Full RCode

```
#LSM3241 CA1 R Code

#install Bioconductor
if (!requireNamespace("BiocManager"))
  install.packages("BiocManager")
BiocManager::install()

#install Bioconductor packages
BiocManager::install('foo')

#install relevant packages needed
if (!requireNamespace("BiocManager", quietly = TRUE))
  install.packages("BiocManager")
BiocManager::install("GEOquery", version = "3.8")

BiocManager::install("affy")
BiocManager::install("limma")
BiocManager::install("hgu133plus2.db")
BiocManager::install("org.Hs.eg.db")

#load packages installed
library(affy)
library(limma)
library(hgu133plus2.db)
library(org.Hs.eg.db)

#check if packages are loaded
sessionInfo()

#Calling GSE50697
library(GEOquery)
#downloading the gse file
gse <- getGEO('GSE50697',GSEMatrix = FALSE)

#retrieving whole GSE
names(GSMList(gse))
```

```
[1] "GSM1226581" "GSM1226582" "GSM1226583" "GSM1226584" "GSM1226585"
"GSM1226586"
```

```
#Getting the raw data for the series using GEOquery
filePaths <- getGEOSuppFiles('GSE50697')
#generating a vector containing the name of all the CEL files.
list.celfiles('GSE50697_RAW')
```

```
[1] "GSM1226581_S1100958.MDA.01_HG-U133_Plus_2_.CEL.gz"
[2] "GSM1226582_S1100958.MDA.02_HG-U133_Plus_2_.CEL.gz"
[3] "GSM1226583_S1100958.MDA.03_HG-U133_Plus_2_.CEL.gz"
[4] "GSM1226584_S1100958.MDA.04_HG-U133_Plus_2_.CEL.gz"
[5] "GSM1226585_S1100958.MDA.05_HG-U133_Plus_2_.CEL.gz"
[6] "GSM1226586_S1100958.MDA.06_HG-U133_Plus_2_.CEL.gz"
```

```
#finding info from meta data of gse
names(Meta(gse))
```

```
[1] "contact_address"          "contact_city"             "contact_country"
 [4] "contact_department"      "contact_email"            "contact_institute"
 [7] "contact_laboratory"      "contact_name"             "contact_phone"
[10] "contact_state"           "contact_zip/postal_code"  "contributor"
[13] "email"                   "geo_accession"            "institute"
[16] "last_update_date"        "name"                     "overall_design"
[19] "platform_id"             "platform_taxid"           "pubmed_id"
[22] "relation"                "sample_id"                "sample_taxid"
[25] "status"                  "submission_date"          "summary"
[28] "supplementary_file"      "title"                    "type"
[31] "web_link"
```

```
#creating GSM object names
gsm <- GSMList(gse)[[1]]
```

```
#find information we need from the metadata of gsm
names(Meta(gsm))
```

```
[1] "channel_count"            "characteristics_ch1"      "contact_address"
 [4] "contact_city"            "contact_country"          "contact_department"
 [7] "contact_email"           "contact_institute"        "contact_laboratory"
[10] "contact_name"            "contact_phone"            "contact_state"
[13] "contact_zip/postal_code" "data_processing"          "data_row_count"
[16] "extract_protocol_ch1"    "geo_accession"            "hyb_protocol"
[19] "label_ch1"               "label_protocol_ch1"       "last_update_date"
[22] "molecule_ch1"            "organism_ch1"             "platform_id"
```

```
[25] "scan_protocol"                "series_id"                    "source_name_ch1"
[28] "status"                       "submission_date"              "supplementary_file"
[31] "taxid_ch1"                     "title"                        "type"
```

#extract metadata of gsm not in gse

```
names(Meta(gsm))[!(names(Meta(gsm)) %in% names(Meta(gse)))]
```

```
[1] "channel_count"          "characteristics_ch1"  "data_processing"
 [4] "data_row_count"        "extract_protocol_ch1"       "hyb_protocol"
 [7] "label_ch1"             "label_protocol_ch1"         "molecule_ch1"
[10] "organism_ch1"          "scan_protocol"              "series_id"
[13] "source_name_ch1"       "taxid_ch1"
```

```
Meta(gsm)[!(names(Meta(gsm)) %in% names(Meta(gse)))]
```

```
$channel_count
[1] "1"

$characteristics_ch1
[1] "cell line: SUM159"                "treatment: control"
[3] "tissue: Claudin-low breast cancer"

$data_processing
[1] "MAS 5.0"

$data_row_count
[1] "54675"

$extract_protocol_ch1
[1] "standard Affymetrix protocol"

$hyb_protocol
[1] "standard Affymetrix protocol"

$label_ch1
[1] "biotin"

$label_protocol_ch1
[1] "standard Affymetrix protocol"

$molecule_ch1
[1] "total RNA"
```

```
$organism_ch1
[1] "Homo sapiens"

$scan_protocol
[1] "standard Affymetrix protocol"

$series_id
[1] "GSE50697"

$source_name_ch1
[1] "SUM159"

$taxid_ch1
[1] "9606"
```

```
#From the metadata information, we decided to use the following elements
#source_name_ch1 and characteristics_ch1
#getting sample growth conditions into a data frame
culture_medium <- function(gsm) {
  Meta(gsm)[['characteristics_ch1']][2]
}
sapply(GSMList(gse),culture_medium)
```

```
                GSM1226581                 GSM1226582
     "treatment: control"       "treatment: control"
                GSM1226583                 GSM1226584
     "treatment: control"       "treatment: pBabe puro miR-203"
                GSM1226585                 GSM1226586
"treatment: pBabe puro miR-203"  "treatment: pBabe puro miR-203"
```

```
pd <- data.frame(culture=as.factor(sapply(GSMList(gse),culture_medium)))
pd
```

```
                                culture
GSM1226581              treatment: control
GSM1226582              treatment: control
GSM1226583              treatment: control
GSM1226584 treatment: pBabe puro miR-203
GSM1226585 treatment: pBabe puro miR-203
GSM1226586 treatment: pBabe puro miR-203
```

```
#simplifying our dataframe, convert columns to 2 values only
pd$culture <- as.factor(pd$culture)
levels(pd$culture) <- c("control","miR203")
```

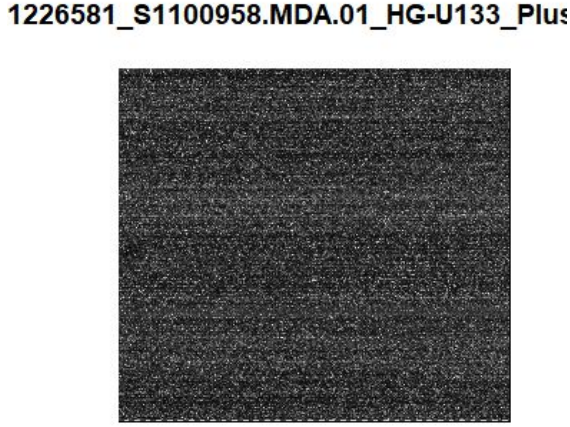
```
#to enable the kable function, we need to install the knitr package
install.packages("knitr")
library(knitr)
kable(pd)
```

	culture
GSM1226581	control
GSM1226582	control
GSM1226583	control
GSM1226584	miR203
GSM1226585	miR203
GSM1226586	miR203

```
#Reading in the CEL files with the phenoData
celfiles <- paste0('GSE50697_RAW/', list.celfiles('GSE50697_RAW/'),'.')
affydata <- read.affybatch(celfiles,phenoData = new("AnnotatedDataFrame",pd))
phenoData(affydata)
```

