Contents

1 Setup

This entire analysis is carried out in R using a variety of packages available at the Comprehensive R Archive Network (CRAN) (1) and at the Bioconductor project (2). Among the latter bundle of packages, Biobase provides a function called `cache()` which we will employ all throughout the entire analysis in order to load pre-computed results when we do not want to calculate them again. For this purpose we will create a directory called `results` and three subdirectories (`QA_dir`, `DE_dir`, `FA_dir` and `FIG_dir`). Pre-computed objects and figures will be stored in the corresponding subdirectory and prefixed by names `QA_`, `DE_` or `FA_`. In order to re-compute everything it suffices then to remove the directories with names under `results` or indiviual files from calculations we want to re-compute.

Our first step in our analysis is to set the directories into the corresponding variables, and create them, if do not exist.

```
> cacheDir_QA <- "results/QA_dir"
> cachePrefix_QA <- "QA_"
> cacheDir_DE <- "results/DE_dir"
> cachePrefix_DE <- "DE_"
> cacheDir_FE <- "results/FE_dir"
> cachePrefix_FE <- "FE_"
> cacheDir_NE <- "results/NE_dir"
> cachePrefix_NE <- "NE_"
> cacheDir_FIG <- "results/FIG_dir"
> cacheDir_TAB <- "results/FIG_dir"
> if (!file.exists("results"))
    dir.create("results")
> if (!file.exists(cacheDir_QA))
    dir.create(cacheDir_QA)
> if (!file.exists(cacheDir_DE))
    dir.create(cacheDir_DE)
> if (!file.exists(cacheDir_FE))
    dir.create(cacheDir_FE)
> if (!file.exists(cacheDir_NE))
    dir.create(cacheDir_NE)
> if (!file.exists(cacheDir_FIG))
    dir.create(cacheDir_FIG)
```

2 Quality assessment and pre-processing

We start pre-processing the raw microarray data through the following steps: import gene expression and phenotypic data, batch effect identification, surrogate variable identification and normalization. After these pre-processing steps we perform a differential expression analysis followed by a functional enrichment analysis.

2.1 Importing gene expression data

We import the publicly available raw expression data from the Cohen *et al.* (2007) study deposited at Gene Expression Omnibus (National Center for Biotechnology Information, Bethesda, MD) (3) with accession number GSE8586.

The raw expression data were obtained with an Affymetrix HG-133plus2 chip, and therefore, it is stored in the so-called CEL format files. We use the function `Readaffy()` from the affy package to read CEL files into an *AffyBatch* object.

```
> cache(affyCELonly <- ReadAffy(celfile.path=file.path(getwd(), "celFiles")),
        dir=cacheDir_QA, prefix=cachePrefix_QA)
> sampleNames(affyCELonly) <- sub(".CEL", "", sampleNames(affyCELonly))
```

2.2 Importing phenotypic data

We start by fetching part of the phenotpypic data from the so-called SOFT file available at GEO under accesion number `GSE8586`, using the the function `getGEO()` from the GEOquery package. This function downloads the corresponding data from the internet and stores the them into the file `phenoData/phenoData.txt`.

```
> cache(infoGEO <- getGEO("GSE8586")[[1]],
        dir=cacheDir_QA, prefix=cachePrefix_QA)
> write.table(pData(infoGEO), file=file.path("phenoData", "phenoData.txt"),
              sep="\t", quote=FALSE)
```

Using the function `read.table()` we read into R the previously downloaded data creating and *data.frame* object called `GSE8586.pData`:

```
> GSE8586.pData <- read.table(file=file.path("phenoData", "phenoData.txt"),
                              sep="\t", header=TRUE, quote="", row.names=1,
                              stringsAsFactors=FALSE)
> head(GSE8586.pData, n=3)
                             title geo_accession                status
GSM213072 umbilical cord_nobpd: 1A     GSM213072 Public on Jul 26 2007
GSM213073   umbilical cord_bpd: 1C     GSM213073 Public on Jul 26 2007
GSM213074   umbilical cord_bpd: 1D     GSM213074 Public on Jul 26 2007
          submission_date last_update_date type channel_count
GSM213072     Jul 25 2007      Aug 14 2011  RNA             1
GSM213073     Jul 25 2007      Aug 14 2011  RNA             1
GSM213074     Jul 25 2007      Aug 14 2011  RNA             1
                                      source_name_ch1 organism_ch1
GSM213072 umbilical cord tissue; high gestational age Homo sapiens
GSM213073 umbilical cord tissue; high gestational age Homo sapiens
GSM213074  umbilical cord tissue; low gestational age Homo sapiens
          characteristics_ch1 molecule_ch1
GSM213072       27 weeks male    total RNA
GSM213073     27 weeks female    total RNA
GSM213074       24 weeks male    total RNA

GSM213072 initial step used Trizol extraction of RNA, following manufacturer's recommended protoc
GSM213073 initial step used Trizol extraction of RNA, following manufacturer's recommended protoc
GSM213074 initial step used Trizol extraction of RNA, following manufacturer's recommended protoc
          label_ch1
GSM213072    biotin
GSM213073    biotin
GSM213074    biotin

GSM213072 GeneChipÂ■ One-Cycle Target Labeling and Control Reagents according to the manufacturer
GSM213073 GeneChipÂ■ One-Cycle Target Labeling and Control Reagents according to the manufacturer
GSM213074 GeneChipÂ■ One-Cycle Target Labeling and Control Reagents according to the manufacturer
          taxid_ch1
GSM213072      9606
GSM213073      9606
GSM213074      9606

GSM213072 Affymetrix Hybridization Oven 640 and GeneChipÂ■ Fluidics Station 450 according to the
GSM213073 Affymetrix Hybridization Oven 640 and GeneChipÂ■ Fluidics Station 450 according to the
GSM213074 Affymetrix Hybridization Oven 640 and GeneChipÂ■ Fluidics Station 450 according to the
                                                                                    scan_protocol
GSM213072 Affymetrix GeneChipÂ■ Scanner 3000 according to the manufacturer's recommended protocol
GSM213073 Affymetrix GeneChipÂ■ Scanner 3000 according to the manufacturer's recommended protocol
```

```
GSM213074 Affymetrix GeneChipÂ■ Scanner 3000 according to the manufacturer's recommended protocol
                   description       data_processing platform_id
GSM213072 high gestational age MAS_5.0 scaled to 1500      GPL570
GSM213073 high gestational age MAS_5.0 scaled to 1500      GPL570
GSM213074  low gestational age MAS_5.0 scaled to 1500      GPL570
              contact_name                        contact_email contact_phone
GSM213072 Jennifer,N,Cohen jennifer.cohen@childrens.harvard.edu    6177328263
GSM213073 Jennifer,N,Cohen jennifer.cohen@childrens.harvard.edu    6177328263
GSM213074 Jennifer,N,Cohen jennifer.cohen@childrens.harvard.edu    6177328263
          contact_department            contact_institute   contact_address
GSM213072   Newborn Medicine Brigham and Women's Hospital 75 Francis Street
GSM213073   Newborn Medicine Brigham and Women's Hospital 75 Francis Street
GSM213074   Newborn Medicine Brigham and Women's Hospital 75 Francis Street
          contact_city contact_state contact_zip.postal_code contact_country
GSM213072       Boston            MA                    2115             USA
GSM213073       Boston            MA                    2115             USA
GSM213074       Boston            MA                    2115             USA
                                                                               supplem
GSM213072 ftp://ftp.ncbi.nlm.nih.gov/pub/geo/DATA/supplementary/samples/GSM213nnn/GSM213072/GSM21
GSM213073 ftp://ftp.ncbi.nlm.nih.gov/pub/geo/DATA/supplementary/samples/GSM213nnn/GSM213073/GSM21
GSM213074 ftp://ftp.ncbi.nlm.nih.gov/pub/geo/DATA/supplementary/samples/GSM213nnn/GSM213074/GSM21
          data_row_count
GSM213072          54675
GSM213073          54675
GSM213074          54675
> GSE8586.pData$title
 [1] "umbilical cord_nobpd: 1A"    "umbilical cord_bpd: 1C"
 [3] "umbilical cord_bpd: 1D"      "umbilical cord_nobpd: 1E"
 [5] "umbilical cord_nobpd: 2A"    "umbilical cord_bpd: 2B"
 [7] "umbilical cord_nobpd: 2C"    "umbilical cord_nobpd: 2D"
 [9] "umbilical cord_nobpd: 2E"    "umbilical cord_nobpd: 2G"
[11] "umbilical cord_nobpd: 3A"    "umbilical cord_bpd: 3B"
[13] "umbilical cord_nobpd: 3C"    "umbilical cord_nobpd: 3D"
[15] "umbilical cord_bpd: 3E"      "umbilical cord_nobpd: 3F"
[17] "umbilical cord_nobpd: 3G"    "umbilical cord_nobpd: 4A"
[19] "umbilical cord_nobpd: 4B"    "umbilical cord_nobpd: 4C"
[21] "umbilical cord_nobpd: 4D"    "umbilical cord_nobpd: 5A"
[23] "umbilical cord_bpd: 5B"      "umbilical cord_bpd: 5C"
[25] "umbilical cord_nobpd: 5D"    "umbilical cord_nobpd: 5F"
[27] "umbilical cord_bpd: 5H"      "umbilical cord_nobpd: 6A"
[29] "umbilical cord_nobpd: 6D"    "umbilical cord_nobpd: 6E/2F"
[31] "umbilical cord_bpd: 6F"      "umbilical cord_bpd: 6H"
[33] "umbilical cord_bpd: 7C"      "umbilical cord_nobpd: 7D"
[35] "umbilical cord_bpd: 7F"      "umbilical cord_bpd: 7H"
[37] "umbilical cord_bpd: 8A"      "umbilical cord_nobpd: 8D"
[39] "umbilical cord_nobpd: 8F"    "umbilical cord_nobpd: 9A"
[41] "umbilical cord_bpd: 9C"      "umbilical cord_nobpd: 9D"
[43] "umbilical cord_bpd: 9F"      "umbilical cord_bpd: 9I"
[45] "umbilical cord_bpd: 2I"      "umbilical cord_nobpd: 3I"
[47] "umbilical cord_nobpd: 4F"    "umbilical cord_bpd: 4I"
[49] "umbilical cord_nobpd: 5I"    "umbilical cord_nobpd: 7I"
[51] "umbilical cord_nobpd: 8B"    "umbilical cord_nobpd: 8C"
[53] "umbilical cord_bpd: 8I"      "umbilical cord_nobpd: 9B"
```

Some of the variables in `GSE8586.pData` are irrelevant or require a transformation for the purpose of our investigation. We update the `GSE8586.pData` object by modifying or creating new variables that will be incorporated in following steps. Explanation of the variables is found later in the analysis. We check first that all sample identifiers between the phenotypic data and the expression data match.

```
> stopifnot(all(rownames(GSE8586.pData) == sampleNames(affyCELonly))) ## QC
```

We proceed to code the phenotypic variables in three steps, specifying their role as exogenous, endogenous or response:

- Explanatory/exogenous/context variables:

```
> GSE8586.pData$LinkId <- sub(".+ ", "", as.character(GSE8586.pData$title))
> GSE8586.pData$Batch <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Bw <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Elbw <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Ga <- sub("([0-9]+).+", "\\1",
                          as.character(GSE8586.pData$characteristics_ch1))
> GSE8586.pData$Ga <- as.numeric(GSE8586.pData$Ga)
> GSE8586.pData$GaGE27weeks <- sub("(high|low).+", "\\1",
                                   as.character(GSE8586.pData$description))
> GSE8586.pData$GaGE27weeks <- gsub("high", "yes", GSE8586.pData$GaGE27weeks)
> GSE8586.pData$GaGE27weeks <- gsub("low", "no", GSE8586.pData$GaGE27weeks)
> GSE8586.pData$Sex <- sub(".+ (male|female)", "\\1",
                          as.character(GSE8586.pData$characteristics_ch1))
> GSE8586.pData$Male <- rep(NA,dim(GSE8586.pData)[1])
> GSE8586.pData$Nonwhite <- rep(NA,dim(GSE8586.pData)[1])
```

- Intermediate/endogenous variables:

```
> GSE8586.pData$Ptl <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Pih <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Cs <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Acs <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Plate<- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Cord<- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Chorio<- rep(NA, dim(GSE8586.pData)[1])
```

- Response variables:

```
> GSE8586.pData$FIR <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Bpd <- sub("cord_", "",
                          sub(".+(cord_bpd|cord_nobpd).+", "\\1",
                              as.character(GSE8586.pData$title)))
> GSE8586.pData$DaysVent <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$VentGT7days <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$DaysO2 <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$O2GE28days  <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Nec <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$NecEnt<- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Rop <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Pda <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Sepsis <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Funisitis<- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Ivh <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Wmd <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Cp <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Mdi <- rep(NA, dim(GSE8586.pData)[1])
> GSE8586.pData$Pdi <- rep(NA, dim(GSE8586.pData)[1])
```

The phenotypic information stored in GEO did not contain all the clinical data from the ELGAN cohort summarized in the paper by Cohen *et al.* (2007). We obtained additional phenotypic data directly from authors in the form of an Excel file stored in `phenoData/phenotype_data_sheet_9.20.06.xls`. The following lines of code read these phenotypic data and incorporate them into a *data.frame* called `pDataCohen` filtering out those records that did not match individuals in the microarray data.

```
> fname <- file.path(getwd(), "phenoData",
                     "phenotype_data_sheet_9.20.06.xls")
> cache(pDataCohen <- read.xlsx(fname, 1),
        dir=cacheDir_QA, prefix=cachePrefix_QA)
> head(pDataCohen)
       id bpd   ga  bw sex acs iopd daysvent dayso2 fsvneut uc fv linkid number
1 2302221   0 23.4 532   0   1    1        3      3      NA NA NA     3i     24
2 1200692   0 25.0 720   1   1    1       38     67       0  0  0     6d     41
3 1200691   0 25.0 790   1   1    1       36     72       0  0  0     5d     36
4 1200711   0 25.4 713   1   2    2       33     70       0 NA NA     3d     20
5 1301551   0 26.0 840   1   2    6        1     60       1  1  1     3a     17
6 1301631   0 26.0 930   1   2    5       32     67       0  0  0     4f     30
  outcome hi.low
1      no    low
2      no    low
3      no    low
4      no    low
5      no    low
6      no    low
> dim(pDataCohen)
[1] 59 16
> colnames(pDataCohen)
 [1] "id"       "bpd"      "ga"       "bw"       "sex"      "acs"
 [7] "iopd"     "daysvent" "dayso2"   "fsvneut"  "uc"       "fv"
[13] "linkid"   "number"   "outcome"  "hi.low"
> id <- pDataCohen$linkid
> id <- as.character(as.vector(id))
> id <- toupper(id)
> pDataCohen$linkid <- id
> pDataCohen <- pDataCohen[, c("linkid", "bw", "daysvent", "dayso2")]
> dim(pDataCohen)
[1] 59  4
> head(pDataCohen)
  linkid  bw daysvent dayso2
1     3I 532        3      3
2     6D 720       38     67
3     5D 790       36     72
4     3D 713       33     70
5     3A 840        1     60
6     4F 930       32     67
> tail(pDataCohen)
   linkid   bw daysvent dayso2
54     3E 1080       15     58
55     1G  630       NA     NA
56     7B  710       NA     NA
57     8E  760       NA     NA
58     7A  690       NA     NA
59   <NA>   NA       NA     NA
```

```
> pDataCohen <- pDataCohen[-59,] ## remove entry 59 whose values are all NAs
> dim(pDataCohen)
[1] 58  4
> pDataCohen$linkid
 [1] "3I"       "6D"       "5D"       "3D"       "3A"       "4F"      
 [7] "4D"       "3G"       "9A"       "4C"       "2C"       "5I"      
[13] "8B"       "5F"       "1A/4E/5E" "6E/2F/1B" "7D"       "7I"      
[19] "3F"       "9B"       "5A"       "6A"       "8C"       "8F"      
[25] "2G"       "2E"       "1E"       "3C"       "2D"       "4A"      
[31] "4B"       "9D"       "8D"       "2A"       "4I"       "9I"      
[37] "6F"       "1D"       "3B"       "7C"       "2B"       "8A"      
[43] "5C"       "9F"       "7H"       "6H"       "5B"       "9C"      
[49] "2I"       "5H"       "8I"       "1C"       "7F"       "3E"      
[55] "1G"       "7B"       "8E"       "7A"      
> GSE8586.pData$LinkId
 [1] "1A"    "1C"    "1D"    "1E"    "2A"    "2B"    "2C"    "2D"    "2E"   
[10] "2G"    "3A"    "3B"    "3C"    "3D"    "3E"    "3F"    "3G"    "4A"   
[19] "4B"    "4C"    "4D"    "5A"    "5B"    "5C"    "5D"    "5F"    "5H"   
[28] "6A"    "6D"    "6E/2F" "6F"    "6H"    "7C"    "7D"    "7F"    "7H"   
[37] "8A"    "8D"    "8F"    "9A"    "9C"    "9D"    "9F"    "9I"    "2I"   
[46] "3I"    "4F"    "4I"    "5I"    "7I"    "8B"    "8C"    "8I"    "9B"   
> idGSE8586 <- pDataCohen$linkid %in% GSE8586.pData$LinkId
> idFalse <- which(!idGSE8586)
> idFalse
[1] 15 16 55 56 57 58
> pDataCohen[idFalse, ] ## check entries that do not match to the microarray data
     linkid  bw daysvent dayso2
15 1A/4E/5E 770       17     46
16 6E/2F/1B 770        1     29
55       1G 630       NA     NA
56       7B 710       NA     NA
57       8E 760       NA     NA
58       7A 690       NA     NA
> # Change the linkid of row 15 and 16
> pDataCohen[15,]$linkid <- c("1A") ## select one of the multiple IDs to match the microarray dat
> pDataCohen[16,]$linkid <- c("6E/2F") ## select one of the multiple IDs to match the microarray
> idGSE8586 <-pDataCohen$linkid %in% GSE8586.pData$LinkId
> idTrue <- which(idGSE8586)
> pDataCohen <- pDataCohen[idTrue, ]  ## select phenotypic data matching the microarray data
> dim(pDataCohen)
[1] 54  4
```

Further clinical data from the paper by Cohen *et al.* (2007) was missing from all previous files. We applied to the ELGAN data access committee for access to them and stored them into a plain text file in `phenoData/phenoData2.txt`. These phenotypic data, and the previous one stored in the *data.frame* `pDataCohen` are incorporated into the `GSE8586.pData` object with the following lines of code:

```
> phenoData2 <- read.table(file=file.path("phenoData", "phenoData2.txt"),
                           sep="\t", header=TRUE, quote="", row.names=NULL,
                           stringsAsFactors=FALSE)
> pDataSampleNameLinkId <- cbind(rownames(GSE8586.pData), GSE8586.pData$LinkId)
> phenoData2SampleNameLinkId <- cbind(phenoData2$gsm, phenoData2$linkid)
> ## QC that every GEO sample has the same 'LinkId' assignment in both data sets
> stopifnot(identical(phenoData2SampleNameLinkId[order(phenoData2SampleNameLinkId[, 1]), ],
                      pDataSampleNameLinkId[order(pDataSampleNameLinkId[, 1]), ]))
```

```
> # Assign to corresponding columns of GSE8586.pData corresponding columns of phenoData2
> for (phenovar in c("Ivh", "Nec", "Rop", "Pda", "Sepsis", "Ptl", "Pih",
                     "Cs", "Acs", "Nonwhite", "Plate", "Cord")) {
    mt <- match(GSE8586.pData$LinkId, phenoData2$linkid)
    GSE8586.pData[[phenovar]] <- phenoData2[[tolower(phenovar)]][mt]
  }
> # Assign to corresponding columns of GSE8586.pData corresponding columns pDataCohen
> for (phenovar in c("Bw", "DaysVent", "DaysO2")) {
    mt <- match(GSE8586.pData$LinkId, pDataCohen$linkid)
    GSE8586.pData[[phenovar]] <- pDataCohen[[tolower(phenovar)]][mt]
  }
```

Some of the variables define missing values with the string `unkn`. We replace these strings by the R missing value `NA`:

```
> whunkn <- which(apply(GSE8586.pData, 2, function(x) any(x == "unkn", na.rm=FALSE)))
> whunkn
 Plate   Cord    Rop    Pda Sepsis    Ivh 
    46     47     57     58     59     61 
> for(i in whunkn) print(table(GSE8586.pData[[i]]))

  no unkn  yes 
  36    2   16 

  no unkn  yes 
  29    7   18 

  no unkn  yes 
  27    1   26 

  no unkn  yes 
  14    2   38 

  no unkn  yes 
  48    1    5 

  no unkn  yes 
  41    1   12 
> for (i in whunkn)
    GSE8586.pData[[i]] <- ifelse(GSE8586.pData[[i]] == "unkn", NA, GSE8586.pData[[i]])
> for(i in whunkn) print(table(GSE8586.pData[[i]], useNA="always"))

  no  yes <NA> 
  36   16    2 

  no  yes <NA> 
  29   18    7 

  no  yes <NA> 
  27   26    1 

  no  yes <NA> 
  14   38    2 

  no  yes <NA> 
  48    5    1 
```

```
no  yes <NA>
41   12    1
```

We define the phenotypic variables `Bpd`, `Male`, `Chorio` and `Funisitis` in the `GSE8586.pData` object as binary varibles with levels `yes` and `no`:

```
> bpd <- which(GSE8586.pData$Bpd == "bpd")
> nobpd <- which(GSE8586.pData$Bpd == "nobpd")
> GSE8586.pData$Bpd[bpd] <- "yes"
> GSE8586.pData$Bpd[nobpd] <- "no"
> male <- which(GSE8586.pData$Sex == "male")
> female <- which(GSE8586.pData$Sex == "female")
> GSE8586.pData$Male[male]<- "yes"
> GSE8586.pData$Male[female] <- "no"
> plate <- which(GSE8586.pData$Plate == "yes")
> cord <- which(GSE8586.pData$Cord == "yes")
> plateNo <- which(GSE8586.pData$Plate == "no")
> cordNo <- which(GSE8586.pData$Cord == "no")
> chorio <- intersect(plate,cordNo)
> fun <- intersect(plateNo,cord)
> choFun <- union(plate, cord)
> chorioNo <- union(plateNo, cord)
> funNo <- union(cordNo, plate)
> choFunNo <- intersect(cordNo, plateNo)
> GSE8586.pData$Chorio[chorio] <- "yes"
> GSE8586.pData$Chorio[chorioNo] <- "no"
> GSE8586.pData$Funisitis[fun] <- "yes"
> GSE8586.pData$Funisitis[funNo] <- "no"
```

The fetal inflammatory response (FIR) status variable is defined by the presence of neutrophils in either the chorionic plate or in the umbilical chord.

```
> GSE8586.pData$FIR[choFun] <- "yes"
> GSE8586.pData$FIR[choFunNo] <- "no"
```

We create now the variable corresponding to necrotizing enterocolitis.

```
> stopifnot(all(!is.na(GSE8586.pData$Nec))) ## QC
> necNo <- which(GSE8586.pData$Nec== "no")
> nec <- which(!GSE8586.pData$Nec== "no")
> GSE8586.pData$NecEnt[nec] <- "yes"
> GSE8586.pData$NecEnt[necNo] = "no"
```

We store the days of mechanical ventilation and supplemental oxygen, as well as binarized versions of them using 7 and 27 days as cutoffs, respectively.

```
> GSE8586.pData$DaysVent <- as.numeric(GSE8586.pData$DaysVent)
> ventmore7days <- which(GSE8586.pData$DaysVent > 7)
> ventlLessEqualThan7days <- which(GSE8586.pData$DaysVent <= 7)
> GSE8586.pData$VentGT7days[ventmore7days] <- "yes"
> GSE8586.pData$VentGT7days[ventlLessEqualThan7days] <- "no"
> table(GSE8586.pData$VentGT7days, useNA="ifany")
  no  yes <NA>
  30   23    1
> sapply(split(GSE8586.pData$DaysVent, GSE8586.pData$VentGT7days),
         function(x) do.call("names<-", list(range(x), c("min", "max"))))
```

```
    no yes
min  0  10
max  7  77
> GSE8586.pData$DaysO2 <- as.numeric(GSE8586.pData$DaysO2)
> oxygen28days <- which(GSE8586.pData$DaysO2 > 27)
> oxygenLess28days <-  which(GSE8586.pData$DaysO2 < 27)
> GSE8586.pData$O2GE28days[oxygen28days] <- "yes"
> GSE8586.pData$O2GE28days[oxygenLess28days] <- "no"
> table(GSE8586.pData$O2GE28days, useNA="ifany")
  no  yes <NA>
  14   39    1
> sapply(split(GSE8586.pData$DaysO2, GSE8586.pData$O2GE28days),
         function(x) do.call("names<-", list(range(x), c("min", "max"))))
    no yes
min  2  28
max 26  91
```

We store the body weight and a binarized version for indicating extremely low body weight (ELBW) with 1000 grams as cutoff value.

```
> GSE8586.pData$Bw <- as.numeric(GSE8586.pData$Bw)
> bwLess1000 <- which(GSE8586.pData$Bw < 1000)
> bwmoreEqual1000 <- which(GSE8586.pData$Bw >= 1000)
> GSE8586.pData$Elbw[bwLess1000] <- "yes"
> GSE8586.pData$Elbw[bwmoreEqual1000] <- "no"
> table(GSE8586.pData$Elbw, useNA="ifany")
 no yes
 19  35
> sapply(split(GSE8586.pData$Bw, GSE8586.pData$Elbw),
         function(x) do.call("names<-", list(range(x), c("min", "max"))))
      no yes
min 1015 460
max 1360 980
```

We also import through the ELGAN data access committee clinical information about the neurological condition of the infants.

```
> cache(FIRstatusNeuro <- read.xlsx(file.path("phenoData", "FIRstatusNeuro.xlsx"),
                                    sheetIndex=1, stringsAsFactors=FALSE),
        dir=cacheDir_QA, prefix=cachePrefix_QA)
> stopifnot(all(FIRstatusNeuro$LinkId %in% GSE8586.pData$LinkId))
> mt <- match(FIRstatusNeuro$LinkId, GSE8586.pData$LinkId)
> stopifnot(all(!is.na(mt)))
> stopifnot(FIRstatusNeuro$GA == GSE8586.pData$Ga[mt])
> stopifnot(FIRstatusNeuro$Sex == GSE8586.pData$Sex[mt])
> GSE8586.pData$Wmd[mt] <- c("no", "yes")[FIRstatusNeuro$wmd+1]
> GSE8586.pData$Cp[mt] <- c("no", "yes")[FIRstatusNeuro$cp+1]
> GSE8586.pData$Mdi[mt] <- c("no", "yes")[FIRstatusNeuro$mdi+1]
> GSE8586.pData$Pdi[mt] <- c("no", "yes")[FIRstatusNeuro$pdi+1]
```

Finally, select the phenotypic variables that are relevant for the analysis.

```
> explanatoryVariables <- c("LinkId", "Batch", "Bw", "Elbw", "Ga", "GaGE27weeks",
                            "Sex", "Male", "Nonwhite")
> intermediateVariables <- c("Ptl", "Pih", "Cs", "Acs", "Plate", "Cord", "Chorio")
> responseVariables <- c("FIR", "Bpd", "DaysVent", "VentGT7days", "DaysO2",
```

```
                             "O2GE28days", "Nec", "NecEnt", "Rop", "Pda", "Sepsis",
                             "Funisitis", "Ivh", "Wmd", "Cp", "Mdi", "Pdi")
> GSE8586.pData <- GSE8586.pData[ , c(explanatoryVariables,
                                      intermediateVariables,
                                      responseVariables)]
```

2.3 Batch processing information

Since the article of Cohen *et al.* (2007) contains no information about the experimental design, we have taken the scanning date stored in the raw CEL files as surrogate for the possible processing of the samples into batches underlying the experimental design. We build the batch indicator variable using the time difference in days with respect to the earliest scanning date. Considering hybridizations closer in time together leads to three batch groups. This batch indicator variable becomes part of the phenotypic data stored in the *data.frame* `GSE8586.pData`:

```
> scanDate <- pData(protocolData(affyCELonly))$ScanDate
> scanDate <- gsub(" .*", "", scanDate)
> scanDate <- as.Date(scanDate, "%m/%d/%Y")
> batch <- scanDate - min(scanDate)
> sort(batch)
Time differences in days
 [1]   0   0   0   0  13  33  33  33  33  33  33  33  33  33  33  35  35  35  35
[20]  35  67  67  67  67  67  76  76  76  76  76  76  82  82  82  82  89  89  89
[39]  89  89  95  95  95  95  95  95 101 101 101 101 116 116 116 116
> batch <- cut(as.numeric(batch),c(-1,59,89,119))
> batch
 [1] (59,89]  (89,119] (59,89]  (59,89]  (59,89]  (-1,59]  (-1,59]  (-1,59]
 [9] (-1,59]  (-1,59]  (59,89]  (-1,59]  (-1,59]  (-1,59]  (-1,59]  (59,89]
[17] (59,89]  (59,89]  (-1,59]  (59,89]  (-1,59]  (89,119] (59,89]  (-1,59]
[25] (-1,59]  (89,119] (89,119] (59,89]  (-1,59]  (59,89]  (59,89]  (89,119]
[33] (-1,59]  (-1,59]  (59,89]  (89,119] (59,89]  (-1,59]  (-1,59]  (89,119]
[41] (-1,59]  (59,89]  (-1,59]  (89,119] (89,119] (89,119] (59,89]  (89,119]
[49] (89,119] (89,119] (59,89]  (59,89]  (89,119] (59,89]
Levels: (-1,59] (59,89] (89,119]
> table(batch)
batch
 (-1,59]  (59,89] (89,119]
      20       20       14
> batch <- as.numeric(batch)
> batch
 [1] 2 3 2 2 2 1 1 1 1 1 2 1 1 1 1 2 2 2 1 2 1 3 2 1 1 3 3 2 1 2 2 3 1 1 2 3 2 1
[39] 1 3 1 2 1 3 3 3 2 3 3 3 2 2 3 2
> GSE8586.pData$Batch <- batch
> table(GSE8586.pData$Batch)
 1  2  3
20 20 14
```

2.4 Addition of metadata

We build a *data.frame* object with the corresponding metadata information.

```
> GSE8586.Metadata.df <- data.frame(labelDescription = c(
    ## explanatory variables
```

```
        "Internal link identity", "Batch indicator", "Body weight at birth",
        "Extremely low birth weight (< 1000gr.)",
        "Gestational Age (GA) in weeks", "GA greater or equal than 27 weeks", "Sex", "Sex is male",
        "Race is nonwhite",

        ## intermediate variables
        "Preterm labor", "Pregnancy induced hypertension", "Cesarean section",
        "Antenatal glucocorticoids", "Neutrophils in chorionic plate", "Neutrophils in chord",
        "Neutrophils only in plate",

        ## reponse variables
        "FIR/chorioamnionitis, i.e., neutrophils in chorionic plate and/or neutrophils in chord",
        "Presence of bronchopulmonary dysplasia", "Days of ventilation",
        "Days of ventilation greater than 7 days", "Days of supplemental oxygen",
        "Days of supplemental oxygen greater or equal than 28 days",
        "Grade of necrotizing enterocolitis", "Presence of necrotizing enterocolitis",
        "Presence of retinopathy of prematurity", "Presence of patent ductus arteriosus",
        "Presence of Sepsis", "Neutrophils only in cord", "Presence of intraventricular hemorrhage",
        "White matter abnormality (ventriculomegaly and/or echolucency) seen on head ultrasound whil
        "Cerebral palsy diagnosis at age 2",
        "Bayley scales of infant development II, Mental development index < 70 at age 2",
        "Bayley scales of infant development II, Psychomotor development index < 70 at age 2"
        ), row.names=c(explanatoryVariables, intermediateVariables, responseVariables))
> stopifnot(rownames(GSE8586.Metadata.df) == colnames(GSE8586.pData)) ## QC
```

Finally, we combine the metadata information with the `GSE8586.pData` object into an *AnnotatedDataFrame* object, that we add to the *affyBatch* object `affyCELonly` and store it into the file `results/QA_dir/QA_affy.RData` as an object called `affy`:

```
> GSE8586.phenoData <- new("AnnotatedDataFrame", data=GSE8586.pData,
                           varMetadata=GSE8586.Metadata.df)
> affy <- affyCELonly
> phenoData(affy) <- GSE8586.phenoData
> stopifnot(rownames(pData(protocolData(affy))) == rownames(pData(affy))) ## QC
> stopifnot(colnames(exprs(affy)) == rownames(pData(affy))) ## QC
> save(affy, file=file.path(cacheDir_QA, paste0(cachePrefix_QA, "affy.RData")))
```

2.5 Quality Assesment and normalization

The goal of the quality assessment (QA) is to identify, and eventually remove, samples resulting from low quality chips. As a first exploration of the expression data we look at the distribution of the raw intensity values in $\log_2$ scale, distinguishing samples by batch, shown in Figure 1.

We do not observe a correlation between batch and raw intensity distribution and while most of the samples have similar distribution a few of them have a distinctive shape from the rest, concretely GSM213079, GSM213094, GSM213090, GSM213073 from which we take note here:

```
> badSamplesRawDist <- rnk[1:4] ## rnk is a ranking by mode height derived
>                               ## from the plotting hidden code
> badSamplesRawDist
[1] "GSM213079" "GSM213094" "GSM213090" "GSM213073"
```

We examine now the presence of intensity-dependence biases in the $\log_2$ values of raw intensity expression data by means of MA-plots. Their presence would be indicative of possible quality problems, although only after normalization has taken place this diagnostic is fully informative. We produced the MA-plots shown in Figures 2 and 3 removing control probes prior to calculating the plotted quantities.

