

day15_bulk_v_floating_adipoDE

Sarah Hazell Pickering

2/9/2022

Summary

The donor 1 floating adipocytes are quite different between replicates, but still very distinct from bulk adipocyte samples (see PCA plot). Floating adipocytes are enriched for mitochondrial metabolism genes or adipogenesis hallmarks, whilst bulk samples are enriched for genes related to the extracellular matrix and/or blood vessel morphogenesis.

Output file:

- day15_bulk.v.float.tab
which shows pvalues and logFCs for DE between bulk and floating adipocytes (D1 only and D2 only stats are given in the right most columns as well).

```
library(limma)
library(biomaRt)
library(ggplot2)
library(RUVSeq)
library(clusterProfiler)
```

Load read counts

```
fc = read.delim(file.path("../02featureCounts/day15_bulk_v_float.counts"),
                header=T, stringsAsFactors = F, skip = 1)
colnames(fc) = gsub("output.01hisat.", "", gsub(".sorted.bam", "", colnames(fc)))
head(fc)

##          Geneid      Chr
## 1 ENSG00000223972    1;1;1;1;1;1;1;1;1
## 2 ENSG00000227232 1;1;1;1;1;1;1;1;1;1
## 3 ENSG00000278267           1
## 4 ENSG00000243485    1;1;1;1;1
## 5 ENSG00000284332           1
## 6 ENSG00000237613    1;1;1;1;1
##                                         Start
## 1                         11869;12010;12179;12613;12613;12975;13221;13221;13453
## 2 14404;15005;15796;16607;16858;17233;17606;17915;18268;24738;29534
## 3                               17369
## 4                         29554;30267;30564;30976;30976
## 5                           30366
```

```

## 6                               34554;35245;35277;35721;35721
##                                         End
## 1           12227;12057;12227;12721;12697;13052;13374;14409;13670
## 2 14501;15038;15947;16765;17055;17368;17742;18061;18366;24891;29570
## 3                                         17436
## 4                               30039;30667;30667;31109;31097
## 5                                         30503
## 6                               35174;35481;35481;36073;36081
##          Strand Length 13.21417_S53 14.21418_S2 15.21419_S9
## 1     +;+;+;+;+;+;+;+;+  1735      2.98      0.60      4.02
## 2   -;-;-;-;-;-;-;-;-  1351     71.95     63.77    116.12
## 3           -       68      5.00      3.00      1.00
## 4     +;+;+;+;+;+  1021      0.00      1.00      0.00
## 5           +      138      0.00      0.00      0.00
## 6   -;-;-;-;-;-  1219      5.38     11.66     18.16
## 16.21420_S16 17.21421_S24 18.21422_S32 19.21423_S40 20.21424_S47 21.21425_S54
## 1      6.99      8.54      5.99      2.88      1.37      1.70
## 2     83.60     95.42     60.59     86.75    169.35    144.65
## 3      1.00      1.00      1.00      6.75      6.90      5.25
## 4      0.45      0.50      0.00      0.79      1.75      0.70
## 5      0.00      0.00      0.00      0.00      0.00      0.00
## 6      0.00      0.00      0.00     11.62     31.77     17.51
## 22.21426_S3 23.21427_S10 24.21428_S17
## 1      4.13      8.44      5.39
## 2     83.21    112.73    265.86
## 3      8.84      3.50     11.40
## 4      0.00      0.50      0.53
## 5      0.00      0.00      0.00
## 6      0.00      0.00      0.50

```

Load sample information

```

sample_info = read.delim("../sample_info/day15_bulk_v_float_info.csv", sep=",",
                        header=T)
sample_info$sample_id = gsub("-", ".", sample_info$sample_id)
sample_info$bio.condition = gsub(" ", ".", sample_info$bio.condition)
sample_info

```

```

##   time rep donor separation bio.condition   sample_id
## 1 D15  P6   D1G   floating D1G.floating 13.21417_S53
## 2 D15 P8_1  D1G   floating D1G.floating 14.21418_S2
## 3 D15 P8_2  D1G   floating D1G.floating 15.21419_S9
## 4 D15 P6_1  D2A   floating D2A.floating 16.21420_S16
## 5 D15 P6_2  D2A   floating D2A.floating 17.21421_S24
## 6 D15  P7   D2A   floating D2A.floating 18.21422_S32
## 7 D15  P7   D1G     bulk      D1G 19.21423_S40
## 8 D15 P8_2  D1G     bulk      D1G 20.21424_S47
## 9 D15 P8_1  D1G     bulk      D1G 21.21425_S54
## 10 D15  P7   D2A     bulk      D2A 22.21426_S3
## 11 D15  P6   D2A     bulk      D2A 23.21427_S10
## 12 D15  P7   D2A     bulk      D2A 24.21428_S17

```

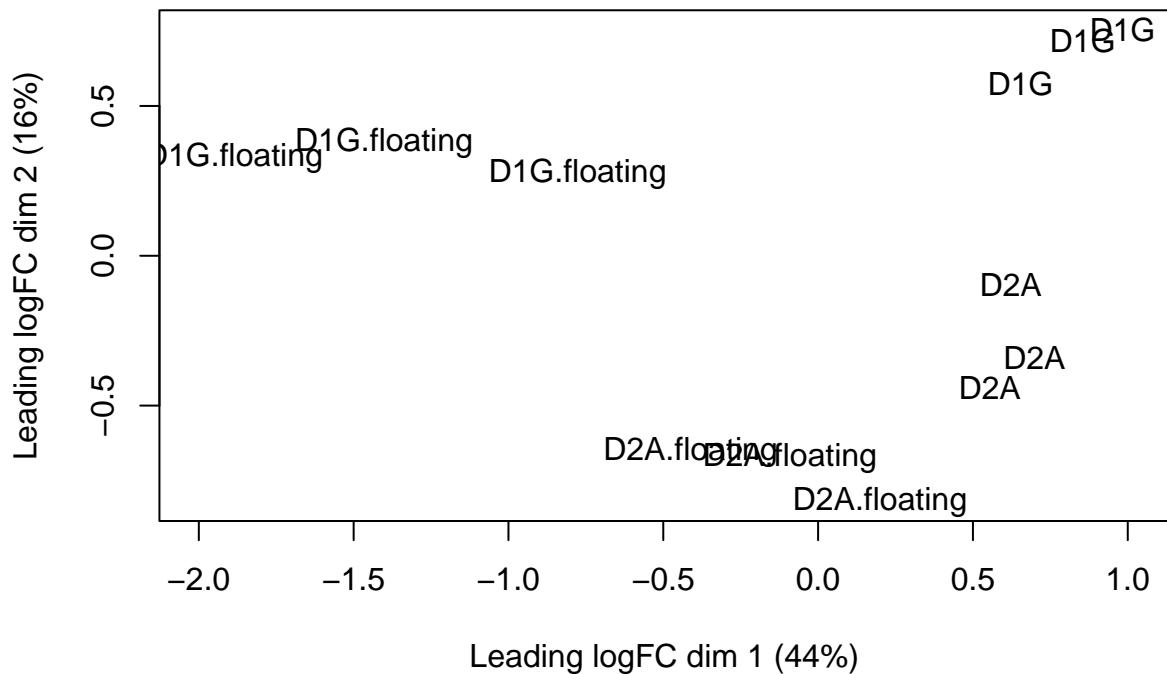
```
###Combine into DGE object

ob = DGEList(counts = data.matrix(fc[7:ncol(fc)]),
             samples = sample_info,
             group = sample_info$bio.condition,
             genes = fc[c(1,6)])
rownames(ob$counts) = ob$genes$Geneid
```

PCA

Bulk and floating transcriptomes are more different for donor1 than donor2. This first dimension does separate bulk and floating (though better for donor1) and the second dimension seems to separate donors. Bulk adipocytes have similarities between donors, but the floating adipocytes seem quite different from one another, even with donor1 there's a lot of variability.

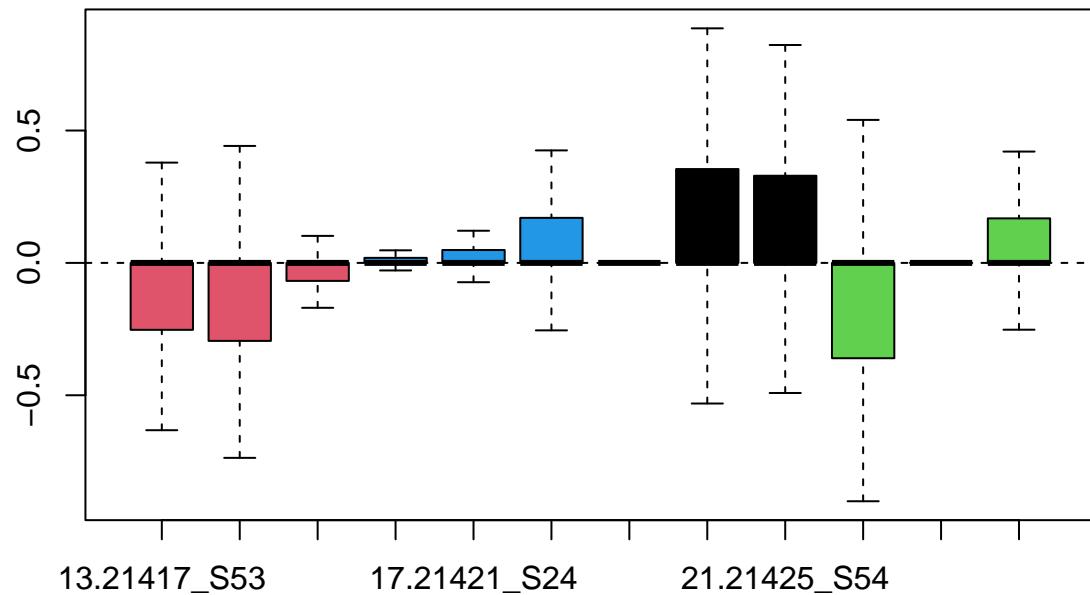
```
mds = plotMDS(ob, labels = ob$samples$group, top=5000)
```



Filtering

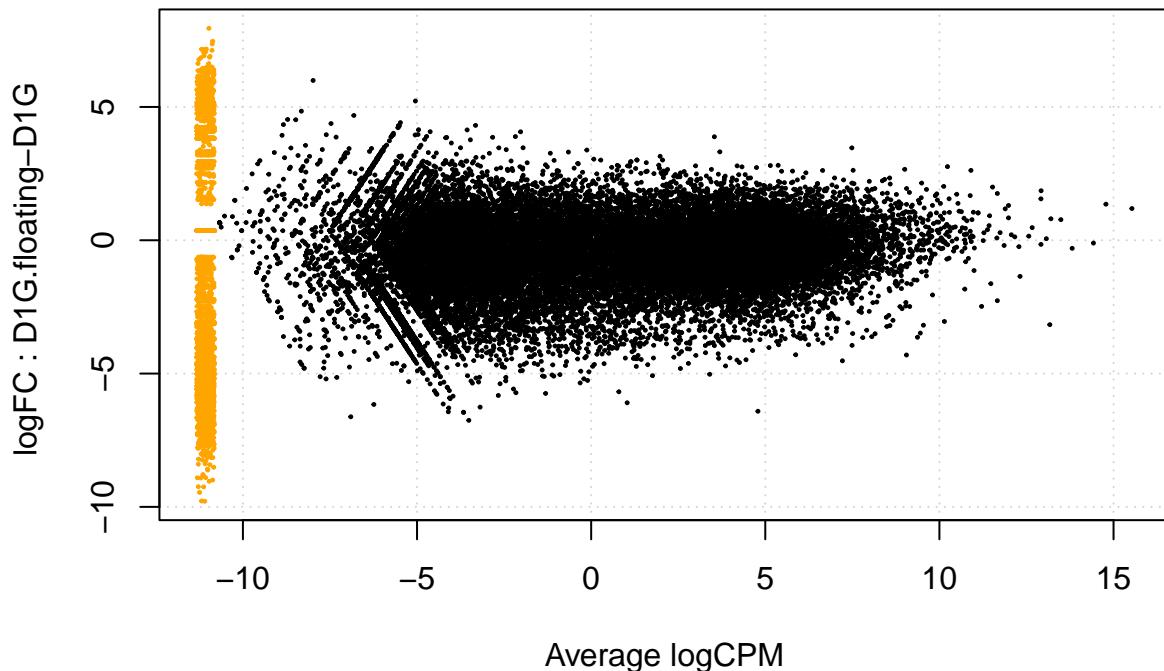
```
plotRLE(ob$counts, outline=FALSE, col=ob$samples$group, main="Before Filtering")
```

Before Filtering



```
plotSmear(ob, main = "Before Filtering")
```

Before Filtering



```
nrow(ob) #58 735 genes before filtering
```

```
## [1] 58735
```

```
filter = ob[filterByExpr(ob), , keep.lib.sizes=FALSE]
nrow(filter) #17 882 genes kept after filtering
```

```
## [1] 17882
```

The exact filter is approxiamtely 10/ libsize in millions (now about 0.2 CPM). This is a lower filter than previous, I guess because our library sizes are bigger.