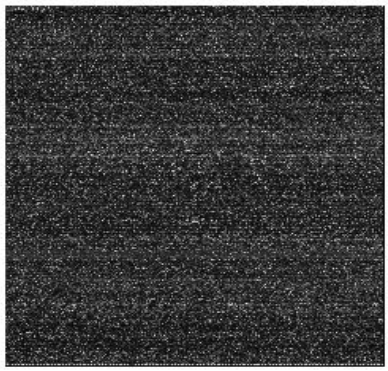
```
An object of class 'AnnotatedDataFrame'
  sampleNames: GSM1226581 GSM1226582 ... GSM1226586 (6 total)
  varLabels: culture
  varMetadata: labelDescription
```

```
#pseudo images of chips
image(data[,1])
```



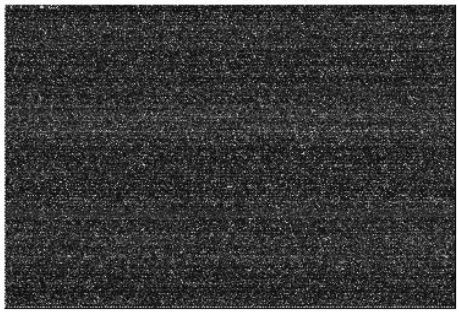
```
image(data[,2])
```

```
226582_S1100958.MDA.02_HG-U133_Pl
```



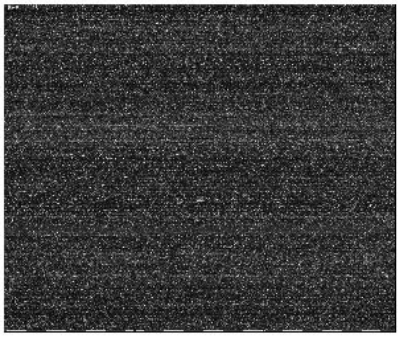
```
image(data[,3])
```

```
GSM1226583_S1100958.MDA.03_HG-U133_Plus_2_.
```



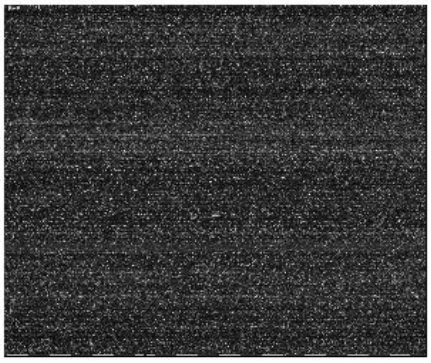
```
image(data[,4])
```

```
M1226584_S1100958.MDA.04_HG-U133_Plus_
```



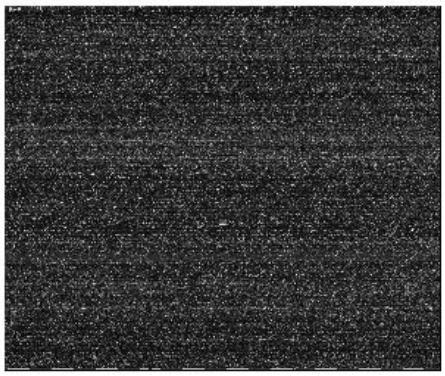
```
image(data[,5])
```

M1226585_S1100958.MDA.05_HG-U133_Plus_

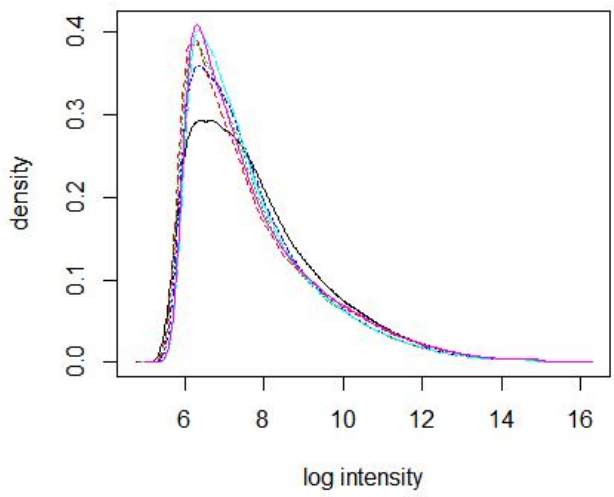


```
image(data[,6])
```

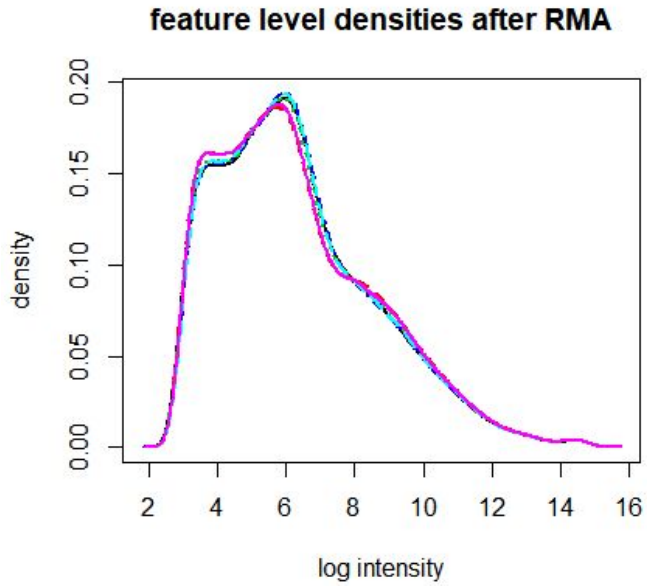
M1226586_S1100958.MDA.06_HG-U133_Plus_



```
#CEL file densities before normalisation or background correction
plotDensity.AffyBatch(affydata)
```




```
#perform RMA normalisation
eset <- rma(affydata)
plotDensity(exprs(eset),xlab='log intensity',main="feature level densities after
RMA",lwd=2)
```




```
#phenotype data for each samples after rma is retained
pData(eset)
```

```
                culture
GSM1226581 control
GSM1226582 control
```

```
GSM1226583 control
GSM1226584  miR203
GSM1226585  miR203
GSM1226586  miR203
```

```
#generate model matrix
model <- model.matrix( ~ 0 + eset$culture)
#rename the model columns to correspond to the different growth conditions
colnames(model) <- levels(eset$culture)
model
```

```
     control miR203
1       1     0
2       1     0
3       1     0
4       0     1
5       0     1
6       0     1
attr(,"assign")
[1] 1 1
attr(,"contrasts")
attr(,"contrasts")$`eset$culture`
[1] "contr.treatment"
```

```
#look at when the growth conditions differ, we create contrast
contrasts <- makeContrasts(control - miR203, levels=model)
contrasts
```

```
          Contrasts
Levels    control - miR203
control                  1
miR203                  -1
```

```
#fit model and contrast matrix into a data
fit <- lmFit(eset, model)
```

fit

An object of class "MArrayLM"
$coefficients

	control	miR203
1007_s_at	9.245338	9.408269
1053_at	9.586208	9.605038
117_at	6.404494	6.500503
121_at	8.496605	8.577745
1255_g_at	3.425654	3.360337

54670 more rows ...