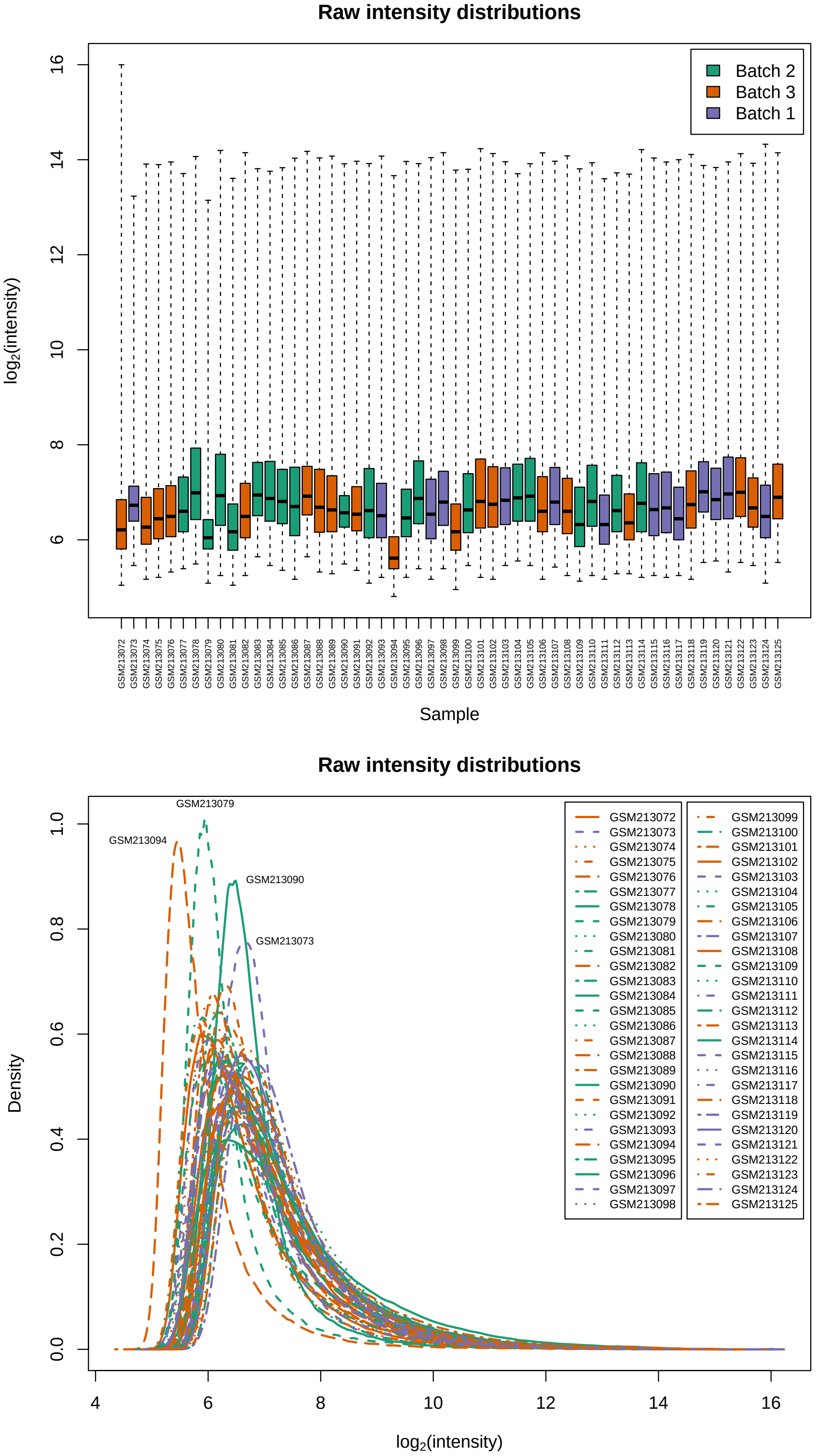

Figure 1. Distributions of raw intensity values in $\log_2$ scale. In the bottom panel, distributions with a distinctive shape from the bulk of the data are labeled by the corresponding sample identifier.

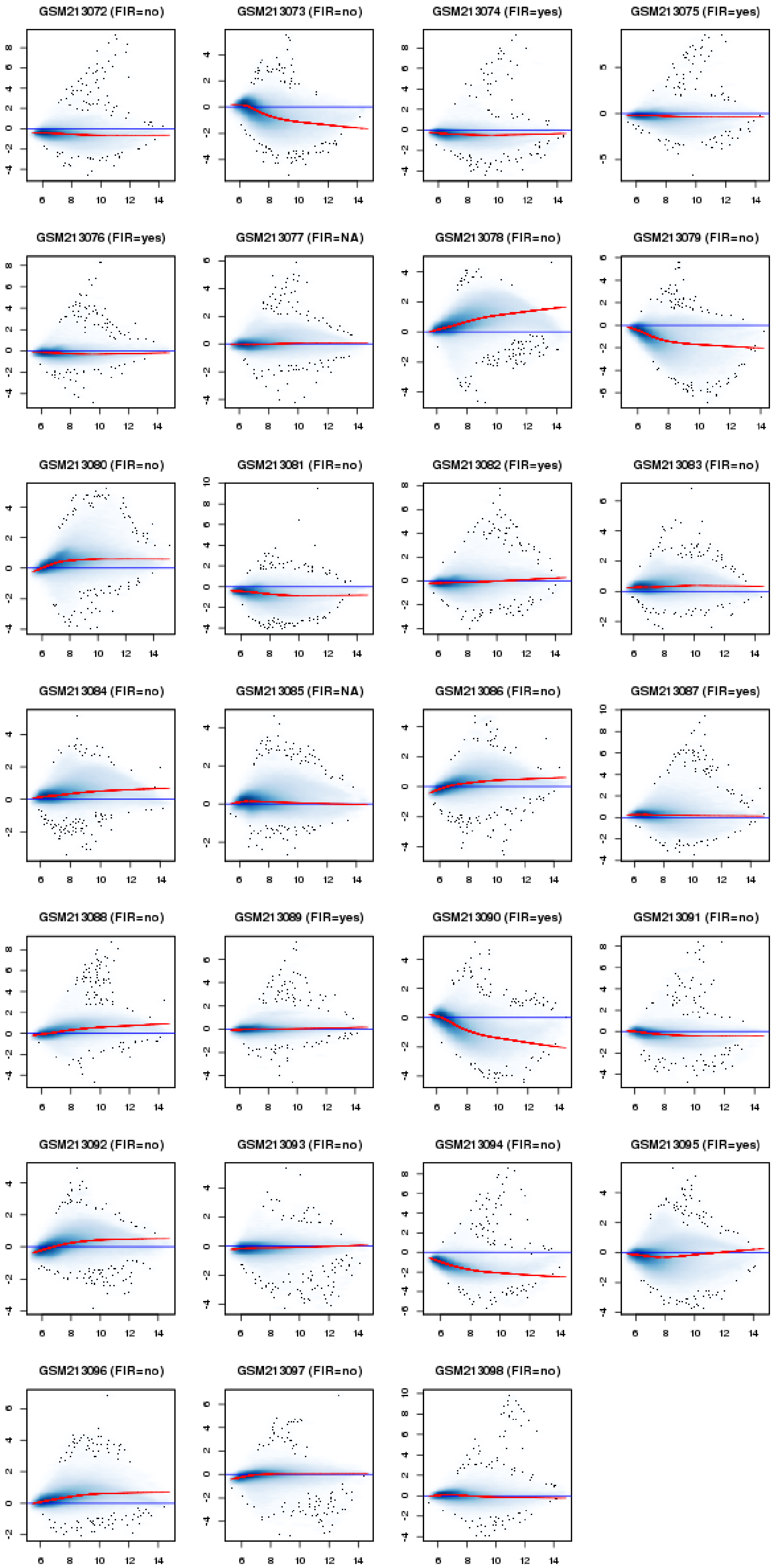

Figure 2. MA-plots of the raw expression values for the first half of the samples.

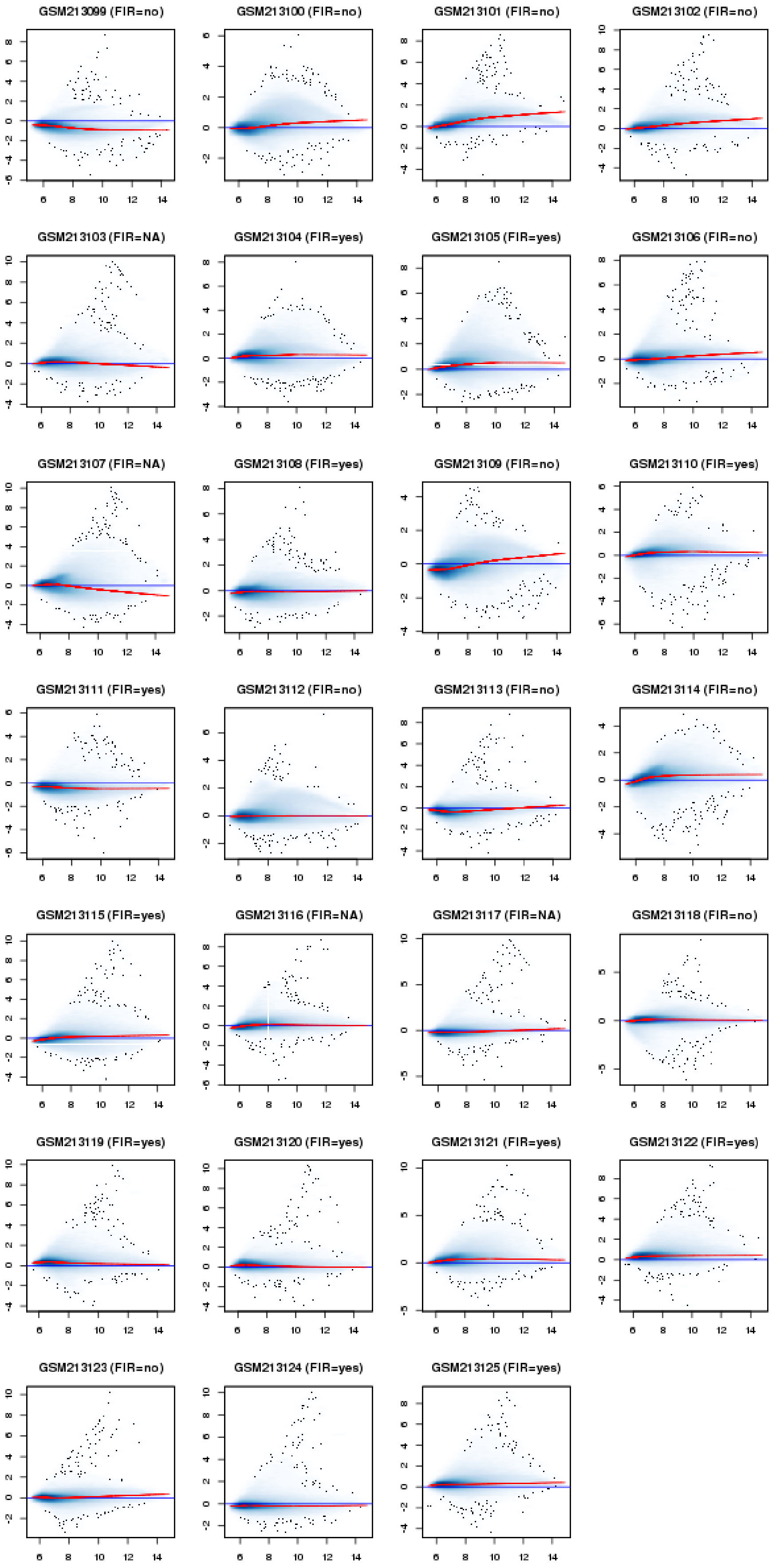

Figure 3. MA-plots of the raw expression values for the second half of the samples.

The next exploration of the data consists of fitting a probe-level linear model (PLM) and using it to produce some QA diagnostics. We begin by fitting this model through the function `fitPLM()` from the oligo package.

```
> cache(plm <- fitPLM(affy), dir=cacheDir_QA, prefix=cachePrefix_QA)
```

The first QA diagnostic from PLM approach are the images of the residuals of the PLM which are shown in Figure 4. Samples showing a distinctive pattern of residuals with respect to the rest of the samples could arise from problems with the RNA source, hybridization or scanning of the corresponding microarray chip. We observe such a pattern in the following samples:

```
> badSamplesPLMresids <- c("GSM213073", "GSM213079", "GSM213081", "GSM213090",
                           "GSM213094", "GSM213107", "GSM213110", "GSM213114", "GSM213118")
```

Figure 4. Images of the residuals of the probe-level model (PLM) per sample. Samples with a distinctive pattern of residuals have their identifier highlighted in red.

The third and fourth diagnostics are the Normalized Unscaled Standard Errors (NUSE) and the Relative Log Expression (RLE) plots (4), shown in Figure 5.

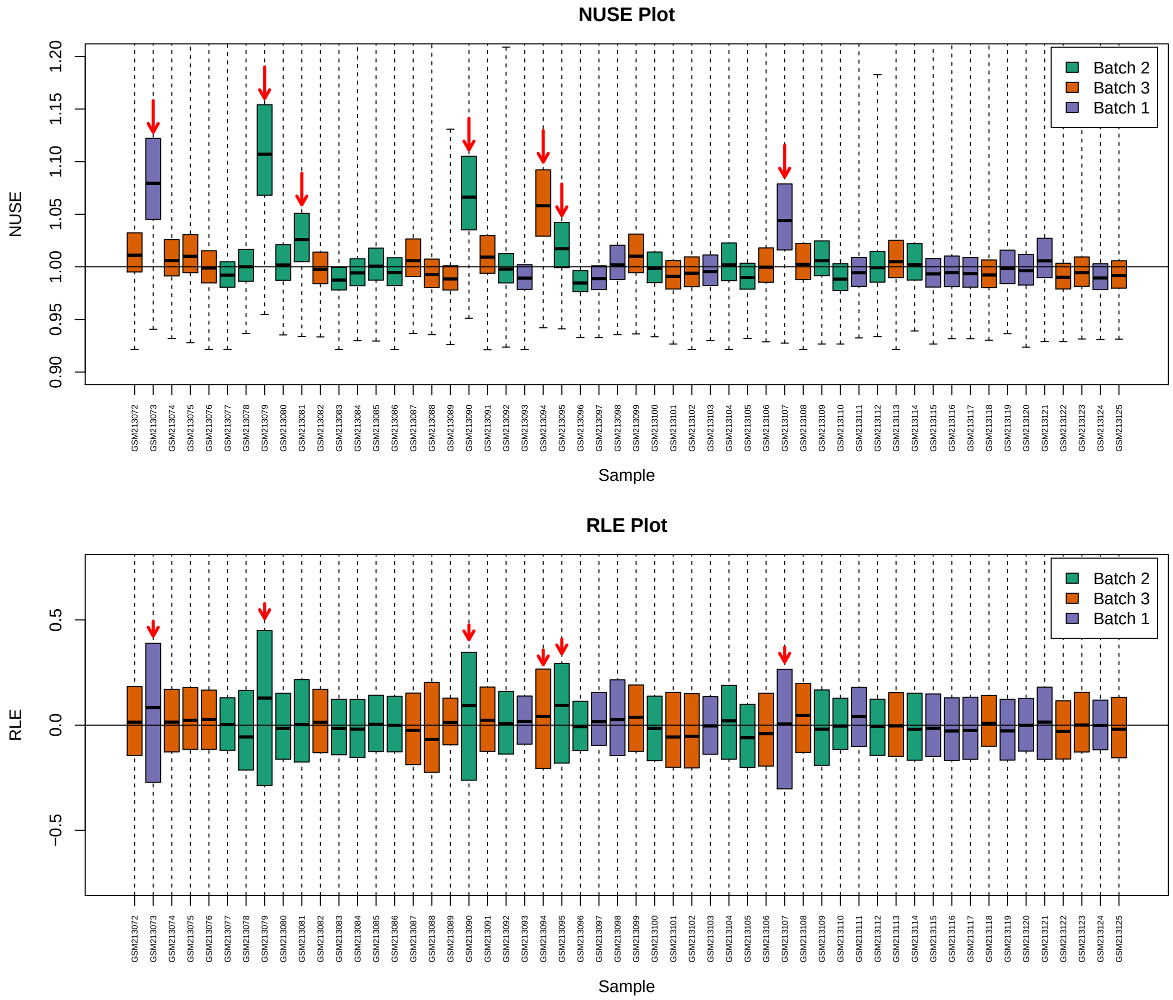

Figure 5. NUSE and RLE plots. Samples with abnormally high median NUSE value or abnormally interquartile range RLE value are indicated by red arrows.

These two diagnostics show problems with the following samples:

```
> ## rnkNUSE and rnkRLE are defined in the hidden code to produce the plots
> badSamplesNUSE <- sampleNames(plm)[rnkNUSE[1:7]]
> badSamplesNUSE
[1] "GSM213079" "GSM213073" "GSM213090" "GSM213094" "GSM213107" "GSM213081"
[7] "GSM213095"
> badSamplesRLE <- sampleNames(plm)[rnkRLE[1:6]]
> badSamplesRLE
[1] "GSM213079" "GSM213073" "GSM213090" "GSM213107" "GSM213094" "GSM213095"
```

We apply the recently developed frozen robust multichip analysis (fRMA) normalization algorithm (5) implemented in the function `frma()` from the frma to do background correction, quantile normalization and summarization of probe-level intensities into probeset-level expression values.

```
> library(frma)
> cache(frmaEsetAll <- frma(affy, summarize="random_effect"),
        dir=cacheDir_QA, prefix=cachePrefix_QA)
```

We perform MA-plots again on the *ExpressionSet* object to verify whether intensity-dependent biases observed in the raw expression intensity data have been removed. They are shown in Figures 6 and 7 and calculations have been made excluding Affymetrix control probesets. The following samples show still some degree of fRMA expression-level dependent bias:

```
> badSamplesMAplot1 <- c("GSM213073", "GSM213079", "GSM213081",
                         "GSM213090", "GSM213095")
> badSamplesMAplot2 <- c("GSM213104", "GSM213107", "GSM213112")
> badSamplesMAplot <- c(badSamplesMAplot1, badSamplesMAplot2)
```

2.6 Batch Effect identification

Using the batch effect indicator variable previously built on the basis of scanning dates stored in CEL files, we explore the distribution of samples in the primary outcome of interest (phenotypic variable `FIR`) across batches. An unbalanced distribution could lead to a confounding effect between batch and primary outcome. However, as shown in Table 1, the experimental design is balanced between FIR status and the batch indicator variable.

Table 1. Distribution of samples across batches (columns) and FIR status (rows).

	1	2	3
no	13	11	4
yes	5	9	6

Next, we examine how samples cluster together by the primary outcome of interest and by the batch indicator variable. In the top panel of Figure 8 we have the hierarchical clustering annotated by the batch indicator variable `Batch` while in the bottom panel we have a multidimensional scaling (MDS) plot.

The hierarchical clustering shows some tendency of samples from the same batch to cluster together. More importantly, some of the samples are very dissimilar to the rest, concretely, the following ones:

```
> sampleMainClusters <- cutree(sampleClustering, k=4)
> table(sampleMainClusters)
sampleMainClusters
 1  2  3  4
45  7  1  1
> badSamplesHC <- names(sampleMainClusters)[sampleMainClusters > 1]
> badSamplesHC
[1] "GSM213073" "GSM213079" "GSM213090" "GSM213094" "GSM213095" "GSM213098"
[7] "GSM213104" "GSM213107" "GSM213121"
```

Analgously, the following samples in the MDS plot are very dissimilar to the rest:

```
> badSamplesMDS <- rownames(cmd)[cmd[, 1] < -0.05]
> badSamplesMDS
[1] "GSM213073" "GSM213079" "GSM213090" "GSM213095" "GSM213107"
```

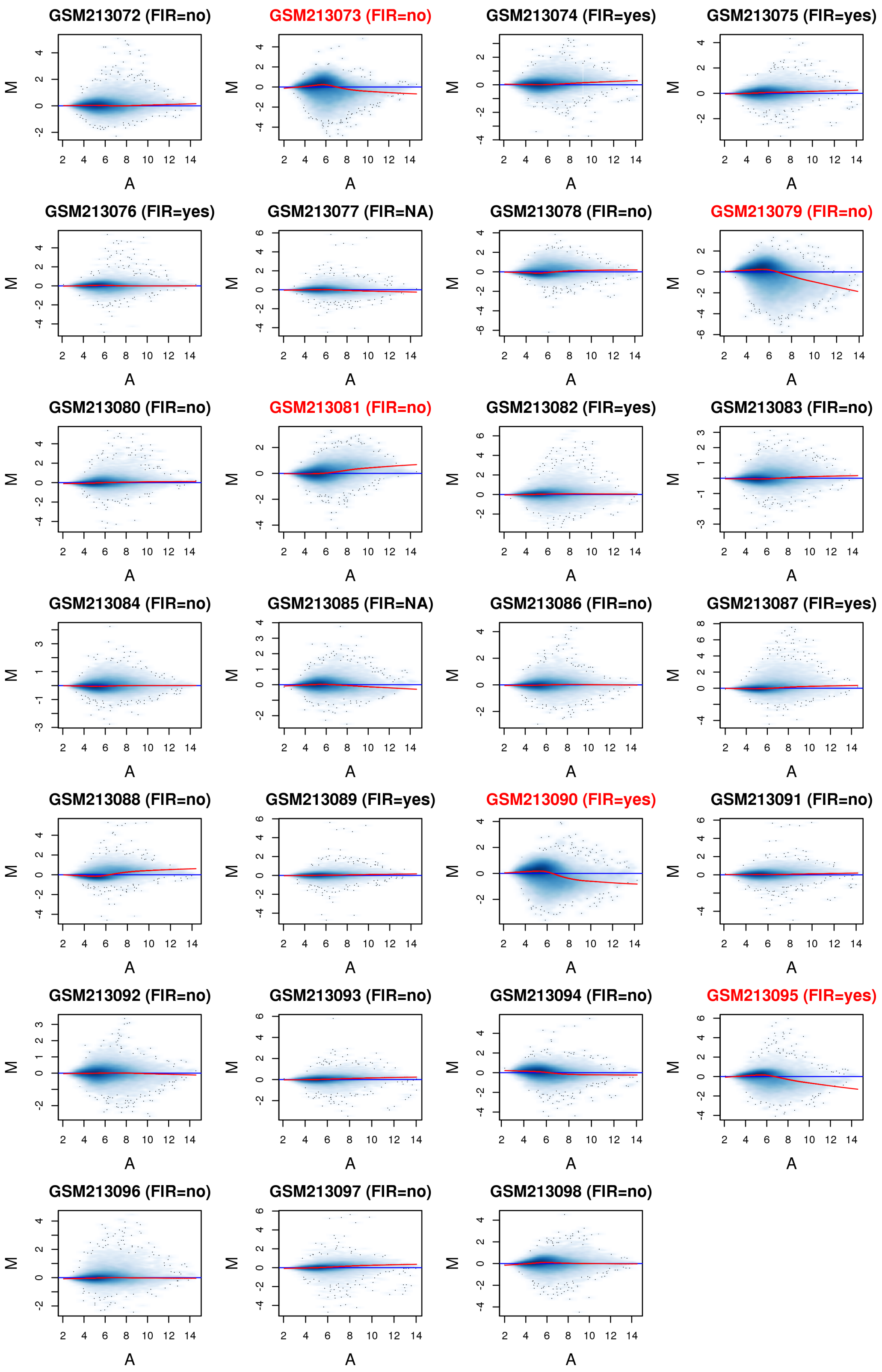

Figure 6. MA-plots of the fRMA expression values for the first half of the samples. Samples with some sizeable degree of expression-level dependence have their identifier highlighted in red.

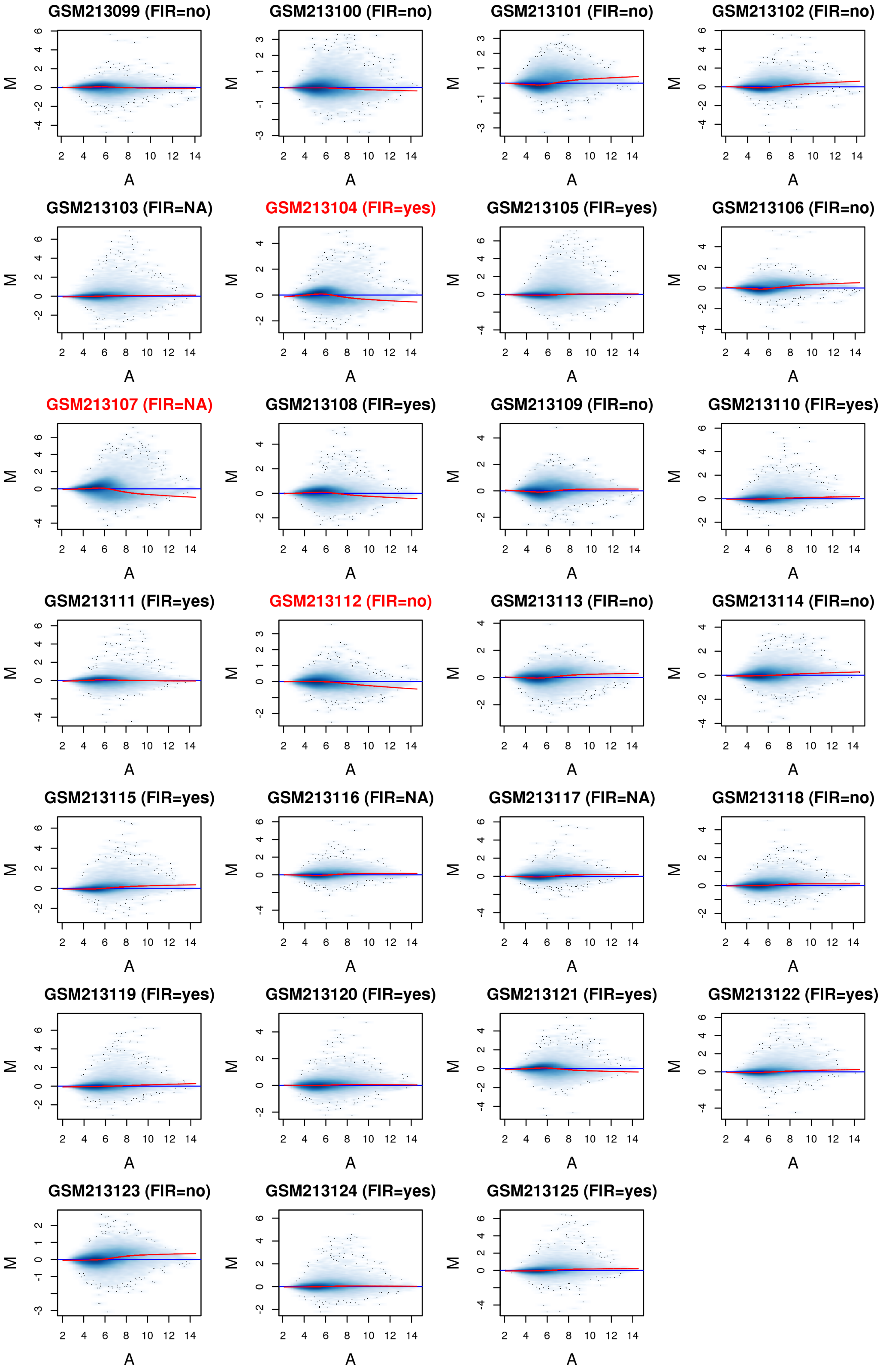

Figure 7. MA-plots of the fRMA expression values for the second half of the samples. Samples with some sizeable degree of expression-level dependence have their identifier highlighted in red.

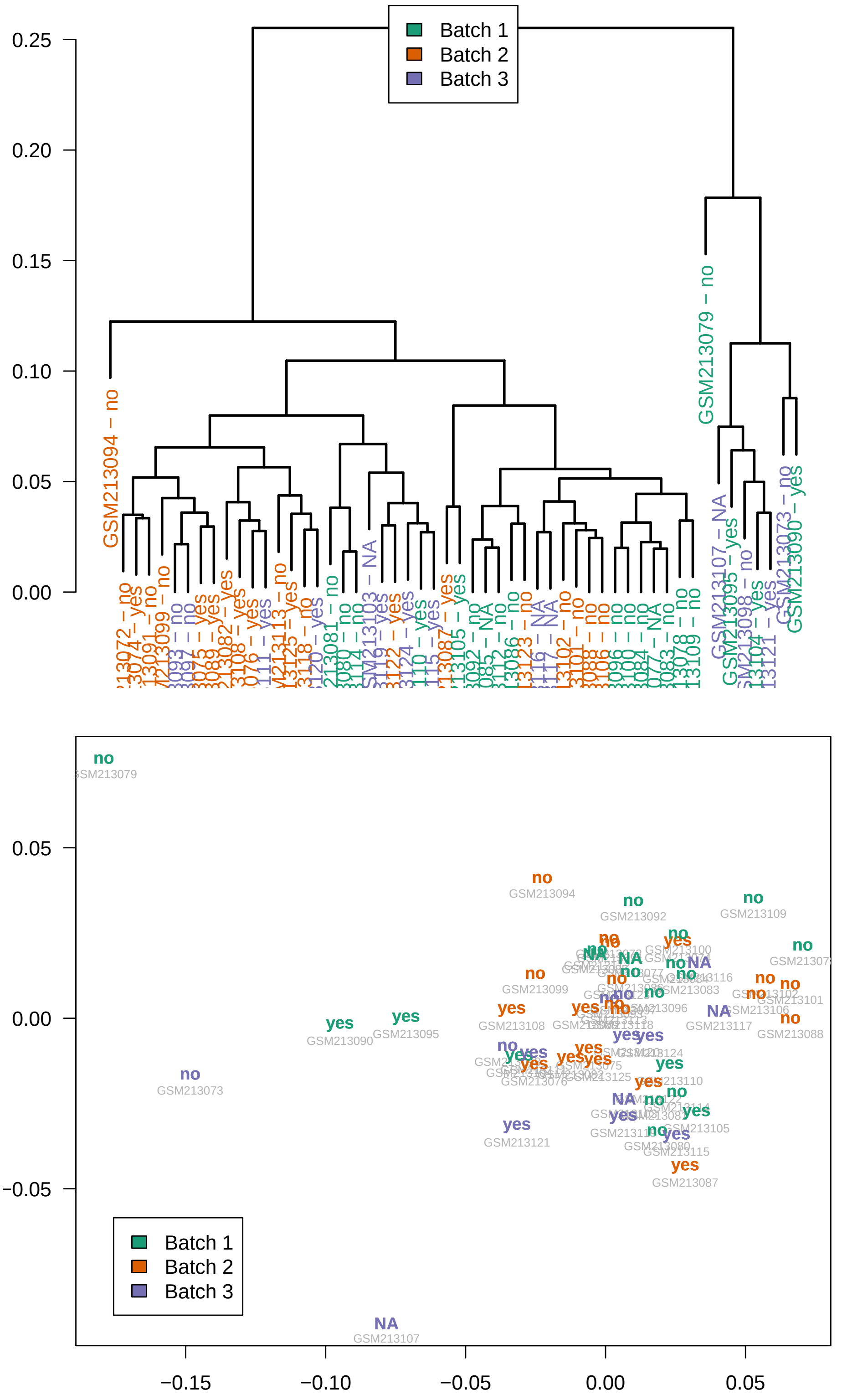

Figure 8. Hierarchical clustering (top) and multidimensional scaling (bottom) plots of samples labeled by FIR status and colored batch indicator variable. Some samples are very dissimilar to the rest.

2.7 Summary of quality assessment diagnostics

We will build now a table summarizing the result of all diagnostics by flagging every bad diagnostic at every sample and rank the samples by summing all the failed diagnostics.

```
> qaDiagnostics <- data.frame(RawDist=rep(FALSE, ncol(frmaEsetAll)),
                              Resids=rep(FALSE, ncol(frmaEsetAll)),
                              NUSE=rep(FALSE, ncol(frmaEsetAll)),
                              RLE=rep(FALSE, ncol(frmaEsetAll)),
                              MA=rep(FALSE, ncol(frmaEsetAll)),
                              HC=rep(FALSE, ncol(frmaEsetAll)),
                              MDS=rep(FALSE, ncol(frmaEsetAll)),
                              Failed=rep(0, ncol(frmaEsetAll)))
> rownames(qaDiagnostics) <- sampleNames(frmaEsetAll)
> qaDiagnostics[badSamplesRawDist, "RawDist"] <- TRUE
> qaDiagnostics[badSamplesPLMresids, "Resids"] <- TRUE
> qaDiagnostics[badSamplesNUSE, "NUSE"] <- TRUE
> qaDiagnostics[badSamplesRLE, "RLE"] <- TRUE
> qaDiagnostics[badSamplesMAplot, "MA"] <- TRUE
> qaDiagnostics[badSamplesHC, "HC"] <- TRUE
> qaDiagnostics[badSamplesMDS, "MDS"] <- TRUE
> qaDiagnostics$Failed <- rowSums(qaDiagnostics)
> qaDiagnostics <- qaDiagnostics[order(qaDiagnostics$Failed, decreasing=TRUE), ]
> head(qaDiagnostics, n=13)
          RawDist Resids  NUSE   RLE    MA    HC   MDS Failed
GSM213073    TRUE   TRUE  TRUE  TRUE  TRUE  TRUE  TRUE      7
GSM213079    TRUE   TRUE  TRUE  TRUE  TRUE  TRUE  TRUE      7
GSM213090    TRUE   TRUE  TRUE  TRUE  TRUE  TRUE  TRUE      7
GSM213107   FALSE   TRUE  TRUE  TRUE  TRUE  TRUE  TRUE      6
GSM213094    TRUE   TRUE  TRUE  TRUE FALSE  TRUE FALSE      5
GSM213095   FALSE  FALSE  TRUE  TRUE  TRUE  TRUE  TRUE      5
GSM213081   FALSE   TRUE  TRUE FALSE  TRUE FALSE FALSE      3
GSM213104   FALSE  FALSE FALSE FALSE  TRUE  TRUE FALSE      2
GSM213098   FALSE  FALSE FALSE FALSE FALSE  TRUE FALSE      1
GSM213110   FALSE   TRUE FALSE FALSE FALSE FALSE FALSE      1
GSM213112   FALSE  FALSE FALSE FALSE  TRUE FALSE FALSE      1
GSM213114   FALSE   TRUE FALSE FALSE FALSE FALSE FALSE      1
GSM213118   FALSE   TRUE FALSE FALSE FALSE FALSE FALSE      1
```

The full ranking of samples is shown in Table 2 and in the light of this ranking we decide to discard the samples for which the majority of the diagnostics fail, i.e., more than 4, concretely the following ones:

```
> badSamples <- rownames(qaDiagnostics)[qaDiagnostics$Failed > (ncol(qaDiagnostics)-1)/2]
> badSamples
[1] "GSM213073" "GSM213079" "GSM213090" "GSM213107" "GSM213094" "GSM213095"
```

We normalize the data again excluding the selected samples:

```
> cache(frmaEset <- frma(affy[, -match(badSamples, sampleNames(affy))],
                         summarize="random_effect"),
        dir=cacheDir_QA, prefix=cachePrefix_QA)
> dim(frmaEset)
Features  Samples
   54675       48
```

This resulting normalized data set of 54675 features and 48 samples will be the one we use in the rest of the analysis. We check again in Table 3 below that the experimental design is still balanced with respect to the batch indicator variable and FIR status.

Table 2. Summary of quality assessment diagnostics for all microarray samples. Every row corresponds to a microarray sample and every column to a different diagnostic. Rows are ordered by the decreasing number of failed diagnostics indicated in the column "Failed". Samples which failed in more than half of the diagnostics have their identifiers highlighted in red and were discarded from further analysis. Diagnostics key: distribution of raw intensities (RawDist), residuals of a probe-level model (Resids), normalized unscaled standard errors (NUSE), relative log expression (RLE), MA plot (MA), hierarchical clustering (HC), multidimensional scaling (MDS).