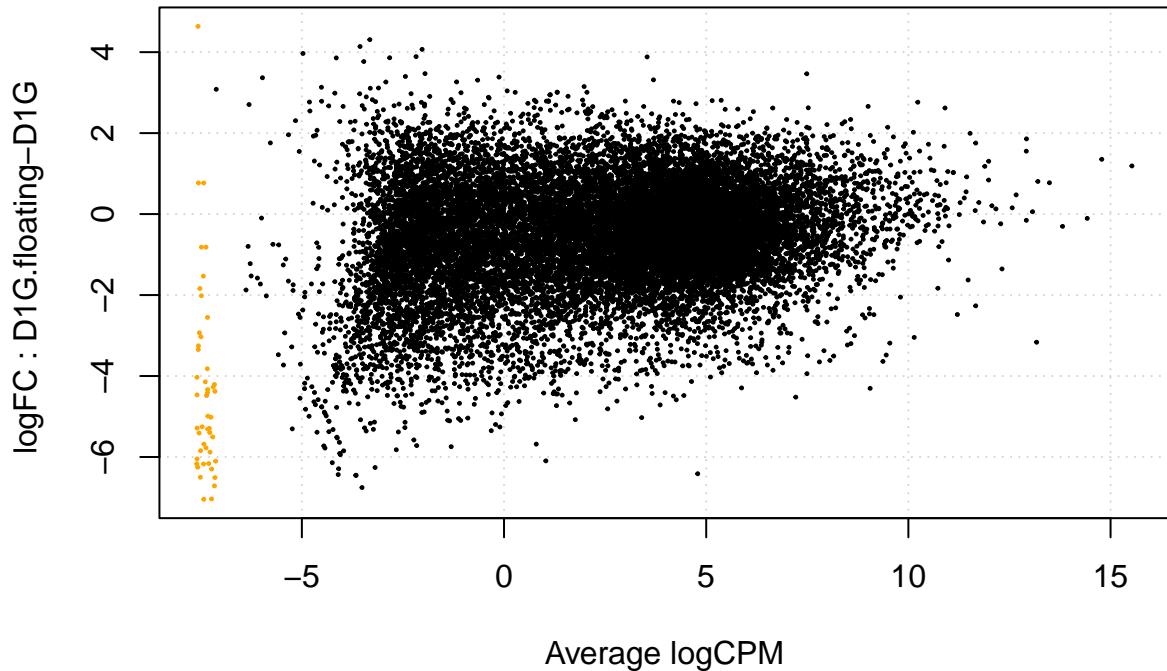
```
10/(median(ob$samples$lib.size)/1000000) #0.190 minimum CPM
```

```
## [1] 0.1901267
```

Hmm interesting there's higher expression in the floating samples, especially for lower CPM, higher sensitivity perhaps?

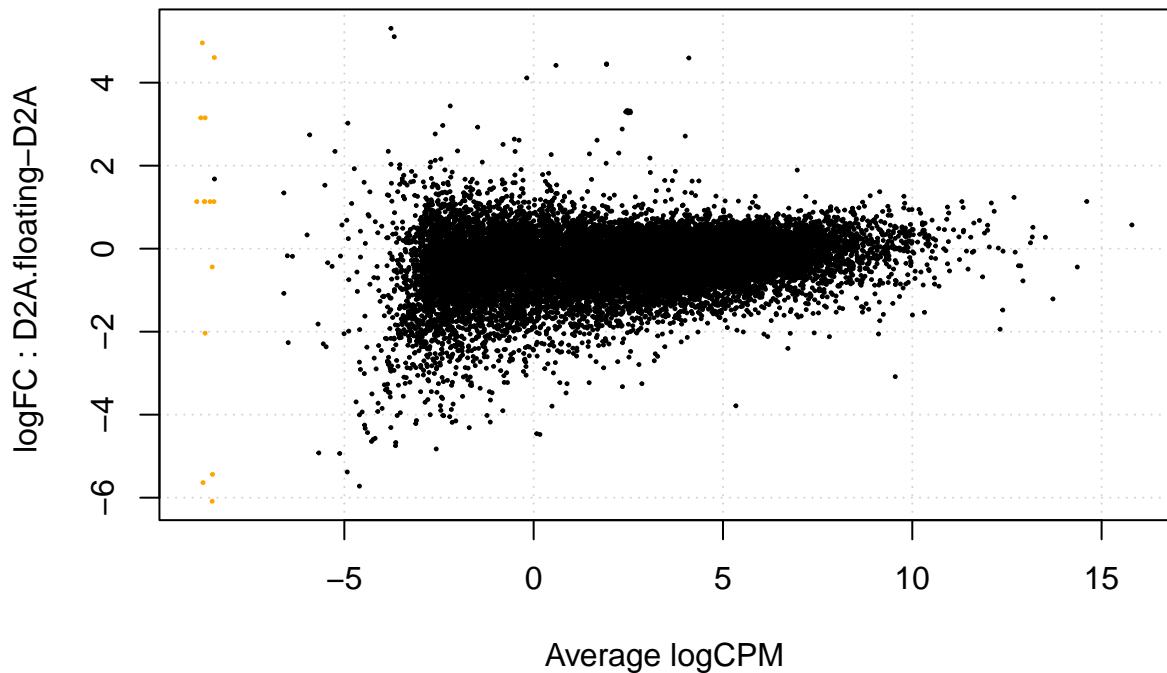
```
plotSmear(filter, main = "After Filtering, D1G")
```

After Filtering, D1G



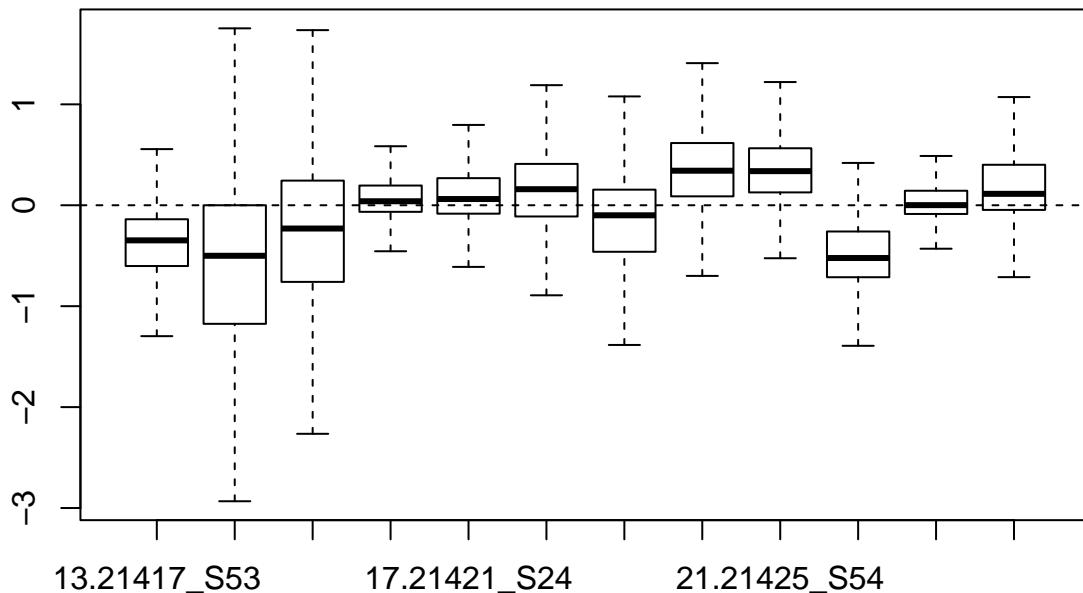
```
plotSmear(filt, main = "After Filtering, D2A", pair=c("D2A","D2A.floating"))
```

After Filtering, D2A



```
plotRLE(filt$counts, outline=FALSE,  
        col=filt$samples$treat, main="After filtering")
```

After filtering



Sanity check that the geneids haven't been mixed up.

```
filt$genes[filt$genes$Geneid == "ENSG00000228630",] #Hotair is still included
```

```
##           Geneid Length
## 37431 ENSG00000228630    2602
```

```
filt$genes[filt$genes$Geneid == "ENSG00000132170",] #PPARG is included
```

```
##           Geneid Length
## 9485 ENSG00000132170    4535
```

Normalise

Hmm these libraries have quite a bit of substructure, the means are not around zero, and the variances have quite large differences, particularly for two of those donor1 floating samples. You can see they sit differently in the PCA.

```
filt = calcNormFactors(filt,method = "TMM")
filt$samples[1:5]
```

```
##           group lib.size norm.factors time   rep
## 13.21417_S53 D1G.floating 39508175    0.9340652  D15   P6
```

```

## 14.21418_S2 D1G.floating 43186506      0.7584613 D15 P8_1
## 15.21419_S9 D1G.floating 53177563      0.7813403 D15 P8_2
## 16.21420_S16 D2A.floating 58737144      0.9856742 D15 P6_1
## 17.21421_S24 D2A.floating 64198765      0.9363287 D15 P6_2
## 18.21422_S32 D2A.floating 65275068      0.9953401 D15 P7
## 19.21423_S40           D1G 43492838      1.1040517 D15 P7
## 20.21424_S47           D1G 68497143      1.1196543 D15 P8_2
## 21.21425_S54           D1G 67310638      1.1531789 D15 P8_1
## 22.21426_S3            D2A 28511179      1.1061111 D15 P7
## 23.21427_S10           D2A 51385275      1.0600427 D15 P6
## 24.21428_S17           D2A 51968129      1.1765865 D15 P7

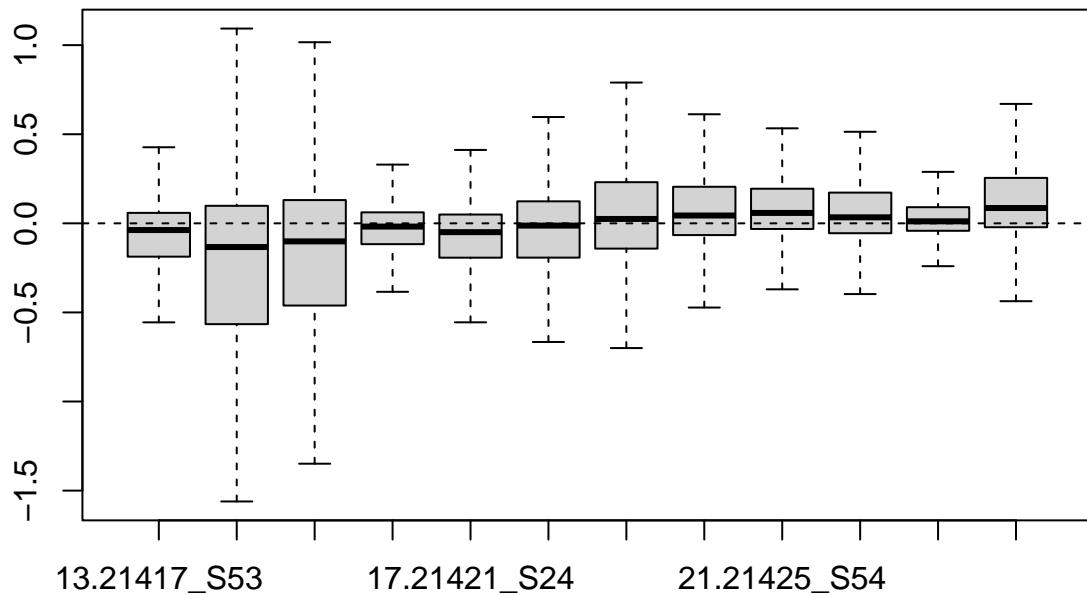
```

```

plotRLE(cpm(filt$counts, normalized.lib.sizes = FALSE), outline=FALSE,
        main="After TMM normalisation (no libsize)")

```

After TMM normalisation (no libsize)

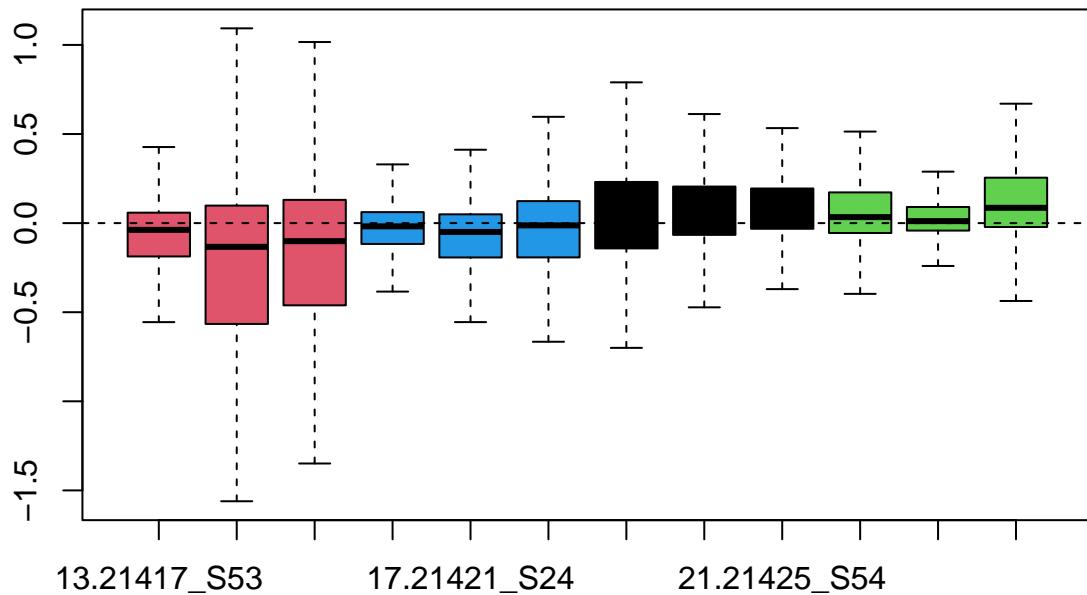


```

plotRLE(cpm(filt$counts, normalized.lib.sizes = TRUE), outline=FALSE,
        col=filt$samples$group, main="After TMM normalisation (with libsize)")

```

After TMM normalisation (with libsize)