$rank
[1] 2

$assign
[1] 1 1

$qr
$qr

	control	miR203
1	-1.7320508	0.0000000
2	0.5773503	-1.7320508
3	0.5773503	0.0000000
4	0.0000000	0.5773503
5	0.0000000	0.5773503
6	0.0000000	0.5773503

attr(,"assign")
[1] 1 1
attr(,"contrasts")
attr(,"contrasts")$`eset$culture`
[1] "contr.treatment"

$qraux
[1] 1.57735 1.00000

$pivot
[1] 1 2

$tol
[1] 1e-07

$rank
[1] 2

$df.residual
[1] 4 4 4 4 4
54670 more elements ...

$sigma
 1007_s_at   1053_at   117_at   121_at 1255_g_at
0.07378465 0.13302616 0.18946888 0.14112115 0.06639880
54670 more elements ...

$cov.coefficients
          control   miR203
control 0.3333333   0.0000000
miR203 0.0000000   0.3333333

$stdev.unscaled
                  control   miR203
1007_s_at   0.5773503   0.5773503
1053_at   0.5773503   0.5773503
117_at   0.5773503   0.5773503
121_at   0.5773503   0.5773503
1255_g_at   0.5773503   0.5773503
54670 more rows ...

$pivot
[1] 1 2

$Amean
1007_s_at   1053_at   117_at   121_at 1255_g_at
 9.326804 9.595623 6.452499 8.537175 3.392996
54670 more elements ...

$method
[1] "ls"

$design
  control miR203
1   1   0
2   1   0
3   1   0

```
4       0      1
5       0      1
6       0      1
attr(,"assign")
[1] 1 1
attr(,"contrasts")
attr(,"contrasts")$`eset$culture`
[1] "contr.treatment"
```

fitted.contrast <- contrasts.fit(fit,contrasts)
fitted.contrast

```
An object of class "MArrayLM"
$coefficients
          Contrasts
                     control - miR203
  1007_s_at      -0.16293107
  1053_at        -0.01883039
  117_at         -0.09600894
  121_at         -0.08113934
  1255_g_at       0.06531688
54670 more rows ...

$rank
[1] 2

$assign
[1] 1 1

$qr
$qr
     control         miR203
1 -1.7320508   0.0000000
2  0.5773503 -1.7320508
3  0.5773503  0.0000000
4  0.0000000  0.5773503
5  0.0000000  0.5773503
6  0.0000000  0.5773503
attr(,"assign")
[1] 1 1
attr(,"contrasts")
attr(,"contrasts")$`eset$culture`
[1] "contr.treatment"
```

```
$qraux
[1] 1.57735 1.00000

$pivot
[1] 1 2

$tol
[1] 1e-07

$rank
[1] 2


$df.residual
[1] 4 4 4 4 4
54670 more elements ...

$sigma
 1007_s_at    1053_at     117_at     121_at  1255_g_at
0.07378465 0.13302616 0.18946888 0.14112115 0.06639880
54670 more elements ...

$cov.coefficients
                Contrasts
Contrasts          control - miR203
  control - miR203        0.6666667

$stdev.unscaled
           Contrasts
                      control - miR203
  1007_s_at        0.8164966
  1053_at          0.8164966
  117_at           0.8164966
  121_at           0.8164966
  1255_g_at        0.8164966
54670 more rows ...

$Amean
1007_s_at   1053_at    117_at    121_at 1255_g_at
 9.326804  9.595623  6.452499  8.537175  3.392996
54670 more elements ...
```

```
$method
[1] "ls"

$design
     control   miR203
1       1           0
2       1           0
3       1           0
4       0           1
5       0           1
6       0           1
attr(,"assign")
[1] 1 1
attr(,"contrasts")
attr(,"contrasts")$`eset$culture`
[1] "contr.treatment"


$contrasts
         Contrasts
Levels    control - miR203
 control                1
 miR203                -1
```

```
#calculate test statistics via performing eBayes correction
fitted.ebayes <- eBayes(fitted.contrast)
fitted.ebayes
```

```
An object of class "MArrayLM"
$coefficients
           Contrasts
                     control - miR203
  1007_s_at     -0.16293107
  1053_at       -0.01883039
  117_at        -0.09600894
  121_at        -0.08113934
  1255_g_at      0.06531688
54670 more rows ...

$rank
[1] 2
```

$assign
[1] 1 1

$qr
$qr
    control   miR203
1 -1.7320508 0.0000000
2 0.5773503 -1.7320508
3 0.5773503 0.0000000
4 0.0000000 0.5773503
5 0.0000000 0.5773503
6 0.0000000 0.5773503
attr(,"assign")
[1] 1 1
attr(,"contrasts")
attr(,"contrasts")$`eset$culture`
[1] "contr.treatment"

$qraux
[1] 1.57735 1.00000

$pivot
[1] 1 2

$tol
[1] 1e-07

$rank
[1] 2

$df.residual
[1] 4 4 4 4 4
54670 more elements ...

$sigma
 1007_s_at   1053_at   117_at   121_at 1255_g_at
0.07378465 0.13302616 0.18946888 0.14112115 0.06639880
54670 more elements ...

$cov.coefficients

```
                Contrasts
Contrasts          control - miR203
control - miR203        0.6666667

$stdev.unscaled
          Contrasts
                     control - miR203
  1007_s_at        0.8164966
  1053_at          0.8164966
  117_at           0.8164966
  121_at           0.8164966
  1255_g_at        0.8164966
54670 more rows ...

$Amean
1007_s_at   1053_at    117_at    121_at 1255_g_at
 9.326804  9.595623  6.452499  8.537175  3.392996
54670 more elements ...

$method
[1] "ls"

$design
  control miR203
1       1      0
2       1      0
3       1      0
4       0      1
5       0      1
6       0      1
attr(,"assign")
[1] 1 1
attr(,"contrasts")
attr(,"contrasts")$`eset$culture`
[1] "contr.treatment"


$contrasts
        Contrasts
Levels    control - miR203
control              1
miR203              -1
```

$df.prior
[1] 5.163163

$s2.prior
[1] 0.01315895

$var.prior
[1] 51.55685

$proportion
[1] 0.01

$s2.post
  1007_s_at   1053_at   117_at   121_at   1255_g_at
0.009791214 0.015139491 0.023085437 0.016108249 0.009339242
54670 more elements ...