	RawDist	Resids	NUSE	RLE	MA	HC	MDS	Failed
GSM213073	✗	✗	✗	✗	✗	✗	✗	7
GSM213079	✗	✗	✗	✗	✗	✗	✗	7
GSM213090	✗	✗	✗	✗	✗	✗	✗	7
GSM213107	✓	✗	✗	✗	✗	✗	✗	6
GSM213094	✗	✗	✗	✗	✓	✗	✓	5
GSM213095	✓	✓	✗	✗	✗	✗	✗	5
GSM213081	✓	✗	✗	✓	✗	✓	✓	3
GSM213104	✓	✓	✓	✓	✗	✗	✓	2
GSM213098	✓	✓	✓	✓	✓	✗	✓	1
GSM213110	✓	✗	✓	✓	✓	✓	✓	1
GSM213112	✓	✓	✓	✓	✗	✓	✓	1
GSM213114	✓	✗	✓	✓	✓	✓	✓	1
GSM213118	✓	✗	✓	✓	✓	✓	✓	1
GSM213121	✓	✓	✓	✓	✓	✗	✓	1
GSM213072	✓	✓	✓	✓	✓	✓	✓	0
GSM213074	✓	✓	✓	✓	✓	✓	✓	0
GSM213075	✓	✓	✓	✓	✓	✓	✓	0
GSM213076	✓	✓	✓	✓	✓	✓	✓	0
GSM213077	✓	✓	✓	✓	✓	✓	✓	0
GSM213078	✓	✓	✓	✓	✓	✓	✓	0
GSM213080	✓	✓	✓	✓	✓	✓	✓	0
GSM213082	✓	✓	✓	✓	✓	✓	✓	0
GSM213083	✓	✓	✓	✓	✓	✓	✓	0
GSM213084	✓	✓	✓	✓	✓	✓	✓	0
GSM213085	✓	✓	✓	✓	✓	✓	✓	0
GSM213086	✓	✓	✓	✓	✓	✓	✓	0
GSM213087	✓	✓	✓	✓	✓	✓	✓	0
GSM213088	✓	✓	✓	✓	✓	✓	✓	0
GSM213089	✓	✓	✓	✓	✓	✓	✓	0
GSM213091	✓	✓	✓	✓	✓	✓	✓	0
GSM213092	✓	✓	✓	✓	✓	✓	✓	0
GSM213093	✓	✓	✓	✓	✓	✓	✓	0
GSM213096	✓	✓	✓	✓	✓	✓	✓	0
GSM213097	✓	✓	✓	✓	✓	✓	✓	0
GSM213099	✓	✓	✓	✓	✓	✓	✓	0
GSM213100	✓	✓	✓	✓	✓	✓	✓	0
GSM213101	✓	✓	✓	✓	✓	✓	✓	0
GSM213102	✓	✓	✓	✓	✓	✓	✓	0
GSM213103	✓	✓	✓	✓	✓	✓	✓	0
GSM213105	✓	✓	✓	✓	✓	✓	✓	0
GSM213106	✓	✓	✓	✓	✓	✓	✓	0
GSM213108	✓	✓	✓	✓	✓	✓	✓	0
GSM213109	✓	✓	✓	✓	✓	✓	✓	0
GSM213111	✓	✓	✓	✓	✓	✓	✓	0
GSM213113	✓	✓	✓	✓	✓	✓	✓	0
GSM213115	✓	✓	✓	✓	✓	✓	✓	0
GSM213116	✓	✓	✓	✓	✓	✓	✓	0
GSM213117	✓	✓	✓	✓	✓	✓	✓	0
GSM213119	✓	✓	✓	✓	✓	✓	✓	0
GSM213120	✓	✓	✓	✓	✓	✓	✓	0
GSM213122	✓	✓	✓	✓	✓	✓	✓	0
GSM213123	✓	✓	✓	✓	✓	✓	✓	0
GSM213124	✓	✓	✓	✓	✓	✓	✓	0
GSM213125	✓	✓	✓	✓	✓	✓	✓	0

Table 3. Distribution of samples across batches (columns) and FIR status (rows). The batch indicator variable was derived on the basis of the microarray chip scanning timestemp.

Batch FIR	1	2	3
No	12	10	3
Yes	3	9	6

2.8 Session Information

```
> toLatex(sessionInfo())
```

- R version 3.1.3 (2015-03-09), `x86_64-unknown-linux-gnu`
- Locale: `LC_CTYPE=en_US.UTF8`, `LC_NUMERIC=C`, `LC_TIME=en_US.UTF8`, `LC_COLLATE=en_US.UTF8`, `LC_MONETARY=en_US.UTF8`, `LC_MESSAGES=en_US.UTF8`, `LC_PAPER=en_US.UTF8`, `LC_NAME=en_US.UTF8`, `LC_ADDRESS=en_US.UTF8`, `LC_TELEPHONE=en_US.UTF8`, `LC_MEASUREMENT=en_US.UTF8`, `LC_IDENTIFICATION=en_US.UTF8`
- Base packages: base, datasets, graphics, grDevices, methods, parallel, stats, stats4, utils
- Other packages: affy 1.44.0, affyPLM 1.42.0, annotate 1.44.0, AnnotationDbi 1.28.2, Biobase 2.26.0, BiocGenerics 0.12.1, DBI 0.3.1, frma 1.18.0, gcrma 2.38.0, genefilter 1.48.1, GenomeInfoDb 1.2.5, GEOquery 2.32.0, GO.db 3.0.0, hgu133plus2cdf 2.15.0, hgu133plus2.db 3.0.0, IRanges 2.0.1, limma 3.22.7, mgcv 1.8-6, nlme 3.1-120, org.Hs.eg.db 3.0.0, preprocessCore 1.28.0, RColorBrewer 1.1-2, rJava 0.9-6, RSQLite 1.0.0, S4Vectors 0.4.0, sva 3.12.0, xlsx 0.5.7, xlsxjars 0.6.1, XML 3.98-1.1, xtable 1.7-4
- Loaded via a namespace (and not attached): affxparser 1.38.0, affyio 1.34.0, BiocInstaller 1.16.4, Biostrings 2.34.1, bit 1.1-12, bitops 1.0-6, codetools 0.2-11, ff 2.2-13, foreach 1.4.2, GenomicRanges 1.18.4, grid 3.1.3, iterators 1.0.7, lattice 0.20-31, MASS 7.3-40, Matrix 1.2-0, oligo 1.30.0, oligoClasses 1.28.0, RCurl 1.95-4.5, splines 3.1.3, survival 2.38-1, tools 3.1.3, XVector 0.6.0, zlibbioc 1.12.0

3 Differential gene expression analysis

The aim of the differential expression (DE) analysis in this study is to identify genes that significantly change their expression between absence or presence of a fetal inflamatory response (FIR) as defined by the phenotypic variable `FIR`. Due to missing clinical data, the FIR status could not be established for some of the samples. For the purpose of this DE analysis we have to discard these samples. We start the analysis on the expression data which has been already normalized with fRMA.

```
> dim(frmaEset)
Features  Samples
   54675       48
> frmaEset <- frmaEset[, !is.na(frmaEset$FIR)]
> dim(frmaEset)
Features  Samples
   54675       43
```

We use the Bioconductor package limma (6) for DE analysis using linear models and moderated t-statistics. Throughout the forthcoming variants on this analysis we will use the following cutoffs on maximum FDR and minimum fold-change:

```
> FDRcutoff <- 0.01
> fcCutoff <- 1.5
```

3.1 Population study analysis

Before we start the DE analysis we study baseline characteristics of labor/mother and infants in relationship with the FIR indicator variable (neutrophils in plate and/or neutrophils in cord).

```
> laborVar <- c("Nonwhite", "Ptl", "Pih", "Cs", "Acs")
> laborDf <- cbind(Description=varMetadata(frmaEset)[laborVar, ],
                 as.data.frame(do.call("rbind",
                   lapply(pData(frmaEset)[, laborVar],
                     function(phe, outcome, absenceOutcome, presenceOutcome) {
                       xtab <- table(phe, outcome)
                       ft <- fisher.test(xtab)
                       outcomeOverVar <- c(sprintf("%d (%.1f\\%%)", xtab["yes", ],
                                                   100 * xtab["yes", ] / colSums(xtab)),
                                           sprintf("%.3f", ft$p.value))
                       names(outcomeOverVar) <- c(absenceOutcome, presenceOutcome, "P")
                       outcomeOverVar
                     }, frmaEset$FIR, "No FIR", "FIR")),
                 stringsAsFactors=FALSE), stringsAsFactors=FALSE)
```

Table 4. Characteristics of labor and mother related to FIR. Number of infants whose birth matched the characteristics listed on the rows stratified by FIR outcome specified on the columns. Between parenthesis is given the percentage of infants with that characteristic over the total number of infants with the given FIR outcome. The P column gives the two-sided Fisher's exact p-value for the null hypothesis of no-association between characteristic and FIR outcome.

	Description	No FIR	FIR	P
Nonwhite	Race is nonwhite	9 (36.0%)	7 (38.9%)	1.000
Ptl	Preterm labor	16 (64.0%)	8 (44.4%)	0.230
Pih	Pregnancy induced hypertension	4 (16.0%)	0 (0.0%)	0.127
Cs	Cesarean section	23 (92.0%)	9 (50.0%)	0.004
Acs	Antenatal glucocorticoids	14 (56.0%)	12 (66.7%)	0.541

The characteristics of labor and mother are summarized in Table 4. We proceed to analyse the characteristics of the infants.

```
> infantVar <- c("Elbw", "GaGE27weeks", "Male", "Bpd", "VentGT7days",
                 "O2GE28days", "NecEnt", "Rop", "Pda", "Sepsis",
                 "Ivh", "Wmd")
> infantDf <- cbind(Description=gsub("<", "$<$", varMetadata(frmaEset)[infantVar, ]),
                  as.data.frame(do.call("rbind",
                    lapply(pData(frmaEset)[, infantVar],
                      function(phe, outcome, absenceOutcome, presenceOutcome) {
                        xtab <- table(phe, outcome)
                        ft <- fisher.test(xtab)
                        outcomeOverVar <- c(sprintf("%d (%.1f\\%%)", xtab["yes", ],
                                                    100 * xtab["yes", ] / colSums(xtab)),
                                            sprintf("%.3f", ft$p.value))
                        names(outcomeOverVar) <- c(absenceOutcome, presenceOutcome, "P")
                        outcomeOverVar
                      }, frmaEset$FIR, "No FIR", "FIR")),
                  stringsAsFactors=FALSE), stringsAsFactors=FALSE)
```

The characteristics of infants are summarized in Table 5.

We analyze resulting characteristics of infants, conditioned on the administration of antenatal glucocorticoids, which have been described as an effective therapy that improves neonatal outcome (7). The results are summarized in Table 6.

Table 5. Characteristics of infants related to FIR. Number of infants whose birth matched the characteristics listed on the rows stratified by FIR outcome specified on the columns. Between parenthesis is given the percentage of infants with that characteristic over the total number of infants with the given FIR outcome. The P column gives the two-sided Fisher's exact p-value for the null hypothesis of no-association between characteristic and FIR outcome.

	Description	No FIR	FIR	P
Elbw	Extremely low birth weight (< 1000gr.)	18 (72.0%)	8 (44.4%)	0.114
GaGE27weeks	GA greater or equal than 27 weeks	14 (56.0%)	9 (50.0%)	0.763
Male	Sex is male	17 (68.0%)	10 (55.6%)	0.526
Bpd	Presence of bronchopulmonary dysplasia	7 (28.0%)	6 (33.3%)	0.747
VentGT7days	Days of ventilation greater than 7 days	13 (52.0%)	5 (29.4%)	0.208
O2GE28days	Days of supplemental oxygen greater or equal than 28 days	18 (72.0%)	12 (70.6%)	1.000
NecEnt	Presence of necrotizing enterocolitis	6 (24.0%)	3 (16.7%)	0.712
Rop	Presence of retinopathy of prematurity	11 (44.0%)	8 (47.1%)	1.000
Pda	Presence of patent ductus arteriosus	22 (88.0%)	10 (58.8%)	0.062
Sepsis	Presence of Sepsis	1 (4.0%)	2 (11.8%)	0.556
Ivh	Presence of intraventricular hemorrhage	3 (12.0%)	7 (38.9%)	0.067
Wmd	White matter abnormality (ventriculomegaly and/or echolucency) seen on head ultrasound while the infant was in the NICU	2 (8.0%)	5 (27.8%)	0.110

```
> infantVarOut <- c("Elbw", "GaGE27weeks", "Bpd", "NecEnt", "Rop", "Pda", "Sepsis", "Ivh", "Wmd"
> infantAcsDf <- cbind(Description=gsub("<", "$<$", varMetadata(frmaEset)[infantVarOut, ]),
                       as.data.frame(do.call("rbind",
                         lapply(pData(frmaEset)[, infantVarOut],
                           function(phe, outcome, absenceOutcome, presenceOutcome,
                                    covariate, absenceCovariate, presenceCovariate) {
                             xtab <- table(phe, outcome, covariate)
                             xtabCovNo <- xtab[, , "no"]
                             xtabCovYes <- xtab[, , "yes"]
                             ftCovNo <- fisher.test(xtabCovNo)
                             ftCovYes <- fisher.test(xtabCovYes)
                             outcomePheCovNo <- c(sprintf("%d (%.1f\\%%)", xtabCovNo["yes", ],
                                                          100 * xtabCovNo["yes", ] / colSums(xtab
                                                  sprintf("%.3f", ftCovNo$p.value))
                             names(outcomePheCovNo) <- c(sprintf("%s|%s", absenceOutcome, absence
                                                         sprintf("%s|%s", presenceOutcome, absenc
                                                         "P")
                             outcomePheCovYes <- c(sprintf("%d (%.1f\\%%)", xtabCovYes["yes", ],
                                                           100 * xtabCovYes["yes", ] / colSums(xt
                                                  sprintf("%.3f", ftCovYes$p.value))
                             names(outcomePheCovYes) <- c(sprintf("%s|%s", absenceOutcome, presen
                                                          sprintf("%s|%s", presenceOutcome, prese
                                                          "P")
```

```
                    c(outcomePheCovNo, outcomePheCovYes)
                  }, frmaEset$FIR, "$\\neg$FIR", "FIR", frmaEset$Acs, "$\\neg$ACS", "AC
              stringsAsFactors=FALSE), stringsAsFactors=FALSE)
```

Table 6. Characteristics of infants related to FIR and conditioned on the administration of antenatal glucocortidcoids (ACS). Number of infants whose birth matched the characteristics listed on the rows stratified by FIR outcome specified on the columns. Between parenthesis is given the percentage of infants with that characteristic over the total number of infants with the FIR outcome given the ACS therapy. The P column gives the two-sided Fisher's exact p-value for the null hypothesis of no-association between characteristic and FIR outcome given ACS.

	Description	$\neg$FIR$\mid\neg$ACS	FIR$\mid\neg$ACS	P	$\neg$FIR$\mid$ACS	FIR$\mid$ACS	P
Elbw	Extremely low birth weight ($<$ 1000gr.)	7 (63.6%)	3 (50.0%)	0.644	11 (78.6%)	5 (41.7%)	0.105
GaGE27weeks	GA greater or equal than 27 weeks	5 (45.5%)	3 (50.0%)	1.000	9 (64.3%)	6 (50.0%)	0.692
Bpd	Presence of bronchopulmonary dysplasia	3 (27.3%)	2 (33.3%)	1.000	4 (28.6%)	4 (33.3%)	1.000
NecEnt	Presence of necrotizing enterocolitis	3 (27.3%)	1 (16.7%)	1.000	3 (21.4%)	2 (16.7%)	1.000
Rop	Presence of retinopathy of prematurity	5 (45.5%)	3 (50.0%)	1.000	6 (42.9%)	5 (45.5%)	1.000
Pda	Presence of patent ductus arteriosus	10 (90.9%)	4 (66.7%)	0.515	12 (85.7%)	6 (54.5%)	0.177
Sepsis	Presence of Sepsis	1 (9.1%)	2 (33.3%)	0.515	0 (0.0%)	0 (0.0%)	1.000
Ivh	Presence of intraventricular hemorrhage	1 (9.1%)	5 (83.3%)	0.005	2 (14.3%)	2 (16.7%)	1.000
Wmd	White matter abnormality (ventriculomegaly and/or echolucency) seen on head ultrasound while the infant was in the NICU	1 (9.1%)	3 (50.0%)	0.099	1 (7.1%)	2 (16.7%)	0.580

3.2 Non-specific filtering by functional features

Probesets for which there is no known gene annotated and chip control probesets cannot be directly used in any downstream functional analysis. For this reason we are going to remove these probesets from further analysis. Likewise, multiple probesets annotated to the same gene are, in principle, providing a redundant signal which may arise from different isoforms where only one of them is the one being expressed. In this case we want to remove probesets corresponding to duplicated gene annotations by keeping only one per gene, the one with largest variability measured by the interquartile range (IQR). For this purpose we use the function `nsfilter()` from the genefilter package as follows:

```
> frmaEsetFilt <- nsFilter(frmaEset, require.entrez=TRUE, remove.dupEntrez=TRUE,
                           var.func=IQR, var.filter=FALSE,
                           feature.exclude="^AFFX")$eset
> dim(frmaEsetFilt)
```

```
Features  Samples
   20155       43
```

Since we have now a one to one correspondence between probesets and Entrez Gene identifiers, we replace the Affymetrix probeset identifiers in the *ExpressionSet* object by their Entrez Gene counterparts and update the `annotation` slot to the human organism-level annotation package in Bioconductor. Prior to this step, we store the selected Affymetrix probeset identifiers to keep track of what probesets finally participate in the analysis.

```
> filteredProbesets <- featureNames(frmaEsetFilt)
> featureNames(frmaEsetFilt) <- select(hgu133plus2.db,
                                       keys=filteredProbesets,
                                       columns="ENTREZID",
                                       keytype="PROBEID")$ENTREZID
> annotation(frmaEsetFilt) <- "org.Hs.eg.db"
```

3.3 Estimation of hidden factors of expression heterogeneity

Expression data generally contains a substantial amount of heterogeneity due to a number of biological and non-biological sources of variability unrelated to the primary outcome of interest, which in our case is the FIR indicator variable `FIR`. To address this fact and adjust for these unknown confounders we run a surrogate variable analysis (SVA) among the top-10% of genes with largest variability, to estimate these unknown confounders and incorporate them into the linear model we fit with limma to the expression profile of each gene.

```
> IQRs <- esApply(frmaEsetFilt, 1, IQR)
> mod <- model.matrix(~ factor(frmaEsetFilt$FIR) + factor(frmaEsetFilt$Batch))
> mod0 <- model.matrix(~ factor(frmaEsetFilt$Batch))
> cache(sv <- sva(dat=exprs(frmaEsetFilt[IQRs > quantile(IQRs, prob=0.90), ]),
                  mod=mod, mod0=mod0),
        dir=cacheDir_DE, prefix=cachePrefix_DE)
> sv$n
[1] 7
```

The SVA analysis estimates up to 7 surrogate variables that capture variability unrelated to the primary outcome of interest, the FIR indicator variable.

3.4 Non-specific filtering by variability

Genes that do not change significantly across the experimental condition of interest generally display little variability across samples. Therefore, a straightforward way to alleviate the multiple testing problem and augment the statistical power in a DE analysis is to filter out those genes with little variability. However, since the moderated t-statistics from limma are calculated using all available probesets, we calculate now the variablity per probeset using the IQR and we will evaluate later the amount of filtering we perform as function of the power to detect DE genes (8). To avoid measuring variability unrelated to our outcome of interest, we will calculate the IQR on the expression data after removing the effect of the batch and surrogate variables.

```
> cleanExps <- removeBatchEffect(exprs(frmaEsetFilt), batch=frmaEsetFilt$Batch,
                                covariates=sv$sv,
                                design=model.matrix(~ factor(frmaEsetFilt$FIR)))
> IQRs <- apply(cleanExps, 1, IQR)
> filteringQuantiles <- seq(0.1, 0.9, by=0.1)
> filteringQuantilesIQRs <- quantile(IQRs, probs=filteringQuantiles)
> genesByFilterCutoff <- sapply(filteringQuantilesIQRs,
                               function(q, IQRs) names(IQRs)[IQRs > q], IQRs)
```

3.5 Differentially expressed gene calling

We perform the differential expression analysis using limma, fitting a linear model to each expression profile where the `FIR` indicator variable, the `Batch` indicator variable and the surrogate variables estimated with SVA are the explanatory variables acting as main effects.

```
> design <- cbind(mod, sv$sv)
> colnames(design) <- c("FIR", "FIRvsNoFIR",
                        paste0("Batch", seq_len(length(unique(frmaEsetFilt$Batch))-1)),
                        paste0("SV", seq_len(sv$n)))
> fit <- lmFit(frmaEsetFilt, design)
> fit$genes <- data.frame(ID=featureNames(frmaEsetFilt),
                          Symbol=getSYMBOL(featureNames(frmaEsetFilt),
                                           "org.Hs.eg.db"), stringsAsFactors=FALSE)
> fit<- eBayes(fit, trend=TRUE)
> ttAll <- topTable(fit, coef="FIRvsNoFIR", number=Inf, adjust.method="fdr")
> dim(ttAll)
[1] 20155     8
```

We select different subsets of DE genes according to different thresholds on the non-specific filtering on expression variability. Within each non-specific filter, adjusted *p*-values by FDR are re-calculated.

```
> DEgenesByFilterCutoff <- lapply(genesByFilterCutoff,
                                  function(genes, tt, FDRcutoff) {
                                    tt <- tt[tt$ID %in% genes, ]
                                    tt$adj.P.Val <- p.adjust(tt$P.Value, method="fdr")
                                    tt$ID[tt$adj.P.Val <= FDRcutoff]
                                  }, ttAll, FDRcutoff)
> nDEgenesByFilterCutoff <- sapply(DEgenesByFilterCutoff, length)
> nDEgenesByFilterCutoff
 10%  20%  30%  40%  50%  60%  70%  80%  90%
3800 3851 3892 3896 3797 3559 3178 2504 1465
> ## rev() is used here below to select highest quantile with ties
> whqmaxDE <- length(nDEgenesByFilterCutoff)-which.max(rev(nDEgenesByFilterCutoff))+1
> whqmaxDE
40%
  4
```

In Figure 9 we have all different amounts of DE genes corresponding to each variability filter as well as the cumulative distribution of IQR values previously employed to define those filters. From that figure it follows that the most stringent filter that provides the largest number of DE genes is the one that discards 40% of the genes with lower variability defined in terms IQR values.

According to Figure 9 discarding the 40genes with lower variability yields the highest number of DE genes (3896). Using this non-specific filter we re-adjust the raw *p*-values and obtain the final list of DE genes we will use in the rest of the analysis.

```
> ttAllIQRfilt <- ttAll[ttAll$ID %in% genesByFilterCutoff[[whqmaxDE]], ]
> ttAllIQRfilt$adj.P.Val <- p.adjust(ttAllIQRfilt$P.Value, method="fdr")
> dim(ttAllIQRfilt)
[1] 12093     8
> stopifnot(identical(ttAllIQRfilt$ID[ttAllIQRfilt$adj.P.Val <= FDRcutoff],
                      DEgenesByFilterCutoff[[whqmaxDE]])) ## QC
> ttDEgenes <- ttAllIQRfilt[abs(ttAllIQRfilt$logFC) > log2(fcCutoff) &
                            ttAllIQRfilt$adj.P.Val <= FDRcutoff, ]
> dim(ttDEgenes)
[1] 1097    8
```

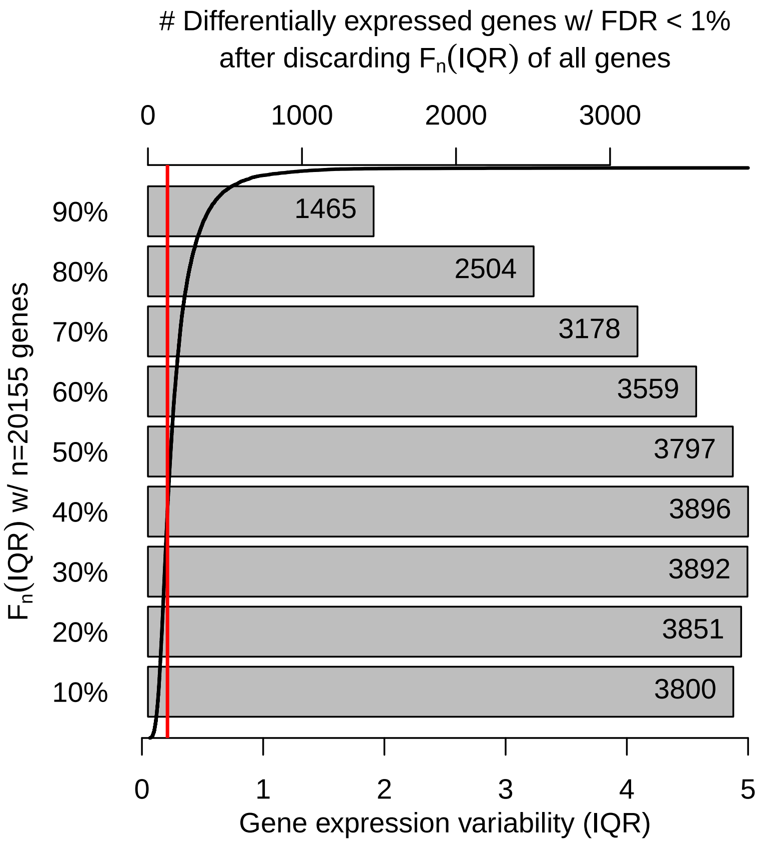

Figure 9. Empirical cumulative distribution function (CDF) of interquartile range (IQR) values for all n =20155 genes (black line) on the y-axis as function of IQR values specified on the bottom x-axis. Horizontal bars correspond to the number of differentially expressed (DE) genes (top x-axis) called at 1% FDR as function of $F_n(IQR)$, the empirical CDF in the y-axis. The empirical CDF $F_n(IQR)$ is employed to select the fraction of genes with lower variability to discard before doing the multiple test correction by FDR. Numbers inside bars indicate the precise value of the bar. In these data $F_n(IQR) = 30\%$ provides the highest statistical power to detect DE genes with FDR$< 1\%$.

```
> DEgenes <- ttDEgenes$ID
> length(DEgenes)
[1] 1097
> stopifnot(DEgenes == ttDEgenes$ID) ## QC
```

Figure 10 shows a volcano plot highlighting the subset of 1097 genes called DE.

In Tables 7 and 8 we have the list of 592 and 505 significantly up and down regulated genes, respectively, between FIR and non-FIR infants.

Table 7. List of 592 significantly upregulated genes between FIR and nonFIR affected infants. Genes are ordered by absolute $\log_2$ fold-change in expression. Columns "LFC" and "FC" contain the fold-change in $\log_2$ and absolute scales, respectively. The column "P rnk" provides the rank of each gene by raw p-value within the entire list of 1097 genes called differentially expressed.

	Entrez ID	Symbol	Full name	LFC	FC	P value	P rnk
1	6279	S100A8	S100 calcium binding protein A8	5.03	32.7	3.0560e-13	17
2	5806	PTX3	pentraxin 3, long	4.19	18.2	1.4863e-12	24
3	6283	S100A12	S100 calcium binding protein A12	3.93	15.2	5.0176e-13	20
4	6280	S100A9	S100 calcium binding protein A9	3.82	14.1	2.7028e-14	5

Table 7 – Continued on next page

Table 7 – continued from previous page

5	6347	CCL2	chemokine (C-C motif) ligand 2	3.73	13.2	1.5112e-16	1
6	6374	CXCL5	chemokine (C-X-C motif) ligand 5	3.64	12.5	5.8123e-11	75
7	4069	LYZ	lysozyme	3.56	11.8	5.7026e-11	74
8	1116	CHI3L1	chitinase 3-like 1 (cartilage glycoprotein-39)	3.46	11.0	1.6255e-09	182
9	2919	CXCL1	chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)	3.32	10.0	6.1136e-13	21
10	3576	CXCL8	chemokine (C-X-C motif) ligand 8	3.31	9.9	2.8776e-08	379
11	6372	CXCL6	chemokine (C-X-C motif) ligand 6	3.22	9.3	1.0058e-08	292
12	6648	SOD2	superoxide dismutase 2, mitochondrial	2.96	7.8	1.2876e-12	23
13	629	CFB	complement factor B	2.95	7.7	1.0873e-11	46
14	10135	NAMPT	nicotinamide phosphoribosyltransferase	2.93	7.6	1.7379e-14	3
15	7850	IL1R2	interleukin 1 receptor, type II	2.80	7.0	4.6215e-10	114
16	597	BCL2A1	BCL2-related protein A1	2.61	6.1	2.2950e-09	202
17	7128	TNFAIP3	tumor necrosis factor, alpha-induced protein 3	2.60	6.1	7.3964e-14	7
18	240	ALOX5	arachidonate 5-lipoxygenase	2.55	5.9	5.0715e-11	72
19	3371	TNC	tenascin C	2.51	5.7	2.4620e-09	206
20	7130	TNFAIP6	tumor necrosis factor, alpha-induced protein 6	2.48	5.6	1.8018e-10	102
21	5265	SERPINA1	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1	2.36	5.1	3.5532e-10	109
22	9173	IL1RL1	interleukin 1 receptor-like 1	2.19	4.6	7.9713e-06	829
23	960	CD44	CD44 molecule (Indian blood group)	2.14	4.4	4.4566e-09	237
24	199675	MCEMP1	mast cell-expressed membrane protein 1	2.08	4.2	3.8828e-11	56
25	2878	GPX3	glutathione peroxidase 3 (plasma)	2.07	4.2	4.5739e-04	1059
26	5054	SERPINE1	serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1	2.07	4.2	3.2195e-12	36
27	5743	PTGS2	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	2.04	4.1	8.6033e-11	87
28	4332	MNDA	myeloid cell nuclear differentiation antigen	2.04	4.1	2.4671e-08	366
29	929	CD14	CD14 molecule	2.01	4.0	8.9297e-10	147
30	6364	CCL20	chemokine (C-C motif) ligand 20	2.00	4.0	2.0482e-06	718
31	2920	CXCL2	chemokine (C-X-C motif) ligand 2	1.99	4.0	6.4042e-08	433
32	7412	VCAM1	vascular cell adhesion molecule 1	1.98	3.9	1.3349e-05	863
33	4067	LYN	LYN proto-oncogene, Src family tyrosine kinase	1.94	3.8	4.5876e-10	113
34	5788	PTPRC	protein tyrosine phosphatase, receptor type, C	1.93	3.8	7.8762e-10	141
35	2921	CXCL3	chemokine (C-X-C motif) ligand 3	1.91	3.8	1.4551e-06	685
36	5156	PDGFRA	platelet-derived growth factor receptor, alpha polypeptide	1.89	3.7	1.6472e-12	27
37	2643	GCH1	GTP cyclohydrolase 1	1.89	3.7	6.9300e-09	267
38	5069	PAPPA	pregnancy-associated plasma protein A, pappalysin 1	1.89	3.7	8.4548e-10	144

Table 7 – Continued on next page

Table 7 – continued from previous page

39	3976	LIF	leukemia inhibitory factor	1.88	3.7	9.0824e-10	149
40	7127	TNFAIP2	tumor necrosis factor, alpha-induced protein 2	1.87	3.6	2.4667e-08	365
41	1440	CSF3	colony stimulating factor 3 (granulocyte)	1.86	3.6	6.5639e-10	127
42	9945	GFPT2	glutamine-fructose-6-phosphate transaminase 2	1.84	3.6	2.5789e-13	15
43	4982	TNFRSF11B	tumor necrosis factor receptor superfamily, member 11b	1.84	3.6	8.7613e-11	88
44	3055	HCK	HCK proto-oncogene, Src family tyrosine kinase	1.83	3.6	7.4955e-09	272
45	1439	CSF2RB	colony stimulating factor 2 receptor, beta, low-affinity (granulocyte-macrophage)	1.83	3.5	1.2935e-09	169
46	1475	CSTA	cystatin A (stefin A)	1.82	3.5	1.4274e-09	174
47	3290	HSD11B1	hydroxysteroid (11-beta) dehydrogenase 1	1.81	3.5	9.4514e-08	465
48	5918	RARRES1	retinoic acid receptor responder (tazarotene induced) 1	1.81	3.5	1.8282e-07	524
49	3557	IL1RN	interleukin 1 receptor antagonist	1.79	3.5	2.3623e-06	725
50	25884	CHRDL2	chordin-like 2	1.79	3.5	2.2190e-08	358
51	84419	C15orf48	chromosome 15 open reading frame 48	1.78	3.4	1.9603e-05	895
52	6402	SELL	selectin L	1.77	3.4	1.0587e-07	479
53	9235	IL32	interleukin 32	1.75	3.4	6.1226e-12	40
54	29923	HILPDA	hypoxia inducible lipid droplet-associated	1.73	3.3	1.5177e-08	319
55	5473	PPBP	pro-platelet basic protein (chemokine (C-X-C motif) ligand 7)	1.72	3.3	8.0501e-08	453
56	715	C1R	complement component 1, r subcomponent	1.71	3.3	1.7813e-12	29
57	9021	SOCS3	suppressor of cytokine signaling 3	1.70	3.3	8.9597e-10	148
58	2180	ACSL1	acyl-CoA synthetase long-chain family member 1	1.70	3.3	2.0251e-11	52
59	7037	TFRC	transferrin receptor	1.70	3.3	9.0516e-09	286
60	7357	UGCG	UDP-glucose ceramide glucosyltransferase	1.70	3.3	9.9466e-16	2
61	4688	NCF2	neutrophil cytosolic factor 2	1.68	3.2	6.9667e-07	623
62	2357	FPR1	formyl peptide receptor 1	1.68	3.2	3.5700e-11	54
63	3684	ITGAM	integrin, alpha M (complement component 3 receptor 3 subunit)	1.68	3.2	9.8127e-08	468
64	7076	TIMP1	TIMP metallopeptidase inhibitor 1	1.67	3.2	1.9846e-13	11
65	3569	IL6	interleukin 6	1.66	3.2	5.3242e-06	797
66	2215	FCGR3B	Fc fragment of IgG, low affinity IIIb, receptor (CD16b)	1.65	3.1	9.4431e-09	290
67	59269	HIVEP3	human immunodeficiency virus type I enhancer binding protein 3	1.65	3.1	2.0118e-08	345
68	1326	MAP3K8	mitogen-activated protein kinase kinase kinase 8	1.64	3.1	1.1202e-09	159
69	79689	STEAP4	STEAP family member 4	1.64	3.1	1.1587e-08	299
70	1117	CHI3L2	chitinase 3-like 2	1.63	3.1	1.4136e-06	682
71	23516	SLC39A14	solute carrier family 39 (zinc transporter), member 14	1.62	3.1	2.2316e-13	13