Annnote with gene names

```
mart <- biomaRt::useMart(biomart = "ensembl",
  dataset = "hsapiens_gene_ensembl",
  host = "jan2019.archive.ensembl.org")

## Warning: Ensembl will soon enforce the use of https.
## Ensure the 'host' argument includes "https://"

annot = getBM(c("external_gene_name", "description", "ensembl_gene_id", "gene_biotype"),
  filters = "ensembl_gene_id",
  values = as.character(filt$genes$Geneid),
  mart = mart,
  useCache = F)
head(annot)

##   external_gene_name
## 1             SCYL3
## 2         C1orf112
## 3           FGR
## 4           CFH
## 5          STPG1
## 6          NIPAL3
```

```

##                                     description
## 1 SCY1 like pseudokinase 3 [Source:HGNC Symbol;Acc:HGNC:19285]
## 2 chromosome 1 open reading frame 112 [Source:HGNC Symbol;Acc:HGNC:25565]
## 3 FGR proto-oncogene, Src family tyrosine kinase [Source:HGNC Symbol;Acc:HGNC:3697]
## 4 complement factor H [Source:HGNC Symbol;Acc:HGNC:4883]
## 5 sperm tail PG-rich repeat containing 1 [Source:HGNC Symbol;Acc:HGNC:28070]
## 6 NIPA like domain containing 3 [Source:HGNC Symbol;Acc:HGNC:25233]
##   ensembl_gene_id   gene_biotype
## 1 ENSG00000000457 protein_coding
## 2 ENSG00000000460 protein_coding
## 3 ENSG00000000938 protein_coding
## 4 ENSG00000000971 protein_coding
## 5 ENSG00000001460 protein_coding
## 6 ENSG00000001461 protein_coding

#Tidying up the annot table
annot$description = gsub("\\[Source:.+\\]", "", annot$description)
colnames(annot)[1] = "gene_name"
colnames(annot)[3] = "gene"

#Add gene names to data object
filt$genes = merge(filt$genes, annot, by.x= "Geneid",
                   by.y = "gene", sort=FALSE)
colnames(filt$genes) = c("gene", "length", "gene_name", "description", "gene_biotype")
head(filt$genes)

##           gene length gene_name
## 1 ENSG0000227232    1351   WASH7P
## 2 ENSG0000237613    1219   FAM138A
## 3 ENSG0000238009    3726 AL627309.1
## 4 ENSG0000241860    6195 AL627309.5
## 5 ENSG0000279928     570 F0538757.1
## 6 ENSG0000279457    1397   WASH9P
##                                     description      gene_biotype
## 1   WAS protein family homolog 7, pseudogene unprocessed_pseudogene
## 2 family with sequence similarity 138 member A                  lincRNA
## 3                               novel transcript                  lincRNA
## 4                                         processed_transcript
## 5   DEAD/H-box helicase 11 (DDX11) pseudogene unprocessed_pseudogene
## 6   WAS protein family homolog 9, pseudogene unprocessed_pseudogene

```

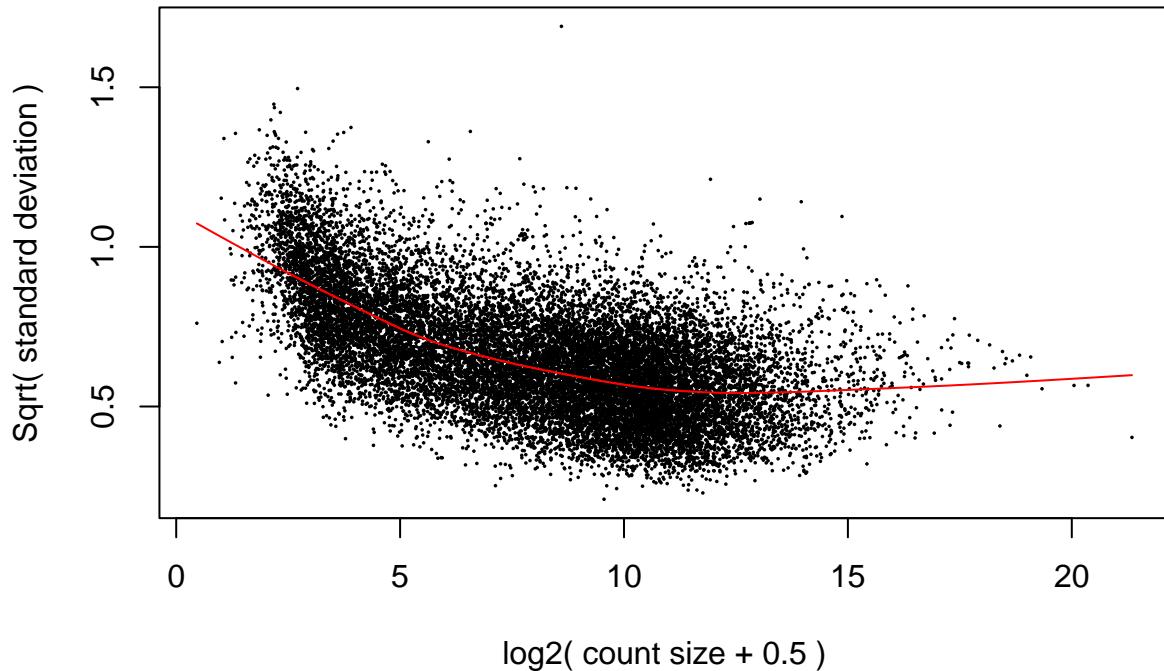
Limma DE

```

filt$samples$group = factor(filt$samples$group)
norm_ob = voom(filt, plot=T)

```

voom: Mean–variance trend



Simple separate design

```
design = model.matrix(~0+norm_ob$targets$group)
colnames(design) = levels(norm_ob$targets$group)
design

##      D1G D1G.floating D2A D2A.floating
## 1      0            1    0
## 2      0            1    0
## 3      0            1    0
## 4      0            0    0
## 5      0            0    0
## 6      0            0    0
## 7      1            0    0
## 8      1            0    0
## 9      1            0    0
## 10     0            0    1
## 11     0            0    1
## 12     0            0    1

## attr(),"assign")
## [1] 1 1 1 1
## attr(),"contrasts")
## attr(),"contrasts")$`norm_ob$targets$group`
## [1] "contr.treatment"
```

```

tests = makeContrasts(D1.bulk.v.float = D1G.floating - D1G,
                      D2.bulk.v.float = D2A.floating - D2A,
                      bulk.v.float = (D1G.floating - D1G) + (D2A.floating - D2A),
                      D1.v.D2 = D1G - D2A,
                      floating.D1.v.D2 = D1G.floating - D2A.floating,
                      levels = design) #test for which genes change the same

simple_fit = lmFit(norm_ob, design=design)
simple_fit = contrasts.fit(simple_fit, tests)
simple_fit = eBayes(simple_fit, robust=T)

results= decideTests(simple_fit, method="separate", adjust.method = "BH")
summary(results)

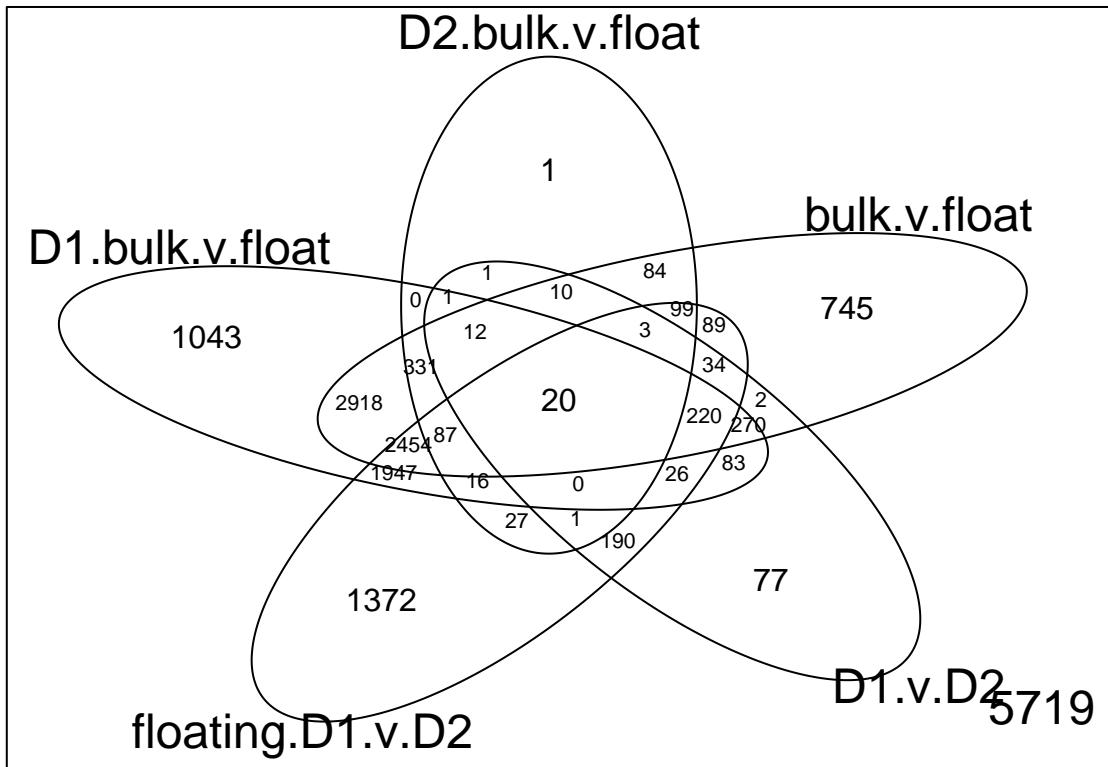
##          D1.bulk.v.float D2.bulk.v.float bulk.v.float D1.v.D2 floating.D1.v.D2
## Down           4809            427       4022      542        3385
## NotSig         8454           17189      10504     16932       11297
## Up             4619            266       3356      408        3200

summary(decideTests(simple_fit, method="separate", adjust.method = "BH", p.value = 0.01))

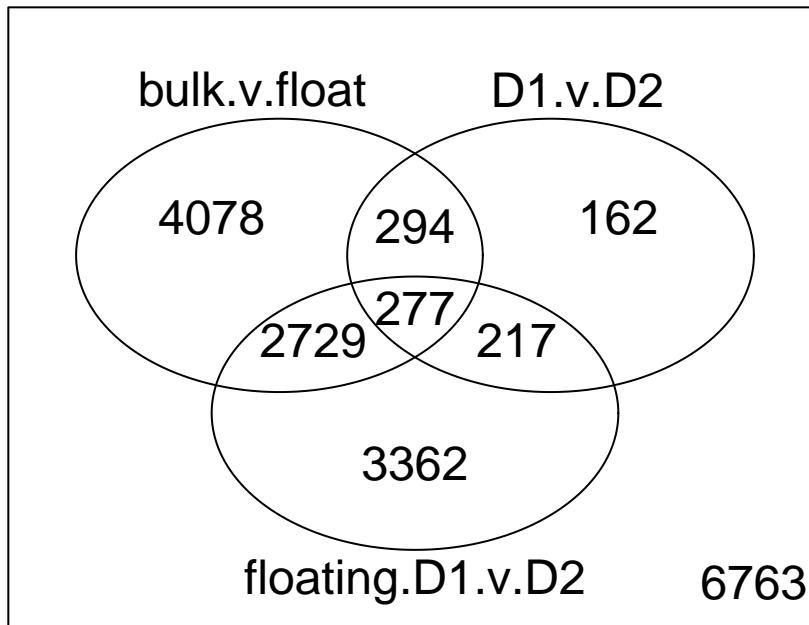
##          D1.bulk.v.float D2.bulk.v.float bulk.v.float D1.v.D2 floating.D1.v.D2
## Down           2915              0       2358      189        876
## NotSig         12193            17882      13774     17543       16364
## Up             2774              0       1750      150        642

vennDiagram(results)

```



```
vennDiagram(results[,3:5])
```



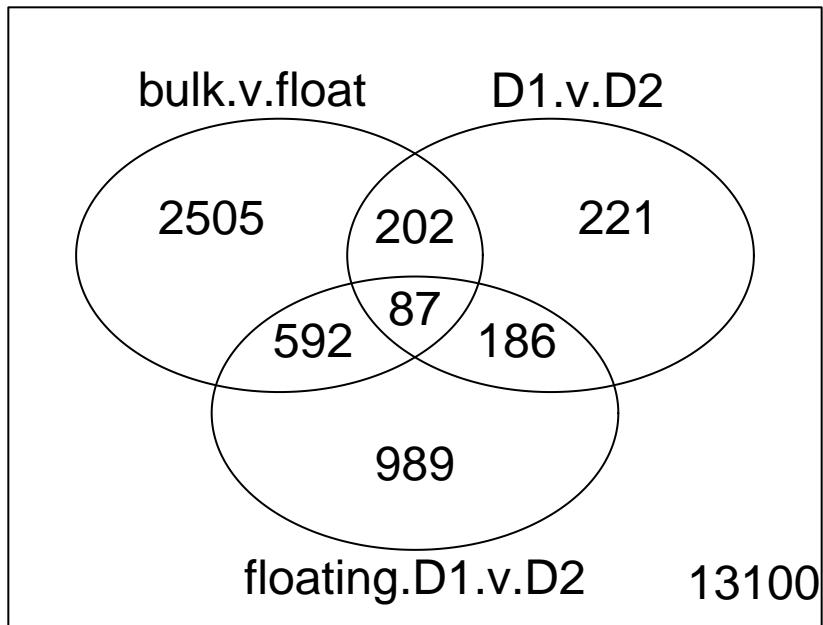
It's worth being a little more stringent on the pvalue threshold as well, considering how many D1 genes are showing up.