$t
          Contrasts
                    control - miR203
  1007_s_at   -2.0166534
  1053_at   -0.1874344
  117_at   -0.7739061
  121_at   -0.7829841
  1255_g_at   0.8277806
54670 more rows ...

$df.total
[1] 9.163163 9.163163 9.163163 9.163163 9.163163
54670 more elements ...

$p.value
          Contrasts
                    control - miR203
  1007_s_at   0.07396289
  1053_at   0.85540486
  117_at   0.45848725
  121_at   0.45339651
  1255_g_at   0.42882619
54670 more rows ...

$lods
          Contrasts

```
                control - miR203
1007_s_at       -4.937866
1053_at         -6.756422
117_at          -6.458114
121_at          -6.450863
1255_g_at       -6.414005
54670 more rows ...

$F
[1] 4.06689075 0.03513165 0.59893067 0.61306414 0.68522080
54670 more elements ...

$F.p.value
[1] 0.07396289 0.85540486 0.45848725 0.45339651 0.42882619
54670 more elements ...
```

#extracting differentially expressed genes

topTable(fitted.ebayes)

```
                logFC      AveExpr         t          P.Value
209719_x_at -2.096174  9.668246 -27.98136 3.449900e-10
210413_x_at -1.984099  9.649669 -25.77291 7.265063e-10
201721_s_at  2.254152  9.710288  25.72106 7.398668e-10
211906_s_at -2.143763  9.071134 -24.70101 1.066850e-09
206172_at   -1.789371  7.706033 -24.07788 1.343931e-09
209720_s_at -2.432950  8.729486 -23.12911 1.931967e-09
211756_at   -3.067195  4.671276 -21.91533 3.140852e-09
202949_s_at -1.502988 10.092484 -20.84188 4.935440e-09
202007_at    3.494914  5.646599  20.62198 5.429179e-09
206002_at    1.704752  8.907492  20.45564 5.839018e-09
                adj.P.Val        B
209719_x_at 1.348407e-05 12.13796
210413_x_at 1.348407e-05 11.73165
201721_s_at 1.348407e-05 11.72129
211906_s_at 1.458251e-05 11.50898
206172_at   1.469588e-05 11.37091
209720_s_at 1.760505e-05 11.14745
211756_at   2.453230e-05 10.83610
202949_s_at 3.192483e-05 10.53422
202007_at   3.192483e-05 10.46905
206002_at   3.192483e-05 10.41895
```

#Get a limited number of probesets

```
ps <- rownames(topTable(fitted.ebayes))
ps
```

```
[1] "209719_x_at" "210413_x_at" "201721_s_at" "211906_s_at"
 [5] "206172_at"   "209720_s_at" "211756_at"   "202949_s_at"
 [9] "202007_at"   "206002_at"
```

```
#The AnnotationDbi interface
#look at available columns for our chip
columns(hgu133plus2.db)
```

```
 [1] "ACCNUM"       "ALIAS"        "ENSEMBL"      "ENSEMBLPROT"
 [5] "ENSEMBLTRANS" "ENTREZID"     "ENZYME"       "EVIDENCE"
 [9] "EVIDENCEALL"  "GENENAME"     "GO"           "GOALL"
[13] "IPI"          "MAP"          "OMIM"         "ONTOLOGY"
[17] "ONTOLOGYALL"  "PATH"         "PFAM"         "PMID"
[21] "PROBEID"      "PROSITE"      "REFSEQ"       "SYMBOL"
[25] "UCSCKG"       "UNIGENE"      "UNIPROT"
```

```
#Look at which that can be used as keys
keytypes(hgu133plus2.db)
```

```
[1] "ACCNUM"       "ALIAS"        "ENSEMBL"      "ENSEMBLPROT"
 [5] "ENSEMBLTRANS" "ENTREZID"     "ENZYME"       "EVIDENCE"
 [9] "EVIDENCEALL"  "GENENAME"     "GO"           "GOALL"
[13] "IPI"          "MAP"          "OMIM"         "ONTOLOGY"
[17] "ONTOLOGYALL"  "PATH"         "PFAM"         "PMID"
[21] "PROBEID"      "PROSITE"      "REFSEQ"       "SYMBOL"
[25] "UCSCKG"       "UNIGENE"      "UNIPROT"
```

```
#realise all can be used as keys, choose "PROBEID" as most suitable
head(keys(hgu133plus2.db,keytype="PROBEID"))
```

```
[1] "1007_s_at" "1053_at"   "117_at"    "121_at"    "1255_g_at"
[6] "1294_at"
```

```
AnnotationDbi::select(hgu133plus2.db,ps,c("SYMBOL","ENTREZID","GENENAME"),keyty
pe="PROBEID")
```

```
  PROBEID        SYMBOL    ENTREZID
1  209719_x_at    SERPINB3  6317
2  210413_x_at    SERPINB4  6318
3  210413_x_at    SERPINB3  6317
```

```
4  201721_s_at        LAPTM5       7805
5  211906_s_at        SERPINB4     6318
6    206172_at        IL13RA2      3598
7  209720_s_at        SERPINB3     6317
8    211756_at        PTHLH        5744
9  202949_s_at        FHL2         2274
10   202007_at        NID1         4811
11   206002_at        ADGRG2       10149
                          GENENAME
1               serpin family B member 3
2               serpin family B member 4
3               serpin family B member 3
4       lysosomal protein transmembrane 5
5               serpin family B member 4
6  interleukin 13 receptor subunit alpha 2
7               serpin family B member 3
8        parathyroid hormone like hormone
9          four and a half LIM domains 2
10                        nidogen 1
11  adhesion G protein-coupled receptor G2
```

#Restrict to upregulated genes

```
differentially_expressed <- topTable(fitted.ebayes,number=Inf,p.value=0.1,lfc=1)
differentially_expressed
```

```
                logFC      AveExpr         t          P.Value
209719_x_at -2.096174  9.668246 -27.981362 3.449900e-10
210413_x_at -1.984099  9.649669 -25.772906 7.265063e-10
201721_s_at  2.254152  9.710288  25.721059 7.398668e-10
211906_s_at -2.143763  9.071134 -24.701007 1.066850e-09
206172_at   -1.789371  7.706033 -24.077878 1.343931e-09
209720_s_at -2.432950  8.729486 -23.129114 1.931967e-09
211756_at   -3.067195  4.671276 -21.915333 3.140852e-09
202949_s_at -1.502988 10.092484 -20.841881 4.935440e-09
202007_at    3.494914  5.646599  20.621979 5.429179e-09
206002_at    1.704752  8.907492  20.455638 5.839018e-09
228946_at    2.850568  7.433498  19.353438 9.598769e-09
203296_s_at  1.436208  7.128925  18.624289 1.353915e-08
200953_s_at  2.114548  6.413324  18.518165 1.424938e-08
201681_s_at  1.343568 10.848433  18.048727 1.792645e-08
200618_at    1.396831 10.601469  17.304137 2.610933e-08
214051_at    1.266167  6.486634  16.511752 3.963213e-08
```