Table 7 – Continued on next page

Table 7 – continued from previous page

72	4311	MME	membrane metallo-endopeptidase	1.62	3.1	2.3551e-08	362
73	2219	FCN1	ficolin (collagen/fibrinogen domain containing) 1	1.62	3.1	3.7870e-08	391
74	3554	IL1R1	interleukin 1 receptor, type I	1.61	3.1	1.1987e-10	96
75	4318	MMP9	matrix metallopeptidase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)	1.61	3.0	2.2601e-09	201
76	64332	NFKBIZ	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta	1.60	3.0	3.0204e-12	34
77	133	ADM	adrenomedullin	1.60	3.0	1.1708e-06	664
78	5320	PLA2G2A	phospholipase A2, group IIA (platelets, synovial fluid)	1.59	3.0	1.5613e-06	693
79	1520	CTSS	cathepsin S	1.58	3.0	6.1642e-07	617
80	7805	LAPTM5	lysosomal protein transmembrane 5	1.57	3.0	6.2766e-09	262
81	2268	FGR	FGR proto-oncogene, Src family tyrosine kinase	1.55	2.9	5.2624e-10	121
82	56729	RETN	resistin	1.55	2.9	1.7915e-09	186
83	602	BCL3	B-cell CLL/lymphoma 3	1.55	2.9	8.9480e-14	8
84	4542	MYO1F	myosin IF	1.54	2.9	9.9507e-11	92
85	1441	CSF3R	colony stimulating factor 3 receptor (granulocyte)	1.54	2.9	1.9883e-08	344
86	50486	G0S2	G0/G1 switch 2	1.54	2.9	1.4081e-07	505
87	3689	ITGB2	integrin, beta 2 (complement component 3 receptor 3 and 4 subunit)	1.53	2.9	6.6581e-08	437
88	79930	DOK3	docking protein 3	1.53	2.9	9.2441e-09	288
89	5552	SRGN	serglycin	1.52	2.9	1.1839e-09	164
90	7852	CXCR4	chemokine (C-X-C motif) receptor 4	1.51	2.8	3.4421e-04	1046
91	5329	PLAUR	plasminogen activator, urokinase receptor	1.50	2.8	9.2817e-12	43
92	64092	SAMSN1	SAM domain, SH3 domain and nuclear localization signals 1	1.50	2.8	7.3542e-07	625
93	56999	ADAMTS9	ADAM metallopeptidase with thrombospondin type 1 motif, 9	1.50	2.8	5.3105e-11	73
94	5270	SERPINE2	serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 2	1.49	2.8	1.9335e-09	194
95	27242	TNFRSF21	tumor necrosis factor receptor superfamily, member 21	1.48	2.8	3.6846e-09	229
96	7305	TYROBP	TYRO protein tyrosine kinase binding protein	1.47	2.8	6.3263e-08	429
97	64581	CLEC7A	C-type lectin domain family 7, member A	1.47	2.8	4.7277e-08	410
98	1890	TYMP	thymidine phosphorylase	1.46	2.8	2.4787e-13	14
99	140885	SIRPA	signal-regulatory protein alpha	1.45	2.7	8.3689e-11	85
100	5209	PFKFB3	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3	1.44	2.7	6.0030e-09	255
101	3383	ICAM1	intercellular adhesion molecule 1	1.44	2.7	1.0433e-09	157
102	728	C5AR1	complement component 5a receptor 1	1.43	2.7	2.9651e-07	557
103	54898	ELOVL2	ELOVL fatty acid elongase 2	1.42	2.7	6.2446e-12	41
104	11213	IRAK3	interleukin-1 receptor-associated kinase 3	1.42	2.7	1.2649e-09	167

Table 7 – Continued on next page

Table 7 – continued from previous page

105	123	PLIN2	perilipin 2	1.39	2.6	1.0321e-07	476
106	3718	JAK3	Janus kinase 3	1.39	2.6	4.9021e-13	19
107	1052	CEBPD	CCAAT/enhancer binding protein (C/EBP), delta	1.39	2.6	9.1179e-10	150
108	51316	PLAC8	placenta-specific 8	1.38	2.6	9.1238e-08	462
109	6781	STC1	stanniocalcin 1	1.38	2.6	1.0412e-05	850
110	7133	TNFRSF1B	tumor necrosis factor receptor superfamily, member 1B	1.37	2.6	2.3175e-09	203
111	727936	GXYLT2	glucoside xylosyltransferase 2	1.37	2.6	2.1940e-08	355
112	51129	ANGPTL4	angiopoietin-like 4	1.36	2.6	1.5731e-07	516
113	219972	MPEG1	macrophage expressed 1	1.36	2.6	3.0200e-06	748
114	6286	S100P	S100 calcium binding protein P	1.36	2.6	1.6657e-05	880
115	8972	MGAM	maltase-glucoamylase (alpha-glucosidase)	1.36	2.6	8.6381e-08	458
116	6653	SORL1	sortilin-related receptor, L(DLR class) A repeats containing	1.36	2.6	1.0532e-06	654
117	4084	MXD1	MAX dimerization protein 1	1.36	2.6	1.4538e-08	315
118	23514	SPIDR	scaffolding protein involved in DNA repair	1.35	2.6	4.7056e-11	67
119	330	BIRC3	baculoviral IAP repeat containing 3	1.35	2.5	5.2686e-06	795
120	24145	PANX1	pannexin 1	1.35	2.5	6.3084e-10	126
121	11027	LILRA2	leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 2	1.34	2.5	2.4986e-09	207
122	290	ANPEP	alanyl (membrane) aminopeptidase	1.34	2.5	5.0953e-08	415
123	2207	FCER1G	Fc fragment of IgE, high affinity I, receptor for; gamma polypeptide	1.32	2.5	5.7669e-08	422
124	1827	RCAN1	regulator of calcineurin 1	1.32	2.5	8.4228e-09	280
125	2124	EVI2B	ecotropic viral integration site 2B	1.32	2.5	7.6602e-06	824
126	4689	NCF4	neutrophil cytosolic factor 4, 40kDa	1.32	2.5	8.9140e-10	146
127	11151	CORO1A	coronin, actin binding protein, 1A	1.32	2.5	9.5430e-10	152
128	7078	TIMP3	TIMP metallopeptidase inhibitor 3	1.31	2.5	5.3421e-07	607
129	729230	CCR2	chemokine (C-C motif) receptor 2	1.31	2.5	4.2369e-05	940
130	9535	GMFG	glia maturation factor, gamma	1.31	2.5	3.7306e-09	231
131	51279	C1RL	complement component 1, r subcomponent-like	1.30	2.5	1.3345e-08	309
132	5055	SERPINB2	serpin peptidase inhibitor, clade B (ovalbumin), member 2	1.30	2.5	2.8186e-04	1031
133	9536	PTGES	prostaglandin E synthase	1.30	2.5	3.0955e-08	385
134	10288	LILRB2	leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 2	1.29	2.5	2.1388e-08	348
135	6401	SELE	selectin E	1.29	2.4	4.3298e-04	1056
136	7057	THBS1	thrombospondin 1	1.27	2.4	2.4122e-09	205
137	54541	DDIT4	DNA-damage-inducible transcript 4	1.27	2.4	3.0755e-06	751
138	23601	CLEC5A	C-type lectin domain family 5, member A	1.27	2.4	2.0302e-07	531
139	6303	SAT1	spermidine/spermine N1-acetyltransferase 1	1.26	2.4	2.9119e-09	216
140	5880	RAC2	ras-related C3 botulinum toxin substrate 2 (rho family, small GTP binding protein Rac2)	1.25	2.4	2.9295e-09	217

Table 7 – Continued on next page

Table 7 – continued from previous page

141	19	ABCA1	ATP-binding cassette, sub-family A (ABC1), member 1	1.25	2.4	6.7423e-10	128
142	716	C1S	complement component 1, s subcomponent	1.25	2.4	4.9732e-10	116
143	366	AQP9	aquaporin 9	1.23	2.4	2.3158e-07	541
144	1301	COL11A1	collagen, type XI, alpha 1	1.23	2.3	2.7346e-06	738
145	5359	PLSCR1	phospholipid scramblase 1	1.23	2.3	6.4080e-09	265
146	3937	LCP2	lymphocyte cytosolic protein 2 (SH2 domain containing leukocyte protein of 76kDa)	1.22	2.3	6.2864e-05	960
147	6850	SYK	spleen tyrosine kinase	1.22	2.3	1.2600e-08	304
148	4501	MT1X	metallothionein 1X	1.22	2.3	1.1749e-06	666
149	6556	SLC11A1	solute carrier family 11 (proton-coupled divalent metal ion transporter), member 1	1.22	2.3	4.9878e-08	414
150	6515	SLC2A3	solute carrier family 2 (facilitated glucose transporter), member 3	1.22	2.3	1.9222e-07	527
151	397	ARHGDIB	Rho GDP dissociation inhibitor (GDI) beta	1.21	2.3	1.7325e-09	184
152	25801	GCA	grancalcin, EF-hand calcium binding protein	1.21	2.3	1.0152e-07	473
153	120425	AMICA1	adhesion molecule, interacts with CXADR antigen 1	1.21	2.3	5.6137e-08	421
154	3772	KCNJ15	potassium inwardly-rectifying channel, subfamily J, member 15	1.21	2.3	2.6131e-07	549
155	9180	OSMR	oncostatin M receptor	1.21	2.3	1.8432e-09	188
156	7378	UPP1	uridine phosphorylase 1	1.21	2.3	6.2550e-10	124
157	2532	ACKR1	atypical chemokine receptor 1 (Duffy blood group)	1.21	2.3	1.0467e-06	653
158	10562	OLFM4	olfactomedin 4	1.20	2.3	3.2296e-04	1041
159	115123	MARCH3	membrane-associated ring finger (C3HC4) 3, E3 ubiquitin protein ligase	1.20	2.3	1.4504e-10	99
160	1043	CD52	CD52 molecule	1.20	2.3	1.5139e-04	1007
161	3037	HAS2	hyaluronan synthase 2	1.20	2.3	1.2493e-06	671
162	2321	FLT1	fms-related tyrosine kinase 1	1.20	2.3	2.1533e-08	350
163	128346	C1orf162	chromosome 1 open reading frame 162	1.19	2.3	1.0796e-07	480
164	10261	IGSF6	immunoglobulin superfamily, member 6	1.19	2.3	1.0914e-06	656
165	55022	PID1	phosphotyrosine interaction domain containing 1	1.19	2.3	1.9257e-08	342
166	27286	SRPX2	sushi-repeat containing protein, X-linked 2	1.19	2.3	1.0262e-07	475
167	2152	F3	coagulation factor III (thromboplastin, tissue factor)	1.19	2.3	1.1455e-09	160
168	57126	CD177	CD177 molecule	1.18	2.3	1.6232e-08	325
169	55784	MCTP2	multiple C2 domains, transmembrane 2	1.18	2.3	3.1566e-09	224
170	3385	ICAM3	intercellular adhesion molecule 3	1.17	2.2	7.1569e-08	444
171	7940	LST1	leukocyte specific transcript 1	1.17	2.2	6.3543e-09	263
172	3310	HSPA6	heat shock 70kDa protein 6 (HSP70B')	1.16	2.2	2.2941e-07	540
173	5352	PLOD2	procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2	1.16	2.2	2.1215e-10	105
174	10468	FST	follistatin	1.16	2.2	2.9733e-05	922

Table 7 – Continued on next page

Table 7 – continued from previous page

175	118788	PIK3AP1	phosphoinositide-3-kinase adaptor protein 1	1.16	2.2	1.0112e-07	470
176	604	BCL6	B-cell CLL/lymphoma 6	1.16	2.2	3.6837e-08	389
177	5618	PRLR	prolactin receptor	1.16	2.2	6.0810e-06	803
178	51311	TLR8	toll-like receptor 8	1.15	2.2	4.1359e-07	585
179	84935	MEDAG	mesenteric estrogen-dependent adipogenesis	1.15	2.2	1.1816e-07	489
180	5791	PTPRE	protein tyrosine phosphatase, receptor type, E	1.15	2.2	4.6040e-07	594
181	7422	VEGFA	vascular endothelial growth factor A	1.15	2.2	4.9246e-09	242
182	22936	ELL2	elongation factor, RNA polymerase II, 2	1.14	2.2	1.5323e-07	515
183	22822	PHLDA1	pleckstrin homology-like domain, family A, member 1	1.14	2.2	9.1469e-09	287
184	25939	SAMHD1	SAM domain and HD domain 1	1.14	2.2	2.1132e-05	899
185	92610	TIFA	TRAF-interacting protein with forkhead-associated domain	1.14	2.2	2.2634e-07	538
186	1604	CD55	CD55 molecule, decay accelerating factor for complement (Cromer blood group)	1.14	2.2	2.1715e-08	351
187	3099	HK2	hexokinase 2	1.13	2.2	4.3432e-07	589
188	29887	SNX10	sorting nexin 10	1.13	2.2	3.5355e-06	765
189	1230	CCR1	chemokine (C-C motif) receptor 1	1.13	2.2	6.1110e-06	804
190	10560	SLC19A2	solute carrier family 19 (thiamine transporter), member 2	1.13	2.2	3.6151e-09	228
191	1806	DPYD	dihydropyrimidine dehydrogenase	1.13	2.2	3.4360e-05	933
192	54518	APBB1IP	amyloid beta (A4) precursor protein-binding, family B, member 1 interacting protein	1.13	2.2	8.2492e-08	454
193	79887	PLBD1	phospholipase B domain containing 1	1.12	2.2	1.9152e-07	526
194	6696	SPP1	secreted phosphoprotein 1	1.12	2.2	2.7098e-06	737
195	272	AMPD3	adenosine monophosphate deaminase 3	1.12	2.2	9.0149e-09	285
196	113730	KLHDC7B	kelch domain containing 7B	1.12	2.2	7.1421e-10	135
197	963	CD53	CD53 molecule	1.12	2.2	2.6948e-06	735
198	4502	MT2A	metallothionein 2A	1.11	2.2	3.6425e-08	388
199	1545	CYP1B1	cytochrome P450, family 1, subfamily B, polypeptide 1	1.11	2.2	8.0221e-05	973
200	9589	WTAP	Wilms tumor 1 associated protein	1.09	2.1	1.0299e-06	652
201	9332	CD163	CD163 molecule	1.09	2.1	5.0446e-05	948
202	2212	FCGR2A	Fc fragment of IgG, low affinity IIa, receptor (CD32)	1.09	2.1	8.2286e-06	832
203	79827	CLMP	CXADR-like membrane protein	1.08	2.1	6.1027e-09	256
204	3553	IL1B	interleukin 1, beta	1.08	2.1	5.6223e-07	611
205	718	C3	complement component 3	1.08	2.1	2.5218e-05	914
206	81704	DOCK8	dedicator of cytokinesis 8	1.07	2.1	3.2903e-04	1042
207	6581	SLC22A3	solute carrier family 22 (organic cation transporter), member 3	1.07	2.1	3.3285e-09	226
208	84106	PRAM1	PML-RARA regulated adaptor molecule 1	1.07	2.1	1.9035e-09	191
209	1805	DPT	dermatopontin	1.07	2.1	3.1671e-03	1096
210	8870	IER3	immediate early response 3	1.07	2.1	1.8198e-10	103

Table 7 – Continued on next page

Table 7 – continued from previous page

211	241	ALOX5AP	arachidonate 5-lipoxygenase-activating protein	1.07	2.1	2.4133e-06	727
212	6503	SLA	Src-like-adaptor	1.06	2.1	8.2381e-09	278
213	11182	SLC2A6	solute carrier family 2 (facilitated glucose transporter), member 6	1.06	2.1	1.7941e-08	338
214	4837	NNMT	nicotinamide N-methyltransferase	1.05	2.1	8.4249e-11	86
215	9123	SLC16A3	solute carrier family 16 (monocarboxylate transporter), member 3	1.05	2.1	5.6118e-09	248
216	2752	GLUL	glutamate-ammonia ligase	1.05	2.1	4.5249e-06	782
217	83706	FERMT3	fermitin family member 3	1.04	2.1	9.3892e-10	151
218	12	SERPINA3	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3	1.04	2.1	3.7920e-08	392
219	4155	MBP	myelin basic protein	1.04	2.1	2.0766e-07	533
220	2123	EVI2A	ecotropic viral integration site 2A	1.04	2.1	2.8809e-05	920
221	51312	SLC25A37	solute carrier family 25 (mitochondrial iron transporter), member 37	1.04	2.0	1.8438e-06	707
222	3101	HK3	hexokinase 3 (white cell)	1.03	2.0	6.8310e-10	131
223	23213	SULF1	sulfatase 1	1.03	2.0	2.5348e-08	368
224	10673	TNFSF13B	tumor necrosis factor (ligand) superfamily, member 13b	1.03	2.0	2.8841e-06	744
225	64333	ARHGAP9	Rho GTPase activating protein 9	1.03	2.0	1.3472e-08	311
226	81671	VMP1	vacuole membrane protein 1	1.03	2.0	1.4457e-05	869
227	10630	PDPN	podoplanin	1.03	2.0	1.2743e-08	306
228	1051	CEBPB	CCAAT/enhancer binding protein (C/EBP), beta	1.03	2.0	5.0076e-10	117
229	84708	LNX1	ligand of numb-protein X 1, E3 ubiquitin protein ligase	1.02	2.0	5.4001e-07	608
230	4651	MYO10	myosin X	1.02	2.0	4.6749e-11	66
231	3659	IRF1	interferon regulatory factor 1	1.02	2.0	2.9628e-07	556
232	222487	GPR97	G protein-coupled receptor 97	1.02	2.0	7.7835e-09	275
233	257106	ARHGAP30	Rho GTPase activating protein 30	1.02	2.0	9.2872e-08	464
234	60675	PROK2	prokineticin 2	1.02	2.0	1.1824e-05	856
235	5971	RELB	v-rel avian reticuloendotheliosis viral oncogene homolog B	1.01	2.0	1.6377e-09	183
236	4496	MT1H	metallothionein 1H	1.01	2.0	3.3999e-07	567
237	10095	ARPC1B	actin related protein 2/3 complex, subunit 1B, 41kDa	1.01	2.0	1.1807e-10	95
238	8942	KYNU	kynureninase	1.01	2.0	6.8167e-05	966
239	137835	TMEM71	transmembrane protein 71	1.01	2.0	2.8359e-05	919
240	5341	PLEK	pleckstrin	1.00	2.0	7.8477e-08	450
241	57758	SCUBE2	signal peptide, CUB domain, EGF-like 2	1.00	2.0	6.3681e-08	430
242	2192	FBLN1	fibulin 1	1.00	2.0	5.1500e-07	605
243	80149	ZC3H12A	zinc finger CCCH-type containing 12A	1.00	2.0	1.9623e-09	195
244	25975	EGFL6	EGF-like-domain, multiple 6	1.00	2.0	9.3219e-04	1077
245	1809	DPYSL3	dihydropyrimidinase-like 3	0.99	2.0	5.2704e-07	606
246	55332	DRAM1	DNA-damage regulated autophagy modulator 1	0.99	2.0	2.7371e-08	374
247	50856	CLEC4A	C-type lectin domain family 4, member A	0.99	2.0	7.6848e-04	1070
248	6999	TDO2	tryptophan 2,3-dioxygenase	0.99	2.0	4.9057e-07	602

Table 7 – Continued on next page

Table 7 – continued from previous page

249	54102	CLIC6	chloride intracellular channel 6	0.99	2.0	1.8407e-06	705
250	3934	LCN2	lipocalin 2	0.99	2.0	8.4476e-06	833
251	6446	SGK1	serum/glucocorticoid regulated kinase 1	0.98	2.0	4.5917e-08	409
252	642273	FAM110C	family with sequence similarity 110, member C	0.98	2.0	5.4506e-06	799
253	8935	SKAP2	src kinase associated phosphoprotein 2	0.98	2.0	1.5571e-04	1009
254	146722	CD300LF	CD300 molecule-like family member f	0.98	2.0	3.1559e-06	755
255	3486	IGFBP3	insulin-like growth factor binding protein 3	0.98	2.0	8.2929e-05	978
256	1311	COMP	cartilage oligomeric matrix protein	0.98	2.0	4.4003e-05	943
257	969	CD69	CD69 molecule	0.98	2.0	3.2163e-03	1097
258	54210	TREM1	triggering receptor expressed on myeloid cells 1	0.98	2.0	4.2099e-08	402
259	4792	NFKBIA	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha	0.97	2.0	6.3246e-08	428
260	7421	VDR	vitamin D (1,25- dihydroxyvitamin D3) receptor	0.97	2.0	3.3768e-07	566
261	80183	KIAA0226L	KIAA0226-like	0.97	2.0	6.9787e-06	817
262	25953	PNKD	paroxysmal nonkinesigenic dyskinesia	0.97	2.0	7.3820e-09	270
263	79180	EFHD2	EF-hand domain family, member D2	0.97	2.0	1.6337e-08	326
264	4791	NFKB2	nuclear factor of kappa light polypeptide gene enhancer in B-cells 2 (p49/p100)	0.97	2.0	5.1431e-07	604
265	117247	SLC16A10	solute carrier family 16 (aromatic amino acid transporter), member 10	0.97	2.0	3.2797e-06	760
266	2769	GNA15	guanine nucleotide binding protein (G protein), alpha 15 (Gq class)	0.96	1.9	2.1395e-08	349
267	6004	RGS16	regulator of G-protein signaling 16	0.96	1.9	5.1065e-06	792
268	3587	IL10RA	interleukin 10 receptor, alpha	0.95	1.9	9.2274e-07	644
269	7462	LAT2	linker for activation of T cells family, member 2	0.95	1.9	5.7693e-07	614
270	121512	FGD4	FYVE, RhoGEF and PH domain containing 4	0.95	1.9	2.5551e-06	731
271	4494	MT1F	metallothionein 1F	0.95	1.9	5.9469e-07	616
272	9766	KIAA0247	KIAA0247	0.95	1.9	6.8377e-11	77
273	168455	CCDC71L	coiled-coil domain containing 71-like	0.95	1.9	8.7905e-09	284
274	85027	SMIM3	small integral membrane protein 3	0.94	1.9	1.6543e-08	329
275	4015	LOX	lysyl oxidase	0.94	1.9	1.0097e-05	846
276	8875	VNN2	vanin 2	0.94	1.9	5.6355e-09	249
277	4495	MT1G	metallothionein 1G	0.94	1.9	2.3848e-07	543
278	4257	MGST1	microsomal glutathione S-transferase 1	0.93	1.9	1.8919e-05	891
279	1512	CTSH	cathepsin H	0.93	1.9	1.3572e-05	864
280	645745	MT1HL1	metallothionein 1H-like 1	0.93	1.9	3.9127e-07	581
281	10318	TNIP1	TNFAIP3 interacting protein 1	0.93	1.9	1.4792e-09	177
282	3696	ITGB8	integrin, beta 8	0.93	1.9	2.7359e-05	918
283	8673	VAMP8	vesicle-associated membrane protein 8	0.93	1.9	3.8388e-07	578
284	58475	MS4A7	membrane-spanning 4-domains, subfamily A, member 7	0.93	1.9	2.9340e-04	1034
285	83593	RASSF5	Ras association (RalGDS/AF-6) domain family member 5	0.93	1.9	1.2682e-06	672

Table 7 – Continued on next page

Table 7 – continued from previous page

286	5771	PTPN2	protein tyrosine phosphatase, non-receptor type 2	0.93	1.9	1.4067e-07	504
287	8553	BHLHE40	basic helix-loop-helix family, member e40	0.92	1.9	4.5887e-06	784
288	64859	NABP1	nucleic acid binding protein 1	0.92	1.9	4.6604e-06	788
289	6404	SELPLG	selectin P ligand	0.92	1.9	1.4744e-07	510
290	2619	GAS1	growth arrest-specific 1	0.92	1.9	1.3308e-07	495
291	3936	LCP1	lymphocyte cytosolic protein 1 (L-plastin)	0.92	1.9	4.2187e-06	776
292	124460	SNX20	sorting nexin 20	0.92	1.9	1.4891e-07	511
293	8741	TNFSF13	tumor necrosis factor (ligand) superfamily, member 13	0.91	1.9	9.0116e-07	641
294	4778	NFE2	nuclear factor, erythroid 2	0.91	1.9	2.2374e-06	722
295	912	CD1D	CD1d molecule	0.91	1.9	6.6788e-06	812
296	10221	TRIB1	tribbles pseudokinase 1	0.91	1.9	1.4159e-08	314
297	10008	KCNE3	potassium voltage-gated channel, Isk-related family, member 3	0.91	1.9	1.0528e-07	477
298	29108	PYCARD	PYD and CARD domain containing	0.91	1.9	1.3576e-07	498
299	23526	HMHA1	histocompatibility (minor) HA-1	0.91	1.9	8.7560e-09	283
300	3572	IL6ST	interleukin 6 signal transducer	0.90	1.9	4.1063e-10	111
301	5008	OSM	oncostatin M	0.90	1.9	1.6639e-07	519
302	11010	GLIPR1	GLI pathogenesis-related 1	0.90	1.9	3.9492e-07	582
303	8840	WISP1	WNT1 inducible signaling pathway protein 1	0.90	1.9	1.9570e-04	1018
304	135112	NCOA7	nuclear receptor coactivator 7	0.90	1.9	4.7492e-08	411
305	30001	ERO1L	ERO1-like (S. cerevisiae)	0.90	1.9	3.0179e-06	747
306	79627	OGFRL1	opioid growth factor receptor-like 1	0.90	1.9	2.0307e-05	896
307	9473	THEMIS2	thymocyte selection associated family member 2	0.90	1.9	3.1440e-06	754
308	4891	SLC11A2	solute carrier family 11 (proton-coupled divalent metal ion transporter), member 2	0.90	1.9	1.7783e-08	337
309	202	AIM1	absent in melanoma 1	0.90	1.9	2.1731e-05	905
310	284207	METRNL	meteorin, glial cell differentiation regulator-like	0.89	1.9	2.9431e-09	218
311	9966	TNFSF15	tumor necrosis factor (ligand) superfamily, member 15	0.89	1.9	8.7964e-05	979
312	89796	NAV1	neuron navigator 1	0.89	1.9	5.6501e-06	800
313	115908	CTHRC1	collagen triple helix repeat containing 1	0.89	1.8	2.9779e-05	923
314	860	RUNX2	runt-related transcription factor 2	0.88	1.8	1.5119e-06	689
315	5293	PIK3CD	phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit delta	0.88	1.8	1.0621e-08	293
316	415116	PIM3	Pim-3 proto-oncogene, serine/threonine kinase	0.88	1.8	7.3272e-11	82
317	1902	LPAR1	lysophosphatidic acid receptor 1	0.87	1.8	2.6974e-06	736
318	9595	CYTIP	cytohesin 1 interacting protein	0.87	1.8	9.1538e-05	981
319	9060	PAPSS2	3'-phosphoadenosine 5'-phosphosulfate synthase 2	0.87	1.8	1.8179e-07	523
320	4319	MMP10	matrix metallopeptidase 10 (stromelysin 2)	0.87	1.8	1.7237e-04	1015
321	56243	KIAA1217	KIAA1217	0.87	1.8	8.4864e-04	1074

Table 7 – Continued on next page

Table 7 – continued from previous page

322	91653	BOC	BOC cell adhesion associated, oncogene regulated	0.87	1.8	1.0121e-05	847
323	101	ADAM8	ADAM metallopeptidase domain 8	0.86	1.8	6.1560e-08	425
324	53829	P2RY13	purinergic receptor P2Y, G-protein coupled, 13	0.86	1.8	1.8747e-05	890
325	8660	IRS2	insulin receptor substrate 2	0.86	1.8	3.5604e-05	935
326	2023	ENO1	enolase 1, (alpha)	0.86	1.8	4.1942e-07	588
327	6813	STXBP2	syntaxin binding protein 2	0.86	1.8	5.7731e-09	252
328	3726	JUNB	jun B proto-oncogene	0.85	1.8	3.9769e-08	396
329	199	AIF1	allograft inflammatory factor 1	0.85	1.8	5.4352e-08	417
330	6355	CCL8	chemokine (C-C motif) ligand 8	0.85	1.8	6.3700e-08	431
331	58484	NLRC4	NLR family, CARD domain containing 4	0.85	1.8	2.9898e-06	746
332	54504	CPVL	carboxypeptidase, vitellogenic-like	0.85	1.8	1.7454e-05	885
333	23135	KDM6B	lysine (K)-specific demethylase 6B	0.84	1.8	1.4029e-07	502
334	10507	SEMA4D	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4D	0.84	1.8	3.0227e-06	749
335	83937	RASSF4	Ras association (RalGDS/AF-6) domain family member 4	0.84	1.8	2.3378e-07	542
336	1992	SERPINB1	serpin peptidase inhibitor, clade B (ovalbumin), member 1	0.84	1.8	7.7976e-08	449
337	255488	RNF144B	ring finger protein 144B	0.84	1.8	2.0187e-06	716
338	3687	ITGAX	integrin, alpha X (complement component 3 receptor 4 subunit)	0.84	1.8	9.0708e-08	461
339	79772	MCTP1	multiple C2 domains, transmembrane 1	0.83	1.8	1.0461e-03	1082
340	3459	IFNGR1	interferon gamma receptor 1	0.83	1.8	3.7832e-06	769
341	664	BNIP3	BCL2/adenovirus E1B 19kDa interacting protein 3	0.83	1.8	4.0594e-05	937
342	409	ARRB2	arrestin, beta 2	0.83	1.8	1.0234e-07	474
343	4046	LSP1	lymphocyte-specific protein 1	0.83	1.8	9.1279e-07	643
344	29931	LINC00312	long intergenic non-protein coding RNA 312	0.82	1.8	5.7113e-07	612
345	30011	SH3KBP1	SH3-domain kinase binding protein 1	0.82	1.8	8.6169e-07	640
346	29015	SLC43A3	solute carrier family 43, member 3	0.82	1.8	1.3599e-08	312
347	284996	RNF149	ring finger protein 149	0.82	1.8	4.8905e-06	790
348	56938	ARNTL2	aryl hydrocarbon receptor nuclear translocator-like 2	0.82	1.8	1.5298e-08	320
349	4170	MCL1	myeloid cell leukemia 1	0.82	1.8	4.9562e-11	69
350	6352	CCL5	chemokine (C-C motif) ligand 5	0.82	1.8	2.7022e-04	1028
351	5337	PLD1	phospholipase D1, phosphatidylcholine-specific	0.82	1.8	1.2390e-04	996
352	624	BDKRB2	bradykinin receptor B2	0.82	1.8	2.3694e-06	726
353	89790	SIGLEC10	sialic acid binding Ig-like lectin 10	0.81	1.8	2.7611e-06	740
354	23643	LY96	lymphocyte antigen 96	0.81	1.8	1.1982e-04	992
355	3059	HCLS1	hematopoietic cell-specific Lyn substrate 1	0.81	1.8	5.2521e-05	952
356	3579	CXCR2	chemokine (C-X-C motif) receptor 2	0.81	1.8	3.8200e-06	771
357	10154	PLXNC1	plexin C1	0.81	1.8	2.6157e-04	1027
358	90853	SPOCD1	SPOC domain containing 1	0.81	1.8	3.6061e-07	573
359	8506	CNTNAP1	contactin associated protein 1	0.81	1.8	4.4017e-12	38

Table 7 – Continued on next page

Table 7 – continued from previous page

360	11026	LILRA3	leukocyte immunoglobulin-like receptor, subfamily A (without TM domain), member 3	0.81	1.8	1.3516e-04	998
361	5464	PPA1	pyrophosphatase (inorganic) 1	0.81	1.8	1.2230e-10	97
362	3455	IFNAR2	interferon (alpha, beta and omega) receptor 2	0.81	1.7	7.1080e-10	134
363	23764	MAFF	v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog F	0.81	1.7	1.1231e-04	989
364	115701	ALPK2	alpha-kinase 2	0.80	1.7	4.6308e-06	787
365	1536	CYBB	cytochrome b-245, beta polypeptide	0.80	1.7	2.2059e-06	720
366	7706	TRIM25	tripartite motif containing 25	0.80	1.7	7.0936e-08	443
367	80315	CPEB4	cytoplasmic polyadenylation element binding protein 4	0.80	1.7	6.3807e-05	962
368	4057	LTF	lactotransferrin	0.80	1.7	5.8989e-04	1064
369	6675	UAP1	UDP-N-acetylglucosamine pyrophosphorylase 1	0.80	1.7	2.6428e-08	372
370	8843	HCAR3	hydroxycarboxylic acid receptor 3	0.79	1.7	1.9105e-04	1016
371	1075	CTSC	cathepsin C	0.79	1.7	3.3325e-06	761
372	116844	LRG1	leucine-rich alpha-2-glycoprotein 1	0.79	1.7	7.0802e-08	441
373	3673	ITGA2	integrin, alpha 2 (CD49B, alpha 2 subunit of VLA-2 receptor)	0.79	1.7	3.9683e-05	936
374	3908	LAMA2	laminin, alpha 2	0.79	1.7	2.1803e-08	353
375	221178	SPATA13	spermatogenesis associated 13	0.79	1.7	2.2935e-06	723
376	822	CAPG	capping protein (actin filament), gelsolin-like	0.79	1.7	1.6058e-06	696
377	10659	CELF2	CUGBP, Elav-like family member 2	0.79	1.7	8.9747e-06	838
378	951	CD37	CD37 molecule	0.79	1.7	7.8886e-06	828
379	4609	MYC	v-myc avian myelocytomatosis viral oncogene homolog	0.79	1.7	1.4050e-10	98
380	157506	RDH10	retinol dehydrogenase 10 (all-trans)	0.79	1.7	1.9352e-04	1017
381	1839	HBEGF	heparin-binding EGF-like growth factor	0.79	1.7	5.2640e-04	1062
382	283209	PGM2L1	phosphoglucomutase 2-like 1	0.78	1.7	3.4792e-05	934
383	10320	IKZF1	IKAROS family zinc finger 1 (Ikaros)	0.78	1.7	1.2462e-05	859
384	11067	C10orf10	chromosome 10 open reading frame 10	0.78	1.7	1.7244e-05	884
385	84695	LOXL3	lysyl oxidase-like 3	0.77	1.7	8.4110e-08	456
386	10312	TCIRG1	T-cell, immune regulator 1, ATPase, H+ transporting, lysosomal V0 subunit A3	0.77	1.7	1.6230e-07	518
387	4327	MMP19	matrix metallopeptidase 19	0.77	1.7	7.9120e-09	276
388	84002	B3GNT5	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 5	0.77	1.7	1.6854e-04	1013
389	148932	MOB3C	MOB kinase activator 3C	0.77	1.7	7.0859e-08	442
390	79168	LILRA6	leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 6	0.77	1.7	1.6458e-06	698
391	2701	GJA4	gap junction protein, alpha 4, 37kDa	0.77	1.7	8.8012e-10	145
392	6195	RPS6KA1	ribosomal protein S6 kinase, 90kDa, polypeptide 1	0.77	1.7	1.3848e-07	500
393	634	CEACAM1	carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein)	0.77	1.7	3.0578e-06	750
394	118932	ANKRD22	ankyrin repeat domain 22	0.77	1.7	1.8537e-05	889