This is not a bad test, we have 213 genes that are found in all three tests and so seems robust. Only two genes seem to be significant in opposite directions which is a good sign. Interestingly the combined test picks up 563 genes that were ignored in the individual tests. There are less genes found uniquely in each individual test than in the combined, which I also take as a good sign.

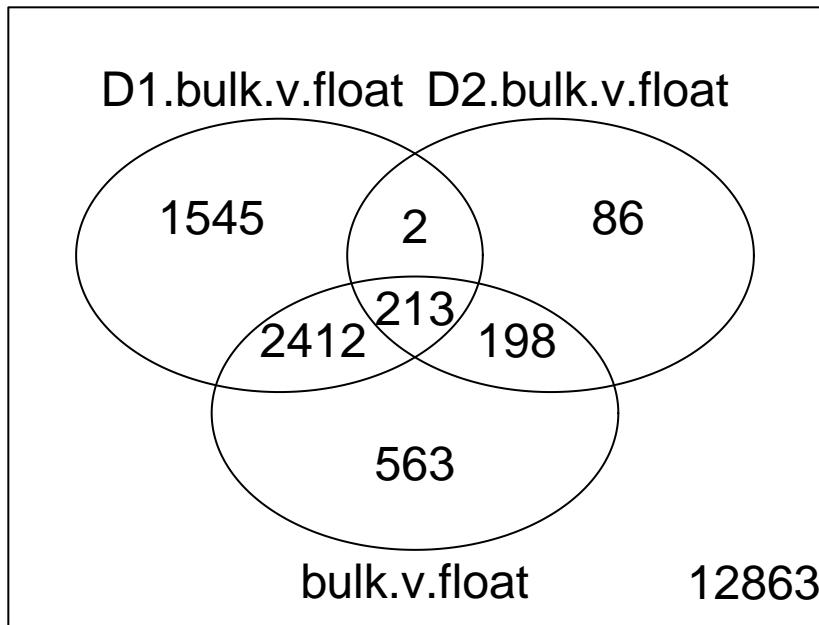
```
stringent = decideTests(simple_fit, method="global", adjust.method = "BH", p.value = 0.01)
summary(stringent)

##          D1.bulk.v.float D2.bulk.v.float bulk.v.float D1.v.D2 floating.D1.v.D2
## Down           2208          304       1978      399          1068
## NotSig        13710         17383      14496     17186          16028
## Up            1964           195       1408      297           786

vennDiagram(stringent[,3:5])
```



```
vennDiagram(stringent[,1:3])
```



Make a combined table

In an ANOVA-style combined test 4.5k genes are different between bulk and floating (not shown). Instead I opted for the combined test, which presents 4.1k DE genes at a 0.01 pvalue threshold.

Overwhelmingly significant genes are downregulated in floating cells, which makes sense with the loss of population diversity.

Top gene preferentially expressed in bulk samples is MEG3, a potential lncRNA tumor suppressor, via proliferation control. Another top gene is WEE1, a negative regulator of mitosis entry. There is also CDH11. These genes are also pretty highly expressed.

Genes preferentially expressed in floating adipocytes include a lot of TCA cycle and metabolism genes. Though not in the top 10 galectin 12 is expressed by floating adipocytes, as well as some other galectins.

```
common = topTable(simple_fit, coef = "bulk.v.float", number=nrow(simple_fit$genes))
head(common[3:ncol(common)], n=10) #bulk genes
```

```
##                      gene_name
## ENSG00000214548      MEG3
## ENSG00000059804      SLC2A3
## ENSG00000135318      NT5E
## ENSG00000163661      PTX3
## ENSG00000172716      SLFN11
## ENSG00000166483      WEE1
## ENSG00000235770 LINC00607
```

```

## ENSG00000140937      CDH11
## ENSG00000069535      MAOB
## ENSG00000115380      EFEMP1
##                                         description
## ENSG00000214548      maternally expressed 3
## ENSG00000059804      solute carrier family 2 member 3
## ENSG00000135318      5'-nucleotidase ecto
## ENSG00000163661      pentraxin 3
## ENSG00000172716      schlafen family member 11
## ENSG00000166483      WEE1 G2 checkpoint kinase
## ENSG00000235770      long intergenic non-protein coding RNA 607
## ENSG00000140937      cadherin 11
## ENSG00000069535      monoamine oxidase B
## ENSG00000115380      EGF containing fibulin extracellular matrix protein 1
##   gene_biotype      logFC    AveExpr      t      P.Value
## ENSG00000214548      lincRNA -9.633367 4.930369 -12.79926 1.254808e-08
## ENSG00000059804      protein_coding -3.893367 4.450288 -12.53753 1.432461e-08
## ENSG00000135318      protein_coding -3.627524 5.556880 -12.36862 1.671937e-08
## ENSG00000163661      protein_coding -3.770692 9.850227 -12.27791 1.987216e-08
## ENSG00000172716      protein_coding -3.867124 4.074533 -12.23848 1.883417e-08
## ENSG00000166483      protein_coding -3.176580 3.993851 -11.98935 2.314855e-08
## ENSG00000235770      lincRNA -4.591394 1.602233 -12.16860 1.939781e-08
## ENSG00000140937      protein_coding -3.937758 7.160479 -11.68409 3.668827e-08
## ENSG00000069535      protein_coding  2.958731 4.715113  11.53557 3.658506e-08
## ENSG00000115380      protein_coding -3.376068 8.219264 -11.38710 4.758911e-08
##   adj.P.Val      B
## ENSG00000214548  5.913461e-05 10.156846
## ENSG00000059804  5.913461e-05 10.082682
## ENSG00000135318  5.913461e-05 9.980404
## ENSG00000163661  5.913461e-05 9.829568
## ENSG00000172716  5.913461e-05 9.808846
## ENSG00000166483  5.913461e-05 9.617041
## ENSG00000235770  5.913461e-05 9.419036
## ENSG00000140937  7.289551e-05 9.266695
## ENSG00000069535  7.289551e-05 9.239335
## ENSG00000115380  7.352105e-05 9.017122

```

```
head(common[common$logFC > 0,3:ncol(common)], n=10)
```

```

##           gene_name
## ENSG00000069535      MAOB
## ENSG00000114054      PCCB
## ENSG00000077420      APBB1IP
## ENSG00000109107      ALDOC
## ENSG00000160752      FDPS
## ENSG00000143149      ALDH9A1
## ENSG00000104325      DECR1
## ENSG00000158571      PFKFB1
## ENSG00000278540      ACACA
## ENSG00000120437      ACAT2
##                                         description
## ENSG00000069535      monoamine oxidase B
## ENSG00000114054      propionyl-CoA carboxylase subunit beta
## ENSG00000077420      amyloid beta precursor protein binding family B member 1 interacting protein

```

```

## ENSG00000109107                               aldolase, fructose-bisphosphate C
## ENSG00000160752                             farnesyl diphosphate synthase
## ENSG00000143149                            aldehyde dehydrogenase 9 family member A1
## ENSG00000104325                           2,4-dienoyl-CoA reductase 1
## ENSG00000158571                           6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 1
## ENSG00000278540                           acetyl-CoA carboxylase alpha
## ENSG00000120437                          acetyl-CoA acetyltransferase 2

##          gene_biotype    logFC   AveExpr      t     P.Value
## ENSG0000069535 protein_coding 2.958731 4.715113 11.535566 3.658506e-08
## ENSG00000114054 protein_coding 2.397404 6.013270 10.707883 8.775489e-08
## ENSG0000077420 protein_coding 3.088500 7.325403 10.743580 9.202086e-08
## ENSG00000109107 protein_coding 3.776913 5.555454 10.773617 9.410744e-08
## ENSG00000160752 protein_coding 3.479902 9.443267 10.620699 1.109451e-07
## ENSG00000143149 protein_coding 2.358275 8.388481 10.338698 1.319895e-07
## ENSG00000104325 protein_coding 2.397319 7.913023 10.314143 1.356780e-07
## ENSG00000158571 protein_coding 2.627636 4.918301 10.234162 1.484761e-07
## ENSG00000278540 protein_coding 3.104517 8.794839 10.279698 1.581577e-07
## ENSG00000120437 protein_coding 2.856351 7.046501 9.894321 2.388206e-07

##          adj.P.Val      B
## ENSG0000069535 7.289551e-05 9.239335
## ENSG00000114054 7.352105e-05 8.423523
## ENSG0000077420 7.352105e-05 8.384341
## ENSG00000109107 7.352105e-05 8.359260
## ENSG00000160752 7.352105e-05 8.200432
## ENSG00000143149 7.352105e-05 8.027212
## ENSG00000104325 7.352105e-05 8.001176
## ENSG00000158571 7.506704e-05 7.906557
## ENSG00000278540 7.506704e-05 7.858165
## ENSG00000120437 7.700079e-05 7.455468

```

```
common[grep("galectin", common$description),]
```

| | gene | length | gene_name | description | |
|----|-----------------|-----------------|-------------|-------------|----------------------------|
| ## | ENSG00000133317 | ENSG00000133317 | 2270 | LGALS12 | galectin 12 |
| ## | ENSG00000131981 | ENSG00000131981 | 2397 | LGALS3 | galectin 3 |
| ## | ENSG00000168961 | ENSG00000168961 | 5714 | LGALS9 | galectin 9 |
| ## | ENSG00000100097 | ENSG00000100097 | 1099 | LGALS1 | galectin 1 |
| ## | ENSG00000108679 | ENSG00000108679 | 3584 | LGALS3BP | galectin 3 binding protein |
| ## | ENSG00000119862 | ENSG00000119862 | 4003 | LGALSL | galectin like |
| ## | ENSG00000116977 | ENSG00000116977 | 9710 | LGALS8 | galectin 8 |
| | gene_biotype | logFC | AveExpr | t | P.Value |
| ## | ENSG00000133317 | protein_coding | 3.24842219 | 8.361482 | 9.1688164 5.864689e-07 |
| ## | ENSG00000131981 | protein_coding | 1.32498326 | 6.970505 | 5.0856727 2.146017e-04 |
| ## | ENSG00000168961 | protein_coding | 3.07349954 | 1.583066 | 4.4295113 7.186445e-04 |
| ## | ENSG00000100097 | protein_coding | 2.10441483 | 10.029597 | 2.8892815 1.292895e-02 |
| ## | ENSG00000108679 | protein_coding | 1.29332756 | 7.784121 | 2.5297310 2.550644e-02 |
| ## | ENSG00000119862 | protein_coding | 0.16587060 | 2.433679 | 0.3465680 7.345919e-01 |
| ## | ENSG00000116977 | protein_coding | -0.09327851 | 5.666180 | -0.2653062 7.950321e-01 |
| | adj.P.Val | B | | | |
| ## | ENSG00000133317 | 8.337763e-05 | 6.5756143 | | |
| ## | ENSG00000131981 | 1.974026e-03 | 0.6183539 | | |
| ## | ENSG00000168961 | 4.488579e-03 | -0.3547616 | | |
| ## | ENSG00000100097 | 3.539719e-02 | -3.4271000 | | |
| ## | ENSG00000108679 | 5.847514e-02 | -4.1194544 | | |

```

## ENSG00000119862 7.957819e-01 -6.5656407
## ENSG00000116977 8.453302e-01 -6.8240437

summary(common$adj.P.Val < 0.01)

##      Mode     FALSE      TRUE
## logical    13774     4108

d1 = topTable(simple_fit, coef = "D1.bulk.v.float", number=nrow(simple_fit$genes))

d2 = topTable(simple_fit, coef = "D2.bulk.v.float", number=nrow(simple_fit$genes))

both = merge(d1, d2, suffixes=c("_d1", "_d2"), by=c("gene", "length", "gene_name", "description", "gene_biotype"))
head(both)

##           gene length gene_name
## 1 ENSG00000000003   4535   TSPAN6
## 2 ENSG00000000005   1610    TNMD
## 3 ENSG00000000419   1207    DPM1
## 4 ENSG00000000457   6883   SCYL3
## 5 ENSG00000000460   5967 C1orf112
## 6 ENSG00000000938   3474    FGR

##                                     description gene_biotype
## 1                               tetraspanin 6 protein_coding
## 2                               tenomodulin protein_coding
## 3 dolichyl-phosphate mannosyltransferase subunit 1, catalytic protein_coding
## 4                               SCY1 like pseudokinase 3 protein_coding
## 5                               chromosome 1 open reading frame 112 protein_coding
## 6 FGR proto-oncogene, Src family tyrosine kinase protein_coding

##      logFC_d1 AveExpr_d1      t_d1 P.Value_d1 adj.P.Val_d1      B_d1
## 1  0.67547430  5.514915  3.89487543 1.872510e-03 0.0069182267 -1.645111
## 2 -0.02997169  0.616374 -0.06191358 9.515956e-01 0.9617064152 -6.408588
## 3 -0.34481733  5.230867 -2.02247538 6.438590e-02 0.1040944560 -5.056931
## 4 -0.22910794  3.348528 -1.12317095 2.818515e-01 0.3559622158 -6.179222
## 5  0.46905196  1.671116  1.95110165 7.312595e-02 0.1158740158 -4.920744
## 6  2.47196326  2.527861  7.51165881 5.095736e-06 0.0001874937  4.464759

##      logFC_d2 AveExpr_d2      t_d2 P.Value_d2 adj.P.Val_d2      B_d2
## 1  0.8405987  5.514915  4.8299570 0.0003367825 0.0286158  0.2992962
## 2  0.3168268  0.616374  0.8607211 0.4053521278 0.6389727 -6.0156622
## 3  0.1961160  5.230867  1.1778541 0.2601463871 0.5047128 -5.9876391
## 4  0.2047768  3.348528  1.0288597 0.3224678841 0.5650197 -6.0196316
## 5  0.5926893  1.671116  2.4166014 0.0312488679 0.1664746 -3.8739835
## 6  0.6055041  2.527861  1.9922708 0.0682922514 0.2482793 -4.6900573

all = merge(common, both, by=c("gene", "length", "gene_name", "description", "gene_biotype"))
head(all)