205207_at	-1.972971	10.543937	-16.302947	4.437863e-08
215785_s_at	1.310530	6.548376	16.145121	4.838383e-08
218499_at	2.660353	6.810620	16.064949	5.057022e-08
225295_at	1.134359	9.370925	15.777075	5.936971e-08
210445_at	1.244751	7.713232	15.766920	5.970961e-08
225237_s_at	1.424176	7.916869	15.762484	5.985878e-08
204602_at	-1.266821	11.377001	-15.706611	6.177312e-08
204019_s_at	1.947556	8.145434	15.701004	6.196895e-08
201739_at	-1.138479	9.582154	-15.278721	7.889207e-08
222288_at	1.827582	7.897300	14.987890	9.350461e-08
217388_s_at	-1.340457	9.048417	-14.945053	9.589930e-08
1556773_at	-1.928811	5.156095	-14.816192	1.035213e-07
227609_at	-1.245758	7.672937	-14.484025	1.264433e-07
215189_at	1.468563	10.044900	14.282840	1.430238e-07
222857_s_at	-1.132656	8.121989	-14.208997	1.497009e-07
213711_at	1.159206	9.053300	14.131621	1.570690e-07
241864_x_at	1.489112	7.380173	14.128284	1.573956e-07
209946_at	-1.781273	6.908853	-13.616536	2.176355e-07
226134_s_at	1.136151	9.222280	13.608353	2.187858e-07
219049_at	-1.948935	5.572982	-13.575335	2.234962e-07
212230_at	1.313188	8.909360	13.492343	2.358382e-07
202242_at	1.407202	5.471476	13.412674	2.483975e-07
206343_s_at	-1.609212	8.373871	-13.352914	2.583030e-07
209739_s_at	1.121705	8.529516	13.309525	2.657682e-07
218162_at	1.045352	6.832097	13.283405	2.703769e-07
207802_at	1.667464	5.074460	13.166901	2.920360e-07
1557883_a_at	1.806322	4.515782	12.969677	3.332030e-07
224901_at	1.490742	6.113577	12.888866	3.518878e-07
201720_s_at	1.174496	8.359157	12.888647	3.519399e-07
212543_at	-1.530326	6.498998	-12.838209	3.641887e-07
227963_at	1.444880	6.061685	12.834804	3.650323e-07
226571_s_at	1.196636	8.630486	12.813410	3.703829e-07
213436_at	1.271657	9.635054	12.756578	3.850216e-07
1553995_a_at	-1.253015	7.654807	-12.658433	4.118291e-07
203910_at	-1.315174	9.064571	-12.634820	4.185819e-07
227156_at	1.328361	5.991962	12.582181	4.340796e-07
228863_at	1.472731	6.357374	12.506484	4.574876e-07
223395_at	1.068056	10.465127	12.441449	4.787137e-07
206068_s_at	1.424994	5.780641	12.400733	4.925544e-07
208891_at	-1.385639	8.398753	-12.318390	5.219150e-07
212942_s_at	-1.516025	6.738569	-12.259489	5.441012e-07
1553602_at	1.363603	6.596712	12.236775	5.529329e-07
205534_at	-1.336143	8.974132	-12.149474	5.883802e-07

205767_at	-1.716322	8.767721	-12.118055	6.017450e-07
228742_at	1.041995	7.583651	12.046972	6.332343e-07
223557_s_at	-1.363652	7.888034	-12.009754	6.504438e-07
235274_at	1.005532	7.657250	11.748894	7.865815e-07
216870_x_at	-1.296386	6.322829	-11.714764	8.066076e-07
211964_at	-1.063786	6.932677	-11.585661	8.876065e-07
225240_s_at	1.297440	8.599465	11.531470	9.242444e-07
243309_at	1.127696	5.892718	11.521002	9.315123e-07
213947_s_at	1.096488	7.688141	11.481711	9.593578e-07
219874_at	-1.032381	7.306612	-11.477603	9.623221e-07
243880_at	1.122249	5.598890	11.461119	9.743176e-07
227020_at	-1.008496	8.595696	-11.460947	9.744438e-07
224856_at	1.191607	9.560194	11.401104	1.019403e-06
207623_at	1.160545	7.651684	11.384154	1.032549e-06
205404_at	-1.277894	6.526299	-11.351684	1.058258e-06
202712_s_at	1.239059	7.502293	11.345660	1.063105e-06
205110_s_at	-1.128768	9.124357	-11.343864	1.064554e-06
226463_at	1.665975	7.020946	11.258157	1.136338e-06
1560225_at	1.186996	6.322784	11.230714	1.160437e-06
217996_at	-1.340874	9.870791	-11.174406	1.211669e-06
217997_at	-1.471840	8.227516	-11.142785	1.241528e-06
221731_x_at	-1.255271	7.842270	-11.133802	1.250158e-06
214803_at	1.349396	6.911944	11.108018	1.275297e-06
213201_s_at	1.095576	7.502540	11.021185	1.364143e-06
210663_s_at	-1.162039	8.866615	-11.019233	1.366217e-06
226921_at	1.040167	8.608740	11.009422	1.376696e-06
224691_at	1.015055	10.545175	11.003052	1.383545e-06
222920_s_at	1.051959	6.425689	10.994649	1.392639e-06
208892_s_at	-1.349587	7.710335	-10.978803	1.409967e-06
227061_at	1.270223	7.406550	10.851655	1.557932e-06
201150_s_at	1.744768	7.016096	10.795555	1.628587e-06
203836_s_at	-1.173029	6.334483	-10.701144	1.755598e-06
226757_at	-1.156531	7.348931	-10.666205	1.805336e-06
205476_at	-1.026426	9.118523	-10.493921	2.074396e-06
209355_s_at	1.283780	7.946635	10.412755	2.216216e-06
203837_at	-1.664224	6.336915	-10.408882	2.223246e-06
204194_at	-1.476294	7.693366	-10.378921	2.278458e-06
213316_at	1.024644	5.038984	10.317175	2.397069e-06
226189_at	-1.494698	6.607014	-10.265638	2.501282e-06
215629_s_at	-1.390720	5.298695	-10.262250	2.508306e-06
210503_at	1.028117	7.913849	10.241691	2.551394e-06
227486_at	-1.368408	7.642176	-10.239413	2.556218e-06
214467_at	-1.382421	4.825823	-10.213686	2.611405e-06