Table 7 – Continued on next page

Table 7 – continued from previous page

395	2331	FMOD	fibromodulin	0.77	1.7	1.4593e-05	871
396	64115	C10orf54	chromosome 10 open reading frame 54	0.76	1.7	1.3795e-05	865
397	6039	RNASE6	ribonuclease, RNase A family, k6	0.76	1.7	2.8360e-04	1033
398	1362	CPD	carboxypeptidase D	0.76	1.7	1.5147e-08	318
399	4790	NFKB1	nuclear factor of kappa light polypeptide gene enhancer in B-cells 1	0.76	1.7	4.4009e-07	592
400	64127	NOD2	nucleotide-binding oligomerization domain containing 2	0.76	1.7	1.7237e-05	883
401	51616	TAF9B	TAF9B RNA polymerase II, TATA box binding protein (TBP)-associated factor, 31kDa	0.76	1.7	5.1469e-05	950
402	2706	GJB2	gap junction protein, beta 2, 26kDa	0.76	1.7	2.5636e-05	915
403	5327	PLAT	plasminogen activator, tissue	0.75	1.7	3.9222e-04	1051
404	57580	PREX1	phosphatidylinositol-3,4,5-trisphosphate-dependent Rac exchange factor 1	0.75	1.7	6.6067e-07	619
405	8651	SOCS1	suppressor of cytokine signaling 1	0.75	1.7	4.7140e-07	599
406	5580	PRKCD	protein kinase C, delta	0.75	1.7	9.8774e-06	842
407	5734	PTGER4	prostaglandin E receptor 4 (subtype EP4)	0.75	1.7	8.9168e-06	837
408	64005	MYO1G	myosin IG	0.75	1.7	1.6604e-06	700
409	2210	FCGR1B	Fc fragment of IgG, high affinity Ib, receptor (CD64)	0.75	1.7	7.2832e-05	968
410	11314	CD300A	CD300a molecule	0.75	1.7	1.0084e-05	845
411	7097	TLR2	toll-like receptor 2	0.75	1.7	3.4299e-06	763
412	23765	IL17RA	interleukin 17 receptor A	0.74	1.7	1.6127e-09	180
413	5777	PTPN6	protein tyrosine phosphatase, non-receptor type 6	0.74	1.7	6.4213e-08	434
414	7187	TRAF3	TNF receptor-associated factor 3	0.74	1.7	6.1348e-09	258
415	51015	ISOC1	isochorismatase domain containing 1	0.74	1.7	1.3829e-04	1000
416	3240	HP	haptoglobin	0.74	1.7	6.3020e-06	806
417	2004	ELK3	ELK3, ETS-domain protein (SRF accessory protein 2)	0.74	1.7	5.2979e-06	796
418	4794	NFKBIE	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, epsilon	0.74	1.7	2.8614e-08	378
419	55640	FLVCR2	feline leukemia virus subgroup C cellular receptor family, member 2	0.74	1.7	2.3000e-05	907
420	51363	CHST15	carbohydrate (N-acetylgalactosamine 4-sulfate 6-O) sulfotransferase 15	0.73	1.7	1.8746e-06	708
421	9111	NMI	N-myc (and STAT) interactor	0.73	1.7	1.3910e-05	867
422	5698	PSMB9	proteasome (prosome, macropain) subunit, beta type, 9	0.73	1.7	1.5844e-05	877
423	8837	CFLAR	CASP8 and FADD-like apoptosis regulator	0.73	1.7	3.8004e-04	1049
424	10855	HPSE	heparanase	0.73	1.7	3.3401e-05	931
425	6091	ROBO1	roundabout, axon guidance receptor, homolog 1 (Drosophila)	0.73	1.7	6.4283e-06	809
426	5210	PFKFB4	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 4	0.73	1.7	2.0038e-07	530
427	4050	LTB	lymphotoxin beta (TNF superfamily, member 3)	0.73	1.7	2.6234e-06	734

Table 7 – Continued on next page

Table 7 – continued from previous page

428	7980	TFPI2	tissue factor pathway inhibitor 2	0.72	1.7	1.0148e-04	986
429	5199	CFP	complement factor properdin	0.72	1.7	7.0734e-06	820
430	8807	IL18RAP	interleukin 18 receptor accessory protein	0.72	1.6	6.1819e-09	260
431	9953	HS3ST3B1	heparan sulfate (glucosamine) 3-O-sulfotransferase 3B1	0.72	1.6	1.2036e-04	993
432	2710	GK	glycerol kinase	0.72	1.6	4.2670e-06	777
433	1281	COL3A1	collagen, type III, alpha 1	0.72	1.6	2.7837e-04	1029
434	50515	CHST11	carbohydrate (chondroitin 4) sulfotransferase 11	0.72	1.6	2.7145e-09	213
435	3038	HAS3	hyaluronan synthase 3	0.72	1.6	9.7875e-05	984
436	29995	LMCD1	LIM and cysteine-rich domains 1	0.72	1.6	9.1857e-08	463
437	23150	FRMD4B	FERM domain containing 4B	0.72	1.6	1.3933e-04	1001
438	2847	MCHR1	melanin-concentrating hormone receptor 1	0.72	1.6	1.6975e-04	1014
439	3958	LGALS3	lectin, galactoside-binding, soluble, 3	0.72	1.6	1.2735e-06	674
440	1026	CDKN1A	cyclin-dependent kinase inhibitor 1A (p21, Cip1)	0.72	1.6	1.7601e-07	521
441	2615	LRRC32	leucine rich repeat containing 32	0.72	1.6	5.0424e-09	244
442	126014	OSCAR	osteoclast associated, immunoglobulin-like receptor	0.71	1.6	2.5299e-06	730
443	5226	PGD	phosphogluconate dehydrogenase	0.71	1.6	7.7382e-06	826
444	4217	MAP3K5	mitogen-activated protein kinase kinase kinase 5	0.71	1.6	1.2126e-05	857
445	55013	CCDC109B	coiled-coil domain containing 109B	0.71	1.6	2.0765e-05	897
446	51411	BIN2	bridging integrator 2	0.71	1.6	5.6028e-08	419
447	83641	FAM107B	family with sequence similarity 107, member B	0.71	1.6	2.1235e-05	900
448	2683	B4GALT1	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 1	0.71	1.6	8.1284e-05	974
449	23092	ARHGAP26	Rho GTPase activating protein 26	0.71	1.6	1.6865e-08	331
450	53831	GPR84	G protein-coupled receptor 84	0.71	1.6	4.9845e-07	603
451	6542	SLC7A2	solute carrier family 7 (cationic amino acid transporter, y+ system), member 2	0.70	1.6	2.1058e-05	898
452	6774	STAT3	signal transducer and activator of transcription 3 (acute-phase response factor)	0.70	1.6	5.4618e-09	246
453	8795	TNFRSF10B	tumor necrosis factor receptor superfamily, member 10b	0.70	1.6	1.9294e-07	528
454	124935	SLC43A2	solute carrier family 43 (amino acid system L transporter), member 2	0.70	1.6	1.6657e-06	701
455	1535	CYBA	cytochrome b-245, alpha polypeptide	0.70	1.6	7.0016e-06	818
456	5294	PIK3CG	phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit gamma	0.70	1.6	9.2249e-06	839
457	54440	SASH3	SAM and SH3 domain containing 3	0.70	1.6	3.6511e-07	575
458	948	CD36	CD36 molecule (thrombospondin receptor)	0.70	1.6	2.6896e-05	917
459	11177	BAZ1A	bromodomain adjacent to zinc finger domain, 1A	0.70	1.6	5.3941e-06	798
460	84888	SPPL2A	signal peptide peptidase like 2A	0.70	1.6	3.2047e-11	53
461	9991	PTBP3	polypyrimidine tract binding protein 3	0.69	1.6	2.1361e-09	198

Table 7 – Continued on next page

Table 7 – continued from previous page

462	637	BID	BH3 interacting domain death agonist	0.69	1.6	3.2117e-06	756
463	79415	C17orf62	chromosome 17 open reading frame 62	0.69	1.6	3.7964e-09	232
464	9961	MVP	major vault protein	0.69	1.6	2.1978e-08	356
465	2213	FCGR2B	Fc fragment of IgG, low affinity IIb, receptor (CD32)	0.69	1.6	1.0233e-03	1081
466	1436	CSF1R	colony stimulating factor 1 receptor	0.69	1.6	1.1783e-04	990
467	22904	SBNO2	strawberry notch homolog 2 (Drosophila)	0.69	1.6	7.2331e-10	136
468	8793	TNFRSF10D	tumor necrosis factor receptor superfamily, member 10d, decoy with truncated death domain	0.69	1.6	1.2926e-06	676
469	2634	GBP2	guanylate binding protein 2, interferon-inducible	0.69	1.6	3.8486e-04	1050
470	83442	SH3BGRL3	SH3 domain binding glutamate-rich protein like 3	0.69	1.6	2.7364e-06	739
471	57575	PCDH10	protocadherin 10	0.69	1.6	1.5841e-05	876
472	54206	ERRFI1	ERBB receptor feedback inhibitor 1	0.69	1.6	2.2426e-05	906
473	5328	PLAU	plasminogen activator, urokinase	0.68	1.6	9.4556e-05	983
474	5230	PGK1	phosphoglycerate kinase 1	0.68	1.6	3.7697e-07	576
475	23670	TMEM2	transmembrane protein 2	0.68	1.6	1.3807e-09	173
476	6688	SPI1	Spi-1 proto-oncogene	0.68	1.6	6.4882e-06	810
477	51114	ZDHHC9	zinc finger, DHHC-type containing 9	0.68	1.6	7.7918e-07	628
478	9509	ADAMTS2	ADAM metallopeptidase with thrombospondin type 1 motif, 2	0.68	1.6	1.6241e-04	1011
479	8566	PDXK	pyridoxal (pyridoxine, vitamin B6) kinase	0.68	1.6	2.6298e-09	210
480	283897	C16orf54	chromosome 16 open reading frame 54	0.68	1.6	5.7148e-05	956
481	8934	RAB29	RAB29, member RAS oncogene family	0.68	1.6	1.4353e-05	868
482	3091	HIF1A	hypoxia inducible factor 1, alpha subunit (basic helix-loop-helix transcription factor)	0.68	1.6	8.8103e-11	90
483	7360	UGP2	UDP-glucose pyrophosphorylase 2	0.68	1.6	6.9577e-06	816
484	3732	CD82	CD82 molecule	0.68	1.6	7.7174e-10	140
485	441168	FAM26F	family with sequence similarity 26, member F	0.67	1.6	6.1602e-04	1067
486	9997	SCO2	SCO2 cytochrome c oxidase assembly protein	0.67	1.6	1.1351e-11	47
487	571	BACH1	BTB and CNC homology 1, basic leucine zipper transcription factor 1	0.67	1.6	2.9572e-09	219
488	5611	DNAJC3	DnaJ (Hsp40) homolog, subfamily C, member 3	0.67	1.6	1.0865e-05	851
489	151987	PPP4R2	protein phosphatase 4, regulatory subunit 2	0.67	1.6	1.8305e-11	50
490	27128	CYTH4	cytohesin 4	0.67	1.6	1.3706e-06	680
491	89849	ATG16L2	autophagy related 16-like 2 (S. cerevisiae)	0.67	1.6	3.9704e-07	583
492	26585	GREM1	gremlin 1, DAN family BMP antagonist	0.67	1.6	7.9604e-04	1071
493	3995	FADS3	fatty acid desaturase 3	0.67	1.6	9.7034e-08	467
494	5937	RBMS1	RNA binding motif, single stranded interacting protein 1	0.67	1.6	5.7477e-06	801
495	29992	PILRA	paired immunoglobin-like type 2 receptor alpha	0.67	1.6	3.0693e-07	558

Table 7 – Continued on next page

Table 7 – continued from previous page

496	4615	MYD88	myeloid differentiation primary response 88	0.67	1.6	2.8790e-09	215
497	6892	TAPBP	TAP binding protein (tapasin)	0.66	1.6	1.2092e-07	491
498	1959	EGR2	early growth response 2	0.66	1.6	1.2236e-04	994
499	84649	DGAT2	diacylglycerol O-acyltransferase 2	0.66	1.6	2.7400e-07	552
500	50650	ARHGEF3	Rho guanine nucleotide exchange factor (GEF) 3	0.66	1.6	7.1263e-06	821
501	4071	TM4SF1	transmembrane 4 L six family member 1	0.66	1.6	2.8230e-04	1032
502	7048	TGFBR2	transforming growth factor, beta receptor II (70/80kDa)	0.66	1.6	1.8001e-06	704
503	6351	CCL4	chemokine (C-C motif) ligand 4	0.66	1.6	3.1917e-04	1039
504	79572	ATP13A3	ATPase type 13A3	0.66	1.6	1.8527e-05	888
505	1820	ARID3A	AT rich interactive domain 3A (BRIGHT-like)	0.66	1.6	8.6186e-09	282
506	23531	MMD	monocyte to macrophage differentiation-associated	0.66	1.6	8.2252e-07	633
507	57085	AGTRAP	angiotensin II receptor-associated protein	0.66	1.6	1.9815e-08	343
508	4783	NFIL3	nuclear factor, interleukin 3 regulated	0.65	1.6	4.6075e-07	595
509	220	ALDH1A3	aldehyde dehydrogenase 1 family, member A3	0.65	1.6	2.4903e-05	912
510	3566	IL4R	interleukin 4 receptor	0.65	1.6	1.6909e-08	332
511	9052	GPRC5A	G protein-coupled receptor, class C, group 5, member A	0.65	1.6	3.2319e-07	563
512	26253	CLEC4E	C-type lectin domain family 4, member E	0.65	1.6	1.0631e-06	655
513	6676	SPAG4	sperm associated antigen 4	0.65	1.6	2.1474e-05	903
514	22898	DENND3	DENN/MADD domain containing 3	0.65	1.6	3.2040e-08	387
515	4493	MT1E	metallothionein 1E	0.65	1.6	5.2559e-04	1061
516	10797	MTHFD2	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2, methenyltetrahydrofolate cyclohydrolase	0.65	1.6	2.0809e-04	1022
517	1277	COL1A1	collagen, type I, alpha 1	0.65	1.6	1.5830e-03	1088
518	56882	CDC42SE1	CDC42 small effector 1	0.65	1.6	4.1548e-04	1054
519	4862	NPAS2	neuronal PAS domain protein 2	0.65	1.6	8.3481e-07	634
520	6275	S100A4	S100 calcium binding protein A4	0.65	1.6	2.3965e-03	1089
521	9051	PSTPIP1	proline-serine-threonine phosphatase interacting protein 1	0.65	1.6	9.8958e-07	649
522	56833	SLAMF8	SLAM family member 8	0.65	1.6	1.8427e-06	706
523	962	CD48	CD48 molecule	0.65	1.6	4.4183e-05	945
524	6398	SECTM1	secreted and transmembrane 1	0.65	1.6	2.9297e-06	745
525	2908	NR3C1	nuclear receptor subfamily 3, group C, member 1 (glucocorticoid receptor)	0.65	1.6	3.0449e-04	1036
526	8828	NRP2	neuropilin 2	0.65	1.6	1.6188e-06	697
527	8809	IL18R1	interleukin 18 receptor 1	0.64	1.6	1.3675e-06	679
528	100133941	CD24	CD24 molecule	0.64	1.6	4.5777e-04	1060
529	359845	FAM101B	family with sequence similarity 101, member B	0.64	1.6	2.4868e-06	729
530	81603	TRIM8	tripartite motif containing 8	0.64	1.6	1.3567e-07	497
531	6925	TCF4	transcription factor 4	0.64	1.6	4.3414e-04	1057

Table 7 – Continued on next page

Table 7 – continued from previous page

532	3654	IRAK1	interleukin-1 receptor-associated kinase 1	0.64	1.6	8.6415e-08	459
533	64231	MS4A6A	membrane-spanning 4-domains, subfamily A, member 6A	0.64	1.6	7.3865e-05	970
534	3097	HIVEP2	human immunodeficiency virus type I enhancer binding protein 2	0.64	1.6	1.1791e-05	855
535	942	CD86	CD86 molecule	0.64	1.6	1.8103e-05	886
536	80215	RUNX1-IT1	RUNX1 intronic transcript 1 (non-protein coding)	0.64	1.6	2.0056e-06	715
537	83862	TMEM120A	transmembrane protein 120A	0.64	1.6	1.2174e-06	668
538	1903	S1PR3	sphingosine-1-phosphate receptor 3	0.64	1.6	2.0111e-04	1019
539	155038	GIMAP8	GTPase, IMAP family member 8	0.64	1.6	2.0761e-04	1021
540	1794	DOCK2	dedicator of cytokinesis 2	0.64	1.6	1.2479e-05	860
541	89846	FGD3	FYVE, RhoGEF and PH domain containing 3	0.64	1.6	2.1913e-07	535
542	433	ASGR2	asialoglycoprotein receptor 2	0.63	1.6	1.7400e-08	334
543	719	C3AR1	complement component 3a receptor 1	0.63	1.6	1.9562e-05	894
544	2200	FBN1	fibrillin 1	0.63	1.6	1.8129e-08	339
545	3552	IL1A	interleukin 1, alpha	0.63	1.5	3.2059e-04	1040
546	2191	FAP	fibroblast activation protein, alpha	0.63	1.5	5.8982e-08	423
547	6464	SHC1	SHC (Src homology 2 domain containing) transforming protein 1	0.63	1.5	2.5091e-09	208
548	391	RHOG	ras homolog family member G	0.63	1.5	3.1935e-07	561
549	10788	IQGAP2	IQ motif containing GTPase activating protein 2	0.63	1.5	1.3755e-06	681
550	3428	IFI16	interferon, gamma-inducible protein 16	0.63	1.5	3.4346e-04	1045
551	9120	SLC16A6	solute carrier family 16, member 6	0.63	1.5	3.2248e-05	927
552	9510	ADAMTS1	ADAM metallopeptidase with thrombospondin type 1 motif, 1	0.63	1.5	4.3266e-05	941
553	22846	VASH1	vasohibin 1	0.63	1.5	1.4416e-03	1086
554	7474	WNT5A	wingless-type MMTV integration site family, member 5A	0.63	1.5	2.6754e-03	1093
555	56261	GPCPD1	glycerophosphocholine phosphodiesterase GDE1 homolog (S. cerevisiae)	0.62	1.5	1.3032e-05	862
556	56975	FAM20C	family with sequence similarity 20, member C	0.62	1.5	1.0936e-08	296
557	9732	DOCK4	dedicator of cytokinesis 4	0.62	1.5	5.3311e-05	954
558	5196	PF4	platelet factor 4	0.62	1.5	4.2709e-06	778
559	55911	APOBR	apolipoprotein B receptor	0.62	1.5	1.2202e-06	669
560	10645	CAMKK2	calcium/calmodulin-dependent protein kinase kinase 2, beta	0.62	1.5	8.5048e-06	834
561	80853	KDM7A	lysine (K)-specific demethylase 7A	0.62	1.5	8.7914e-06	835
562	9586	CREB5	cAMP responsive element binding protein 5	0.62	1.5	6.7042e-05	965
563	10105	PPIF	peptidylprolyl isomerase F	0.62	1.5	2.7567e-09	214
564	7052	TGM2	transglutaminase 2	0.62	1.5	7.4030e-05	971
565	402483	LINC01000	long intergenic non-protein coding RNA 1000	0.61	1.5	2.3103e-05	908
566	8829	NRP1	neuropilin 1	0.61	1.5	8.2622e-05	977
567	3460	IFNGR2	interferon gamma receptor 2 (interferon gamma transducer 1)	0.61	1.5	3.1167e-09	222

Table 7 – Continued on next page

Table 7 – continued from previous page

568	196051	PPAPDC1A	phosphatidic acid phosphatase type 2 domain containing 1A	0.61	1.5	1.4949e-04	1005
569	4241	MFI2	antigen p97 (melanoma associated) identified by monoclonal antibodies 133.2 and 96.5	0.61	1.5	1.2325e-05	858
570	10123	ARL4C	ADP-ribosylation factor-like 4C	0.61	1.5	1.0636e-04	987
571	1290	COL5A2	collagen, type V, alpha 2	0.61	1.5	3.4473e-07	569
572	3709	ITPR2	inositol 1,4,5-trisphosphate receptor, type 2	0.61	1.5	8.9375e-08	460
573	51537	MTFP1	mitochondrial fission process 1	0.61	1.5	1.3497e-06	678
574	3669	ISG20	interferon stimulated exonuclease gene 20kDa	0.61	1.5	1.4023e-04	1002
575	6890	TAP1	transporter 1, ATP-binding cassette, sub-family B (MDR/TAP)	0.60	1.5	4.7060e-06	789
576	64098	PARVG	parvin, gamma	0.60	1.5	3.3700e-07	564
577	23166	STAB1	stabilin 1	0.60	1.5	1.4734e-05	872
578	4818	NKG7	natural killer cell granule protein 7	0.60	1.5	1.3609e-04	999
579	26996	GPR160	G protein-coupled receptor 160	0.60	1.5	1.3427e-04	997
580	271	AMPD2	adenosine monophosphate deaminase 2	0.60	1.5	4.6506e-09	239
581	51303	FKBP11	FK506 binding protein 11, 19 kDa	0.60	1.5	9.7877e-09	291
582	6793	STK10	serine/threonine kinase 10	0.60	1.5	6.7978e-07	622
583	7351	UCP2	uncoupling protein 2 (mitochondrial, proton carrier)	0.60	1.5	2.1116e-04	1023
584	9050	PSTPIP2	proline-serine-threonine phosphatase interacting protein 2	0.60	1.5	7.6745e-06	825
585	9103	FCGR2C	Fc fragment of IgG, low affinity IIc, receptor for (CD32) (gene/pseudogene)	0.59	1.5	1.5135e-06	690
586	4499	MT1M	metallothionein 1M	0.59	1.5	2.7637e-03	1094
587	10581	IFITM2	interferon induced transmembrane protein 2	0.59	1.5	3.4993e-07	570
588	5873	RAB27A	RAB27A, member RAS oncogene family	0.59	1.5	7.3656e-07	626
589	4739	NEDD9	neural precursor cell expressed, developmentally down-regulated 9	0.59	1.5	4.1125e-04	1053
590	619426	C8orf60	chromosome 8 open reading frame 60	0.59	1.5	4.2076e-05	938
591	200916	RPL22L1	ribosomal protein L22-like 1	0.59	1.5	1.8849e-06	709
592	65220	NADK	NAD kinase	0.59	1.5	2.5721e-07	546

Table 8. List of 505 significantly downregulated genes between FIR and nonFIR affected infants. Genes are ordered by absolute $\log_2$ fold-change in expression. Columns "LFC" and "FC" contain the fold-changes in $\log_2$ and absolute scales, respectively. The latter has been inverted to ease its interpretation. The column "P rnk" provides the rank of each gene by raw p-value within the entire list of 1097 genes called differentially expressed.

	Entrez ID	Symbol	Full name	LFC	FC	P value	P rnk
1	1272	CNTN1	contactin 1	-2.07	4.2	2.4281e-14	4
2	29951	PDZRN4	PDZ domain containing ring finger 4	-1.80	3.5	4.0504e-11	61
3	477	ATP1A2	ATPase, Na+/K+ transporting, alpha 2 polypeptide	-1.78	3.4	3.9614e-13	18
4	1908	EDN3	endothelin 3	-1.68	3.2	1.4510e-09	176

Table 8 – Continued on next page

Table 8 – continued from previous page

5	59353	TMEM35	transmembrane protein 35	-1.62	3.1	7.6245e-11	83
6	151742	PPM1L	protein phosphatase, Mg2+/Mn2+ dependent, 1L	-1.56	2.9	1.7575e-13	10
7	590	BCHE	butyrylcholinesterase	-1.55	2.9	1.4049e-11	49
8	4978	OPCML	opioid binding protein/cell adhesion molecule-like	-1.52	2.9	1.3278e-09	171
9	81578	COL21A1	collagen, type XXI, alpha 1	-1.50	2.8	1.2989e-09	170
10	1305	COL13A1	collagen, type XIII, alpha 1	-1.43	2.7	1.7388e-12	28
11	4621	MYH3	myosin, heavy chain 3, skeletal muscle, embryonic	-1.42	2.7	1.0942e-05	852
12	23171	GPD1L	glycerol-3-phosphate dehydrogenase 1-like	-1.42	2.7	8.7765e-11	89
13	54361	WNT4	wingless-type MMTV integration site family, member 4	-1.39	2.6	5.5440e-09	247
14	4916	NTRK3	neurotrophic tyrosine kinase, receptor, type 3	-1.38	2.6	5.0476e-10	119
15	11155	LDB3	LIM domain binding 3	-1.37	2.6	2.7892e-06	741
16	55228	PNMAL1	paraneoplastic Ma antigen family-like 1	-1.37	2.6	6.3637e-09	264
17	51196	PLCE1	phospholipase C, epsilon 1	-1.34	2.5	2.0096e-10	104
18	94240	EPSTI1	epithelial stromal interaction 1 (breast)	-1.33	2.5	1.1689e-09	161
19	139065	SLITRK4	SLIT and NTRK-like family, member 4	-1.31	2.5	3.7914e-10	110
20	5166	PDK4	pyruvate dehydrogenase kinase, isozyme 4	-1.31	2.5	2.4892e-05	911
21	10974	ADIRF	adipogenesis regulatory factor	-1.29	2.5	8.1627e-09	277
22	5784	PTPN14	protein tyrosine phosphatase, non-receptor type 14	-1.28	2.4	4.1706e-08	400
23	3899	AFF3	AF4/FMR2 family, member 3	-1.27	2.4	1.1104e-07	482
24	8787	RGS9	regulator of G-protein signaling 9	-1.26	2.4	2.0278e-13	12
25	4093	SMAD9	SMAD family member 9	-1.26	2.4	2.0386e-12	31
26	4969	OGN	osteoglycin	-1.25	2.4	6.0804e-05	957
27	55504	TNFRSF19	tumor necrosis factor receptor superfamily, member 19	-1.25	2.4	2.4069e-08	364
28	5354	PLP1	proteolipid protein 1	-1.24	2.4	6.0780e-04	1066
29	83468	GLT8D2	glycosyltransferase 8 domain containing 2	-1.24	2.4	3.9610e-11	59
30	84962	AJUBA	ajuba LIM protein	-1.21	2.3	6.1147e-09	257
31	257194	NEGR1	neuronal growth regulator 1	-1.21	2.3	7.0403e-11	79
32	79148	MMP28	matrix metallopeptidase 28	-1.20	2.3	3.9264e-11	58
33	83543	AIF1L	allograft inflammatory factor 1-like	-1.20	2.3	3.7293e-09	230
34	395	ARHGAP6	Rho GTPase activating protein 6	-1.20	2.3	1.8679e-11	51
35	5308	PITX2	paired-like homeodomain 2	-1.20	2.3	6.9537e-08	438
36	8082	SSPN	sarcospan	-1.20	2.3	1.1819e-07	490
37	65055	REEP1	receptor accessory protein 1	-1.20	2.3	4.0231e-09	236
38	572558	PGM5-AS1	PGM5 antisense RNA 1	-1.20	2.3	7.3691e-10	137
39	1288	COL4A6	collagen, type IV, alpha 6	-1.20	2.3	5.7557e-09	251
40	284	ANGPT1	angiopoietin 1	-1.19	2.3	7.2791e-11	80
41	57631	LRCH2	leucine-rich repeats and calponin homology (CH) domain containing 2	-1.19	2.3	4.2181e-11	62
42	84620	ST6GAL2	ST6 beta-galactosamide alpha-2,6-sialyltranferase 2	-1.18	2.3	3.0230e-05	925
43	23554	TSPAN12	tetraspanin 12	-1.17	2.3	1.4174e-07	506

Table 8 – Continued on next page

Table 8 – continued from previous page

44	56952	PRTFDC1	phosphoribosyl transferase domain containing 1	-1.16	2.2	9.8066e-10	154
45	80144	FRAS1	Fraser extracellular matrix complex subunit 1	-1.16	2.2	1.1144e-07	483
46	2258	FGF13	fibroblast growth factor 13	-1.16	2.2	5.4016e-14	6
47	55607	PPP1R9A	protein phosphatase 1, regulatory subunit 9A	-1.15	2.2	6.9398e-09	268
48	79739	TTLL7	tubulin tyrosine ligase-like family, member 7	-1.13	2.2	3.8315e-11	55
49	83690	CRISPLD1	cysteine-rich secretory protein LCCL domain containing 1	-1.13	2.2	3.5151e-07	571
50	140733	MACROD2	MACRO domain containing 2	-1.12	2.2	8.0674e-12	42
51	339260	LOC339260	uncharacterized LOC339260	-1.12	2.2	1.9051e-08	341
52	8436	SDPR	serum deprivation response	-1.12	2.2	6.7802e-10	129
53	6565	SLC15A2	solute carrier family 15 (oligopeptide transporter), member 2	-1.12	2.2	1.5724e-10	101
54	5195	PEX14	peroxisomal biogenesis factor 14	-1.10	2.1	2.7709e-08	375
55	3671	ISLR	immunoglobulin superfamily containing leucine-rich repeat	-1.09	2.1	2.1060e-07	534
56	2977	GUCY1A2	guanylate cyclase 1, soluble, alpha 2	-1.09	2.1	9.6714e-10	153
57	4919	ROR1	receptor tyrosine kinase-like orphan receptor 1	-1.07	2.1	2.6542e-09	211
58	4329	ALDH6A1	aldehyde dehydrogenase 6 family, member A1	-1.07	2.1	4.6073e-06	785
59	29761	USP25	ubiquitin specific peptidase 25	-1.07	2.1	4.6477e-11	65
60	26033	ATRNL1	attractin-like 1	-1.06	2.1	7.0178e-08	440
61	56977	STOX2	storkhead box 2	-1.06	2.1	5.9408e-07	615
62	347902	AMIGO2	adhesion molecule with Ig-like domain 2	-1.06	2.1	4.1186e-08	399
63	51390	AIG1	androgen-induced 1	-1.06	2.1	7.2976e-11	81
64	27109	ATP5S	ATP synthase, H+ transporting, mitochondrial Fo complex, subunit s (factor B)	-1.06	2.1	1.9809e-09	196
65	23236	PLCB1	phospholipase C, beta 1 (phosphoinositide-specific)	-1.05	2.1	2.7594e-10	108
66	23493	HEY2	hes-related family bHLH transcription factor with YRPW motif 2	-1.05	2.1	2.7912e-13	16
67	5087	PBX1	pre-B-cell leukemia homeobox 1	-1.05	2.1	8.9533e-11	91
68	85004	RERG	RAS-like, estrogen-regulated, growth inhibitor	-1.05	2.1	1.5444e-08	321
69	84417	C2orf40	chromosome 2 open reading frame 40	-1.04	2.1	1.6898e-06	703
70	157638	FAM84B	family with sequence similarity 84, member B	-1.04	2.1	7.8846e-10	142
71	200403	VWA3B	von Willebrand factor A domain containing 3B	-1.04	2.1	3.3951e-06	762
72	143098	MPP7	membrane protein, palmitoylated 7 (MAGUK p55 subfamily member 7)	-1.04	2.1	1.5551e-06	691
73	54749	EPDR1	ependymin related 1	-1.03	2.0	1.8629e-12	30
74	283298	OLFML1	olfactomedin-like 1	-1.03	2.0	1.4055e-07	503
75	100124700	HOTAIR	HOX transcript antisense RNA	-1.02	2.0	4.1214e-06	775
76	114905	C1QTNF7	C1q and tumor necrosis factor related protein 7	-1.02	2.0	6.5955e-06	811

Table 8 – Continued on next page

Table 8 – continued from previous page

77	256691	MAMDC2	MAM domain containing 2	-1.01	2.0	7.8241e-06	827
78	2770	GNAI1	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 1	-1.01	2.0	3.1589e-12	35
79	3590	IL11RA	interleukin 11 receptor, alpha	-1.01	2.0	2.0988e-09	197
80	100507311	LOC100507311	uncharacterized LOC100507311	-1.00	2.0	4.2411e-08	403
81	54985	HCFC1R1	host cell factor C1 regulator 1 (XPO1 dependent)	-1.00	2.0	5.0052e-11	70
82	2018	EMX2	empty spiracles homeobox 2	-1.00	2.0	2.5548e-12	32
83	56849	TCEAL7	transcription elongation factor A (SII)-like 7	-0.99	2.0	5.1342e-08	416
84	5906	RAP1A	RAP1A, member of RAS oncogene family	-0.99	2.0	8.7347e-13	22
85	79966	SCD5	stearoyl-CoA desaturase 5	-0.99	2.0	7.9852e-08	452
86	112770	C1orf85	chromosome 1 open reading frame 85	-0.99	2.0	3.9061e-09	233
87	3680	ITGA9	integrin, alpha 9	-0.99	2.0	3.7711e-08	390
88	7552	ZNF711	zinc finger protein 711	-0.98	2.0	2.2279e-09	199
89	56241	SUSD2	sushi domain containing 2	-0.98	2.0	1.2717e-06	673
90	5789	PTPRD	protein tyrosine phosphatase, receptor type, D	-0.98	2.0	7.1563e-09	269
91	54477	PLEKHA5	pleckstrin homology domain containing, family A member 5	-0.98	2.0	1.5695e-10	100
92	286205	SCAI	suppressor of cancer cell invasion	-0.98	2.0	3.6490e-12	37
93	8395	PIP5K1B	phosphatidylinositol-4-phosphate 5-kinase, type I, beta	-0.98	2.0	7.3460e-05	969
94	8404	SPARCL1	SPARC-like 1 (hevin)	-0.98	2.0	6.5371e-09	266
95	9737	GPRASP1	G protein-coupled receptor associated sorting protein 1	-0.98	2.0	6.5411e-11	76
96	5727	PTCH1	patched 1	-0.97	2.0	3.9074e-11	57
97	220965	FAM13C	family with sequence similarity 13, member C	-0.97	2.0	3.8345e-08	393
98	3033	HADH	hydroxyacyl-CoA dehydrogenase	-0.96	2.0	4.4100e-11	64
99	285759	FLJ34503	uncharacterized FLJ34503	-0.96	1.9	2.3646e-04	1025
100	80318	GKAP1	G kinase anchoring protein 1	-0.96	1.9	5.8992e-09	254
101	2487	FRZB	frizzled-related protein	-0.96	1.9	6.2503e-05	959
102	3670	ISL1	ISL LIM homeobox 1	-0.95	1.9	6.3091e-05	961
103	9547	CXCL14	chemokine (C-X-C motif) ligand 14	-0.95	1.9	2.4720e-03	1091
104	1842	ECM2	extracellular matrix protein 2, female organ and adipocyte specific	-0.95	1.9	1.0887e-07	481
105	401093	MBNL1-AS1	MBNL1 antisense RNA 1	-0.94	1.9	2.9954e-08	383
106	3741	KCNA5	potassium voltage-gated channel, shaker-related subfamily, member 5	-0.94	1.9	2.9019e-07	554
107	8842	PROM1	prominin 1	-0.94	1.9	1.6223e-05	879
108	55034	MOCOS	molybdenum cofactor sulfurase	-0.94	1.9	3.8522e-06	772
109	5295	PIK3R1	phosphoinositide-3-kinase, regulatory subunit 1 (alpha)	-0.93	1.9	1.1724e-08	300
110	1287	COL4A5	collagen, type IV, alpha 5	-0.93	1.9	1.1477e-13	9
111	53354	PANK1	pantothenate kinase 1	-0.93	1.9	2.2413e-09	200
112	100288911	LOC100288911	uncharacterized LOC100288911	-0.93	1.9	5.7216e-07	613
113	51422	PRKAG2	protein kinase, AMP-activated, gamma 2 non-catalytic subunit	-0.93	1.9	4.3981e-05	942