```

```

##           gene length gene_name
## 1 ENSG00000000003   4535   TSPAN6
## 2 ENSG00000000005   1610    TNMD
## 3 ENSG00000000419   1207    DPM1
## 4 ENSG00000000457   6883   SCYL3

```

```

## 5 ENSG00000000460 5967 C1orf112
## 6 ENSG00000000938 3474 FGR
##                                     description gene_biotype
## 1                               tetraspanin 6 protein_coding
## 2                               tenomodulin protein_coding
## 3 dolichyl-phosphate mannosyltransferase subunit 1, catalytic protein_coding
## 4                               SCY1 like pseudokinase 3 protein_coding
## 5                               chromosome 1 open reading frame 112 protein_coding
## 6 FGR proto-oncogene, Src family tyrosine kinase protein_coding
##   logFC AveExpr      t     P.Value    adj.P.Val      B
## 1  1.51607298 5.514915  6.17054342 3.504952e-05 0.0006688960  2.470948
## 2  0.28685510 0.616374  0.47169149 6.451497e-01 0.7218926450 -6.292446
## 3 -0.14870133 5.230867 -0.62398664 5.435233e-01 0.6328795816 -6.638823
## 4 -0.02433117 3.348528 -0.08537351 9.332764e-01 0.9501735805 -6.708540
## 5  1.06174124 1.671116  3.09156767 8.661531e-03 0.0261940635 -2.806212
## 6  3.07746733 2.527861  6.86997408 1.284588e-05 0.0003716772  3.571197
##   logFC_d1 AveExpr_d1      t_d1     P.Value_d1 adj.P.Val_d1      B_d1
## 1  0.67547430 5.514915  3.89487543 1.872510e-03 0.0069182267 -1.645111
## 2 -0.02997169 0.616374 -0.06191358 9.515956e-01 0.9617064152 -6.408588
## 3 -0.34481733 5.230867 -2.02247538 6.438590e-02 0.1040944560 -5.056931
## 4 -0.22910794 3.348528 -1.12317095 2.818515e-01 0.3559622158 -6.179222
## 5  0.46905196 1.671116  1.95110165 7.312595e-02 0.1158740158 -4.920744
## 6  2.47196326 2.527861  7.51165881 5.095736e-06 0.0001874937  4.464759
##   logFC_d2 AveExpr_d2      t_d2     P.Value_d2 adj.P.Val_d2      B_d2
## 1  0.8405987 5.514915  4.8299570 0.0003367825 0.0286158  0.2992962
## 2  0.3168268 0.616374  0.8607211 0.4053521278 0.6389727 -6.0156622
## 3  0.1961160 5.230867  1.1778541 0.2601463871 0.5047128 -5.9876391
## 4  0.2047768 3.348528  1.0288597 0.3224678841 0.5650197 -6.0196316
## 5  0.5926893 1.671116  2.4166014 0.0312488679 0.1664746 -3.8739835
## 6  0.6055041 2.527861  1.9922708 0.0682922514 0.2482793 -4.6900573

all = all[grep("^(t.*|B_d.*|P.Value_d.*)", colnames(all), invert=T)]
all = all[order(all$adj.P.Val, -(all$B)),]
head(all)

##           gene length gene_name                  description
## 13573 ENSG00000214548 23224 MEG3 maternally expressed 3
## 669 ENSG00000059804 6159 SLC2A3 solute carrier family 2 member 3
## 5648 ENSG00000135318 4533 NT5E 5'-nucleotidase ecto
## 8718 ENSG00000163661 1940 PTX3 pentraxin 3
## 10268 ENSG00000172716 5982 SLFN11 schlafen family member 11
## 9295 ENSG00000166483 9239 WEE1 WEE1 G2 checkpoint kinase
##   gene_biotype logFC AveExpr      P.Value    adj.P.Val      B
## 13573 lincRNA -9.633367 4.930369 1.254808e-08 5.913461e-05 10.156846
## 669 protein_coding -3.893367 4.450288 1.432461e-08 5.913461e-05 10.082682
## 5648 protein_coding -3.627524 5.556880 1.671937e-08 5.913461e-05 9.980404
## 8718 protein_coding -3.770692 9.850227 1.987216e-08 5.913461e-05 9.829568
## 10268 protein_coding -3.867124 4.074533 1.883417e-08 5.913461e-05 9.808846
## 9295 protein_coding -3.176580 3.993851 2.314855e-08 5.913461e-05 9.617041
##   logFC_d1 AveExpr_d1 adj.P.Val_d1 logFC_d2 AveExpr_d2 adj.P.Val_d2
## 13573 -5.960323 4.930369 4.265980e-05 -3.6730440 4.930369 0.01374213
## 669 -2.958328 4.450288 2.504413e-05 -0.9350386 4.450288 0.03474772
## 5648 -3.074848 5.556880 2.078967e-05 -0.5526757 5.556880 0.11857481
## 8718 -3.066278 9.850227 2.078967e-05 -0.7044144 9.850227 0.08250291

```

```

## 10268 -3.116012 4.074533 2.504413e-05 -0.7511128 4.074533 0.06438885
## 9295 -3.016260 3.993851 1.869745e-05 -0.1603199 3.993851 0.62390448

write.table(all,"../03limma/day15_bulk.v.float.tab",
            sep="\t", row.names=F, quote=F)

```

Quick GO

```

molsig <- clusterProfiler::read.gmt("../sample_info/msigdb.v7.4.symbols.gmt")
head(molsig); nrow(molsig)

```

```

##      term      gene
## 1 chr1p12    VTCN1
## 2 chr1p12 LINCO1525
## 3 chr1p12    MAN1A2
## 4 chr1p12    VPS25P1
## 5 chr1p12    TENT5C
## 6 chr1p12    VDAC2P3

## [1] 3854442

##Pick gene sets to look in bcos all in a lot
prefixes = c("HALLMARK", "KEGG", "REACTOME", "WP", "GOBP", "GOCC", "GOMF")
colnames(molsig) = c("term", "gene")
some.molsig = molsig[gsub("_.*","", molsig$term) %in% prefixes,]
some.molsig$term = factor(some.molsig$term)
table(gsub("_.*","", some.molsig$term))

```

```

##
##          GOBP      GOCC      GOMF HALLMARK      KEGG REACTOME      WP
## 636699    97202    103847     7321    12797    89457    27711

```

Floating adipocytes are largely enriched for mitochondrial metabolism genes, but also 89 adipogenesis genes.

```

go_up = enricher(gene=common$gene_name[common$logFC > 0 &
                                         common$adj.P.Val < 0.01],
                  TERM2GENE = some.molsig,
                  minGSSize = 10,
                  universe = common$gene_name) #universe all expressed genes
print(head(go_up, n=10)[c(3:7)])

```

| | GeneRatio |
|--|-----------|
| ## HALLMARK_OXIDATIVE_PHOSPHORYLATION | 116/1405 |
| ## GOCC_MITOCHONDRIAL_PROTEIN_CONTAINING_COMPLEX | 130/1405 |
| ## GOCC_MITOCHONDRIAL_MATRIX | 176/1405 |
| ## GOBP_GENERATION_OF_PRECURSOR_METABOLITES_AND_ENERGY | 163/1405 |
| ## REACTOME_THE_CITRIC_ACID_TCA_CYCLE_AND_RESPIRATORY_ELECTRON_TRANSPORT | 94/1405 |
| ## GOBP_CELLULAR_RESPIRATION | 91/1405 |
| ## GOBP_OXIDATIVE_PHOSPHORYLATION | 78/1405 |

```

## GOBP_ENERGY_DERIVATION_BY_OXIDATION_OF_ORGANIC_COMPOUNDS      106/1405
## GOCC_INNER_MITOCHONDRIAL_MEMBRANE_PROTEIN_COMPLEX            75/1405
## GOBP_MONOCARBOXYLIC_ACID_METABOLIC_PROCESS                  152/1405
##
##                                     BgRatio
## HALLMARK_OXIDATIVE_PHOSPHORYLATION                          200/12922
## GOCC_MITOCHONDRIAL_PROTEIN_CONTAINING_COMPLEX              253/12922
## GOCC_MITOCHONDRIAL_MATRIX                                442/12922
## GOBP_GENERATION_OF_PRECURSOR_METABOLITES_AND_ENERGY        452/12922
## REACTOME_THE_CITRIC_ACID_TCA_CYCLE_AND_RESPIRATORY_ELECTRON_TRANSPORT 174/12922
## GOBP_CELLULAR_RESPIRATION                               173/12922
## GOBP_OXIDATIVE_PHOSPHORYLATION                           135/12922
## GOBP_ENERGY_DERIVATION_BY_OXIDATION_OF_ORGANIC_COMPOUNDS   248/12922
## GOCC_INNER_MITOCHONDRIAL_MEMBRANE_PROTEIN_COMPLEX          134/12922
## GOBP_MONOCARBOXYLIC_ACID_METABOLIC_PROCESS                485/12922
##
##                                     pvalue
## HALLMARK_OXIDATIVE_PHOSPHORYLATION                         2.086118e-60
## GOCC_MITOCHONDRIAL_PROTEIN_CONTAINING_COMPLEX             3.648471e-59
## GOCC_MITOCHONDRIAL_MATRIX                                6.540631e-59
## GOBP_GENERATION_OF_PRECURSOR_METABOLITES_AND_ENERGY       1.363852e-47
## REACTOME_THE_CITRIC_ACID_TCA_CYCLE_AND_RESPIRATORY_ELECTRON_TRANSPORT 2.424884e-45
## GOBP_CELLULAR_RESPIRATION                               1.251222e-42
## GOBP_OXIDATIVE_PHOSPHORYLATION                           1.231414e-40
## GOBP_ENERGY_DERIVATION_BY_OXIDATION_OF_ORGANIC_COMPOUNDS 8.766013e-39
## GOCC_INNER_MITOCHONDRIAL_MEMBRANE_PROTEIN_COMPLEX          8.766996e-38
## GOBP_MONOCARBOXYLIC_ACID_METABOLIC_PROCESS                5.870673e-36
##
##                                     p.adjust
## HALLMARK_OXIDATIVE_PHOSPHORYLATION                         1.438587e-56
## GOCC_MITOCHONDRIAL_PROTEIN_CONTAINING_COMPLEX             1.257993e-55
## GOCC_MITOCHONDRIAL_MATRIX                                1.503473e-55
## GOBP_GENERATION_OF_PRECURSOR_METABOLITES_AND_ENERGY       2.351281e-44
## REACTOME_THE_CITRIC_ACID_TCA_CYCLE_AND_RESPIRATORY_ELECTRON_TRANSPORT 3.344401e-42
## GOBP_CELLULAR_RESPIRATION                               1.438072e-39
## GOBP_OXIDATIVE_PHOSPHORYLATION                           1.213119e-37
## GOBP_ENERGY_DERIVATION_BY_OXIDATION_OF_ORGANIC_COMPOUNDS 7.556303e-36
## GOCC_INNER_MITOCHONDRIAL_MEMBRANE_PROTEIN_COMPLEX          6.717467e-35
## GOBP_MONOCARBOXYLIC_ACID_METABOLIC_PROCESS                4.048416e-33
##
##                                     qvalue
## HALLMARK_OXIDATIVE_PHOSPHORYLATION                         1.297126e-56
## GOCC_MITOCHONDRIAL_PROTEIN_CONTAINING_COMPLEX             1.134290e-55
## GOCC_MITOCHONDRIAL_MATRIX                                1.355632e-55
## GOBP_GENERATION_OF_PRECURSOR_METABOLITES_AND_ENERGY       2.120073e-44
## REACTOME_THE_CITRIC_ACID_TCA_CYCLE_AND_RESPIRATORY_ELECTRON_TRANSPORT 3.015535e-42
## GOBP_CELLULAR_RESPIRATION                               1.296662e-39
## GOBP_OXIDATIVE_PHOSPHORYLATION                           1.093829e-37
## GOBP_ENERGY_DERIVATION_BY_OXIDATION_OF_ORGANIC_COMPOUNDS 6.813268e-36
## GOCC_INNER_MITOCHONDRIAL_MEMBRANE_PROTEIN_COMPLEX          6.056917e-35
## GOBP_MONOCARBOXYLIC_ACID_METABOLIC_PROCESS                3.650323e-33

```

```
go_up[grep("ADIPO", go_up$ID), 3:9]
```

```

##                                     GeneRatio   BgRatio      pvalue      p.adjust
## HALLMARK_ADIPOGENESIS      89/1405 196/12922 3.495522e-35 2.008760e-32
## WP_LEPTIN_AND_ADIPONECTIN  5/1405 10/12922 2.377042e-03 3.152323e-02
##                                     qvalue