235199_at	-1.550341	6.769346	-10.167323	2.714197e-06
209687_at	-1.177895	6.825387	-10.165181	2.719053e-06
205003_at	-1.531919	6.640436	-10.141147	2.774194e-06
227289_at	1.331888	5.596073	10.084716	2.908564e-06
204620_s_at	-1.035997	7.368166	-10.031903	3.040828e-06
226370_at	1.157409	7.227893	10.010942	3.095146e-06
210118_s_at	-1.458621	7.811831	-9.897507	3.408250e-06
204385_at	-1.010836	8.197030	-9.896692	3.410621e-06
204533_at	-1.103127	6.696345	-9.846135	3.561377e-06
226302_at	-1.190951	9.588665	-9.774950	3.786216e-06
213391_at	1.199648	7.855011	9.693048	4.064410e-06
206290_s_at	-1.210959	6.322545	-9.671217	4.142303e-06
204114_at	-1.255358	4.847498	-9.647616	4.228366e-06
211506_s_at	-1.199992	10.684143	-9.621588	4.325565e-06
228033_at	-1.075819	8.476069	-9.596823	4.420332e-06
223869_at	-1.461387	9.000998	-9.519697	4.730392e-06
224938_at	1.003431	8.507051	9.506805	4.784520e-06
238937_at	1.512821	5.334978	9.481948	4.890806e-06
219287_at	-1.150889	7.041430	-9.444304	5.056746e-06
226218_at	-1.291518	7.704088	-9.321460	5.642803e-06
239866_at	1.021869	7.209041	9.298515	5.760337e-06
240391_at	1.095573	5.639588	9.267364	5.924232e-06
227875_at	1.088525	5.477780	9.240871	6.067649e-06
212226_s_at	1.141638	8.813665	9.174575	6.443553e-06
1557882_at	1.308155	4.398610	8.996530	7.585896e-06
1553994_at	-1.195990	8.464710	-8.929526	8.071966e-06
222446_s_at	1.109060	6.870658	8.787498	9.219349e-06
238953_at	1.102176	7.213338	8.767421	9.395512e-06
227480_at	-1.008557	6.454705	-8.737691	9.663203e-06
211122_s_at	-1.437891	5.481624	-8.701734	9.998195e-06
204597_x_at	-1.114211	6.475372	-8.669453	1.030983e-05
1552546_a_at	-1.056041	7.462691	-8.640677	1.059662e-05
226492_at	1.047015	6.477821	8.578079	1.125120e-05
208893_s_at	-1.003504	7.551996	-8.572491	1.131176e-05
1553684_at	1.159679	6.705608	8.363742	1.385147e-05
204619_s_at	-1.318002	6.670587	-8.300155	1.474473e-05
238462_at	-1.140790	6.530780	-8.119298	1.764965e-05
219961_s_at	1.026839	9.636927	8.032967	1.925292e-05
231568_at	-1.111633	5.995260	-7.998804	1.993085e-05
1557094_at	-1.013669	8.179558	-7.922769	2.153554e-05
210941_at	-1.137155	4.912168	-7.864326	2.286515e-05
228450_at	-1.335993	5.632521	-7.829498	2.370007e-05
209839_at	-1.148218	6.531914	-7.736205	2.610565e-05

207173_x_at	-1.212647	6.236454	-7.597516	3.019017e-05
211959_at	1.364346	6.466787	7.528684	3.247290e-05
222662_at	-1.274162	6.655163	-7.401749	3.719376e-05
1569157_s_at	1.318162	5.058315	7.357447	3.901403e-05
224367_at	1.063579	6.028615	7.175741	4.756705e-05
205830_at	1.179054	5.879032	7.154081	4.871626e-05
212706_at	1.105566	4.929178	7.120332	5.056777e-05
205535_s_at	-1.191363	7.219878	-6.967438	5.997430e-05
210355_at	-1.479006	4.670382	-6.962900	6.028118e-05
205532_s_at	1.274534	6.604113	6.910359	6.396223e-05
1558501_at	-1.198424	5.294933	-6.885537	6.578596e-05
201147_s_at	1.530476	6.170983	6.812613	7.148008e-05
1555673_at	-1.051042	6.785199	-6.792971	7.310412e-05
235561_at	1.079808	5.430246	6.769854	7.506718e-05
202920_at	1.527176	5.747694	6.668522	8.437675e-05
243009_at	1.039904	4.651447	6.630249	8.821433e-05
220030_at	-1.294684	4.335316	-6.250504	1.384963e-04
206300_s_at	-1.374347	4.566611	-6.212739	1.449942e-04
205533_s_at	1.035088	5.136024	6.189540	1.491497e-04
220301_at	1.291894	5.727456	6.061185	1.746081e-04
1554131_at	-1.084905	4.428248	-5.667805	2.868936e-04

	adj.P.Val	B
209719_x_at	1.348407e-05	12.1379607
210413_x_at	1.348407e-05	11.7316498
201721_s_at	1.348407e-05	11.7212891
211906_s_at	1.458251e-05	11.5089762
206172_at	1.469588e-05	11.3709081
209720_s_at	1.760505e-05	11.1474545
211756_at	2.453230e-05	10.8360961
202949_s_at	3.192483e-05	10.5342229
202007_at	3.192483e-05	10.4690463
206002_at	3.192483e-05	10.4189530
228946_at	4.771024e-05	10.0688951
203296_s_at	5.992960e-05	9.8187527
200953_s_at	5.992960e-05	9.7810275
201681_s_at	7.000917e-05	9.6099414
200618_at	9.516851e-05	9.3238050
214051_at	1.354304e-04	8.9978822
205207_at	1.411730e-04	8.9080757
215785_s_at	1.411730e-04	8.8390616
218499_at	1.411730e-04	8.8036239
225295_at	1.411730e-04	8.6742241
210445_at	1.411730e-04	8.6695970