Table 8 – Continued on next page

Table 8 – continued from previous page

114	8630	HSD17B6	hydroxysteroid (17-beta) dehydrogenase 6	-0.93	1.9	2.4694e-08	367
115	29904	EEF2K	eukaryotic elongation factor-2 kinase	-0.92	1.9	2.9110e-12	33
116	54800	KLHL24	kelch-like family member 24	-0.92	1.9	1.3344e-06	677
117	1031	CDKN2C	cyclin-dependent kinase inhibitor 2C (p18, inhibits CDK4)	-0.92	1.9	2.3436e-09	204
118	90293	KLHL13	kelch-like family member 13	-0.92	1.9	6.8198e-06	813
119	619279	ZNF704	zinc finger protein 704	-0.92	1.9	6.3106e-08	427
120	5348	FXYD1	FXYD domain containing ion transport regulator 1	-0.91	1.9	1.0396e-11	45
121	2549	GAB1	GRB2-associated binding protein 1	-0.91	1.9	5.0259e-10	118
122	130271	PLEKHH2	pleckstrin homology domain containing, family H (with MyTH4 domain) member 2	-0.91	1.9	4.5065e-08	407
123	9975	NR1D2	nuclear receptor subfamily 1, group D, member 2	-0.91	1.9	2.5994e-04	1026
124	286097	MICU3	mitochondrial calcium uptake family, member 3	-0.91	1.9	2.5784e-08	371
125	5099	PCDH7	protocadherin 7	-0.91	1.9	8.2125e-06	831
126	169611	OLFML2A	olfactomedin-like 2A	-0.90	1.9	4.5667e-08	408
127	11034	DSTN	destrin (actin depolymerizing factor)	-0.90	1.9	1.4395e-06	683
128	1825	DSC3	desmocollin 3	-0.90	1.9	1.3397e-08	310
129	2254	FGF9	fibroblast growth factor 9	-0.90	1.9	8.8982e-06	836
130	91624	NEXN	nexilin (F actin binding protein)	-0.90	1.9	1.0777e-08	295
131	2307	FOXS1	forkhead box S1	-0.90	1.9	6.4896e-08	436
132	4675	NAP1L3	nucleosome assembly protein 1-like 3	-0.90	1.9	1.0069e-06	650
133	57188	ADAMTSL3	ADAMTS-like 3	-0.90	1.9	1.7748e-07	522
134	3429	IFI27	interferon, alpha-inducible protein 27	-0.89	1.9	2.9432e-04	1035
135	9783	RIMS3	regulating synaptic membrane exocytosis 3	-0.89	1.9	1.2453e-07	493
136	64344	HIF3A	hypoxia inducible factor 3, alpha subunit	-0.88	1.8	3.1307e-04	1038
137	8863	PER3	period circadian clock 3	-0.88	1.8	8.1878e-07	631
138	10964	IFI44L	interferon-induced protein 44-like	-0.88	1.8	1.1794e-04	991
139	3226	HOXC10	homeobox C10	-0.87	1.8	1.3904e-05	866
140	22881	ANKRD6	ankyrin repeat domain 6	-0.87	1.8	4.7091e-07	598
141	57496	MKL2	MKL/myocardin-like 2	-0.87	1.8	2.9384e-08	381
142	51209	RAB9B	RAB9B, member RAS oncogene family	-0.86	1.8	3.6426e-07	574
143	8324	FZD7	frizzled class receptor 7	-0.86	1.8	1.2002e-08	302
144	1021	CDK6	cyclin-dependent kinase 6	-0.86	1.8	7.6760e-10	139
145	55137	FIGN	fidgetin	-0.86	1.8	9.1163e-07	642
146	167681	PRSS35	protease, serine, 35	-0.85	1.8	9.7118e-04	1079
147	11244	ZHX1	zinc fingers and homeoboxes 1	-0.85	1.8	1.5346e-04	1008
148	90871	TMEM261	transmembrane protein 261	-0.85	1.8	2.0949e-08	347
149	23658	LSM5	LSM5 homolog, U6 small nuclear RNA associated (S. cerevisiae)	-0.85	1.8	1.5613e-08	323
150	51477	ISYNA1	inositol-3-phosphate synthase 1	-0.85	1.8	3.4550e-09	227
151	79974	CPED1	cadherin-like and PC-esterase domain containing 1	-0.84	1.8	1.9132e-09	193
152	23189	KANK1	KN motif and ankyrin repeat domains 1	-0.84	1.8	1.1242e-07	485

Table 8 – Continued on next page

Table 8 – continued from previous page

153	10217	CTDSPL	CTD (carboxy-terminal domain, RNA polymerase II, polypeptide A) small phosphatase-like	-0.84	1.8	4.8893e-11	68
154	145270	PRIMA1	proline rich membrane anchor 1	-0.84	1.8	2.6132e-05	916
155	6934	TCF7L2	transcription factor 7-like 2 (T-cell specific, HMG-box)	-0.84	1.8	6.3744e-08	432
156	9352	TXNL1	thioredoxin-like 1	-0.83	1.8	2.5808e-07	548
157	5136	PDE1A	phosphodiesterase 1A, calmodulin-dependent	-0.83	1.8	1.6693e-04	1012
158	5334	PLCL1	phospholipase C-like 1	-0.83	1.8	2.2053e-08	357
159	102724927	LOC102724927	uncharacterized LOC102724927	-0.83	1.8	1.1061e-05	853
160	23467	NPTXR	neuronal pentraxin receptor	-0.83	1.8	6.2871e-10	125
161	64174	DPEP2	dipeptidase 2	-0.83	1.8	2.6893e-09	212
162	27346	TMEM97	transmembrane protein 97	-0.83	1.8	2.8923e-08	380
163	101929340	LOC101929340	uncharacterized LOC101929340	-0.83	1.8	3.4400e-07	568
164	23089	PEG10	paternally expressed 10	-0.82	1.8	2.0746e-07	532
165	54443	ANLN	anillin, actin binding protein	-0.82	1.8	8.5383e-04	1075
166	79582	SPAG16	sperm associated antigen 16	-0.82	1.8	1.0485e-10	93
167	79070	KDELC1	KDEL (Lys-Asp-Glu-Leu) containing 1	-0.82	1.8	3.1589e-07	560
168	81285	OR51E2	olfactory receptor, family 51, subfamily E, member 2	-0.82	1.8	1.9359e-05	893
169	7296	TXNRD1	thioredoxin reductase 1	-0.82	1.8	3.9332e-08	395
170	25842	ASF1A	anti-silencing function 1A histone chaperone	-0.82	1.8	1.7905e-09	185
171	1760	DMPK	dystrophia myotonica-protein kinase	-0.81	1.8	3.0929e-07	559
172	58494	JAM2	junctional adhesion molecule 2	-0.81	1.8	2.8531e-08	377
173	1122	CHML	choroideremia-like (Rab escort protein 2)	-0.81	1.8	1.2862e-06	675
174	4239	MFAP4	microfibrillar-associated protein 4	-0.81	1.8	1.1302e-06	660
175	84952	CGNL1	cingulin-like 1	-0.81	1.8	1.0295e-06	651
176	23641	LDOC1	leucine zipper, down-regulated in cancer 1	-0.81	1.8	5.1078e-10	120
177	26059	ERC2	ELKS/RAB6-interacting/CAST family member 2	-0.81	1.8	1.2305e-04	995
178	100505576	LINC00672	long intergenic non-protein coding RNA 672	-0.81	1.8	2.8065e-08	376
179	56062	KLHL4	kelch-like family member 4	-0.81	1.8	1.4545e-06	684
180	8622	PDE8B	phosphodiesterase 8B	-0.81	1.7	4.4000e-07	591
181	23635	SSBP2	single-stranded DNA binding protein 2	-0.80	1.7	1.7523e-07	520
182	8988	HSPB3	heat shock 27kDa protein 3	-0.80	1.7	9.1650e-05	982
183	80312	TET1	tet methylcytosine dioxygenase 1	-0.80	1.7	2.5383e-08	369
184	139886	SPIN4	spindlin family, member 4	-0.80	1.7	7.3862e-09	271
185	113263	GLCCI1	glucocorticoid induced transcript 1	-0.80	1.7	2.5138e-07	544
186	80760	ITIH5	inter-alpha-trypsin inhibitor heavy chain family, member 5	-0.80	1.7	1.6786e-06	702
187	320	APBA1	amyloid beta (A4) precursor protein-binding, family A, member 1	-0.80	1.7	6.6299e-07	620
188	120114	FAT3	FAT atypical cadherin 3	-0.80	1.7	4.3688e-04	1058
189	2239	GPC4	glypican 4	-0.80	1.7	5.7280e-09	250
190	729082	OIP5-AS1	OIP5 antisense RNA 1	-0.79	1.7	2.1446e-05	902
191	857	CAV1	caveolin 1, caveolae protein, 22kDa	-0.79	1.7	1.4717e-07	509

Table 8 – Continued on next page

Table 8 – continued from previous page

192	29116	MYLIP	myosin regulatory light chain interacting protein	-0.79	1.7	8.2854e-04	1073
193	818	CAMK2G	calcium/calmodulin-dependent protein kinase II gamma	-0.79	1.7	8.5987e-09	281
194	2296	FOXC1	forkhead box C1	-0.79	1.7	1.1821e-06	667
195	2332	FMR1	fragile X mental retardation 1	-0.79	1.7	1.5830e-12	25
196	170691	ADAMTS17	ADAM metallopeptidase with thrombospondin type 1 motif, 17	-0.79	1.7	4.4745e-06	781
197	6769	STAC	SH3 and cysteine rich domain	-0.79	1.7	8.4682e-07	636
198	128977	C22orf39	chromosome 22 open reading frame 39	-0.79	1.7	3.2060e-09	225
199	5239	PGM5	phosphoglucomutase 5	-0.79	1.7	1.4327e-07	507
200	144165	PRICKLE1	prickle homolog 1 (Drosophila)	-0.79	1.7	1.3053e-08	308
201	32	ACACB	acetyl-CoA carboxylase beta	-0.79	1.7	4.1649e-07	587
202	7026	NR2F2	nuclear receptor subfamily 2, group F, member 2	-0.79	1.7	1.5626e-06	694
203	195828	ZNF367	zinc finger protein 367	-0.78	1.7	2.4547e-05	910
204	657	BMPR1A	bone morphogenetic protein receptor, type IA	-0.78	1.7	7.0464e-10	133
205	10370	CITED2	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 2	-0.78	1.7	6.2959e-06	805
206	133746	JMY	junction mediating and regulatory protein, p53 cofactor	-0.78	1.7	7.9835e-08	451
207	84056	KATNAL1	katanin p60 subunit A-like 1	-0.78	1.7	1.3990e-08	313
208	148534	TMEM56	transmembrane protein 56	-0.78	1.7	1.6004e-12	26
209	26037	SIPA1L1	signal-induced proliferation-associated 1 like 1	-0.78	1.7	1.9038e-09	192
210	26084	ARHGEF26	Rho guanine nucleotide exchange factor (GEF) 26	-0.77	1.7	1.4437e-04	1003
211	84131	CEP78	centrosomal protein 78kDa	-0.77	1.7	3.9863e-09	235
212	94274	PPP1R14A	protein phosphatase 1, regulatory (inhibitor) subunit 14A	-0.77	1.7	8.5608e-08	457
213	6444	SGCD	sarcoglycan, delta (35kDa dystrophin-associated glycoprotein)	-0.77	1.7	1.5069e-07	513
214	119504	ANAPC16	anaphase promoting complex subunit 16	-0.77	1.7	6.1655e-10	123
215	84253	GARNL3	GTPase activating Rap/RanGAP domain-like 3	-0.77	1.7	2.9565e-08	382
216	55672	NBPF1	neuroblastoma breakpoint family, member 1	-0.77	1.7	2.5003e-05	913
217	259230	SGMS1	sphingomyelin synthase 1	-0.77	1.7	1.1495e-06	662
218	28978	TMEM14A	transmembrane protein 14A	-0.77	1.7	2.2620e-07	537
219	2261	FGFR3	fibroblast growth factor receptor 3	-0.77	1.7	4.9295e-08	413
220	26273	FBXO3	F-box protein 3	-0.77	1.7	4.4253e-10	112
221	9891	NUAK1	NUAK family, SNF1-like kinase, 1	-0.77	1.7	5.5375e-07	609
222	5137	PDE1C	phosphodiesterase 1C, calmodulin-dependent 70kDa	-0.77	1.7	7.5922e-04	1069
223	6419	SETMAR	SET domain and mariner transposase fusion gene	-0.76	1.7	1.5819e-09	178
224	282996	RBM20	RNA binding motif protein 20	-0.76	1.7	3.2716e-06	759
225	55857	KIZ	kizuna centrosomal protein	-0.76	1.7	9.8585e-08	469
226	9462	RASAL2	RAS protein activator like 2	-0.76	1.7	2.0291e-06	717

Table 8 – Continued on next page

Table 8 – continued from previous page

227	6920	TCEA3	transcription elongation factor A (SII), 3	-0.76	1.7	3.2572e-05	929
228	1296	COL8A2	collagen, type VIII, alpha 2	-0.76	1.7	4.7990e-07	600
229	151525	WDSUB1	WD repeat, sterile alpha motif and U-box domain containing 1	-0.76	1.7	1.1749e-07	488
230	25907	TMEM158	transmembrane protein 158 (gene/pseudogene)	-0.76	1.7	1.0774e-08	294
231	10912	GADD45G	growth arrest and DNA-damage-inducible, gamma	-0.76	1.7	1.0725e-04	988
232	27303	RBMS3	RNA binding motif, single stranded interacting protein 3	-0.76	1.7	4.6676e-07	597
233	112476	PRRT2	proline-rich transmembrane protein 2	-0.76	1.7	8.2183e-07	632
234	1466	CSRP2	cysteine and glycine-rich protein 2	-0.76	1.7	1.3462e-11	48
235	8515	ITGA10	integrin, alpha 10	-0.76	1.7	2.2763e-07	539
236	7029	TFDP2	transcription factor Dp-2 (E2F dimerization partner 2)	-0.75	1.7	1.5979e-04	1010
237	57544	TXNDC16	thioredoxin domain containing 16	-0.75	1.7	1.6847e-08	330
238	4199	ME1	malic enzyme 1, NADP(+)-dependent, cytosolic	-0.75	1.7	1.1220e-07	484
239	8544	PIR	pirin (iron-binding nuclear protein)	-0.75	1.7	7.7172e-08	448
240	55089	SLC38A4	solute carrier family 38, member 4	-0.75	1.7	3.5937e-06	767
241	55107	ANO1	anoctamin 1, calcium activated chloride channel	-0.75	1.7	2.5724e-06	732
242	8736	MYOM1	myomesin 1	-0.75	1.7	4.2971e-08	404
243	5023	P2RX1	purinergic receptor P2X, ligand-gated ion channel, 1	-0.75	1.7	3.5697e-06	766
244	7275	TUB	tubby bipartite transcription factor	-0.75	1.7	1.1427e-06	661
245	64780	MICAL1	microtubule associated monooxygenase, calponin and LIM domain containing 1	-0.75	1.7	6.1608e-08	426
246	84293	FAM213A	family with sequence similarity 213, member A	-0.75	1.7	1.6563e-06	699
247	8665	EIF3F	eukaryotic translation initiation factor 3, subunit F	-0.75	1.7	1.0963e-08	297
248	25849	PARM1	prostate androgen-regulated mucin-like protein 1	-0.75	1.7	1.2549e-03	1085
249	8801	SUCLG2	succinate-CoA ligase, GDP-forming, beta subunit	-0.74	1.7	1.1735e-07	487
250	862	RUNX1T1	runt-related transcription factor 1; translocated to, 1 (cyclin D-related)	-0.74	1.7	4.0632e-06	774
251	400916	CHCHD10	coiled-coil-helix-coiled-coil-helix domain containing 10	-0.74	1.7	2.2912e-08	359
252	348093	RBPMS2	RNA binding protein with multiple splicing 2	-0.74	1.7	4.6195e-07	596
253	114876	OSBPL1A	oxysterol binding protein-like 1A	-0.74	1.7	5.8206e-09	253
254	10351	ABCA8	ATP-binding cassette, sub-family A (ABC1), member 8	-0.74	1.7	1.1413e-03	1084
255	79822	ARHGAP28	Rho GTPase activating protein 28	-0.74	1.7	4.5534e-06	783
256	55553	SOX6	SRY (sex determining region Y)-box 6	-0.74	1.7	1.5115e-07	514
257	84858	ZNF503	zinc finger protein 503	-0.74	1.7	4.2182e-05	939
258	441094	NR2F1-AS1	NR2F1 antisense RNA 1	-0.74	1.7	3.8709e-06	773

Table 8 – Continued on next page

Table 8 – continued from previous page

259	64172	OSGEPL1	O-sialoglycoprotein endopeptidase-like 1	-0.74	1.7	3.1651e-08	386
260	285671	RNF180	ring finger protein 180	-0.74	1.7	8.5662e-07	638
261	23492	CBX7	chromobox homolog 7	-0.74	1.7	9.8954e-07	648
262	26353	HSPB8	heat shock 22kDa protein 8	-0.74	1.7	4.2852e-06	779
263	57639	CCDC146	coiled-coil domain containing 146	-0.74	1.7	3.1227e-06	752
264	65997	RASL11B	RAS-like, family 11, member B	-0.73	1.7	1.6199e-09	181
265	199964	TMEM61	transmembrane protein 61	-0.73	1.7	3.8180e-06	770
266	10966	RAB40B	RAB40B, member RAS oncogene family	-0.73	1.7	1.0555e-07	478
267	4338	MOCS2	molybdenum cofactor synthesis 2	-0.73	1.7	5.1245e-09	245
268	58155	PTBP2	polypyrimidine tract binding protein 2	-0.73	1.7	3.1450e-09	223
269	23284	LPHN3	latrophilin 3	-0.73	1.7	8.9985e-05	980
270	9079	LDB2	LIM domain binding 2	-0.73	1.7	2.3475e-06	724
271	91612	CHURC1	churchill domain containing 1	-0.73	1.7	3.6066e-04	1048
272	79038	ZFYVE21	zinc finger, FYVE domain containing 21	-0.73	1.7	5.8340e-06	802
273	154791	C7orf55	chromosome 7 open reading frame 55	-0.73	1.7	9.4758e-08	466
274	196294	IMMP1L	IMP1 inner mitochondrial membrane peptidase-like (S. cerevisiae)	-0.72	1.7	1.0131e-07	472
275	3275	PRMT2	protein arginine methyltransferase 2	-0.72	1.6	1.2414e-09	165
276	63920	ZBED8	zinc finger, BED-type containing 8	-0.72	1.6	1.0490e-10	94
277	1123	CHN1	chimerin 1	-0.72	1.6	1.4751e-06	687
278	151556	GPR155	G protein-coupled receptor 155	-0.72	1.6	3.1220e-05	926
279	5108	PCM1	pericentriolar material 1	-0.72	1.6	2.5744e-07	547
280	222663	SCUBE3	signal peptide, CUB domain, EGF-like 3	-0.72	1.6	2.7288e-07	551
281	121227	LRIG3	leucine-rich repeats and immunoglobulin-like domains 3	-0.72	1.6	7.2298e-07	624
282	5999	RGS4	regulator of G-protein signaling 4	-0.71	1.6	1.1219e-03	1083
283	57037	ANKMY2	ankyrin repeat and MYND domain containing 2	-0.71	1.6	6.7895e-10	130
284	84078	KBTBD7	kelch repeat and BTB (POZ) domain containing 7	-0.71	1.6	1.9239e-05	892
285	6210	RPS15A	ribosomal protein S15a	-0.71	1.6	1.0824e-09	158
286	100505687	LINC00888	long intergenic non-protein coding RNA 888	-0.71	1.6	9.3559e-07	645
287	2946	GSTM2	glutathione S-transferase mu 2 (muscle)	-0.71	1.6	2.3767e-05	909
288	5764	PTN	pleiotrophin	-0.71	1.6	4.1942e-04	1055
289	6263	RYR3	ryanodine receptor 3	-0.71	1.6	5.1487e-06	794
290	80333	KCNIP4	Kv channel interacting protein 4	-0.71	1.6	5.0584e-06	791
291	51309	ARMCX1	armadillo repeat containing, X-linked 1	-0.71	1.6	7.9210e-07	629
292	30061	SLC40A1	solute carrier family 40 (iron-regulated transporter), member 1	-0.71	1.6	1.5766e-05	875
293	10395	DLC1	DLC1 Rho GTPase activating protein	-0.71	1.6	7.7395e-09	274
294	2170	FABP3	fatty acid binding protein 3, muscle and heart (mammary-derived growth inhibitor)	-0.70	1.6	1.7209e-05	882
295	81579	PLA2G12A	phospholipase A2, group XIIA	-0.70	1.6	3.9090e-08	394
296	111	ADCY5	adenylate cyclase 5	-0.70	1.6	3.8543e-07	580
297	10559	SLC35A1	solute carrier family 35 (CMP-sialic acid transporter), member A1	-0.70	1.6	1.9652e-07	529

Table 8 – Continued on next page

Table 8 – continued from previous page

298	9132	KCNQ4	potassium voltage-gated channel, KQT-like subfamily, member 4	-0.70	1.6	1.5562e-06	692
299	10480	EIF3M	eukaryotic translation initiation factor 3, subunit M	-0.70	1.6	1.1041e-08	298
300	29028	ATAD2	ATPase family, AAA domain containing 2	-0.70	1.6	1.3745e-09	172
301	55425	GPALPP1	GPALPP motifs containing 1	-0.70	1.6	8.4190e-09	279
302	51449	PCYOX1	prenylcysteine oxidase 1	-0.70	1.6	4.5226e-07	593
303	1429	CRYZ	crystallin, zeta (quinone reductase)	-0.69	1.6	2.9173e-05	921
304	85460	ZNF518B	zinc finger protein 518B	-0.69	1.6	1.5881e-06	695
305	222236	NAPEPLD	N-acyl phosphatidylethanolamine phospholipase D	-0.69	1.6	2.7034e-08	373
306	11215	AKAP11	A kinase (PRKA) anchor protein 11	-0.69	1.6	4.8711e-10	115
307	121536	AEBP2	AE binding protein 2	-0.69	1.6	2.2273e-10	107
308	253782	CERS6	ceramide synthase 6	-0.69	1.6	3.3948e-05	932
309	6262	RYR2	ryanodine receptor 2 (cardiac)	-0.69	1.6	4.4660e-05	946
310	57562	KIAA1377	KIAA1377	-0.69	1.6	3.3749e-07	565
311	57659	ZBTB4	zinc finger and BTB domain containing 4	-0.69	1.6	1.4369e-07	508
312	79694	MANEA	mannosidase, endo-alpha	-0.69	1.6	4.0382e-08	398
313	57584	ARHGAP21	Rho GTPase activating protein 21	-0.69	1.6	8.0014e-07	630
314	100507303	SNHG19	small nucleolar RNA host gene 19 (non-protein coding)	-0.69	1.6	9.4718e-07	646
315	27101	CACYBP	calcyclin binding protein	-0.69	1.6	3.8389e-07	579
316	3206	HOXA10	homeobox A10	-0.69	1.6	1.1549e-06	663
317	51175	TUBE1	tubulin, epsilon 1	-0.69	1.6	3.4171e-04	1044
318	4594	MUT	methylmalonyl CoA mutase	-0.69	1.6	6.8588e-10	132
319	56981	PRDM11	PR domain containing 11	-0.69	1.6	3.0209e-05	924
320	113612	CYP2U1	cytochrome P450, family 2, subfamily U, polypeptide 1	-0.69	1.6	3.9097e-09	234
321	845	CASQ2	calsequestrin 2 (cardiac muscle)	-0.68	1.6	3.0709e-08	384
322	25950	RWDD3	RWD domain containing 3	-0.68	1.6	1.2516e-09	166
323	324	APC	adenomatous polyposis coli	-0.68	1.6	1.4315e-09	175
324	23230	VPS13A	vacuolar protein sorting 13 homolog A (S. cerevisiae)	-0.68	1.6	6.4625e-04	1068
325	5274	SERPINI1	serpin peptidase inhibitor, clade I (neuroserpin), member 1	-0.68	1.6	4.4019e-05	944
326	3752	KCND3	potassium voltage-gated channel, Shal-related subfamily, member 3	-0.68	1.6	6.8950e-05	967
327	6450	SH3BGR	SH3 domain binding glutamate-rich protein	-0.68	1.6	7.2043e-06	822
328	79611	ACSS3	acyl-CoA synthetase short-chain family member 3	-0.68	1.6	1.3291e-07	494
329	4330	MN1	meningioma (disrupted in balanced translocation) 1	-0.68	1.6	1.1187e-06	658
330	11278	KLF12	Kruppel-like factor 12	-0.68	1.6	4.7306e-09	240
331	4091	SMAD6	SMAD family member 6	-0.68	1.6	2.6176e-06	733
332	1152	CKB	creatine kinase, brain	-0.67	1.6	3.9909e-08	397
333	26275	HIBCH	3-hydroxyisobutyryl-CoA hydrolase	-0.67	1.6	1.8310e-08	340
334	123207	C15orf40	chromosome 15 open reading frame 40	-0.67	1.6	7.5936e-09	273
335	51101	ZC2HC1A	zinc finger, C2HC-type containing 1A	-0.67	1.6	2.0034e-06	714

Table 8 – Continued on next page

Table 8 – continued from previous page

336	56890	MDM1	Mdm1 nuclear protein homolog (mouse)	-0.67	1.6	1.5481e-08	322
337	196047	EMX2OS	EMX2 opposite strand/antisense RNA	-0.67	1.6	4.0207e-11	60
338	781	CACNA2D1	calcium channel, voltage-dependent, alpha 2/delta subunit 1	-0.67	1.6	2.2292e-07	536
339	9324	HMGN3	high mobility group nucleosomal binding domain 3	-0.67	1.6	4.3783e-11	63
340	7341	SUMO1	small ubiquitin-like modifier 1	-0.67	1.6	7.4987e-08	447
341	283807	FBXL22	F-box and leucine-rich repeat protein 22	-0.67	1.6	2.1651e-05	904
342	221895	JAZF1	JAZF zinc finger 1	-0.67	1.6	4.6371e-09	238
343	10040	TOM1L1	target of myb1 (chicken)-like 1	-0.66	1.6	6.7034e-07	621
344	51306	FAM13B	family with sequence similarity 13, member B	-0.66	1.6	1.7197e-08	333
345	10916	MAGED2	melanoma antigen family D, 2	-0.66	1.6	1.2705e-08	305
346	2944	GSTM1	glutathione S-transferase mu 1	-0.66	1.6	1.5063e-04	1006
347	92126	DSEL	dermatan sulfate epimerase-like	-0.66	1.6	6.6919e-05	964
348	100652772	NNT-AS1	NNT antisense RNA 1	-0.66	1.6	5.6101e-08	420
349	2036	EPB41L1	erythrocyte membrane protein band 4.1-like 1	-0.66	1.6	2.5694e-07	545
350	55600	ITLN1	intelectin 1 (galactofuranose binding)	-0.66	1.6	5.0720e-05	949
351	57326	PBXIP1	pre-B-cell leukemia homeobox interacting protein 1	-0.66	1.6	4.1911e-08	401
352	100131067	CKMT2-AS1	CKMT2 antisense RNA 1	-0.66	1.6	3.2491e-06	758
353	7088	TLE1	transducin-like enhancer of split 1 (E(sp1) homolog, Drosophila)	-0.66	1.6	4.3451e-07	590
354	7332	UBE2L3	ubiquitin-conjugating enzyme E2L 3	-0.66	1.6	2.6547e-07	550
355	64968	MRPS6	mitochondrial ribosomal protein S6	-0.66	1.6	1.1729e-08	301
356	285237	C3orf38	chromosome 3 open reading frame 38	-0.66	1.6	4.4402e-06	780
357	222865	TMEM130	transmembrane protein 130	-0.66	1.6	2.0504e-08	346
358	2857	GPR34	G protein-coupled receptor 34	-0.66	1.6	2.8642e-07	553
359	90624	LYRM7	LYR motif containing 7	-0.66	1.6	1.0341e-11	44
360	11099	PTPN21	protein tyrosine phosphatase, non-receptor type 21	-0.66	1.6	2.5529e-08	370
361	23066	CAND2	cullin-associated and neddylation-dissociated 2 (putative)	-0.66	1.6	2.8418e-06	743
362	84259	DCUN1D5	DCN1, defective in cullin neddylation 1, domain containing 5	-0.66	1.6	2.9732e-09	220
363	22929	SEPHS1	selenophosphate synthetase 1	-0.65	1.6	7.0139e-11	78
364	10656	KHDRBS3	KH domain containing, RNA binding, signal transduction associated 3	-0.65	1.6	3.5687e-07	572
365	23111	SPG20	spastic paraplegia 20 (Troyer syndrome)	-0.65	1.6	1.4851e-08	316
366	93134	ZNF561	zinc finger protein 561	-0.65	1.6	1.6211e-08	324
367	51318	MRPL35	mitochondrial ribosomal protein L35	-0.65	1.6	3.2009e-07	562
368	84186	ZCCHC7	zinc finger, CCHC domain containing 7	-0.65	1.6	9.9029e-10	155
369	22841	RAB11FIP2	RAB11 family interacting protein 2 (class I)	-0.65	1.6	5.0336e-11	71
370	22873	DZIP1	DAZ interacting zinc finger protein 1	-0.65	1.6	9.9691e-06	843
371	23321	TRIM2	tripartite motif containing 2	-0.65	1.6	4.9355e-09	243
372	375248	ANKRD36	ankyrin repeat domain 36	-0.65	1.6	2.1765e-06	719
373	142679	DUSP19	dual specificity phosphatase 19	-0.65	1.6	4.1422e-07	586

Table 8 – Continued on next page

Table 8 – continued from previous page

374	2195	FAT1	FAT atypical cadherin 1	-0.65	1.6	2.4004e-08	363
375	56899	ANKS1B	ankyrin repeat and sterile alpha motif domain containing 1B	-0.65	1.6	2.3054e-08	360
376	401261	FLJ38717	FLJ38717 protein	-0.65	1.6	1.6154e-05	878
377	1756	DMD	dystrophin	-0.65	1.6	1.4985e-08	317
378	154807	VKORC1L1	vitamin K epoxide reductase complex, subunit 1-like 1	-0.64	1.6	5.9839e-12	39
379	9666	DZIP3	DAZ interacting zinc finger protein 3	-0.64	1.6	1.0125e-07	471
380	29087	THYN1	thymocyte nuclear protein 1	-0.64	1.6	8.2132e-11	84
381	5747	PTK2	protein tyrosine kinase 2	-0.64	1.6	1.1417e-07	486
382	2537	IFI6	interferon, alpha-inducible protein 6	-0.64	1.6	3.3122e-04	1043
383	8706	B3GALNT1	beta-1,3-N-acetylgalactosaminyltransferase 1 (globoside blood group)	-0.64	1.6	3.6579e-06	768
384	79745	CLIP4	CAP-GLY domain containing linker protein family, member 4	-0.64	1.6	1.5739e-05	874
385	116228	COX20	COX20 cytochrome C oxidase assembly factor	-0.64	1.6	1.4886e-06	688
386	114884	OSBPL10	oxysterol binding protein-like 10	-0.64	1.6	2.4686e-06	728
387	9865	TRIL	TLR4 interactor with leucine-rich repeats	-0.64	1.6	2.0561e-04	1020
388	92249	LINC01278	long intergenic non-protein coding RNA 1278	-0.64	1.6	9.5816e-07	647
389	143503	OR51E1	olfactory receptor, family 51, subfamily E, member 1	-0.64	1.6	9.3404e-04	1078
390	2982	GUCY1A3	guanylate cyclase 1, soluble, alpha 3	-0.64	1.6	6.5052e-05	963
391	2926	GRSF1	G-rich RNA sequence binding factor 1	-0.64	1.6	5.9538e-08	424
392	130507	UBR3	ubiquitin protein ligase E3 component n-recognin 3 (putative)	-0.64	1.6	4.4264e-08	406
393	29901	SAC3D1	SAC3 domain containing 1	-0.64	1.6	4.6088e-06	786
394	80036	TRPM3	transient receptor potential cation channel, subfamily M, member 3	-0.64	1.6	9.8509e-06	841
395	92270	ATP6AP1L	ATPase, H+ transporting, lysosomal accessory protein 1-like	-0.64	1.6	3.1252e-06	753
396	57558	USP35	ubiquitin specific peptidase 35	-0.64	1.6	1.1153e-05	854
397	55833	UBAP2	ubiquitin associated protein 2	-0.64	1.6	4.3882e-08	405
398	9055	PRC1	protein regulator of cytokinesis 1	-0.64	1.6	2.4083e-03	1090
399	84553	FAXC	failed axon connections homolog (Drosophila)	-0.63	1.5	1.9253e-06	710
400	57494	RIMKLB	ribosomal modification protein rimK-like family member B	-0.63	1.5	9.8437e-06	840
401	9104	RGN	regucalcin	-0.63	1.5	6.9570e-06	815
402	22903	BTBD3	BTB (POZ) domain containing 3	-0.63	1.5	7.5761e-07	627
403	219287	AMER2	APC membrane recruitment protein 2	-0.63	1.5	2.7907e-04	1030
404	6815	STYX	serine/threonine/tyrosine interacting protein	-0.63	1.5	4.8673e-07	601
405	27247	NFU1	NFU1 iron-sulfur cluster scaffold	-0.63	1.5	1.6428e-08	328
406	79956	ERMP1	endoplasmic reticulum metallopeptidase 1	-0.63	1.5	2.1722e-08	352
407	22859	LPHN1	latrophilin 1	-0.63	1.5	2.1393e-05	901
408	4968	OGG1	8-oxoguanine DNA glycosylase	-0.63	1.5	3.0209e-09	221
409	9240	PNMA1	paraneoplastic Ma antigen 1	-0.63	1.5	1.5044e-07	512

Table 8 – Continued on next page

Table 8 – continued from previous page

410	10675	CSPG5	chondroitin sulfate proteoglycan 5 (neuroglycan C)	-0.62	1.5	3.2551e-05	928
411	348235	SKA2	spindle and kinetochore associated complex subunit 2	-0.62	1.5	1.8926e-09	190
412	9946	CRYZL1	crystallin, zeta (quinone reductase)-like 1	-0.62	1.5	8.3528e-07	635
413	64927	TTC23	tetratricopeptide repeat domain 23	-0.62	1.5	8.6809e-04	1076
414	9767	JADE3	jade family PHD finger 3	-0.62	1.5	1.0349e-09	156
415	79752	ZFAND1	zinc finger, AN1-type domain 1	-0.62	1.5	8.3549e-08	455
416	90843	TCEAL8	transcription elongation factor A (SII)-like 8	-0.62	1.5	1.5841e-09	179
417	3209	HOXA13	homeobox A13	-0.62	1.5	4.4886e-05	947
418	7881	KCNAB1	potassium voltage-gated channel, shaker-related subfamily, beta member 1	-0.62	1.5	1.0344e-05	849
419	5414	SEPT4	septin 4	-0.62	1.5	1.0268e-05	848
420	3417	IDH1	isocitrate dehydrogenase 1 (NADP+), soluble	-0.62	1.5	3.5003e-04	1047
421	54741	LEPROT	leptin receptor overlapping transcript	-0.62	1.5	1.9565e-06	711
422	7103	TSPAN8	tetraspanin 8	-0.62	1.5	7.9667e-04	1072
423	23336	SYNM	synemin, intermediate filament protein	-0.62	1.5	6.9347e-06	814
424	3176	HNMT	histamine N-methyltransferase	-0.62	1.5	3.8160e-07	577
425	412	STS	steroid sulfatase (microsomal), isozyme S	-0.62	1.5	1.9865e-06	712
426	150864	FAM117B	family with sequence similarity 117, member B	-0.62	1.5	1.2702e-05	861
427	3730	KAL1	Kallmann syndrome 1 sequence	-0.62	1.5	1.4655e-03	1087
428	54756	IL17RD	interleukin 17 receptor D	-0.62	1.5	6.0077e-04	1065
429	9474	ATG5	autophagy related 5	-0.62	1.5	7.5753e-10	138
430	152926	PPM1K	protein phosphatase, Mg2+/Mn2+ dependent, 1K	-0.62	1.5	1.1192e-06	659
431	6502	SKP2	S-phase kinase-associated protein 2, E3 ubiquitin protein ligase	-0.62	1.5	5.1397e-06	793
432	6645	SNTB2	syntrophin, beta 2 (dystrophin-associated protein A1, 59kDa, basic component 2)	-0.62	1.5	1.7005e-05	881
433	56271	BEX4	brain expressed, X-linked 4	-0.62	1.5	6.1849e-09	261
434	4833	NME4	NME/NM23 nucleoside diphosphate kinase 4	-0.61	1.5	1.2326e-06	670
435	5701	PSMC2	proteasome (prosome, macropain) 26S subunit, ATPase, 2	-0.61	1.5	2.1834e-08	354
436	80817	CEP44	centrosomal protein 44kDa	-0.61	1.5	1.1771e-09	162
437	79884	MAP9	microtubule-associated protein 9	-0.61	1.5	1.2364e-07	492
438	753	LDLRAD4	low density lipoprotein receptor class A domain containing 4	-0.61	1.5	5.2196e-05	951
439	55120	FANCL	Fanconi anemia, complementation group L	-0.61	1.5	8.5485e-07	637
440	57456	KIAA1143	KIAA1143	-0.61	1.5	1.3836e-07	499
441	29103	DNAJC15	DnaJ (Hsp40) homolog, subfamily C, member 15	-0.61	1.5	1.4564e-05	870
442	79953	SYNDIG1	synapse differentiation inducing 1	-0.61	1.5	6.3695e-06	807