```

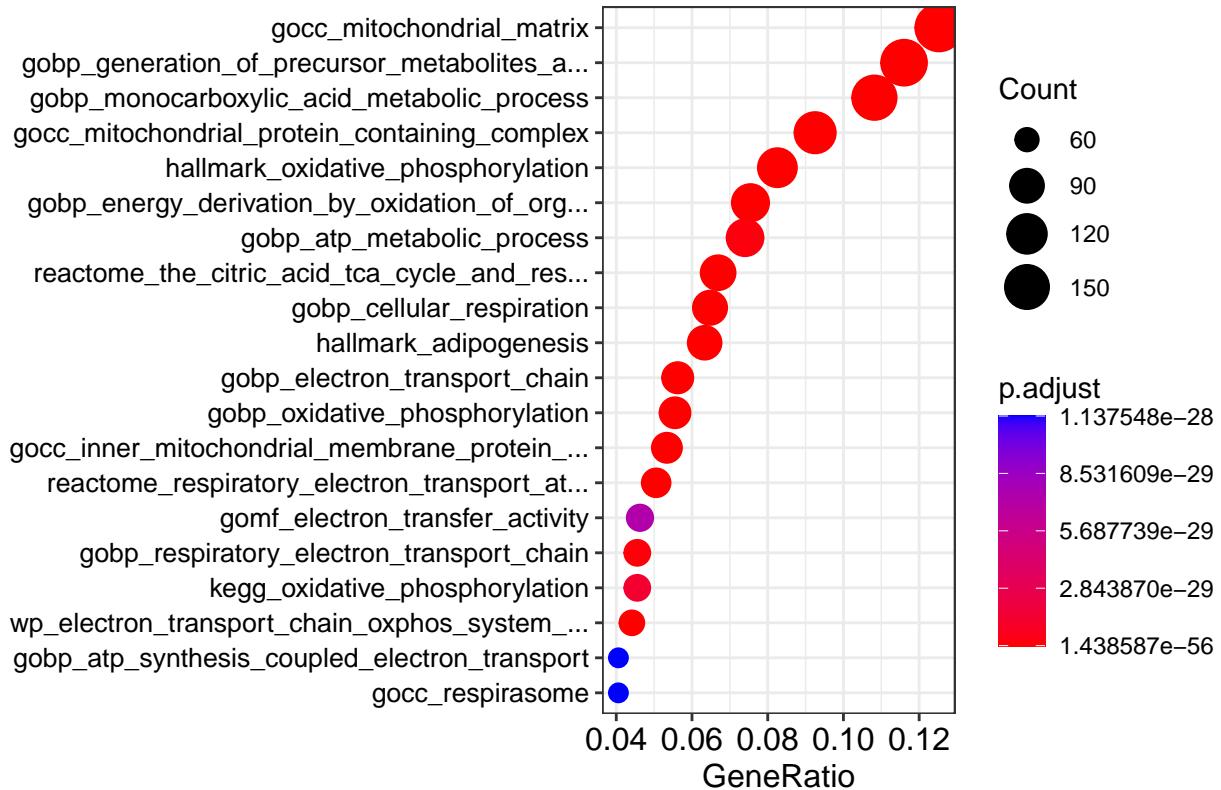
```

## HALLMARKADIPOGENESIS      1.811232e-32
## WPLEPTIN_AND_ADIPONECTIN 2.842345e-02
##
## HALLMARKADIPOGENESIS      DECR1/CAT/PPARG/HADH/CS/CDKN2C/COX7B/QDPR/MGST3/IDH1/LPCAT3/ACAA2/SCARB1/CH
## WPLEPTIN_AND_ADIPONECTIN
##
##                                     Count
## HALLMARKADIPOGENESIS          89
## WPLEPTIN_AND_ADIPONECTIN     5

dotplot(go_up, showCategory=20) + scale_y_discrete(label=function(x) tolower(stringr::str_trunc(x,45)))
theme(axis.text.y = element_text(size=10))

## Scale for y is already present.
## Adding another scale for y, which will replace the existing scale.

```



Bulk adipocytes are enriched for a variety of terms, with extracellular matrix very strong (likely driving the epithelial to mesenchymal gene set) as well as blood vessel and connective tissue GO terms. In fact ~8% of significant protein coding genes are related to blood vessel morphogenesis (GeneRatio = 0.08).

```

go_down = enricher(gene=common$gene_name[common$logFC < 0 &
                                         common$adj.P.Val < 0.01],
                    TERM2GENE = some.molsig,
                    minGSSize = 10,
                    universe = common$gene_name) #universe all expressed genes
print(head(go_down, n=10)[c(3:7)])

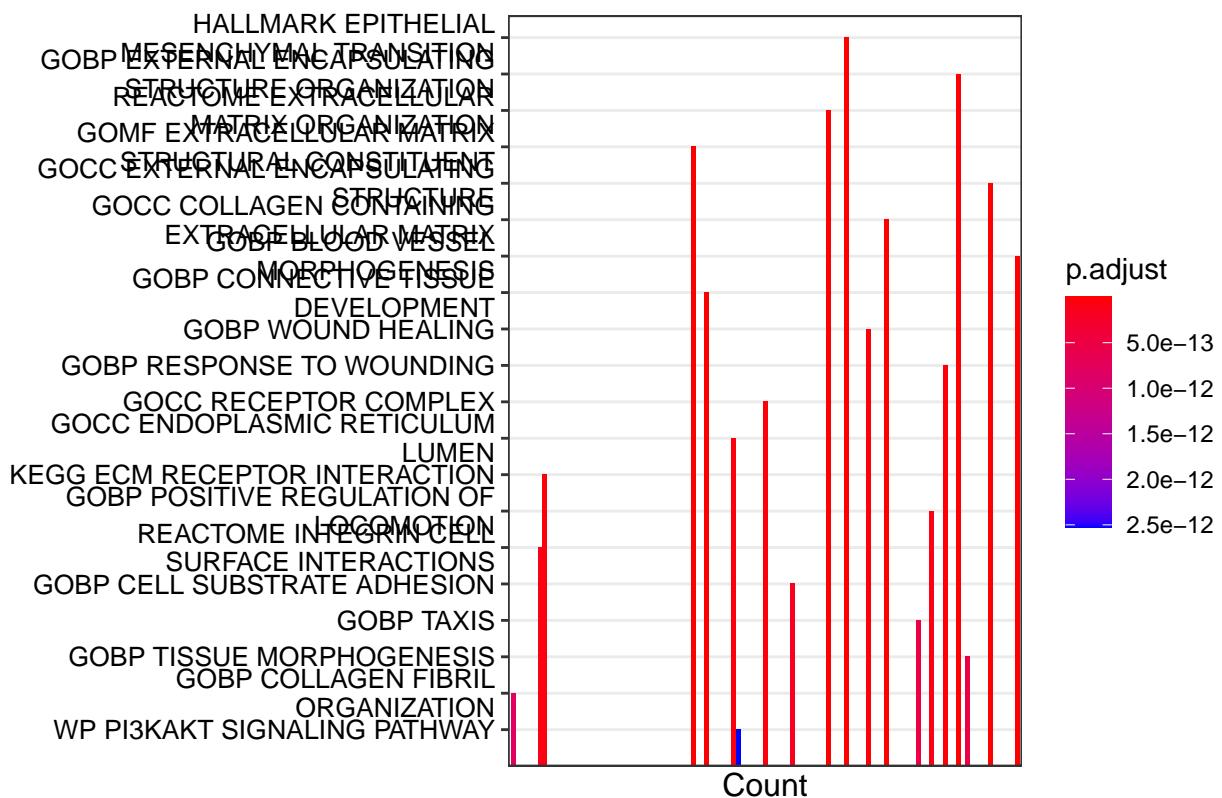
```

```

##                                     GeneRatio   BgRatio
## HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION      105/1751 194/12922
## GOBP_EXTERNAL_ENCAPSULATING_STRUCTURE_ORGANIZATION 130/1751 305/12922
## REACTOME_EXTRACELLULAR_MATRIX_ORGANIZATION       101/1751 228/12922
## GOMF_EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT    71/1751 127/12922
## GOCC_EXTERNAL_ENCAPSULATING_STRUCTURE             137/1751 384/12922
## GOCC_COLLAGEN_CONTAINING_EXTRACELLULAR_MATRIX      114/1751 301/12922
## GOBP_BLOOD_VESSEL_MORPHOGENESIS                   143/1751 484/12922
## GOBP_CONNECTIVE_TISSUE_DEVELOPMENT                 74/1751 193/12922
## GOBP_WOUND_HEALING                                110/1751 363/12922
## GOBP_RESPONSE_TO_WOUNDING                          127/1751 448/12922
##
##                                     pvalue      p.adjust
## HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION      1.796980e-41 1.277653e-37
## GOBP_EXTERNAL_ENCAPSULATING_STRUCTURE_ORGANIZATION 1.138952e-36 4.048975e-33
## REACTOME_EXTRACELLULAR_MATRIX_ORGANIZATION       2.430876e-30 5.761176e-27
## GOMF_EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT 1.459174e-29 2.593681e-26
## GOCC_EXTERNAL_ENCAPSULATING_STRUCTURE             6.436895e-29 9.153264e-26
## GOCC_COLLAGEN_CONTAINING_EXTRACELLULAR_MATRIX      9.038416e-27 1.071052e-23
## GOBP_BLOOD_VESSEL_MORPHOGENESIS                   6.114362e-21 6.210445e-18
## GOBP_CONNECTIVE_TISSUE_DEVELOPMENT                 3.986352e-18 3.542870e-15
## GOBP_WOUND_HEALING                                3.309617e-17 2.614598e-14
## GOBP_RESPONSE_TO_WOUNDING                          4.352932e-17 3.094934e-14
##
##                                     qvalue
## HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION      1.032602e-37
## GOBP_EXTERNAL_ENCAPSULATING_STRUCTURE_ORGANIZATION 3.272389e-33
## REACTOME_EXTRACELLULAR_MATRIX_ORGANIZATION       4.656193e-27
## GOMF_EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT 2.096218e-26
## GOCC_EXTERNAL_ENCAPSULATING_STRUCTURE             7.397686e-26
## GOCC_COLLAGEN_CONTAINING_EXTRACELLULAR_MATRIX      8.656266e-24
## GOBP_BLOOD_VESSEL_MORPHOGENESIS                   5.019294e-18
## GOBP_CONNECTIVE_TISSUE_DEVELOPMENT                 2.863355e-15
## GOBP_WOUND_HEALING                                2.113123e-14
## GOBP_RESPONSE_TO_WOUNDING                          2.501332e-14

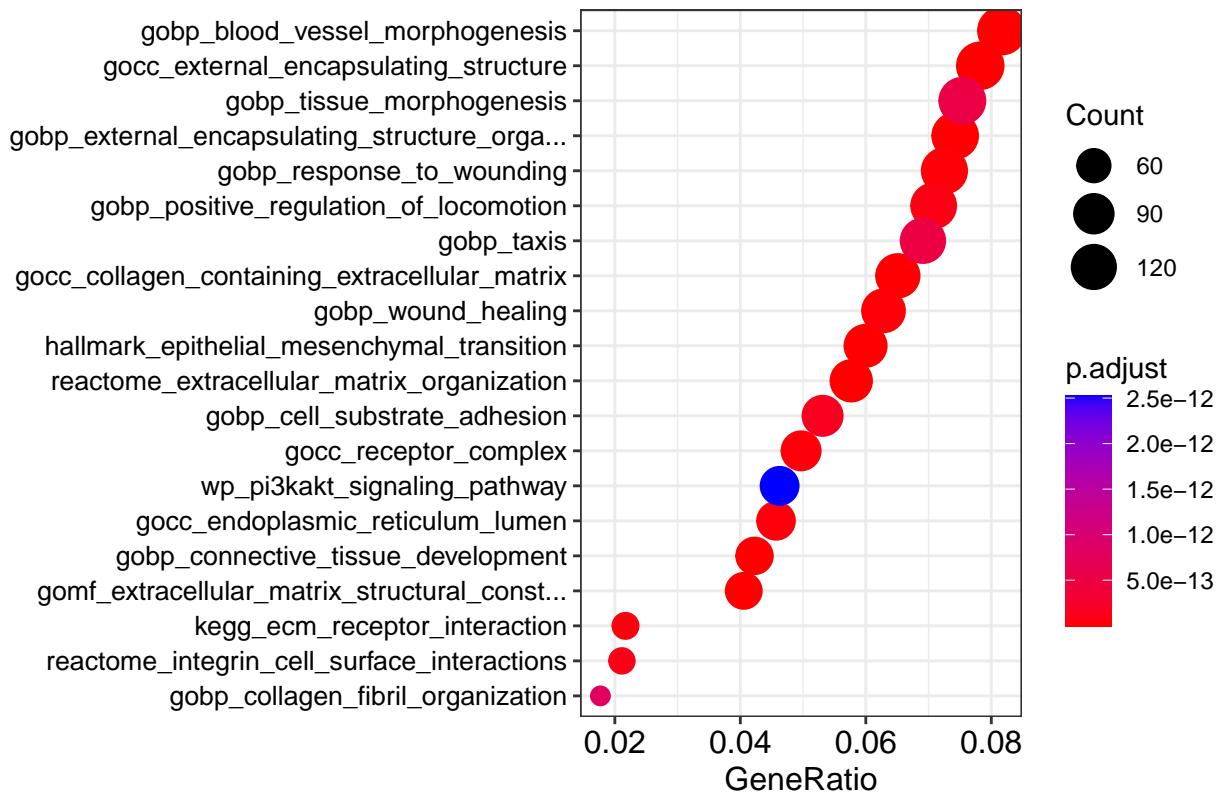
barplot(go_down, showCategory=20)+ scale_x_discrete(label=function(x) tolower(stringr::str_trunc(x,45)))
theme(axis.text.y = element_text(size=10))

