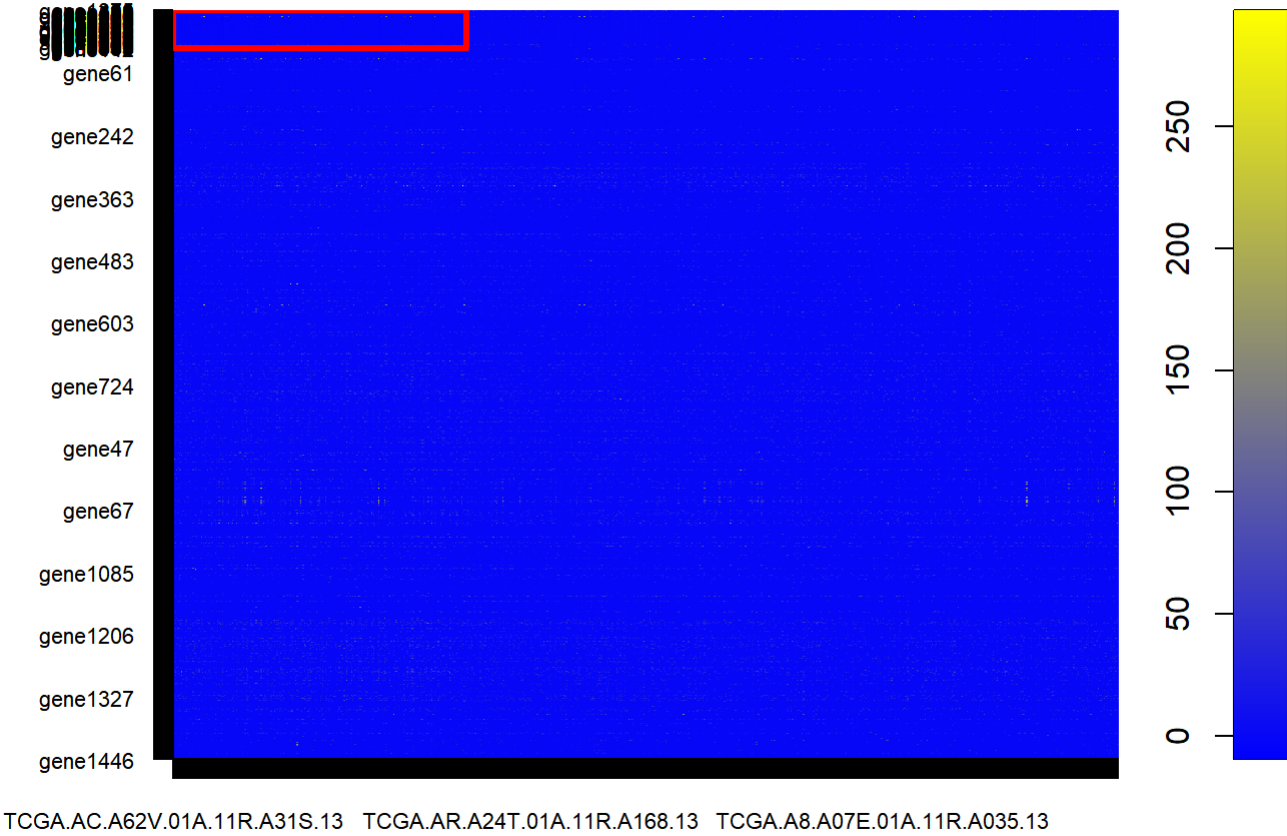
```
#Extracting biclusters
```

```
rb <- extractBic(res)
```

```
#Exploring Biclusters 1
```

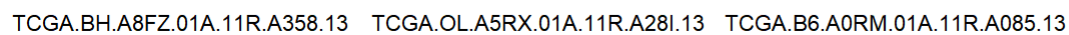
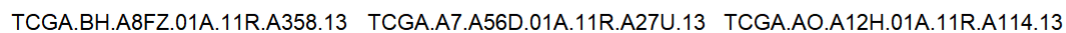
```
plotBiclusters(rb,1)