225237_s_at 1.411730e-04 8.6675742
204602_at 1.411730e-04 8.6420285
204019_s_at 1.411730e-04 8.6394576
201739_at 1.725370e-04 8.4419789
222288_at 1.941961e-04 8.3014157
217388_s_at 1.941961e-04 8.2803884
1556773_at 2.021437e-04 8.2166245
227609_at 2.383893e-04 8.0486591
215189_at 2.606610e-04 7.9443366
222857_s_at 2.607759e-04 7.9055435
213711_at 2.607759e-04 7.8646006
241864_x_at 2.607759e-04 7.8628284
209946_at 3.394349e-04 7.5841819
226134_s_at 3.394349e-04 7.5796142
219049_at 3.394349e-04 7.5611453
212230_at 3.437873e-04 7.5144622
202242_at 3.437873e-04 7.4692929
206343_s_at 3.437873e-04 7.4351813
209739_s_at 3.437873e-04 7.4102897
218162_at 3.437873e-04 7.3952541
207802_at 3.628879e-04 7.3277217
1557883_a_at 3.960407e-04 7.2116284
224901_at 3.970722e-04 7.1634061
201720_s_at 3.970722e-04 7.1632751
212543_at 3.970722e-04 7.1329812
227963_at 3.970722e-04 7.1309309
226571_s_at 3.970722e-04 7.1180314
213436_at 4.048280e-04 7.0836317
1553995_a_at 4.086780e-04 7.0237673
203910_at 4.086780e-04 7.0092773
227156_at 4.163738e-04 6.9768532
228863_at 4.312610e-04 6.9299267
223395_at 4.436215e-04 6.8893262
206068_s_at 4.488402e-04 6.8637730
208891_at 4.602533e-04 6.8117761
212942_s_at 4.722021e-04 6.7743172
1553602_at 4.723689e-04 6.7598130
205534_at 4.874195e-04 6.7037560
205767_at 4.910509e-04 6.6834605
228742_at 5.091483e-04 6.6373060
223557_s_at 5.154060e-04 6.6130070
235274_at 6.057231e-04 6.4401009
216870_x_at 6.125176e-04 6.4171372

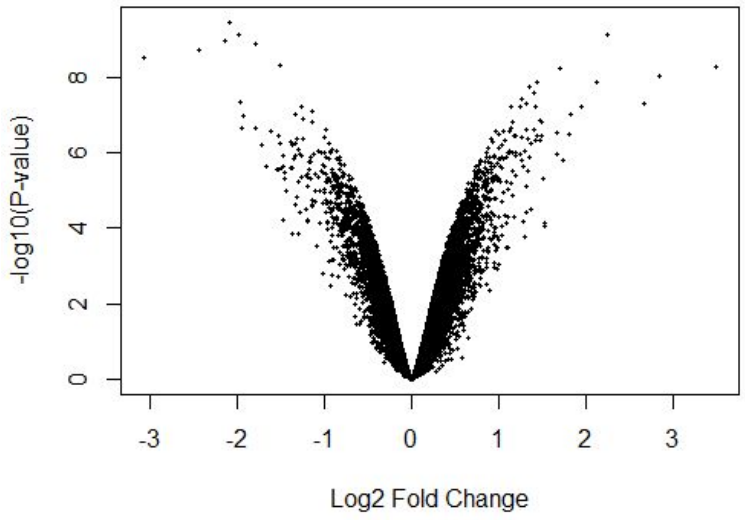
211964_at 6.454816e-04 6.3295445
225240_s_at 6.454816e-04 6.2924306
243309_at 6.454816e-04 6.2852375
213947_s_at 6.454816e-04 6.2581704
219874_at 6.454816e-04 6.2553338
243880_at 6.454816e-04 6.2439405
227020_at 6.454816e-04 6.2438213
224856_at 6.539833e-04 6.2022963
207623_at 6.539833e-04 6.1904884
205404_at 6.539833e-04 6.1678105
202712_s_at 6.539833e-04 6.1635948
205110_s_at 6.539833e-04 6.1623374
226463_at 6.903252e-04 6.1020620
1560225_at 6.972189e-04 6.0826487
217996_at 7.123440e-04 6.0426429
217997_at 7.183258e-04 6.0200746
221731_x_at 7.183258e-04 6.0136493
214803_at 7.183258e-04 5.9951743
213201_s_at 7.183258e-04 5.9325910
210663_s_at 7.183258e-04 5.9311777
226921_at 7.183258e-04 5.9240690
224691_at 7.183258e-04 5.9194505
222920_s_at 7.183258e-04 5.9133522
208892_s_at 7.204667e-04 5.9018388
227061_at 7.673868e-04 5.8087576
201150_s_at 7.879911e-04 5.7672933
203836_s_at 8.134517e-04 5.6969586
226757_at 8.294682e-04 5.6707515
205476_at 9.146580e-04 5.5401028
209355_s_at 9.496562e-04 5.4777211
203837_at 9.496562e-04 5.4747305
204194_at 9.582670e-04 5.4515579
213316_at 9.832287e-04 5.4035676
226189_at 1.001034e-03 5.3632694
215629_s_at 1.001034e-03 5.3606124
210503_at 1.005477e-03 5.3444690
227486_at 1.005477e-03 5.3426781
214467_at 1.018248e-03 5.3224214
235199_at 1.018248e-03 5.2857763
209687_at 1.018248e-03 5.2840786
205003_at 1.031830e-03 5.2650058
227289_at 1.067287e-03 5.2200302
204620_s_at 1.101041e-03 5.1776910

226370_at	1.106059e-03	5.1608199
210118_s_at	1.153201e-03	5.0688550
204385_at	1.153201e-03	5.0681904
204533_at	1.173002e-03	5.0268347
226302_at	1.217714e-03	4.9682196
213391_at	1.267618e-03	4.9002187
206290_s_at	1.267618e-03	4.8819910
204114_at	1.276420e-03	4.8622360
211506_s_at	1.288838e-03	4.8403908
228033_at	1.299364e-03	4.8195480
223869_at	1.341506e-03	4.7542758
224938_at	1.341506e-03	4.7433107
238937_at	1.357385e-03	4.7221270
219287_at	1.358013e-03	4.6899355
226218_at	1.462181e-03	4.5839554
239866_at	1.464867e-03	4.5640009
240391_at	1.485814e-03	4.5368289
227875_at	1.507949e-03	4.5136479
212226_s_at	1.570247e-03	4.4553394
1557882_at	1.750037e-03	4.2966147
1553994_at	1.779576e-03	4.2360641
222446_s_at	1.902143e-03	4.1062131
238953_at	1.916790e-03	4.0876909
227480_at	1.964073e-03	4.0601861
211122_s_at	1.986451e-03	4.0267989
204597_x_at	2.020393e-03	3.9967093
1552546_a_at	2.034249e-03	3.9697950
226492_at	2.106711e-03	3.9109463
208893_s_at	2.110821e-03	3.9056727
1553684_at	2.466869e-03	3.7062751
204619_s_at	2.548792e-03	3.6445963
238462_at	2.863485e-03	3.4667162
219961_s_at	3.051170e-03	3.3805065
231568_at	3.096466e-03	3.3461575
1557094_at	3.234769e-03	3.2692234
210941_at	3.292331e-03	3.2096335
228450_at	3.365718e-03	3.1739324
209839_at	3.586247e-03	3.0775943
207173_x_at	3.826645e-03	2.9324569
211959_at	3.998774e-03	2.8595595
222662_at	4.290230e-03	2.7236000
1569157_s_at	4.404616e-03	2.6756776
224367_at	4.907541e-03	2.4765306