Table 8 – Continued on next page

Table 8 – continued from previous page

443	23142	DCUN1D4	DCN1, defective in cullin neddylation 1, domain containing 4	-0.61	1.5	1.7544e-08	335
444	1410	CRYAB	crystallin, alpha B	-0.61	1.5	2.2314e-06	721
445	81624	DIAPH3	diaphanous-related formin 3	-0.61	1.5	8.0155e-05	972
446	23522	KAT6B	K(lysine) acetyltransferase 6B	-0.61	1.5	1.2928e-08	307
447	142940	TRUB1	TruB pseudouridine (psi) synthase family member 1	-0.61	1.5	4.7591e-09	241
448	57545	CC2D2A	coiled-coil and C2 domain containing 2A	-0.61	1.5	6.1561e-09	259
449	153222	CREBRF	CREB3 regulatory factor	-0.61	1.5	3.2957e-05	930
450	27122	DKK3	dickkopf WNT signaling pathway inhibitor 3	-0.61	1.5	9.9778e-06	844
451	7329	UBE2I	ubiquitin-conjugating enzyme E2I	-0.61	1.5	2.9430e-07	555
452	4041	LRP5	low density lipoprotein receptor-related protein 5	-0.61	1.5	9.9913e-05	985
453	5592	PRKG1	protein kinase, cGMP-dependent, type I	-0.61	1.5	2.1388e-10	106
454	8667	EIF3H	eukaryotic translation initiation factor 3, subunit H	-0.61	1.5	5.4471e-04	1063
455	5549	PRELP	proline/arginine-rich end leucine-rich repeat protein	-0.60	1.5	3.1216e-04	1037
456	56944	OLFML3	olfactomedin-like 3	-0.60	1.5	8.0817e-06	830
457	27018	NGFRAP1	nerve growth factor receptor (TNFRSF16) associated protein 1	-0.60	1.5	6.4727e-08	435
458	3764	KCNJ8	potassium inwardly-rectifying channel, subfamily J, member 8	-0.60	1.5	3.9537e-04	1052
459	54876	DCAF16	DDB1 and CUL4 associated factor 16	-0.60	1.5	8.2429e-05	976
460	51069	MRPL2	mitochondrial ribosomal protein L2	-0.60	1.5	1.8958e-07	525
461	359948	IRF2BP2	interferon regulatory factor 2 binding protein 2	-0.60	1.5	5.6677e-05	955
462	79901	CYBRD1	cytochrome b reductase 1	-0.60	1.5	7.3795e-06	823
463	51661	FKBP7	FK506 binding protein 7	-0.60	1.5	3.9869e-07	584
464	5217	PFN2	profilin 2	-0.60	1.5	6.3294e-07	618
465	11030	RBPMS	RNA binding protein with multiple splicing	-0.60	1.5	7.0092e-06	819
466	57531	HACE1	HECT domain and ankyrin repeat containing E3 ubiquitin protein ligase 1	-0.60	1.5	2.3069e-08	361
467	64426	SUDS3	suppressor of defective silencing 3 homolog (S. cerevisiae)	-0.60	1.5	1.3479e-07	496
468	8175	SF3A2	splicing factor 3a, subunit 2, 66kDa	-0.60	1.5	4.8859e-08	412
469	117583	PARD3B	par-3 family cell polarity regulator beta	-0.60	1.5	1.2312e-08	303
470	57396	CLK4	CDC-like kinase 4	-0.60	1.5	6.2490e-05	958
471	4128	MAOA	monoamine oxidase A	-0.60	1.5	3.1514e-03	1095
472	10632	ATP5L	ATP synthase, H+ transporting, mitochondrial Fo complex, subunit G	-0.60	1.5	9.3612e-09	289
473	81533	ITFG1	integrin alpha FG-GAP repeat containing 1	-0.60	1.5	5.5969e-10	122
474	11168	PSIP1	PC4 and SFRS1 interacting protein 1	-0.60	1.5	6.9862e-08	439
475	5578	PRKCA	protein kinase C, alpha	-0.60	1.5	8.1779e-05	975
476	80176	SPSB1	splA/ryanodine receptor domain and SOCS box containing 1	-0.60	1.5	1.1745e-06	665

Table 8 – Continued on next page

Table 8 – continued from previous page

477	9468	PCYT1B	phosphate cytidylyltransferase 1, choline, beta	-0.60	1.5	9.8659e-04	1080
478	9706	ULK2	unc-51 like autophagy activating kinase 2	-0.60	1.5	1.8753e-09	189
479	80823	BHLHB9	basic helix-loop-helix domain containing, class B, 9	-0.60	1.5	1.7969e-09	187
480	9702	CEP57	centrosomal protein 57kDa	-0.60	1.5	1.4614e-06	686
481	339745	SPOPL	speckle-type POZ protein-like	-0.60	1.5	3.2375e-06	757
482	9910	RABGAP1L	RAB GTPase activating protein 1-like	-0.59	1.5	5.5083e-08	418
483	23588	KLHDC2	kelch domain containing 2	-0.59	1.5	1.9927e-06	713
484	274	BIN1	bridging integrator 1	-0.59	1.5	1.6381e-08	327
485	79710	MORC4	MORC family CW-type zinc finger 4	-0.59	1.5	6.4124e-06	808
486	5648	MASP1	mannan-binding lectin serine peptidase 1 (C4/C2 activating component of Ra-reactive factor)	-0.59	1.5	2.4820e-03	1092
487	10486	CAP2	CAP, adenylate cyclase-associated protein, 2 (yeast)	-0.59	1.5	1.2691e-09	168
488	154043	CNKSR3	CNKSR family member 3	-0.59	1.5	1.8502e-05	887
489	85458	DIXDC1	DIX domain containing 1	-0.59	1.5	2.8309e-06	742
490	25934	NIPSNAP3A	nipsnap homolog 3A (C. elegans)	-0.59	1.5	5.2666e-05	953
491	10426	TUBGCP3	tubulin, gamma complex associated protein 3	-0.59	1.5	1.3881e-07	501
492	4139	MARK1	MAP/microtubule affinity-regulating kinase 1	-0.59	1.5	2.6194e-09	209
493	10314	LANCL1	LanC lantibiotic synthetase component C-like 1 (bacterial)	-0.59	1.5	8.0920e-10	143
494	26227	PHGDH	phosphoglycerate dehydrogenase	-0.59	1.5	8.6082e-07	639
495	3217	HOXB7	homeobox B7	-0.59	1.5	1.4603e-04	1004
496	55082	ARGLU1	arginine and glutamate rich 1	-0.59	1.5	1.1097e-06	657
497	3622	ING2	inhibitor of growth family, member 2	-0.59	1.5	7.1753e-08	445
498	5164	PDK2	pyruvate dehydrogenase kinase, isozyme 2	-0.59	1.5	1.7679e-08	336
499	2791	GNG11	guanine nucleotide binding protein (G protein), gamma 11	-0.59	1.5	2.1346e-04	1024
500	374354	NHLRC2	NHL repeat containing 2	-0.59	1.5	1.1793e-09	163
501	79633	FAT4	FAT atypical cadherin 4	-0.59	1.5	5.6056e-07	610
502	64112	MOAP1	modulator of apoptosis 1	-0.59	1.5	7.3490e-08	446
503	4776	NFATC4	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 4	-0.59	1.5	1.5276e-05	873
504	89944	GLB1L2	galactosidase, beta 1-like 2	-0.59	1.5	3.4536e-06	764
505	51133	KCTD3	potassium channel tetramerization domain containing 3	-0.59	1.5	1.5818e-07	517

3.6 Estimation of the number of DE genes with a smaller sample size

Here we estimate the number of DE genes with a smaller sample size $n = 20$, where 10 samples correspond to FIR-affected infants and the other 10 to non-FIR affected infants, to enable the comparison of the statistical power in our data to detect DE genes with other studies with smaller sample sizes. For this purpose, we generate 100 bootstrapped data sets and perform the DE analysis on each of them.

At each bootstrap iteration we sample first uniformly at random one sample from each batch and FIR status, leading to 3 samples per group, to ensure that FIR outcome is not correlated with the batch indicator variable at the resulting bootstrapped data set. The remaining 7 samples are drawn uniformly

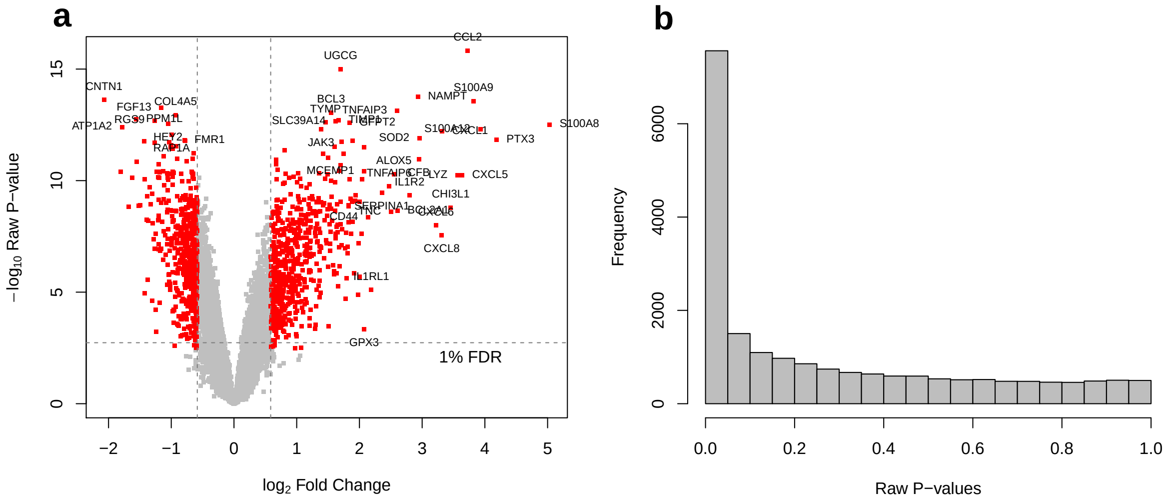

Figure 10. Differential expression between FIR and non-FIR samples.. (a) Volcano plot. (b) Raw *p*-value distribution. The differential expression analysis was carried out using a linear model that included as main effects FIR outcome, the batch indicator variable and surrogate variables estimated by SVA.

at random with replacement among any of the n =43 individuals.

```
> if (!file.exists(file.path(cacheDir_DE,
                              paste0(cachePrefix_DE, "bootstrapSmall_n.RData")))) {
    sample1xbatch <- function(eset, FIR) {
      sapply(unique(eset$Batch),
        function(b, FIR) {
          idx <- which(eset$FIR == FIR & eset$Batch == b)
          if (length(idx) > 1)
            idx <- sample(idx, size=1)
          idx
        }, FIR)
    }

    nbootstrapsamples <- 100
    nbygrp <- 10
    bsamples <- lapply(1:nbootstrapsamples,
                       function(x) list(yes=sample1xbatch(frmaEset, "yes"),
                                        no=sample1xbatch(frmaEset, "no")))


    nb <- length(unique(frmaEset$Batch))
    bsamples <- lapply(bsamples,
                       function(x) list(yes=c(x$yes,
                                              sample(which(frmaEset$FIR == "yes"),
                                                     size=nbygrp-nb, replace=TRUE)),
                                        no=c(x$no,
                                              sample(which(frmaEset$FIR == "no"),
                                                     size=nbygrp-nb, replace=TRUE))))

    stopifnot(all(frmaEset$FIR[as.vector(sapply(bsamples, function(x) x$yes))] == "yes")) ## QC
    stopifnot(all(frmaEset$FIR[as.vector(sapply(bsamples, function(x) x$no))] == "no")) ## QC

    res <- bplapply(bsamples,
             function(bs) {
```

```
## select bootstrapped samples
beset <- frmaEset[, c(bs$yes, bs$no)]
tab <- table(beset$FIR)
stopifnot(all(tab == nbygrp)) ## QC
tab <- table(data.frame(FIR=beset$FIR, Batch=beset$Batch))
stopifnot(all(tab > 0)) ## QC

## select probesets passing functional non-specific filters
beset <- beset[featureNames(beset) %in% filteredProbesets, ]
featureNames(beset) <- featureNames(frmaEsetFilt)

## estimate SVs
IQRs <- esApply(beset, 1, IQR)
mod <- model.matrix(~ factor(beset$FIR) + factor(beset$Batch))
mod0 <- model.matrix(~ factor(beset$Batch))
besetSV <- sva(dat=exprs(beset[IQRs > quantile(IQRs, prob=0.9), ]),
               mod=mod, mod0=mod0)

## calculate IQR over gene expression after cleaning batch
## and surrogate variable effects
besetCleanExps <- removeBatchEffect(exprs(beset), batch=beset$Batch,
                                    covariates=besetSV$sv,
                                    design=model.matrix(~ factor(beset$FIR)))

IQRs <- apply(besetCleanExps, 1, IQR)
## calculate the fraction of genes to discard by IQR at each quantile
filteringQuantiles <- seq(0.1, 0.9, by=0.1)
filteringQuantilesIQRs <- quantile(IQRs, probs=filteringQuantiles)
genesByFilterCutoff <- sapply(filteringQuantilesIQRs,
                              function(q, IQRs) names(IQRs)[IQRs > q], IQRs)

design <- cbind(mod, besetSV$sv)
colnames(design) <- c("FIR", "FIRvsNoFIR",
                      paste0("Batch", seq_len(length(unique(beset$Batch))-1)),
                      paste0("SV", seq_len(besetSV$n)))
fit<- lmFit(beset, design)
fit<- eBayes(fit, trend=TRUE)
tt <- topTable(fit, coef="FIRvsNoFIR", number=Inf, adjust.method="fdr")
tt$ID <- rownames(tt)
DEgenesByFilterCutoff <-
  lapply(genesByFilterCutoff,
    function(genes, tt, fcCutoff, FDRcutoff) {
      tt <- tt[tt$ID %in% genes, ]
      tt$adj.P.Val <- p.adjust(tt$P.Value, method="fdr")
      tt$ID[tt$adj.P.Val <= FDRcutoff]
    }, tt, fcCutoff, FDRcutoff)
nDEgenesByFilterCutoff <- sapply(DEgenesByFilterCutoff, length)
## rev() is used here below to select highest quantile with ties
whqmaxDE <- length(nDEgenesByFilterCutoff) -
            which.max(rev(nDEgenesByFilterCutoff)) + 1
tt <- tt[tt$ID %in% genesByFilterCutoff[[whqmaxDE]], ]
tt$adj.P.Val <- p.adjust(tt$P.Value, method="fdr")
stopifnot(identical(tt$ID[tt$adj.P.Val <= FDRcutoff],
                    DEgenesByFilterCutoff[[whqmaxDE]])) ## QC
pvalues <- tt$P.Value
logfcs <- tt$logFC
```

```
            names(pvalues) <- names(logfcs) <- tt$ID
            list(pvalues=pvalues, logfcs=logfcs)
          }, BPPARAM=MulticoreParam(workers=10))

    bpvalues <- lapply(res, function(x) x$pvalues)
    blogfcs <- lapply(res, function(x) x$logfcs)

    save(bsamples, bpvalues, blogfcs,
         file=file.path(cacheDir_DE, paste0(cachePrefix_DE, "bootstrapSmall_n.RData")))
  } else
    load(file.path(cacheDir_DE, paste0(cachePrefix_DE, "bootstrapSmall_n.RData")))
```

We examine the following distributions of values:

- number of genes called DE at 5% FDR obtained throughout the 100 bootstrapped data sets.

```
> summary(sapply(lapply(bpvalues, function(x) x[p.adjust(x, method="fdr") < 0.05]),
                 length))

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
    890    3127    4125    4532    5985   10300
```

- $-\log_{10}$ raw p-values of all genes called DE at 5% FDR pooled from all bootstrapped data sets.

```
> bpvaluesFDR05 <- lapply(bpvalues, function(x) x[p.adjust(x, method="fdr") < 0.05])
> summary(-log10(unlist(bpvaluesFDR05, use.names=FALSE)))

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.426   2.166   2.872   3.205   3.884  15.240
```

- $|\log_2$ fold-change$|$ of all genes called DE at 5% FDR pooled from all bootstrapped data sets.

```
> blogfcsFDR05 <- mapply(function(lfc, pv) lfc[p.adjust(pv, method="fdr") < 0.05],
                         blogfcs, bpvalues)
> summary(2^abs(unlist(blogfcsFDR05)))

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.060   1.242   1.348   1.534   1.544 151.400
```

In comparison with the corresponding values from the analysis of the entire data set with $n = 43$:

- number of genes called DE at 5% FDR.

```
> sum(ttAllIQRfilt$adj.P.Val < 0.05)

[1] 5461
```

- $-\log_{10}$ raw p-values of genes called DE at 5% FDR.

```
> summary(-log10(ttAllIQRfilt$P.Value[ttAllIQRfilt$adj.P.Val < 0.05]))

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.646   2.353   3.429   4.010   5.107  15.820
```

- $|\log_2$ fold-change$|$ of genes called DE at 5% FDR.

```
> summary(2^abs(ttAllIQRfilt$logFC[ttAllIQRfilt$adj.P.Val < 0.05]))

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.089   1.193   1.278   1.427   1.441  32.700
```

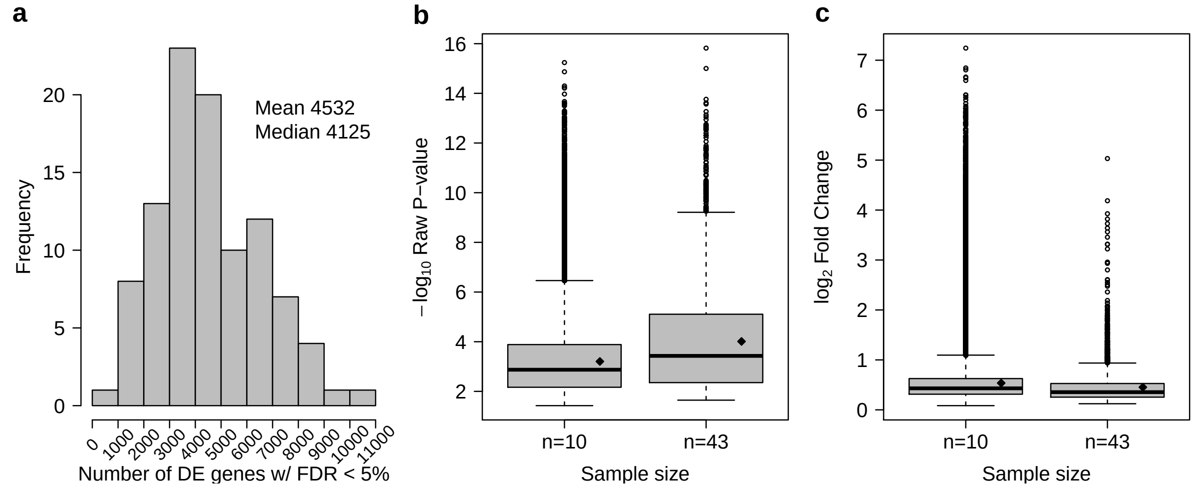

Figure 11. Comparison of the amount of DE between FIR and non-FIR affected infants, obtained from bootstrapped data sets of $n = 10$ individuals per FIR status group, with respect to the original data with n =20155 individuals. (a) Distribution of the number of DE genes. (b) Distribution of raw p-values in $\log_{10}$ scale. (c) Distribution of fold-changes in $\log_2$ scale. In (a) and (b) diamonds indicate mean values.

3.7 Comparison between FIRS and FIR signature

Here we compare our results to the signature of 541 DE genes between FIRS-affected and non-FIRS-affected infants, reported in (9). We start importing the data contained in an excel file kindly sent to us by Sally Madsen-Bouterse.

```
> library(xlsx)
> fname <- file.path("annotData", "Madsen-BouterseRomeroDraghici10_FIRSdeGenes.xls")
> cache(FIRSdeGenesTable <- read.xlsx(fname, sheetName="Gene List", stringsAsFactors=FALSE),
        dir=cacheDir_DE, prefix=cachePrefix_DE)
> dim(FIRSdeGenesTable)
[1] 548   8
> head(FIRSdeGenesTable)
         Target  Transcript   Accession  Symbol       ProbeID      P.Value
1  GI_9966888-S  GI_9966888 NM_020406.1    PRV1  GI_9966888-S 0.0002705998
2 GI_34222165-S GI_34222165 NM_144637.2 ZDHHC19 GI_34222165-S 0.0002705998
3  GI_4504394-S  GI_4504394 NM_002115.1     HK3  GI_4504394-S 0.0002705998
4 GI_32895360-A GI_32895360 NM_181879.1    LIR9 GI_32895360-A 0.0002705998
5 GI_32481214-S GI_32481214 NM_001770.3    CD19 GI_32481214-S 0.0002705998
6  GI_4757873-S  GI_4757873 NM_004334.1    BST1  GI_4757873-S 0.0002705998
  Fold.change in.FIRS
1    8.417706       1
2    8.083178       1
3    2.924218       1
4    2.630003       1
5    2.152937      -1
6    1.864776       1
> tail(FIRSdeGenesTable)
           Target  Transcript   Accession  Symbol       ProbeID    P.Value
543  GI_4557830-S  GI_4557830 NM_000281.1    PCBD  GI_4557830-S 0.04908011
544 GI_28416948-S GI_28416948 NM_018384.2  IAN4L1 GI_28416948-S 0.04919319
545 GI_18426910-S GI_18426910 NM_080792.1  PTPNS1 GI_18426910-S 0.04925700
546 GI_22165423-A GI_22165423 NM_007234.3   DCTN3 GI_22165423-A 0.04925700
547  GI_7661537-S  GI_7661537 NM_015379.1    BRI3  GI_7661537-S 0.04927371
```

```
548  GI_7705778-S  GI_7705778 NM_016013.1 NDUFAF1  GI_7705778-S 0.04927371
    Fold.change in.FIRS
543    1.216383       1
544    1.774117      -1
545    1.744894       1
546    1.209690       1
547    1.578615       1
548    1.220065       1
```

This table contains 548 rows corresponding to targets in the Illumina microarray platform, which in the publication (9) represented 541 DE genes. To enable a comparison with our list of DE genes in FIR we are going to map these targets to Entrez Gene Identifiers using first the RefSeq identifier in the Accession column.

```
> library(org.Hs.eg.db)
> acc2eg <- select(org.Hs.eg.db, keys=gsub(".[0-9]+$", "", FIRSdeGenesTable$Accession),
                   columns="ENTREZID", keytype="REFSEQ")
> stopifnot(nrow(acc2eg) == nrow(FIRSdeGenesTable)) ## QC
> stopifnot(identical(gsub(".[0-9]+$", "", FIRSdeGenesTable$Accession), acc2eg$REFSEQ)) ## QC
> FIRSdeGenesTable <- cbind(FIRSdeGenesTable, ENTREZID=as.character(acc2eg$ENTREZID),
                       stringsAsFactors=FALSE)
> sum(is.na(FIRSdeGenesTable$ENTREZID))
[1] 50
```

We are going to try to map the 50 targets for which their corresponding RefSeq identifiers did not match an Entrez gene using the reported gene symbol.

```
> whisNA <- which(is.na(FIRSdeGenesTable$ENTREZID))
> sym2eg <- select(org.Hs.eg.db, keys=FIRSdeGenesTable$Symbol[whisNA], columns="ENTREZID",
                   keytype="SYMBOL")
> stopifnot(nrow(sym2eg) == length(whisNA)) ## QC
> sum(is.na(sym2eg$ENTREZID))
[1] 45
> alias2eg <- select(org.Hs.eg.db, keys=sym2eg$SYMBOL[is.na(sym2eg$ENTREZID)],
                     columns="ENTREZID", keytype="ALIAS")
> stopifnot(nrow(alias2eg) == sum(is.na(sym2eg$ENTREZID))) ## QC
> sym2eg[is.na(sym2eg$ENTREZID), "ENTREZID"] <- alias2eg$ENTREZID
> FIRSdeGenesTable[whisNA, "ENTREZID"] <- sym2eg$ENTREZID
> sum(is.na(FIRSdeGenesTable$ENTREZID))
[1] 39
> sum(!is.na(FIRSdeGenesTable$ENTREZID) & duplicated(FIRSdeGenesTable$ENTREZID))
[1] 7
```

So, finally we have 39 illumina targets for which no Entrez gene identifier has been found. On the other hand we have 7 duplicated Entrez gene identifiers. To discard duplicates, we will keep those targets with largest absolute fold-change.

```
> dupmask <- FIRSdeGenesTable$ENTREZID %in% FIRSdeGenesTable$ENTREZID[duplicated(FIRSdeGenesTable
> whdups <- which(!is.na(FIRSdeGenesTable$ENTREZID) & dupmask)
> length(whdups)
[1] 14
> fcxeg <- split(whdups, FIRSdeGenesTable[whdups, "ENTREZID"])
> fcxeg <- sapply(fcxeg, function(wh, fc) wh[which.max(fc[wh])],
                  FIRSdeGenesTable$Fold.change)
> dim(FIRSdeGenesTable)
[1] 548   9
```

```
> FIRSdeGenesTable <- FIRSdeGenesTable[-setdiff(whdups, fcxeg), ]
> dim(FIRSdeGenesTable)
[1] 541   9
> stopifnot(!any(!is.na(FIRSdeGenesTable$ENTREZID) & duplicated(FIRSdeGenesTable$ENTREZID))) ## 
> sum(is.na(FIRSdeGenesTable$ENTREZID))
[1] 39
> FIRSdeGenes <- FIRSdeGenesTable$ENTREZID[!is.na(FIRSdeGenesTable$ENTREZID)]
> length(FIRSdeGenes)
[1] 502
> save(FIRSdeGenesTable, file=file.path(cacheDir_DE, paste0(cachePrefix_DE, "FIRSdeGenesTableEntr
```

Therefore, from the initial list of targets reported as DE in FIRS (9) we have been able to map 502 of them to current Entrez Gene identifiers, which we will use for comparison with our results. We start this comparison by first assessing the overlap between this list of FIRS DE genes and our results at different levels:

```
> N <- nrow(ttAll)
> m <- length(intersect(FIRSdeGenes, ttAll$ID))
> n <- sum(ttAllIQRfilt$adj.P.Val < FDRcutoff)
> k <- length(intersect(ttAllIQRfilt$ID[ttAllIQRfilt$adj.P.Val < FDRcutoff], FIRSdeGenes))
> t <- array(c(k, n - k, m - k, N + k - n - m), dim = c(2, 2),
             dimnames=list(FIRS=c("in", "out"), FIR=c("yes", "no")))
> t
     FIR
FIRS   yes    no
  in   185   298
  out 3711 15961
> fisher.test(t, alternative="greater")
        Fisher's Exact Test for Count Data

data:  t
p-value < 2.2e-16
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
 2.270508      Inf
sample estimates:
odds ratio
  2.670074
> n <- length(DEgenes)
> k <- length(intersect(DEgenes, FIRSdeGenes))
> t <- array(c(k, n - k, m - k, N + k - n - m), dim = c(2, 2),
             dimnames=list(FIRS=c("in", "out"), FIR=c("yes", "no")))
> t
     FIR
FIRS   yes    no
  in    79   404
  out 1018 18654
> fisher.test(t, alternative="greater")
        Fisher's Exact Test for Count Data

data:  t
p-value < 2.2e-16
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
```

```
 2.874193      Inf
sample estimates:
odds ratio
  3.582724
```

There is a significant overlap between FIR and FIRS DE genes either when FIR genes are called DE at FDR <1% only or when the are also further restricted to those with a minimum 1.5-fold change. Among the overlapping genes between FIR and FIRS we are going to compare their fold changes.

```
> FIRSupdownregulation <- FIRSdeGenesTable[!is.na(FIRSdeGenesTable$ENTREZID), "in.FIRS"]
> tabFIRS <- table(FIRSupdownregulation) / length(FIRSupdownregulation)
> tabFIRS
FIRSupdownregulation
      -1        1
0.438247 0.561753
> FIRSlfc <- FIRSupdownregulation *
               log2(FIRSdeGenesTable[!is.na(FIRSdeGenesTable$ENTREZID), "Fold.change"])
> names(FIRSlfc) <- FIRSdeGenesTable$ENTREZID[!is.na(FIRSdeGenesTable$ENTREZID)]
> tabFIRfdr05 <- table(sign(ttAllIQRfilt$logFC[ttAllIQRfilt$adj.P.Val < 0.05])) /
                 sum(ttAllIQRfilt$adj.P.Val < 0.05)
> tabFIRfdr05

       -1         1
0.6062992 0.3937008
> tabFIRfdr01 <- table(sign(ttAllIQRfilt$logFC[ttAllIQRfilt$adj.P.Val < 0.01])) /
                 sum(ttAllIQRfilt$adj.P.Val < 0.01)
> tabFIRfdr01

       -1         1
0.6203799 0.3796201
> tabFIRfdr01fc1.5 <- table(sign(ttDEgenes$logFC)) / nrow(ttDEgenes)
> tabFIRfdr01fc1.5

       -1         1
0.4603464 0.5396536
> FIRlfc <- ttAll$logFC
> names(FIRlfc) <- ttAll$ID
> commonGenes <- intersect(names(FIRSlfc), names(FIRlfc))
> length(commonGenes)
[1] 483
> cor.test(FIRSlfc[commonGenes], FIRlfc[commonGenes])
        Pearson's product-moment correlation

data:  FIRSlfc[commonGenes] and FIRlfc[commonGenes]
t = 9.9646, df = 481, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.3368635 0.4849756
sample estimates:
      cor
0.4136527
> commonGenesFDR05 <- intersect(commonGenes, ttAllIQRfilt$ID[ttAllIQRfilt$adj.P.Val < 0.05])
> cor.test(FIRSlfc[commonGenesFDR05], FIRlfc[commonGenesFDR05])
        Pearson's product-moment correlation

data:  FIRSlfc[commonGenesFDR05] and FIRlfc[commonGenesFDR05]
t = 8.6219, df = 224, p-value = 1.332e-15
```

```
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.3943597 0.5911617
sample estimates:
      cor
0.4991707
> commonGenesFDR01 <- intersect(commonGenes, ttAllIQRfilt$ID[ttAllIQRfilt$adj.P.Val < 0.01])
> cor.test(FIRSlfc[commonGenesFDR01], FIRlfc[commonGenesFDR01])
        Pearson's product-moment correlation

data:  FIRSlfc[commonGenesFDR01] and FIRlfc[commonGenesFDR01]
t = 8.1245, df = 183, p-value = 6.484e-14
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.4003305 0.6135575
sample estimates:
      cor
0.5148631
> commonGenesFDR01FC1.5 <- intersect(commonGenes, DEgenes)
> cor.test(FIRSlfc[commonGenesFDR01FC1.5], FIRlfc[commonGenesFDR01FC1.5])
        Pearson's product-moment correlation

data:  FIRSlfc[commonGenesFDR01FC1.5] and FIRlfc[commonGenesFDR01FC1.5]
t = 3.7717, df = 77, p-value = 0.0003164
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.1904058 0.5665347
sample estimates:
     cor
0.394891
```

There are 483 genes among the 20155 that have been tested for DE in the filtered FIR data set, that form part of the 502 in the FIRS gene signature with mapped Entrez Gene identifiers. While the proportions of up and down regulated genes in the FIR data are very different from the ones in the FIRS signature for genes called DE at FDR<5% and FDR<1%, they become very similar for genes called DE at FDR<1% with a minimum 1.5-fold change.

In Figure 13 we can see the comparison of the $\log_2$-fold changes between both data sets. While there are many sources of potential different biases between the two data sets, such as, RNA source material (cord tissue vs. cord blood), sample-preparation protocols or microarray technology (Affymetrix vs. Illumina), there is a positive correlation between the $\log_2$-fold changes of genes called DE with FDR<5%.

3.8 Differential expression analysis in severe FIR

The association of intraventricular hemorraghe (IVH) and white matter disease (WMD) with stronger inflammation allows one to define a surrogate of severe FIR as those FIR-affected infants that also develop IVH and WMD. We search here for a molecular signature of severe FIR by comparing infants affected by all three factors (FIR, IVH and WMD) vs. those who did not have any of them. In Table 9 we can see the number of infants cross-classified by FIR, IVH and WMD status.

Due to the limited available sample size and the loss of statistical power that follows considering 3-way interaction effects we are not going to include the batch indicator variable in the model, letting surrogate variables estimated by SVA to adjust for that effect.

```
> library(Hmisc) ## for generating n-way x-tables
> eset <- frmaEsetFilt
> eset$FIR <- factor(eset$FIR)
> eset$Ivh <- factor(eset$Ivh)
```

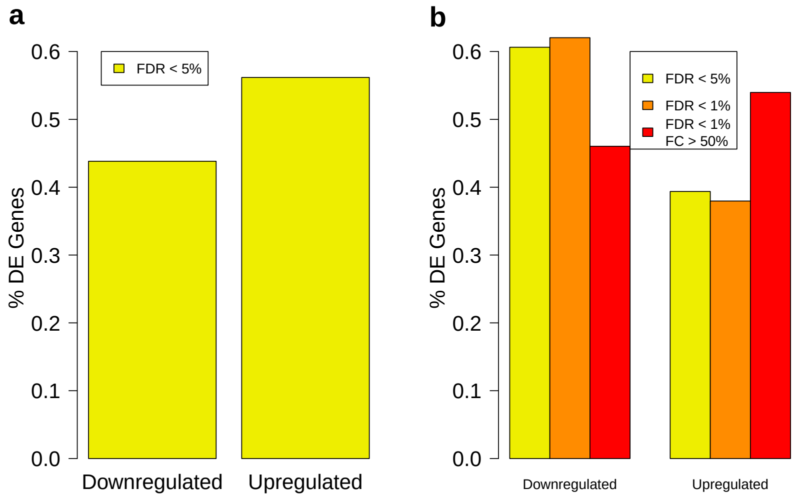

Figure 12. Comparison of the proportion of up and downregulated genes called DE in FIR and FIRS. The y-axis shows the proportion of genes called DE. (a) Proportions for the study of FIRS (9). (b) Proportions for the study of FIR described in this document.