```

```
## Warning in if (is(r) == "Factorization") {: the condition has length > 1  
## and only the first element will be used
```



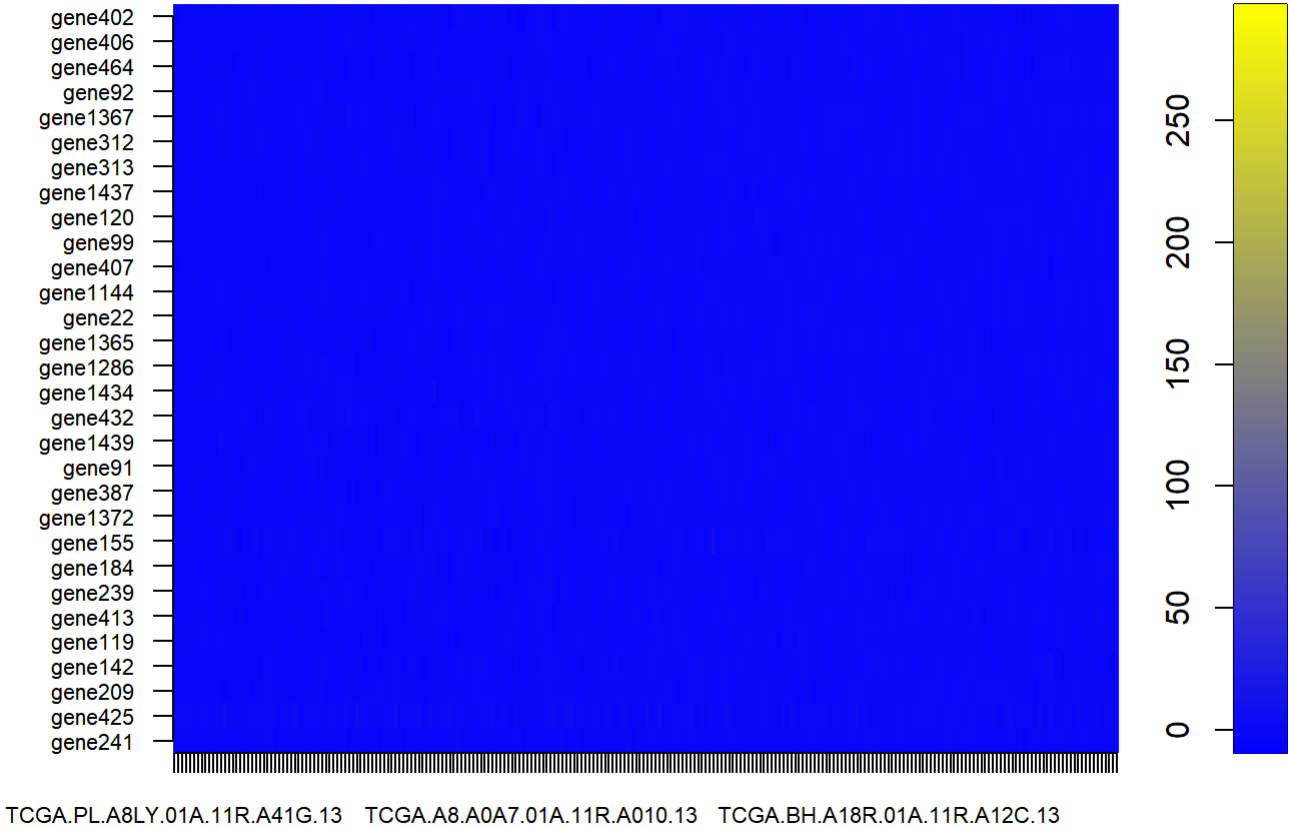
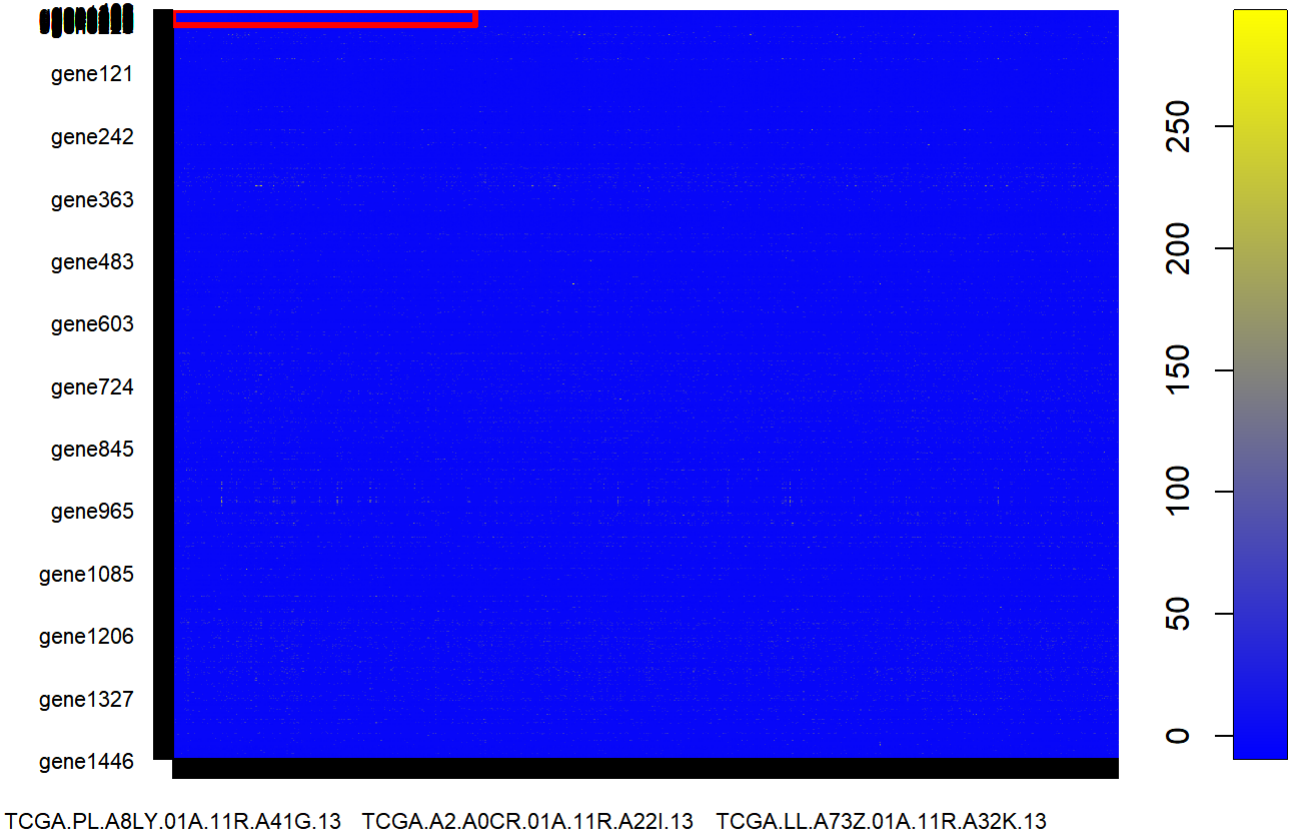
```