```



```
dotplot(go_down, showCategory=20)+ scale_y_discrete(label=function(x) tolower(stringr::str_trunc(x,45)))
theme(axis.text.y = element_text(size=10))
```

```
## Scale for y is already present.
## Adding another scale for y, which will replace the existing scale.
```



```
go_down[grep("MESENCHYMAL", go_down$ID), 3:9]
```

| | GeneRatio |
|---|--------------|
| ## | |
| ## HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION | 105/1751 |
| ## GOBP_MESENCHYMAL_CELL_DIFFERENTIATION | 50/1751 |
| ## WP_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_IN_COLORECTAL_CANCER | 37/1751 |
| ## GOBP_MESENCHYMAL_CELL_PROLIFERATION | 13/1751 |
| ## | BgRatio |
| ## HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION | 194/12922 |
| ## GOBP_MESENCHYMAL_CELL_DIFFERENTIATION | 185/12922 |
| ## WP_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_IN_COLORECTAL_CANCER | 121/12922 |
| ## GOBP_MESENCHYMAL_CELL_PROLIFERATION | 34/12922 |
| ## | pvalue |
| ## HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION | 1.796980e-41 |
| ## GOBP_MESENCHYMAL_CELL_DIFFERENTIATION | 7.951796e-07 |
| ## WP_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_IN_COLORECTAL_CANCER | 8.700126e-07 |
| ## GOBP_MESENCHYMAL_CELL_PROLIFERATION | 2.870155e-04 |
| ## | p.adjust |
| ## HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION | 1.277653e-37 |
| ## GOBP_MESENCHYMAL_CELL_DIFFERENTIATION | 3.385465e-05 |
| ## WP_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_IN_COLORECTAL_CANCER | 3.682018e-05 |
| ## GOBP_MESENCHYMAL_CELL_PROLIFERATION | 4.388560e-03 |
| ## | qvalue |
| ## HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION | 1.032602e-37 |
| ## GOBP_MESENCHYMAL_CELL_DIFFERENTIATION | 2.736140e-05 |

```

## WP_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_IN_COLON_CARCINOMA 2.975814e-05
## GOBP_MESENCHYMAL_CELL_PROLIFERATION 3.546843e-03
##
## HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION NT5E/PTX3/CDH11/WNT5A/THBS2/FSTL1/TPM1
## GOBP_MESENCHYMAL_CELL_DIFFERENTIATION
## WP_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_IN_COLON_CARCINOMA
## GOBP_MESENCHYMAL_CELL_PROLIFERATION
## Count
## HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION 105
## GOBP_MESENCHYMAL_CELL_DIFFERENTIATION 50
## WP_EPITHELIAL_TO_MESENCHYMAL_TRANSITION_IN_COLON_CARCINOMA 37
## GOBP_MESENCHYMAL_CELL_PROLIFERATION 13

go_down[grep("EXTRACELLULAR", go_down$ID), 3:9]

## GeneRatio
## REACTOME_EXTRACELLULAR_MATRIX_ORGANIZATION 101/1751
## GOMF_EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT 71/1751
## GOCC_COLLAGEN_CONTAINING_EXTRACELLULAR_MATRIX 114/1751
## GOMF_EXTRACELLULAR_MATRIX_BINDING 26/1751
## REACTOME_DEGRADATION_OF_THE_EXTRACELLULAR_MATRIX 38/1751
## GOMF_EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT_CONFERRING_TENSILE_STRENGTH 19/1751
## GOMF_EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT_CONFERRING_COMPRESSION_RESISTANCE 11/1751
## WP_EXTRACELLULAR_VESICLEMEDIATED_SIGNALING_IN_RECIPIENT_CELLS 10/1751
## GOBP_EXTRACELLULAR_MATRIX_ASSEMBLY 13/1751
## GOBP_EXTRACELLULAR_REGULATION_OF_SIGNAL_TRANSDUCTION 5/1751
## BgRatio
## REACTOME_EXTRACELLULAR_MATRIX_ORGANIZATION 228/12922
## GOMF_EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT 127/12922
## GOCC_COLLAGEN_CONTAINING_EXTRACELLULAR_MATRIX 301/12922
## GOMF_EXTRACELLULAR_MATRIX_BINDING 49/12922
## REACTOME_DEGRADATION_OF_THE_EXTRACELLULAR_MATRIX 100/12922
## GOMF_EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT_CONFERRING_TENSILE_STRENGTH 33/12922
## GOMF_EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT_CONFERRING_COMPRESSION_RESISTANCE 14/12922
## WP_EXTRACELLULAR_VESICLEMEDIATED_SIGNALING_IN_RECIPIENT_CELLS 25/12922
## GOBP_EXTRACELLULAR_MATRIX_ASSEMBLY 40/12922
## GOBP_EXTRACELLULAR_REGULATION_OF_SIGNAL_TRANSDUCTION 10/12922
## pvalue
## REACTOME_EXTRACELLULAR_MATRIX_ORGANIZATION 2.430876e-30
## GOMF_EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT 1.459174e-29
## GOCC_COLLAGEN_CONTAINING_EXTRACELLULAR_MATRIX 9.038416e-27
## GOMF_EXTRACELLULAR_MATRIX_BINDING 5.646051e-11
## REACTOME_DEGRADATION_OF_THE_EXTRACELLULAR_MATRIX 7.674167e-10
## GOMF_EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT_CONFERRING_TENSILE_STRENGTH 3.596833e-09
## GOMF_EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT_CONFERRING_COMPRESSION_RESISTANCE 6.747927e-08
## WP_EXTRACELLULAR_VESICLEMEDIATED_SIGNALING_IN_RECIPIENT_CELLS 9.550784e-04
## GOBP_EXTRACELLULAR_MATRIX_ASSEMBLY 1.697974e-03
## GOBP_EXTRACELLULAR_REGULATION_OF_SIGNAL_TRANSDUCTION 6.332024e-03
## p.adjust
## REACTOME_EXTRACELLULAR_MATRIX_ORGANIZATION 5.761176e-27
## GOMF_EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT 2.593681e-26
## GOCC_COLLAGEN_CONTAINING_EXTRACELLULAR_MATRIX 1.071052e-23
## GOMF_EXTRACELLULAR_MATRIX_BINDING 7.574231e-09
## REACTOME_DEGRADATION_OF_THE_EXTRACELLULAR_MATRIX 8.024019e-08

```

```
go_down[grep("BLOOD_VESSEL", go_down$ID), 3:9]
```

```
## GOBP_BLOOD_VESSEL_MORPHOGENESIS 143/1751
## GOBP_BLOOD_VESSEL_ENDOTHELIAL_CELL_MIGRATION 30/1751
## GOBP_POSITIVE_REGULATION_OF_BLOOD_VESSEL_ENDOTHELIAL_CELL_MIGRATION 18/1751
## GOBP_VENOUS_BLOOD_VESSEL_DEVELOPMENT 6/1751
## GOBP_BLOOD_VESSEL_ENDOTHELIAL_CELL_PROLIFERATION_INVOLVED_IN_SPROUTING_ANGIOGENESIS 7/1751
## GOBP_BRANCHING_INVOLVED_IN_BLOOD_VESSEL_MORPHOGENESIS 8/1751
##
## GOBP_BLOOD_VESSEL_MORPHOGENESIS 484/12922
## GOBP_BLOOD_VESSEL_ENDOTHELIAL_CELL_MIGRATION 103/12922
## GOBP_POSITIVE_REGULATION_OF_BLOOD_VESSEL_ENDOTHELIAL_CELL_MIGRATION 49/12922
## GOBP_VENOUS_BLOOD_VESSEL_DEVELOPMENT 12/12922
## GOBP_BLOOD_VESSEL_ENDOTHELIAL_CELL_PROLIFERATION_INVOLVED_IN_SPROUTING_ANGIOGENESIS 16/12922
```

```

## GOBP_BRANCHING_INVOLVED_IN_BLOOD_VESSEL_MORPHOGENESIS          21/12922
##                                                               pvalue
## GOBP_BLOOD_VESSEL_MORPHOGENESIS          6.114362e-21
## GOBP_BLOOD_VESSEL_ENDOTHELIAL_CELL_MIGRATION      2.640834e-05
## GOBP_POSITIVE_REGULATION_OF_BLOOD_VESSEL_ENDOTHELIAL_CELL_MIGRATION 3.840855e-05
## GOBP_VENOUS_BLOOD_VESSEL_DEVELOPMENT          2.727905e-03
## GOBP_BLOOD_VESSEL_ENDOTHELIAL_CELL_PROLIFERATION_INVOLVED_IN_SPROUTING_ANGIOGENESIS 3.095192e-03
## GOBP_BRANCHING_INVOLVED_IN_BLOOD_VESSEL_MORPHOGENESIS          4.414016e-03
##                                                               p.adjust
## GOBP_BLOOD_VESSEL_MORPHOGENESIS          6.210445e-18
## GOBP_BLOOD_VESSEL_ENDOTHELIAL_CELL_MIGRATION      6.386506e-04
## GOBP_POSITIVE_REGULATION_OF_BLOOD_VESSEL_ENDOTHELIAL_CELL_MIGRATION 8.575202e-04
## GOBP_VENOUS_BLOOD_VESSEL_DEVELOPMENT          2.603253e-02
## GOBP_BLOOD_VESSEL_ENDOTHELIAL_CELL_PROLIFERATION_INVOLVED_IN_SPROUTING_ANGIOGENESIS 2.836314e-02
## GOBP_BRANCHING_INVOLVED_IN_BLOOD_VESSEL_MORPHOGENESIS          3.731706e-02
##                                                               qvalue
## GOBP_BLOOD_VESSEL_MORPHOGENESIS          5.019294e-18
## GOBP_BLOOD_VESSEL_ENDOTHELIAL_CELL_MIGRATION      5.161587e-04
## GOBP_POSITIVE_REGULATION_OF_BLOOD_VESSEL_ENDOTHELIAL_CELL_MIGRATION 6.930495e-04
## GOBP_VENOUS_BLOOD_VESSEL_DEVELOPMENT          2.103954e-02
## GOBP_BLOOD_VESSEL_ENDOTHELIAL_CELL_PROLIFERATION_INVOLVED_IN_SPROUTING_ANGIOGENESIS 2.292315e-02
## GOBP_BRANCHING_INVOLVED_IN_BLOOD_VESSEL_MORPHOGENESIS          3.015972e-02
##                                                               Count
## GOBP_BLOOD_VESSEL_MORPHOGENESIS          143
## GOBP_BLOOD_VESSEL_ENDOTHELIAL_CELL_MIGRATION      30
## GOBP_POSITIVE_REGULATION_OF_BLOOD_VESSEL_ENDOTHELIAL_CELL_MIGRATION 18
## GOBP_VENOUS_BLOOD_VESSEL_DEVELOPMENT          6
## GOBP_BLOOD_VESSEL_ENDOTHELIAL_CELL_PROLIFERATION_INVOLVED_IN_SPROUTING_ANGIOGENESIS 7
## GOBP_BRANCHING_INVOLVED_IN_BLOOD_VESSEL_MORPHOGENESIS          8

```

```
common[grep("SNAI", common$gene_name),]
```

```

##                               gene length gene_name
## ENSG00000260630 ENSG00000260630    5037 SNAI3-AS1
## ENSG0000019549 ENSG0000019549    3392     SNAI2
## ENSG0000124216 ENSG0000124216    1686     SNAI1
##                               description gene_biotype
## ENSG00000260630           SNAI3 antisense RNA 1      antisense
## ENSG0000019549 snail family transcriptional repressor 2 protein_coding
## ENSG0000124216 snail family transcriptional repressor 1 protein_coding
##             logFC AveExpr      t   P.Value   adj.P.Val
## ENSG00000260630  2.944965 3.602903  5.903158 0.0000573521 0.0008821992
## ENSG0000019549 -1.643245 5.743067 -4.517421 0.0006130985 0.0040144369
## ENSG0000124216  1.570253 1.933057  2.346155 0.0359047304 0.0756062635
##             B
## ENSG00000260630  2.0535178
## ENSG0000019549 -0.4128298

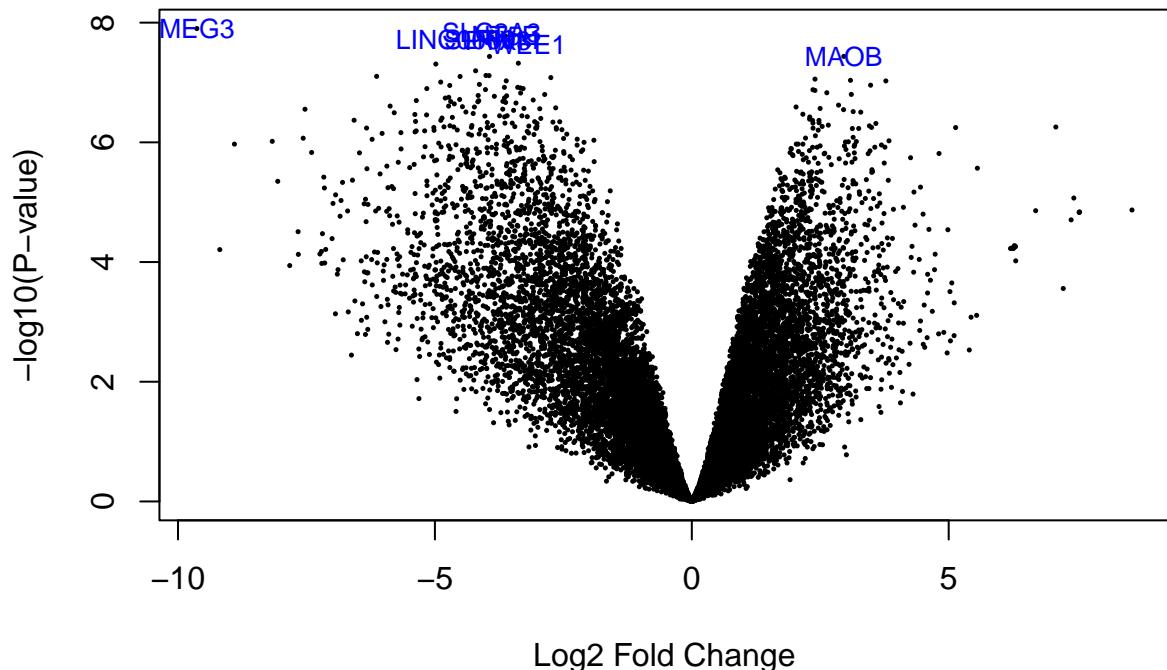
```

```
## ENSG00000124216 -4.1715846
```