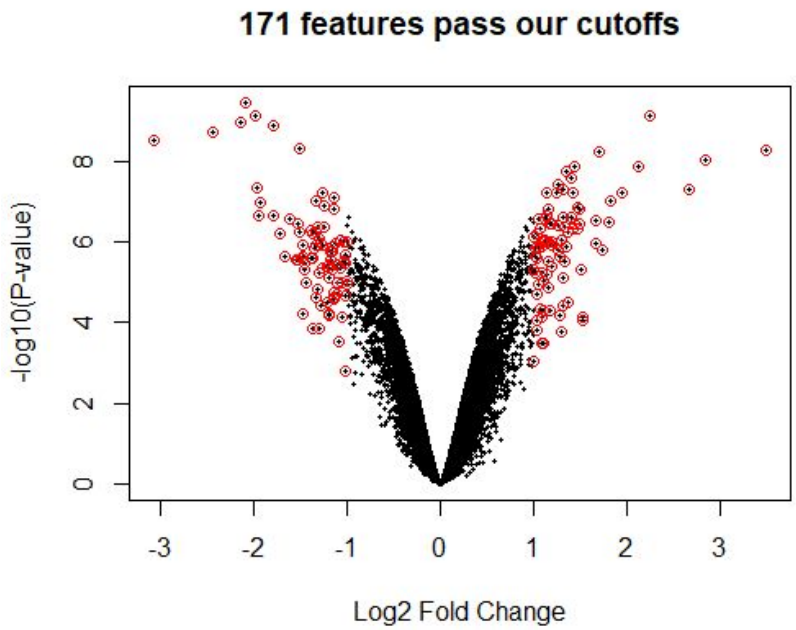
```
205830_at    5.000643e-03  2.4525107
212706_at    5.101094e-03  2.4149637
205535_s_at  5.580465e-03  2.2430129
210355_at    5.586227e-03  2.2378631
205532_s_at  5.808908e-03  2.1780336
1558501_at   5.837822e-03  2.1496409
201147_s_at  6.099969e-03  2.0657510
1555673_at   6.197900e-03  2.0430339
235561_at    6.275685e-03  2.0162314
202920_at    6.724926e-03  1.8978896
243009_at    6.913504e-03  1.8528277
220030_at    8.990789e-03  1.3947082
206300_s_at  9.250357e-03  1.3480368
205533_s_at  9.416580e-03  1.3192648
220301_at    1.041079e-02  1.1586786
1554131_at   1.420825e-02  0.6515165
 [ reached 'max' / getOption("max.print") -- omitted 5 rows ]

upregulated <- differentially_expressed[differentially_expressed$logFC > 0,]
genes_of_interest <- AnnotationDbi::select(hgu133plus2.db,
                                keys=rownames(upregulated),
                                columns=c("SYMBOL","ENTREZID","GENENAME"),
                                keytype="PROBEID")
genes_of_interest
#save data into a csv file
write.csv(genes_of_interest,'genes_of_interest.csv')

#volcanoplots
volcanoplot(fitted.ebayes)
```

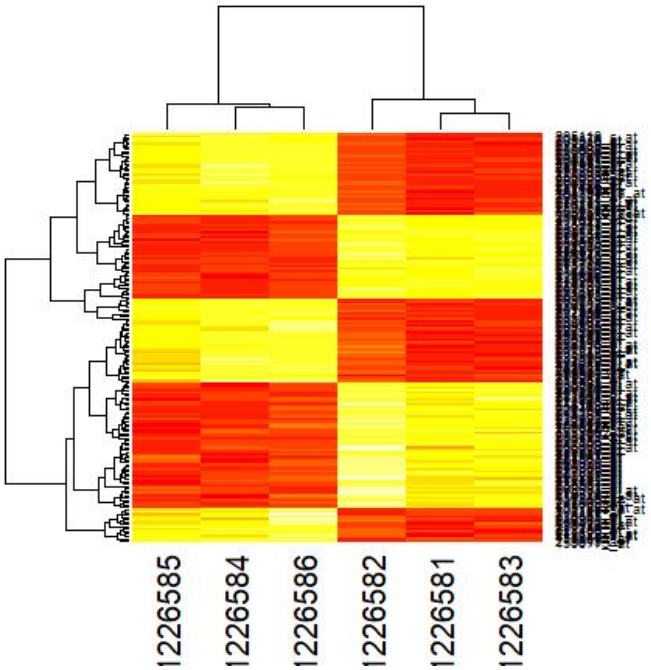



```
interesting_genes <- topTable(fitted.ebayes,number=Inf,p.value = 0.1,lfc=1)
volcanoplot(fitted.ebayes, main=sprintf("%d features pass our
cutoffs",nrow(interesting_genes)))
```

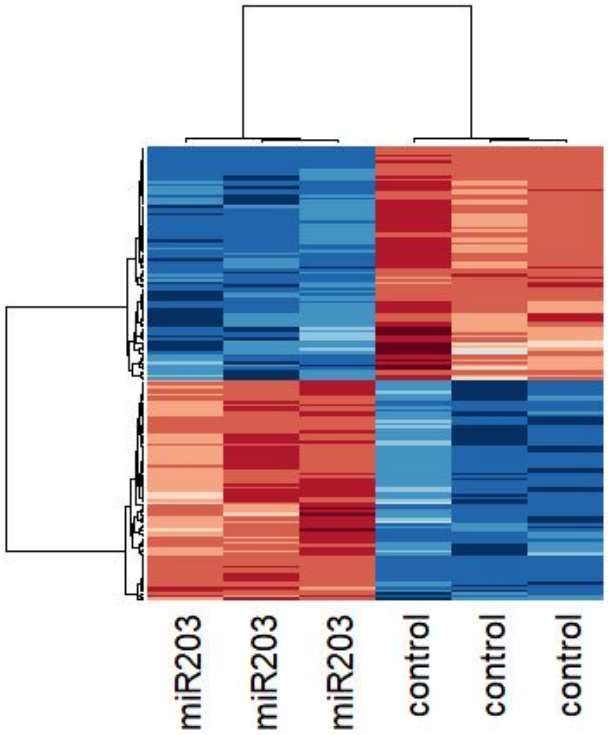



```
points(interesting_genes[['logFC']],-log10(interesting_genes[['P.Value']]),col='red')
```

```
#heatmaps
#normalise expression value
eset_of_interest <- eset[rownames(interesting_genes),]
heatmap(exprs(eset_of_interest))
```




```
#fix and beautify heatmap
install.packages("RColorBrewer")
library(RColorBrewer)
eset_of_interest <- eset[rownames(interesting_genes),]
heatmap(exprs(eset_of_interest),
      labCol=eset$culture,labRow=NA,
      col      = rev(brewer.pal(10, "RdBu")),
      distfun   = function(x) as.dist(1-cor(t(x))))
```




```
#Results
recommendations <- AnnotationDbi::select(hgu133plus2.db,
                                         keys=rownames(interesting_genes),
                                         columns=c("SYMBOL","ENTREZID","GENENAME"),
                                         keytype="PROBEID")
write.csv(recommendations,'recommendations.csv')
#shown in appendix below
```