#Exploring Biclust 2  
plotBiclust(rb,2)
```

```
## Warning in if (is(r) == "Factorization") {: the condition has length > 1  
## and only the first element will be used
```



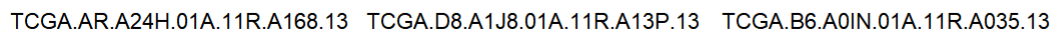
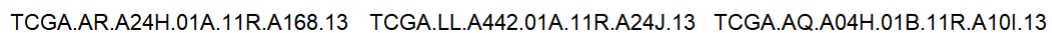
```
#Exploring Biclusters 3  
plotBiclusters(rb,3)
```

```
## Warning in if (is(r) == "Factorization") {: the condition has length > 1  
## and only the first element will be used
```



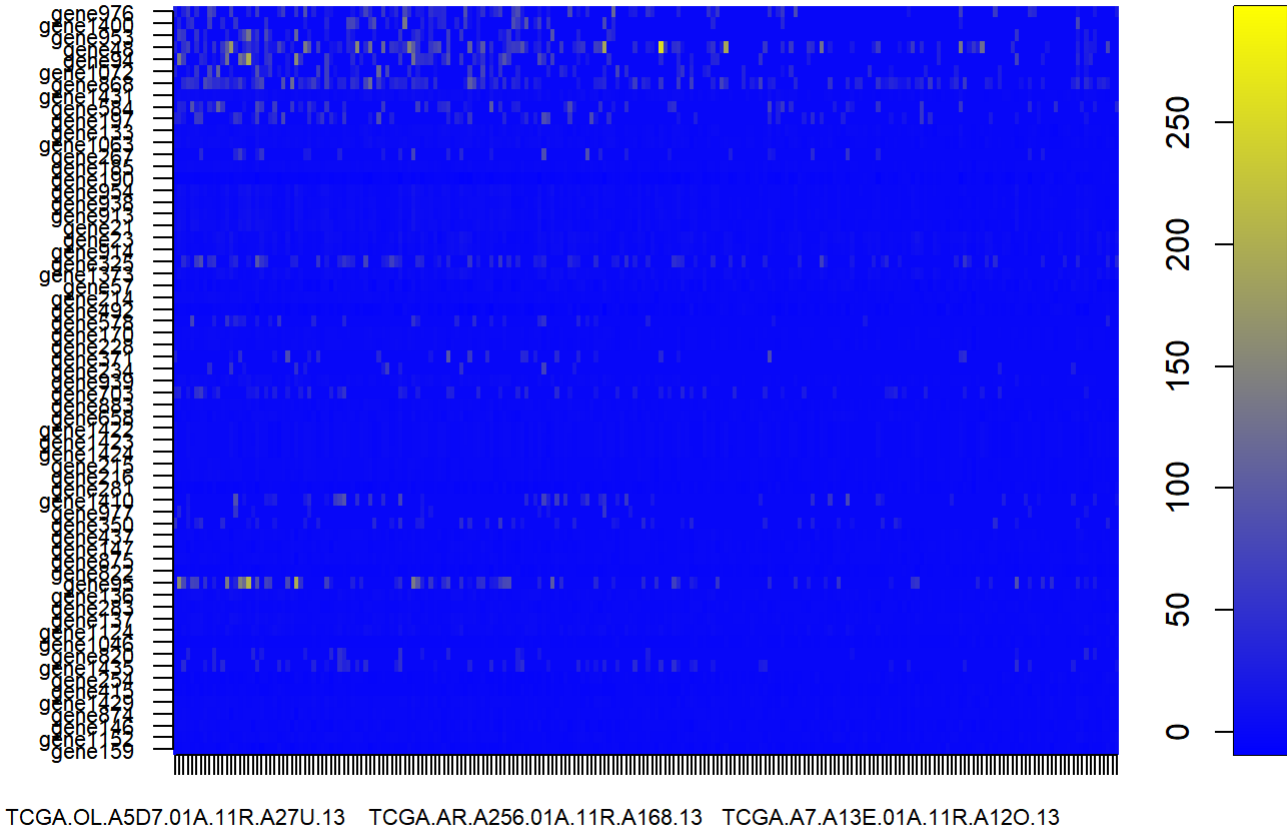