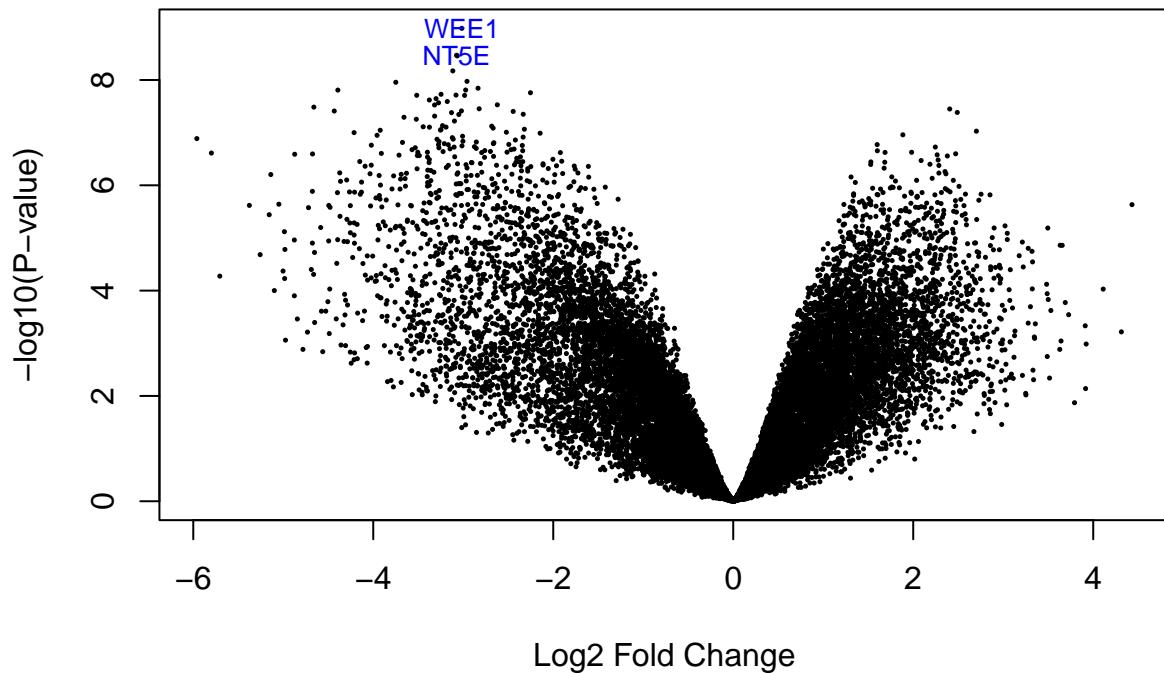
Volcanoplots

(-logfc is more expression in bulk adipocytes)

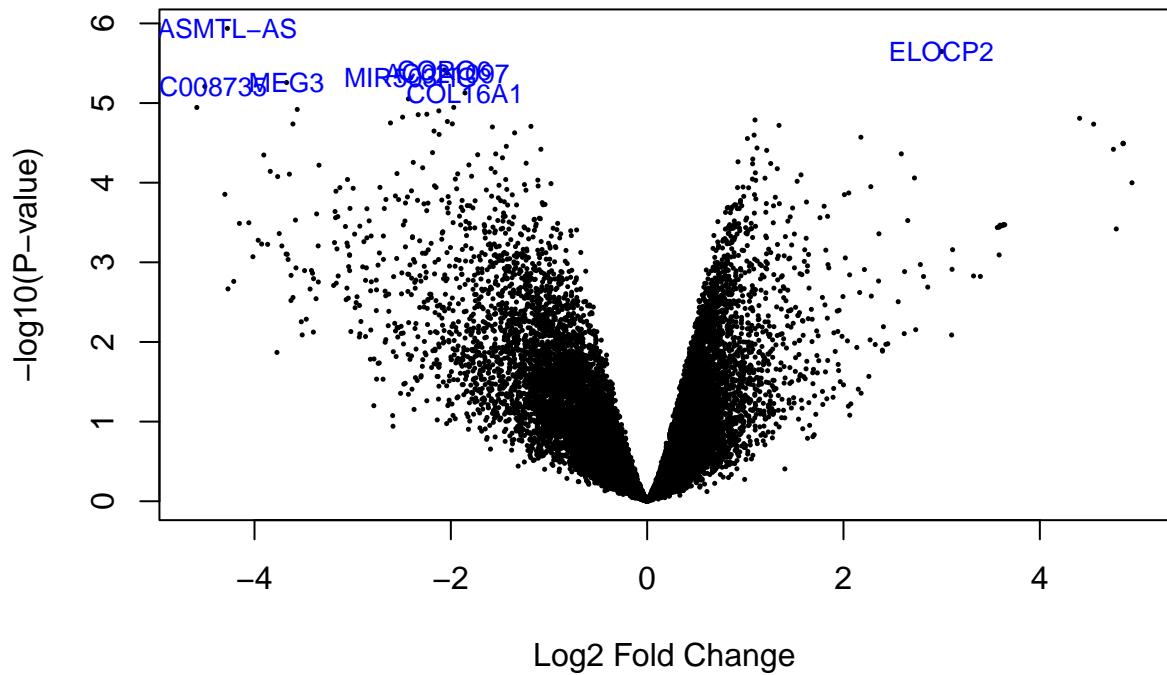
```
volcanoplot(simple_fit, coef = "bulk.v.float", highlight=8, names = simple_fit$genes$gene_name)
```



```
volcanoplot(simple_fit, coef = "D1.bulk.v.float", highlight=2, names = simple_fit$genes$gene_name)
```



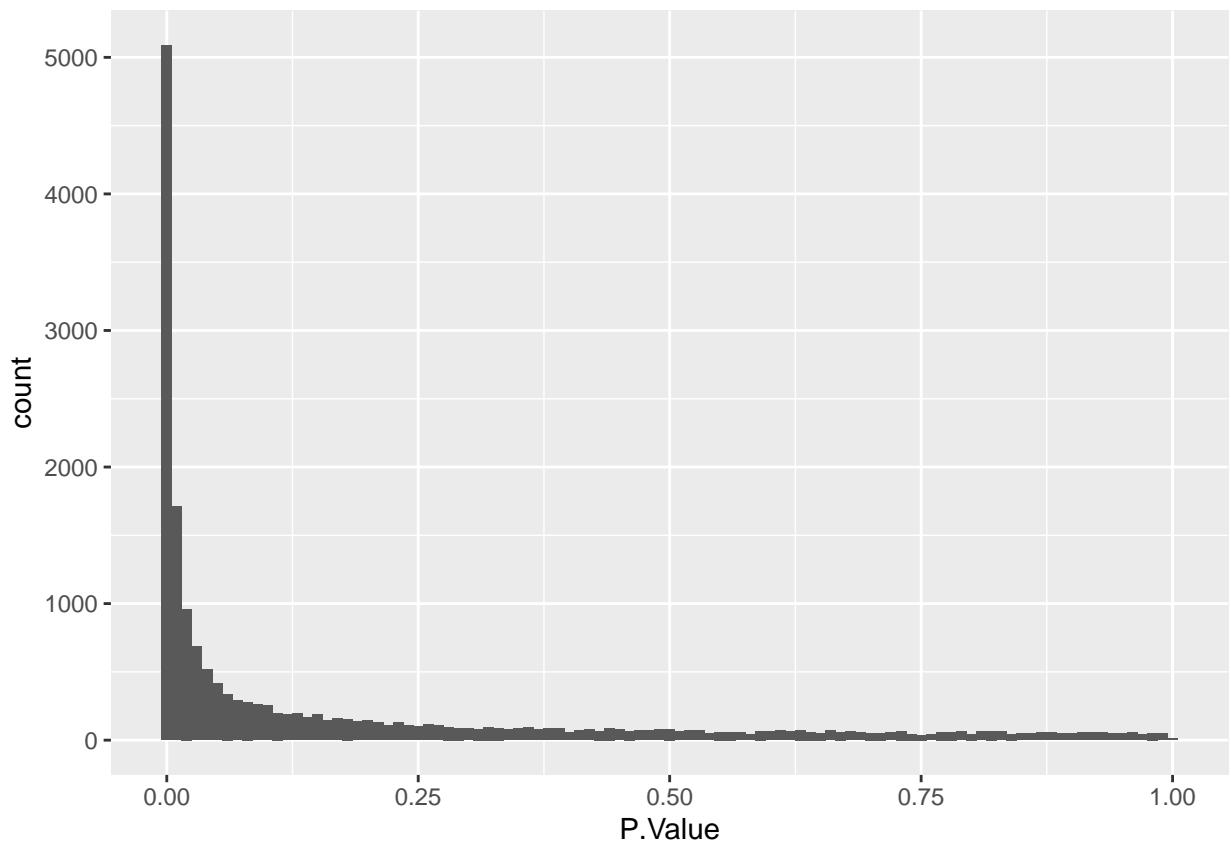
```
volcanoplot(simple_fit, coef = "D2.bulk.v.float", highlight=8, names = simple_fit$genes$gene_name)
```



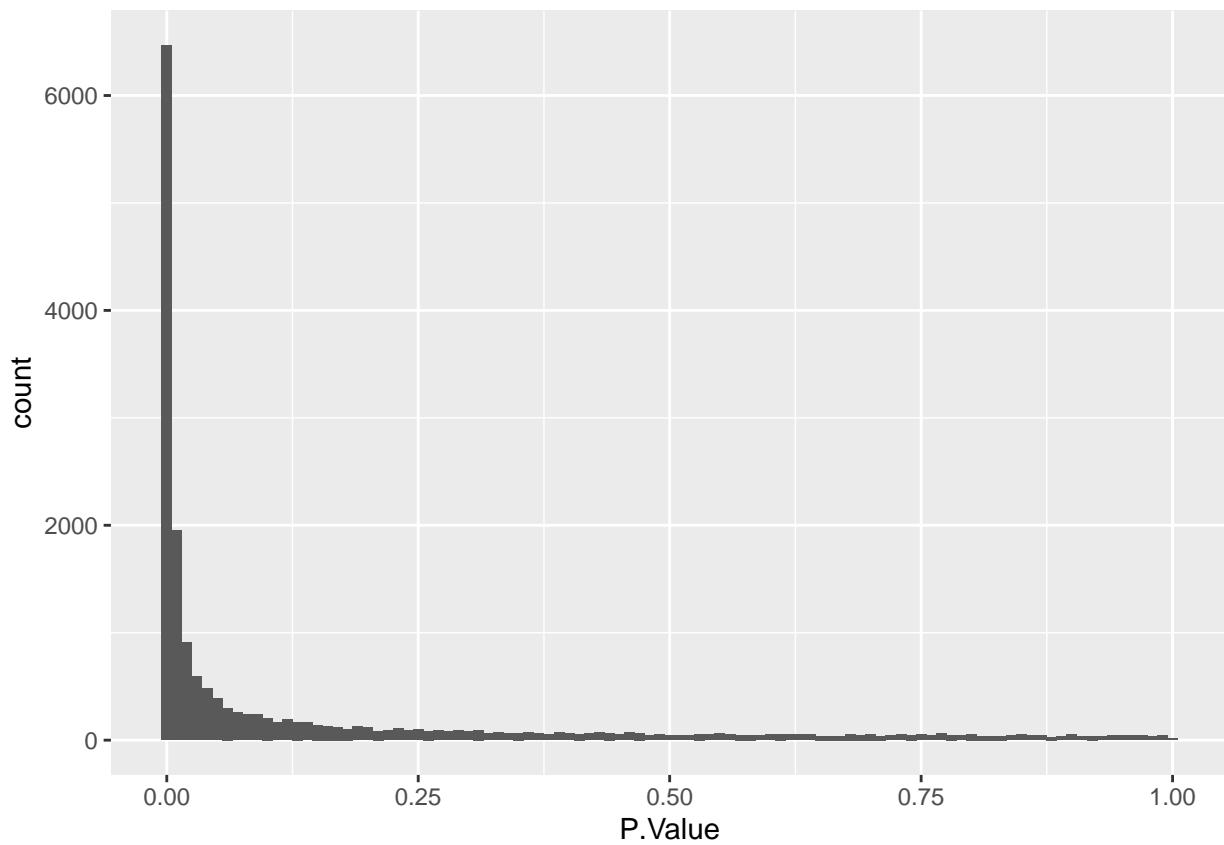
Check p-value distributions

Very strong effect for d1 and combined. d2 looks robust as well.

```
ggplot(common) + geom_histogram(aes(x=P.Value), binwidth=0.01)
```



```
ggplot(d1) + geom_histogram(aes(x=P.Value), binwidth=0.01)
```



```
ggplot(d2) + geom_histogram(aes(x=P.Value), binwidth=0.01)
```