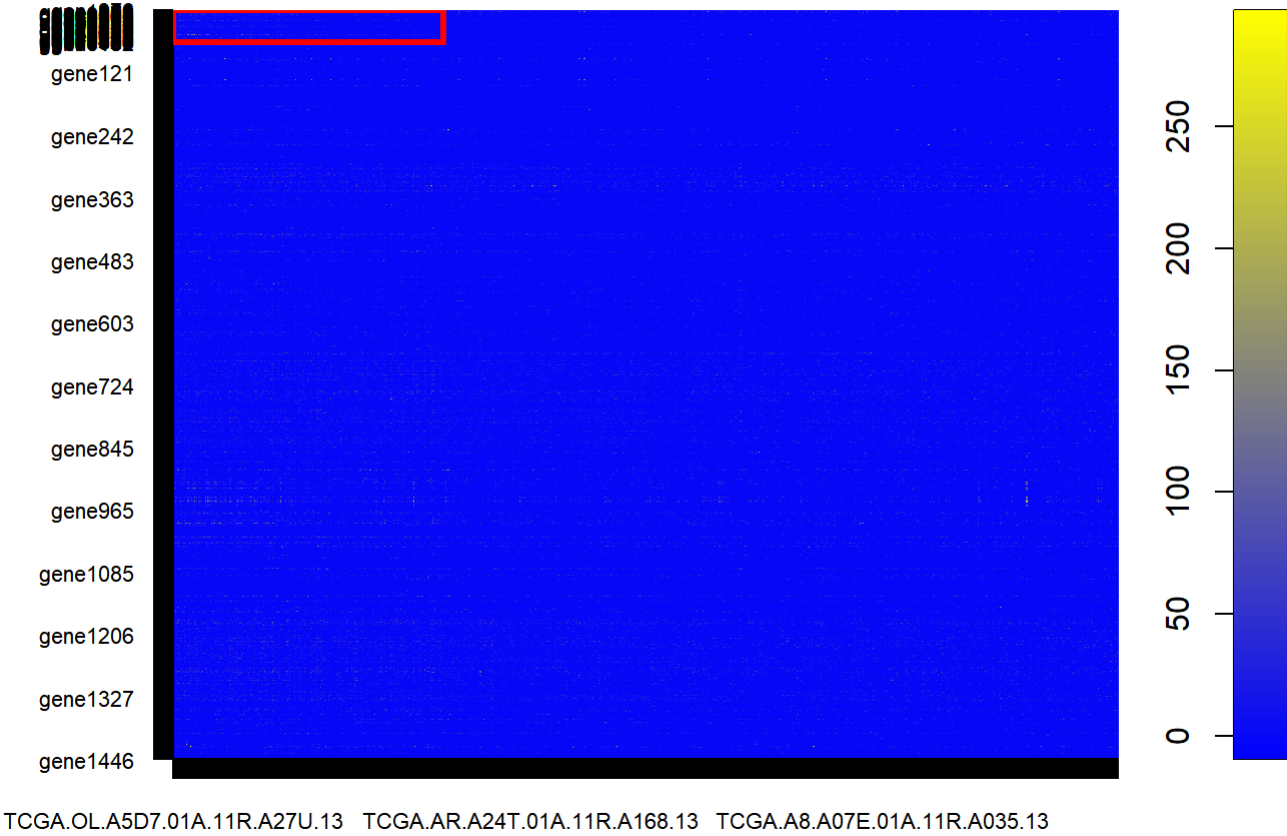
```
#Exploring Biclusters 4  
plotBiclusters(rb,4)
```

```
## Warning in if (is(r) == "Factorization") {: the condition has length > 1  
## and only the first element will be used
```

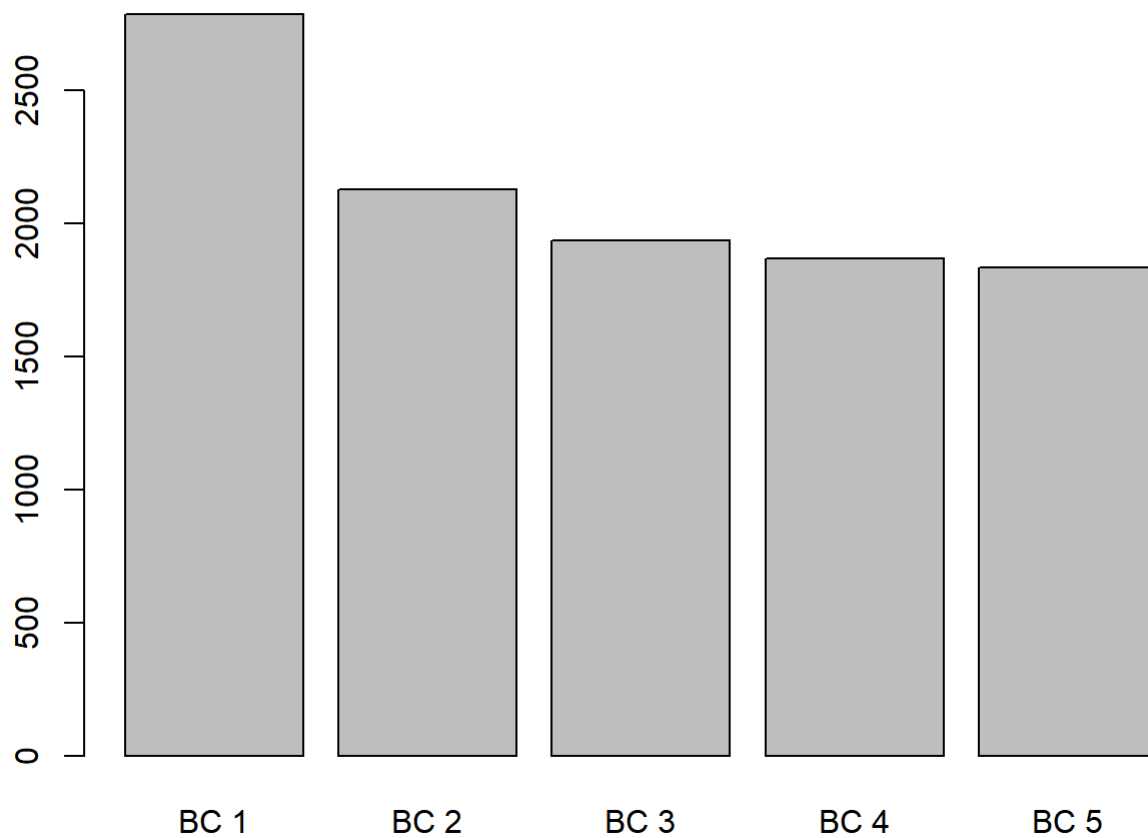
```
#Exploring Biclusters 5  
plotBiclusters(rb,5)
```

```
## Warning in if (is(r) == "Factorization") {: the condition has length > 1  
## and only the first element will be used
```

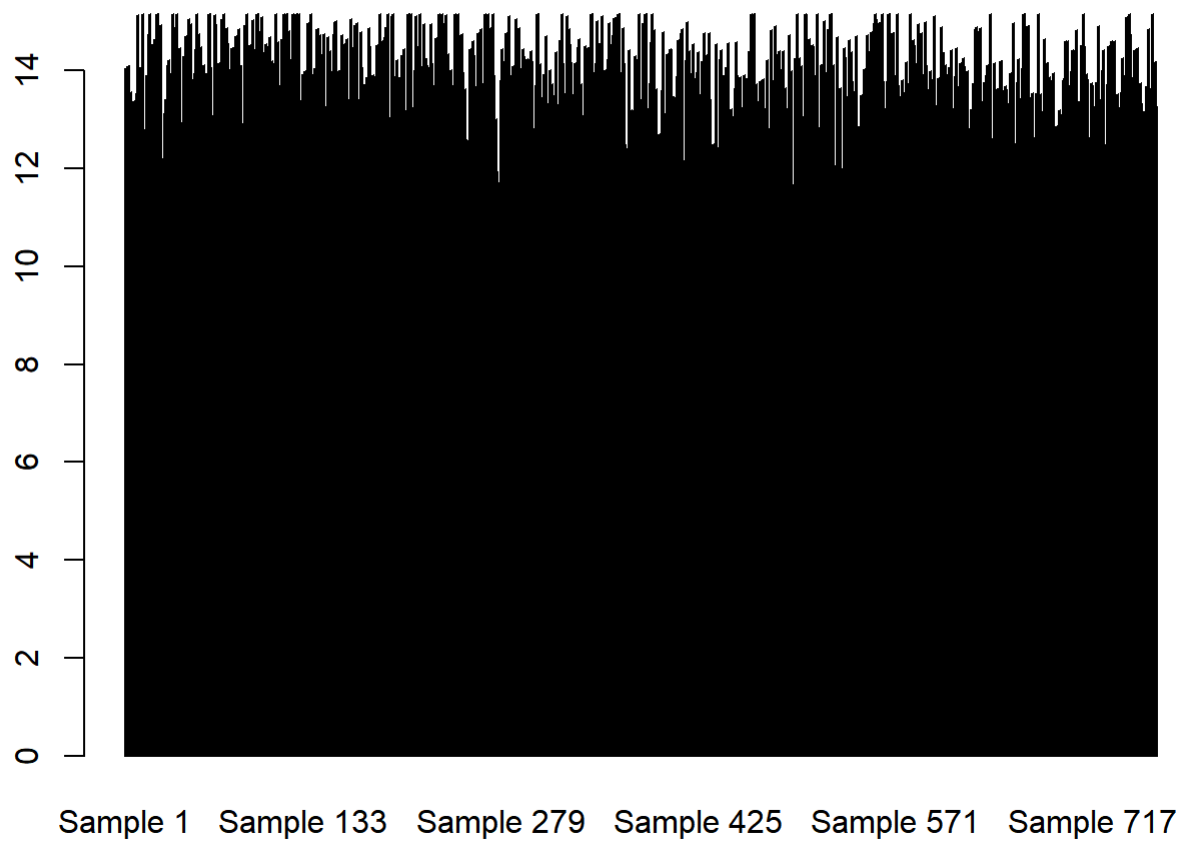


```
print(res)
```

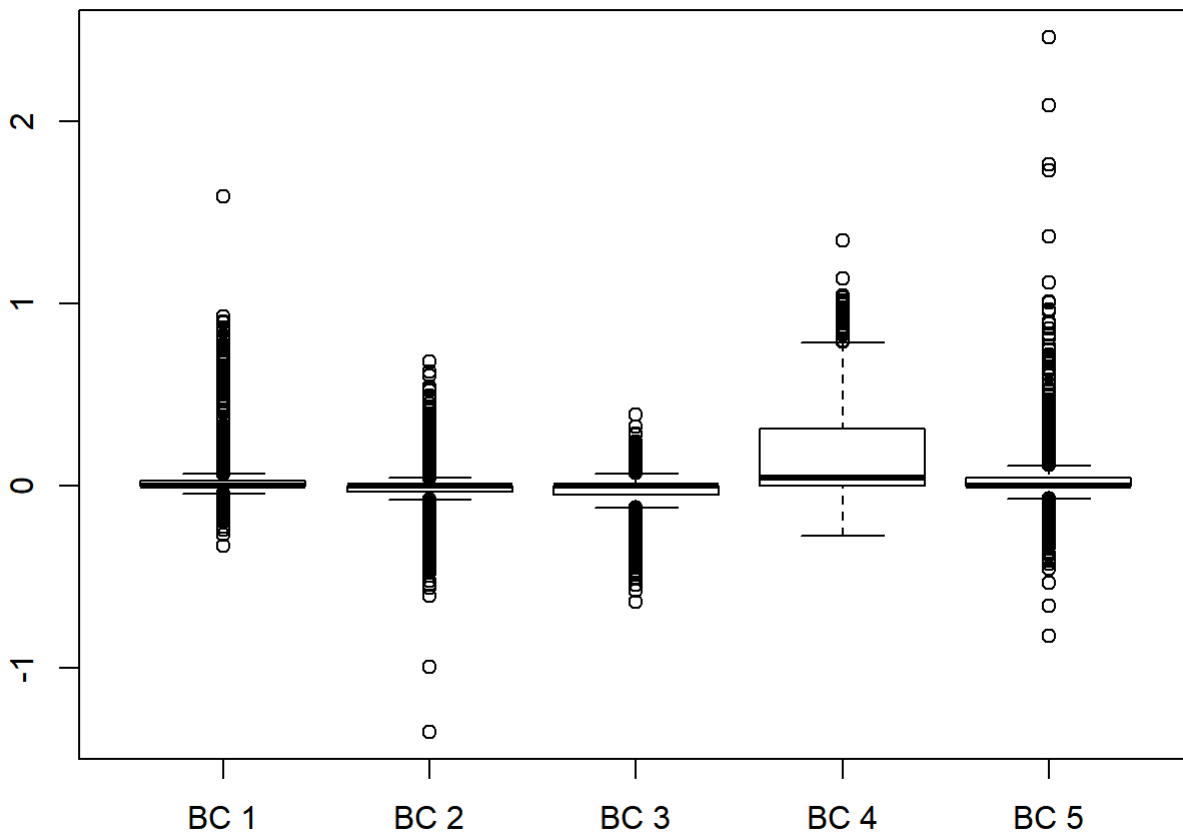
Information Content of Biclusters



Information Content of Samples



Loadings of the Biclusters



Factors of the Biclusters