IVH	WMD	FIR	no	yes
no	no		22	10
	yes		0	1
yes	no		1	3
	yes		2	4

Table 9. Infants cross-classified by FIR, IVH and WMD status.

```
> eset$Wmd <- factor(eset$Wmd)
> df <- data.frame(FIR=eset$FIR, IVH=eset$Ivh, WMD=eset$Wmd)
> xtabFIRIVHWMD <- ftable(FIR ~ IVH + WMD, data=df)
> xtabFIRIVHWMD
        FIR no yes
IVH WMD
no  no      22  10
    yes      0   1
yes no       1   3
    yes      2   4
```

The pipeline for detecting DE genes is the same we use before. We run it here all at once. In this case, we build our linear model by forming a new variable specifying all 3-way combinations of factors FIR, IVH and WMD status, which become the effects employed to predict the expression level of each gene. Then we use the corresponding estimated coefficients to interrogate our contrast of interest, which involves the comparison of $n = 4$ infants affected by FIR, IVH and WMD against $n = 22$ infants not affected by any of these phenotypes.

```
> FIRIvhWmd <- factor(paste(eset$FIR, eset$Ivh, eset$Wmd, sep="."))
> mod <- model.matrix(~ 0 + FIRIvhWmd, eset)
> colnames(mod) <- levels(FIRIvhWmd)
> mod0 <- model.matrix(~ 1, eset)
> IQRs <- esApply(eset, 1, IQR)
> cache(svFIRIVHWMD <- sva(dat=exprs(eset[IQRs > quantile(IQRs, prob=0.90), ]), mod=mod, mod0=mo
        dir=cacheDir_DE, prefix=cachePrefix_DE)
```

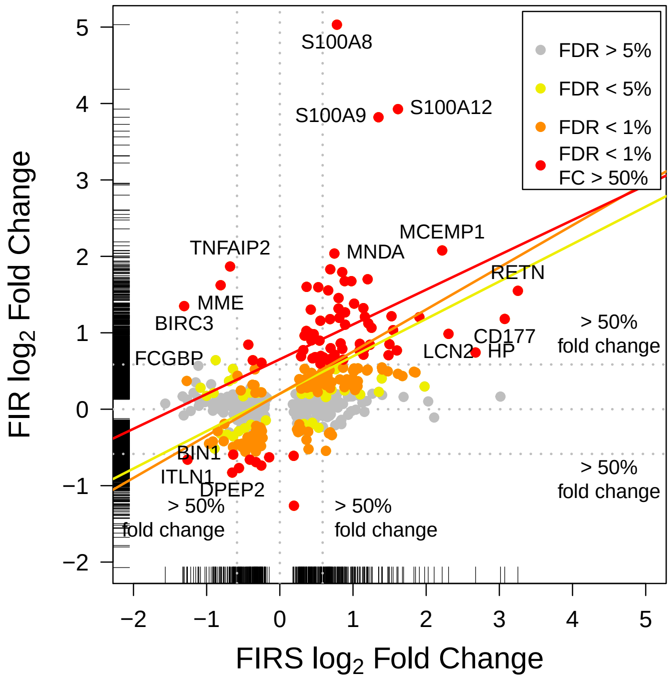

Figure 13. Comparison of fold-changes between genes differentially expressed in FIR and FIRS. In the y-axis the FIR fold changes as function of the FIRS fold-changes in the x-axis, both in $\log_2$ scale. Tick marks show $\log_2$-fold change values for all genes called differentially expressed at FDR<5%, while dots display only genes matching between FIRS and FIR data sets. Regression lines are calculated from the dots with corresponding colors.

```
> cleanExps <- removeBatchEffect(exprs(eset), covariates=svFIRIVHWMD$sv, design=mod)
> cleanEsetIVHWMD <- eset
> exprs(cleanEsetIVHWMD) <- cleanExps
> IQRs <- esApply(cleanEsetIVHWMD, 1, IQR)
> filteringQuantiles <- seq(0.1, 0.9, by=0.1)
> filteringQuantilesIQRs <- quantile(IQRs, probs=filteringQuantiles)
> genesByFilterCutoff <- sapply(filteringQuantilesIQRs,
                               function(q, IQRs) names(IQRs)[IQRs > q], IQRs)
> design <- cbind(mod, svFIRIVHWMD$sv)
> colnames(design) <- c(colnames(mod), paste0("SV", 1:svFIRIVHWMD$n))
> fit <- lmFit(eset, design)
> fit$genes <- data.frame(ID=featureNames(eset),
                         Symbol=getSYMBOL(featureNames(eset),
                                          "org.Hs.eg.db"), stringsAsFactors=FALSE)
```

Our contrast of interest here is infants affected by all, FIR, IVH and WMD against infants without any of these phenotypes.

```
> cont.matrix <- makeContrasts(FIRIvhWmd=yes.yes.yes-no.no.no,
                              levels=design)
> fit2 <- contrasts.fit(fit, cont.matrix)
> fit2 <- eBayes(fit2, trend=TRUE)
> ttAllIVHWMD <- topTable(fit2, coef="FIRIvhWmd", n=Inf)
> DEgenesByFilterCutoff <- lapply(genesByFilterCutoff,
                                 function(genes, tt, FDRcutoff) {
                                   tt <- tt[tt$ID %in% genes, ]
```

```
                                     tt$adj.P.Val <- p.adjust(tt$P.Value, method="fdr")
                                     tt$ID[tt$adj.P.Val <= FDRcutoff]
                                   }, ttAllIVHWMD, FDRcutoff)
> nDEgenesByFilterCutoff <- sapply(DEgenesByFilterCutoff, length)
> nDEgenesByFilterCutoff
10% 20% 30% 40% 50% 60% 70% 80% 90%
 21  21  27  51  58  75 106 131 183
> whqmaxDE <- length(nDEgenesByFilterCutoff) - which.max(rev(nDEgenesByFilterCutoff)) + 1
> ttAllIQRfiltIVHWMD <- ttAllIVHWMD[ttAllIVHWMD$ID %in% genesByFilterCutoff[[whqmaxDE]], ]
> ttAllIQRfiltIVHWMD$adj.P.Val <- p.adjust(ttAllIQRfiltIVHWMD$P.Value, method="fdr")
> dim(ttAllIQRfiltIVHWMD)
[1] 2016    8
> stopifnot(identical(ttAllIQRfiltIVHWMD$ID[ttAllIQRfiltIVHWMD$adj.P.Val <= FDRcutoff],
                      DEgenesByFilterCutoff[[whqmaxDE]])) ## QC
> ttDEgenesIVHWMD <- ttAllIQRfiltIVHWMD[abs(ttAllIQRfiltIVHWMD$logFC) > log2(fcCutoff) &
                                         ttAllIQRfiltIVHWMD$adj.P.Val <= FDRcutoff, ]
> dim(ttDEgenesIVHWMD)
[1] 151   8
> DEgenesIVHWMD <- ttDEgenesIVHWMD$ID
> length(DEgenesIVHWMD)
[1] 151
> stopifnot(DEgenesIVHWMD == ttDEgenesIVHWMD$ID) ## QC
> save(DEgenesIVHWMD, ttDEgenesIVHWMD, ttAllIVHWMD, ttAllIQRfiltIVHWMD, cleanEsetIVHWMD,
       file=file.path(cacheDir_DE, "DEgenesIVHWMD.RData"))
```

There are 151 DE genes at FDR<0.01 and with a minimum 1.5-fold change, between FIR-IVH-WMD affected and non-affected infants. In Table 10 we have the list of these genes ordered by raw *p*-value.

Table 10. List of 151 differentially expressed (DE) genes called with FDR <1% and minimum 1.5-fold change between FIR/IVH/WMD and nonFIR/nonIVH/nonWMD affected infants. Genes are ordered by the raw p-value of the change in expression level. Columns "LFC" and "FC" contain the fold-changes in $\log_2$ and absolute scales, respectively. Columns "FIW1" and "FIW2" indicate whether the gene is only called DE in this FIR/IVH/WMD comparison but not between FIR and nonFIR infants at FDR <1% with minimum 1.5-fold change ("FIW1") or at FDR <1% only ("FIW2").

	Entrez ID	Symbol	Full name	LFC	FC	P value	FIW1	FIW2
1	6347	CCL2	chemokine (C-C motif) ligand 2	3.06	8.3	9.4459e-08	No	No
2	123207	C15orf40	chromosome 15 open reading frame 40	-0.99	2.0	1.7803e-07	No	No
3	8554	PIAS1	protein inhibitor of activated STAT, 1	-0.63	1.5	6.9078e-07	Yes	No
4	7128	TNFAIP3	tumor necrosis factor, alpha-induced protein 3	2.00	4.0	8.6458e-07	No	No
5	6280	S100A9	S100 calcium binding protein A9	2.80	7.0	1.6612e-06	No	No
6	2919	CXCL1	chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)	3.00	8.0	1.7486e-06	No	No
7	7357	UGCG	UDP-glucose ceramide glucosyltransferase	1.12	2.2	2.0843e-06	No	No
8	477	ATP1A2	ATPase, Na+/K+ transporting, alpha 2 polypeptide	-1.59	3.0	2.9288e-06	No	No
9	79974	CPED1	cadherin-like and PC-esterase domain containing 1	-0.92	1.9	4.1895e-06	No	No
10	2018	EMX2	empty spiracles homeobox 2	-0.84	1.8	4.3255e-06	No	No

Table 10 – Continued on next page

Table 10 – continued from previous page

11	64332	NFKBIZ	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta	1.51	2.9	5.3868e-06	No	No
12	2770	GNAI1	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 1	-0.79	1.7	6.5272e-06	No	No
13	657	BMPR1A	bone morphogenetic protein receptor, type IA	-0.77	1.7	6.8938e-06	No	No
14	3275	PRMT2	protein arginine methyltransferase 2	-0.72	1.6	8.3422e-06	No	No
15	324	APC	adenomatous polyposis coli	-0.87	1.8	8.6254e-06	No	No
16	5906	RAP1A	RAP1A, member of RAS oncogene family	-0.82	1.8	8.9998e-06	No	No
17	9945	GFPT2	glutamine-fructose-6-phosphate transaminase 2	1.33	2.5	9.8683e-06	No	No
18	5784	PTPN14	protein tyrosine phosphatase, non-receptor type 14	-1.61	3.1	1.0855e-05	No	No
19	113612	CYP2U1	cytochrome P450, family 2, subfamily U, polypeptide 1	-0.74	1.7	1.1467e-05	No	No
20	4093	SMAD9	SMAD family member 9	-0.98	2.0	1.5690e-05	No	No
21	54898	ELOVL2	ELOVL fatty acid elongase 2	1.27	2.4	1.5841e-05	No	No
22	1070	CETN3	centrin, EF-hand protein, 3	-0.79	1.7	1.7668e-05	Yes	No
23	23493	HEY2	hes-related family bHLH transcription factor with YRPW motif 2	-0.71	1.6	1.8304e-05	No	No
24	10135	NAMPT	nicotinamide phosphoribosyltransferase	2.16	4.5	1.8734e-05	No	No
25	8404	SPARCL1	SPARC-like 1 (hevin)	-1.09	2.1	1.9002e-05	No	No
26	5727	PTCH1	patched 1	-0.87	1.8	2.1214e-05	No	No
27	23516	SLC39A14	solute carrier family 39 (zinc transporter), member 14	1.02	2.0	2.3224e-05	No	No
28	2258	FGF13	fibroblast growth factor 13	-0.74	1.7	2.5970e-05	No	No
29	7076	TIMP1	TIMP metallopeptidase inhibitor 1	1.04	2.1	2.8585e-05	No	No
30	10217	CTDSPL	CTD (carboxy-terminal domain, RNA polymerase II, polypeptide A) small phosphatase-like	-0.70	1.6	2.9753e-05	No	No
31	3718	JAK3	Janus kinase 3	0.93	1.9	3.0193e-05	No	No
32	1272	CNTN1	contactin 1	-1.48	2.8	3.0222e-05	No	No
33	602	BCL3	B-cell CLL/lymphoma 3	0.88	1.8	3.2521e-05	No	No
34	83468	GLT8D2	glycosyltransferase 8 domain containing 2	-1.07	2.1	3.2659e-05	No	No
35	4594	MUT	methylmalonyl CoA mutase	-0.62	1.5	3.4705e-05	No	No
36	58155	PTBP2	polypyrimidine tract binding protein 2	-0.83	1.8	3.4931e-05	No	No
37	1287	COL4A5	collagen, type IV, alpha 5	-0.71	1.6	3.4988e-05	No	No
38	8801	SUCLG2	succinate-CoA ligase, GDP-forming, beta subunit	-0.89	1.9	3.5748e-05	No	No
39	4651	MYO10	myosin X	0.77	1.7	3.6951e-05	No	No
40	629	CFB	complement factor B	2.11	4.3	3.7616e-05	No	No
41	119504	ANAPC16	anaphase promoting complex subunit 16	-0.76	1.7	3.8764e-05	No	No
42	5806	PTX3	pentraxin 3, long	3.16	8.9	3.8950e-05	No	No
43	133746	JMY	junction mediating and regulatory protein, p53 cofactor	-0.91	1.9	4.0290e-05	No	No

Table 10 – Continued on next page

Table 10 – continued from previous page

44	5156	PDGFRA	platelet-derived growth factor receptor, alpha polypeptide	1.40	2.6	4.5793e-05	No	No
45	6279	S100A8	S100 calcium binding protein A8	3.53	11.6	4.7037e-05	No	No
46	1890	TYMP	thymidine phosphorylase	0.89	1.9	5.3684e-05	No	No
47	22841	RAB11FIP2	RAB11 family interacting protein 2 (class I)	-0.60	1.5	6.0239e-05	No	No
48	8630	HSD17B6	hydroxysteroid (17-beta) dehydrogenase 6	-1.01	2.0	6.0597e-05	No	No
49	6286	S100P	S100 calcium binding protein P	1.63	3.1	6.1304e-05	No	No
50	5352	PLOD2	procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2	1.01	2.0	6.8805e-05	No	No
51	284309	ZNF776	zinc finger protein 776	-0.83	1.8	7.0245e-05	Yes	No
52	6648	SOD2	superoxide dismutase 2, mitochondrial	2.17	4.5	7.2837e-05	No	No
53	196513	DCP1B	decapping mRNA 1B	-0.75	1.7	7.3195e-05	Yes	No
54	7187	TRAF3	TNF receptor-associated factor 3	0.73	1.7	7.3394e-05	No	No
55	151742	PPM1L	protein phosphatase, Mg2+/Mn2+ dependent, 1L	-1.02	2.0	7.6146e-05	No	No
56	5348	FXYD1	FXYD domain containing ion transport regulator 1	-0.73	1.7	8.5423e-05	No	No
57	6283	S100A12	S100 calcium binding protein A12	2.48	5.6	8.7003e-05	No	No
58	360023	ZBTB41	zinc finger and BTB domain containing 41	-0.66	1.6	8.9401e-05	Yes	No
59	57531	HACE1	HECT domain and ankyrin repeat containing E3 ubiquitin protein ligase 1	-0.72	1.6	9.0199e-05	No	No
60	590	BCHE	butyrylcholinesterase	-1.20	2.3	9.2066e-05	No	No
61	79611	ACSS3	acyl-CoA synthetase short-chain family member 3	-0.84	1.8	1.0074e-04	No	No
62	29761	USP25	ubiquitin specific peptidase 25	-0.89	1.8	1.0646e-04	No	No
63	83706	FERMT3	fermitin family member 3	0.83	1.8	1.0946e-04	No	No
64	139341	FUNDC1	FUN14 domain containing 1	-0.66	1.6	1.0975e-04	Yes	No
65	7850	IL1R2	interleukin 1 receptor, type II	1.89	3.7	1.1660e-04	No	No
66	79148	MMP28	matrix metallopeptidase 28	-0.91	1.9	1.2163e-04	No	No
67	9910	RABGAP1L	RAB GTPase activating protein 1-like	-0.67	1.6	1.2445e-04	No	No
68	59353	TMEM35	transmembrane protein 35	-1.38	2.6	1.3871e-04	No	No
69	7378	UPP1	uridine phosphorylase 1	1.02	2.0	1.3942e-04	No	No
70	2549	GAB1	GRB2-associated binding protein 1	-0.88	1.8	1.4246e-04	No	No
71	57631	LRCH2	leucine-rich repeats and calponin homology (CH) domain containing 2	-0.97	2.0	1.4835e-04	No	No
72	112770	C1orf85	chromosome 1 open reading frame 85	-0.94	1.9	1.5003e-04	No	No
73	395	ARHGAP6	Rho GTPase activating protein 6	-0.95	1.9	1.5670e-04	No	No
74	29915	HCFC2	host cell factor C2	-0.61	1.5	1.6281e-04	Yes	No
75	29904	EEF2K	eukaryotic elongation factor-2 kinase	-0.66	1.6	1.6790e-04	No	No
76	8436	SDPR	serum deprivation response	-0.87	1.8	1.7912e-04	No	No
77	8082	SSPN	sarcospan	-1.19	2.3	1.9177e-04	No	No
78	79739	TTLL7	tubulin tyrosine ligase-like family, member 7	-0.87	1.8	1.9830e-04	No	No
79	139065	SLITRK4	SLIT and NTRK-like family, member 4	-1.15	2.2	2.1513e-04	No	No
80	1825	DSC3	desmocollin 3	-0.93	1.9	2.2564e-04	No	No
81	5087	PBX1	pre-B-cell leukemia homeobox 1	-0.77	1.7	2.2653e-04	No	No
82	55784	MCTP2	multiple C2 domains, transmembrane 2	1.02	2.0	2.3089e-04	No	No
83	23514	SPIDR	scaffolding protein involved in DNA repair	0.98	2.0	2.4770e-04	No	No

Table 10 – Continued on next page

Table 10 – continued from previous page

84	56952	PRTFDC1	phosphoribosyl transferase domain containing 1	-0.96	1.9	2.5723e-04	No	No
85	6581	SLC22A3	solute carrier family 22 (organic cation transporter), member 3	0.91	1.9	2.5852e-04	No	No
86	27109	ATP5S	ATP synthase, H+ transporting, mitochondrial Fo complex, subunit s (factor B)	-0.93	1.9	2.6324e-04	No	No
87	130507	UBR3	ubiquitin protein ligase E3 component n-recognin 3 (putative)	-0.62	1.5	2.6489e-04	No	No
88	3033	HADH	hydroxyacyl-CoA dehydrogenase	-0.73	1.7	2.6544e-04	No	No
89	7341	SUMO1	small ubiquitin-like modifier 1	-0.71	1.6	2.7382e-04	No	No
90	54741	LEPROT	leptin receptor overlapping transcript	-0.77	1.7	2.7876e-04	No	No
91	5464	PPA1	pyrophosphatase (inorganic) 1	0.60	1.5	2.9179e-04	No	No
92	7127	TNFAIP2	tumor necrosis factor, alpha-induced protein 2	1.67	3.2	2.9852e-04	No	No
93	4792	NFKBIA	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha	0.86	1.8	3.1247e-04	No	No
94	55857	KIZ	kizuna centrosomal protein	-0.83	1.8	3.3092e-04	No	No
95	10426	TUBGCP3	tubulin, gamma complex associated protein 3	-0.63	1.5	3.3205e-04	No	No
96	10370	CITED2	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 2	-1.05	2.1	3.3484e-04	No	No
97	54749	EPDR1	ependymin related 1	-0.67	1.6	3.3538e-04	No	No
98	7130	TNFAIP6	tumor necrosis factor, alpha-induced protein 6	1.80	3.5	3.3742e-04	No	No
99	5329	PLAUR	plasminogen activator, urokinase receptor	0.86	1.8	3.4793e-04	No	No
100	56999	ADAMTS9	ADAM metallopeptidase with thrombospondin type 1 motif, 9	0.96	1.9	3.5479e-04	No	No
101	2036	EPB41L1	erythrocyte membrane protein band 4.1-like 1	-0.75	1.7	3.7892e-04	No	No
102	93134	ZNF561	zinc finger protein 561	-0.63	1.5	3.8110e-04	No	No
103	53354	PANK1	pantothenate kinase 1	-0.72	1.6	3.9414e-04	No	No
104	4199	ME1	malic enzyme 1, NADP(+)-dependent, cytosolic	-0.87	1.8	3.9600e-04	No	No
105	123283	TARSL2	threonyl-tRNA synthetase-like 2	-0.63	1.5	4.0123e-04	Yes	No
106	3148	HMGB2	high mobility group box 2	-0.59	1.5	4.1811e-04	Yes	No
107	51449	PCYOX1	prenylcysteine oxidase 1	-0.74	1.7	4.2812e-04	No	No
108	5069	PAPPA	pregnancy-associated plasma protein A, pappalysin 1	1.40	2.6	4.2933e-04	No	No
109	3709	ITPR2	inositol 1,4,5-trisphosphate receptor, type 2	0.66	1.6	4.4381e-04	No	No
110	64780	MICAL1	microtubule associated monooxygenase, calponin and LIM domain containing 1	-0.76	1.7	4.5005e-04	No	No
111	91526	ANKRD44	ankyrin repeat domain 44	-0.73	1.7	4.6253e-04	Yes	No
112	10672	GNA13	guanine nucleotide binding protein (G protein), alpha 13	-0.69	1.6	4.6613e-04	Yes	No
113	19	ABCA1	ATP-binding cassette, sub-family A (ABC1), member 1	0.95	1.9	4.7447e-04	No	No

Table 10 – Continued on next page

Table 10 – continued from previous page

114	151525	WDSUB1	WD repeat, sterile alpha motif and U-box domain containing 1	-0.77	1.7	4.8723e-04	No	No
115	8988	HSPB3	heat shock 27kDa protein 3	-1.16	2.2	5.0035e-04	No	No
116	715	C1R	complement component 1, r subcomponent	1.02	2.0	5.0631e-04	No	No
117	6374	CXCL5	chemokine (C-X-C motif) ligand 5	2.49	5.6	5.3036e-04	No	No
118	63932	CXorf56	chromosome X open reading frame 56	-0.86	1.8	5.3192e-04	Yes	No
119	23236	PLCB1	phospholipase C, beta 1 (phosphoinositide-specific)	-0.84	1.8	5.4204e-04	No	No
120	199675	MCEMP1	mast cell-expressed membrane protein 1	1.18	2.3	5.4607e-04	No	No
121	4116	MAGOH	mago-nashi homolog, proliferation-associated (Drosophila)	-0.61	1.5	5.5957e-04	Yes	No
122	51133	KCTD3	potassium channel tetramerization domain containing 3	-0.59	1.5	5.9787e-04	No	No
123	29951	PDZRN4	PDZ domain containing ring finger 4	-1.23	2.3	6.1013e-04	No	No
124	121227	LRIG3	leucine-rich repeats and immunoglobulin-like domains 3	-0.84	1.8	6.1407e-04	No	No
125	1027	CDKN1B	cyclin-dependent kinase inhibitor 1B (p27, Kip1)	-0.61	1.5	6.3076e-04	Yes	No
126	84253	GARNL3	GTPase activating Rap/RanGAP domain-like 3	-0.72	1.6	6.4385e-04	No	No
127	284	ANGPT1	angiopoietin 1	-0.90	1.9	6.5218e-04	No	No
128	56890	MDM1	Mdm1 nuclear protein homolog (mouse)	-0.62	1.5	6.6445e-04	No	No
129	57649	PHF12	PHD finger protein 12	-0.66	1.6	6.7611e-04	Yes	No
130	79772	MCTP1	multiple C2 domains, transmembrane 1	1.20	2.3	6.9151e-04	No	No
131	6419	SETMAR	SET domain and mariner transposase fusion gene	-0.61	1.5	7.2037e-04	No	No
132	26273	FBXO3	F-box protein 3	-0.59	1.5	7.2561e-04	No	No
133	4790	NFKB1	nuclear factor of kappa light polypeptide gene enhancer in B-cells 1	0.74	1.7	7.2581e-04	No	No
134	5337	PLD1	phospholipase D1, phosphatidylcholine-specific	1.16	2.2	7.7581e-04	No	No
135	57456	KIAA1143	KIAA1143	-0.59	1.5	7.7668e-04	No	No
136	4209	MEF2D	myocyte enhancer factor 2D	-0.60	1.5	7.8511e-04	Yes	No
137	4794	NFKBIE	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, epsilon	0.61	1.5	7.8863e-04	No	No
138	23171	GPD1L	glycerol-3-phosphate dehydrogenase 1-like	-0.98	2.0	8.0037e-04	No	No
139	80312	TET1	tet methylcytosine dioxygenase 1	-0.76	1.7	8.0665e-04	No	No
140	4675	NAP1L3	nucleosome assembly protein 1-like 3	-0.85	1.8	8.0772e-04	No	No
141	4067	LYN	LYN proto-oncogene, Src family tyrosine kinase	1.32	2.5	8.1580e-04	No	No
142	5195	PEX14	peroxisomal biogenesis factor 14	-0.88	1.8	8.1687e-04	No	No
143	240	ALOX5	arachidonate 5-lipoxygenase	1.49	2.8	8.2014e-04	No	No
144	55332	DRAM1	DNA-damage regulated autophagy modulator 1	0.92	1.9	8.3205e-04	No	No
145	55607	PPP1R9A	protein phosphatase 1, regulatory subunit 9A	-0.93	1.9	8.3909e-04	No	No
146	8175	SF3A2	splicing factor 3a, subunit 2, 66kDa	-0.59	1.5	8.4441e-04	No	No

Table 10 – Continued on next page

IVH	WMD	FIR	no	yes
no	no		10	1
	yes		0	0
yes	no		0	2
	yes		1	3

Table 11. Infants who were not administered with ACS cross-classified by FIR, IVH and WMD status.

Table 10 – continued from previous page

147	84553	FAXC	failed axon connections homolog (Drosophila)	-0.76	1.7	8.5156e-04	No	No
148	59269	HIVEP3	human immunodeficiency virus type I enhancer binding protein 3	1.39	2.6	8.5261e-04	No	No
149	26037	SIPA1L1	signal-induced proliferation-associated 1 like 1	-0.65	1.6	8.5268e-04	No	No
150	148203	ZNF738	zinc finger protein 738	-0.60	1.5	8.6038e-04	Yes	No
151	23336	SYNM	synemin, intermediate filament protein	-0.69	1.6	9.0153e-04	No	No

Since the administration of antenatal glucocorticoids (ACS) has been described to improve the outcome of IVH (7), we are going to repeat the analysis only on those infants who have been not administered with ACS, and described in Table 11.

```
> eset <- frmaEsetFilt
> eset$FIR <- factor(eset$FIR)
> eset$Ivh <- factor(eset$Ivh)
> eset$Wmd <- factor(eset$Wmd)
> eset <- eset[, eset$Acs == "no"]
> df <- data.frame(FIR=eset$FIR, IVH=eset$Ivh, WMD=eset$Wmd)
> xtabFIRIVHWMDnoACS <- ftable(FIR ~ IVH + WMD, data=df)
> xtabFIRIVHWMDnoACS
        FIR no yes
IVH WMD
no  no      10   1
    yes      0   0
yes no       0   2
    yes      1   3
```

In this case, the differential exprsesion analysis involves only the $n = 17$ infants who were not administered with ACS and, more concretely, the comparison between $n = 3$ infants affected by FIR, IVH and WMD against $n = 10$ infants not affected by any of these phenotypes.

```
> FIRIvhWmd <- factor(paste(eset$FIR, eset$Ivh, eset$Wmd, sep="."))
> mod <- model.matrix(~ 0 + FIRIvhWmd, eset)
> colnames(mod) <- levels(FIRIvhWmd)
> mod0 <- model.matrix(~ 1, eset)
> IQRs <- esApply(eset, 1, IQR)
> cache(svFIRIVHWMDnoACS <- sva(dat=exprs(eset[IQRs > quantile(IQRs, prob=0.90), ]), mod=mod, mo
        dir=cacheDir_DE, prefix=cachePrefix_DE)
> cleanExps <- removeBatchEffect(exprs(eset), covariates=svFIRIVHWMDnoACS$sv, design=mod)
> cleanEsetIVHWMDnoACS <- eset
> exprs(cleanEsetIVHWMDnoACS) <- cleanExps
```

```
> IQRs <- esApply(cleanEsetIVHWMDnoACS, 1, IQR)
> filteringQuantiles <- seq(0.1, 0.9, by=0.1)
> filteringQuantilesIQRs <- quantile(IQRs, probs=filteringQuantiles)
> genesByFilterCutoff <- sapply(filteringQuantilesIQRs,
                                function(q, IQRs) names(IQRs)[IQRs > q], IQRs)
> design <- cbind(mod, svFIRIVHWMDnoACS$sv)
> colnames(design) <- c(colnames(mod), paste0("SV", 1:svFIRIVHWMDnoACS$n))
> fit <- lmFit(eset, design)
> fit$genes <- data.frame(ID=featureNames(eset),
                          Symbol=getSYMBOL(featureNames(eset),
                                           "org.Hs.eg.db"), stringsAsFactors=FALSE)
```

Our contrast of interest here is again infants affected by all, FIR, IVH and WMD against infants without any of these phenotypes.

```
> cont.matrix <- makeContrasts(FIRIvhWmd=yes.yes.yes-no.no.no,
                               levels=design)
> fit2 <- contrasts.fit(fit, cont.matrix)
> fit2 <- eBayes(fit2, trend=TRUE)
> ttAllIVHWMDnoACS <- topTable(fit2, coef="FIRIvhWmd", n=Inf)
> DEgenesByFilterCutoff <- lapply(genesByFilterCutoff,
                                  function(genes, tt, FDRcutoff) {
                                    tt <- tt[tt$ID %in% genes, ]
                                    tt$adj.P.Val <- p.adjust(tt$P.Value, method="fdr")
                                    tt$ID[tt$adj.P.Val <= FDRcutoff]
                                  }, ttAllIVHWMDnoACS, FDRcutoff)
> nDEgenesByFilterCutoff <- sapply(DEgenesByFilterCutoff, length)
> nDEgenesByFilterCutoff
10% 20% 30% 40% 50% 60% 70% 80% 90%
 23  36  54  59  91 152 286 498 570
> whqmaxDE <- length(nDEgenesByFilterCutoff) - which.max(rev(nDEgenesByFilterCutoff)) + 1
> ttAllIQRfiltIVHWMDnoACS <- ttAllIVHWMDnoACS[ttAllIVHWMDnoACS$ID %in% genesByFilterCutoff[[whqma
> ttAllIQRfiltIVHWMDnoACS$adj.P.Val <- p.adjust(ttAllIQRfiltIVHWMDnoACS$P.Value, method="fdr")
> dim(ttAllIQRfiltIVHWMDnoACS)
[1] 2016    8
> stopifnot(identical(ttAllIQRfiltIVHWMDnoACS$ID[ttAllIQRfiltIVHWMDnoACS$adj.P.Val <= FDRcutoff],
                      DEgenesByFilterCutoff[[whqmaxDE]])) ## QC
> ttDEgenesIVHWMDnoACS <- ttAllIQRfiltIVHWMDnoACS[abs(ttAllIQRfiltIVHWMDnoACS$logFC) > log2(fcCut
                                        ttAllIQRfiltIVHWMDnoACS$adj.P.Val <= FDRcutoff, ]
> dim(ttDEgenesIVHWMDnoACS)
[1] 566   8
> DEgenesIVHWMDnoACS <- ttDEgenesIVHWMDnoACS$ID
> length(DEgenesIVHWMDnoACS)
[1] 566
> stopifnot(DEgenesIVHWMDnoACS == ttDEgenesIVHWMDnoACS$ID) ## QC
> save(DEgenesIVHWMDnoACS, ttDEgenesIVHWMDnoACS, ttAllIVHWMDnoACS,
       ttAllIQRfiltIVHWMDnoACS, cleanEsetIVHWMDnoACS,
       file=file.path(cacheDir_DE, "DEgenesIVHWMDnoACS.RData"))
```

There are 566 DE genes at FDR<0.01 and with a minimum 1.5-fold change, between FIR-IVH-WMD affected and non-affected infants, who were not administered with ACS. In Table 12 we have the list of these genes ordered by raw *p*-value. There are 256 genes in this list that are also DE between FIR and nonFIR infants, and therefore, 310 that are called DE between FIR/IVH/WMD and nonFIR/nonIVH/nonWMD infants but not between FIR and nonFIR infants. Genes in this latter subset are indicated in columns “FIW1” and “FIW2” from Table 12.

Table 12. List of 566 differentially expressed (DE) genes called with FDR <1% and minimum 1.5-fold change between FIR/IVH/WMD and nonFIR/nonIVH/nonWMD affected infants, who were not administered with antenatal glucocorticoids. Genes are ordered by the raw p-value of the change in expression level. Columns "LFC" and "FC" contain the fold-changes in $\log_2$ and absolute scales, respectively. Columns "FIW1" and "FIW2" indicate whether the gene is only called DE in this FIR/IVH/WMD comparison but not between FIR and nonFIR infants at FDR <1% with minimum 1.5-fold change ("FIW1") or at FDR <1% only ("FIW2").

	Entrez ID	Symbol	Full name	LFC	FC	P value	FIW1	FI
1	1272	CNTN1	contactin 1	-2.65	6.3	3.1703e-09	No	N
2	7128	TNFAIP3	tumor necrosis factor, alpha-induced protein 3	2.28	4.9	5.6835e-08	No	N
3	6283	S100A12	S100 calcium binding protein A12	4.31	19.8	7.6142e-08	No	N
4	6279	S100A8	S100 calcium binding protein A8	6.27	77.1	2.9034e-07	No	N
5	6280	S100A9	S100 calcium binding protein A9	3.95	15.5	3.2978e-07	No	N
6	139886	SPIN4	spindlin family, member 4	-1.34	2.5	6.4689e-07	No	N
7	83706	FERMT3	fermitin family member 3	1.34	2.5	1.5583e-06	No	N
8	2919	CXCL1	chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)	3.11	8.6	1.5822e-06	No	N
9	9737	GPRASP1	G protein-coupled receptor associated sorting protein 1	-1.40	2.6	2.6548e-06	No	N
10	5906	RAP1A	RAP1A, member of RAS oncogene family	-1.49	2.8	2.8351e-06	No	N
11	477	ATP1A2	ATPase, Na+/K+ transporting, alpha 2 polypeptide	-2.05	4.1	3.0335e-06	No	N
12	272	AMPD3	adenosine monophosphate deaminase 3	1.11	2.2	3.3391e-06	No	N
13	8436	SDPR	serum deprivation response	-1.18	2.3	3.8306e-06	No	N
14	25953	PNKD	paroxysmal nonkinesigenic dyskinesia	1.88	3.7	5.1628e-06	No	N
15	2920	CXCL2	chemokine (C-X-C motif) ligand 2	2.71	6.6	5.1842e-06	No	N
16	150000	ABCC13	ATP-binding cassette, sub-family C (CFTR/MRP), member 13, pseudogene	2.54	5.8	6.0462e-06	Yes	Ye
17	5727	PTCH1	patched 1	-1.09	2.1	6.2643e-06	No	N
18	8787	RGS9	regulator of G-protein signaling 9	-1.22	2.3	7.1268e-06	No	N
19	139065	SLITRK4	SLIT and NTRK-like family, member 4	-1.91	3.8	9.2066e-06	No	N
20	4689	NCF4	neutrophil cytosolic factor 4, 40kDa	1.03	2.0	1.0001e-05	No	N
21	84186	ZCCHC7	zinc finger, CCHC domain containing 7	-0.93	1.9	1.1973e-05	No	N
22	257218	SHPRH	SNF2 histone linker PHD RING helicase, E3 ubiquitin protein ligase	-1.01	2.0	1.2032e-05	Yes	N
23	6648	SOD2	superoxide dismutase 2, mitochondrial	3.07	8.4	1.3951e-05	No	N
24	79739	TTLL7	tubulin tyrosine ligase-like family, member 7	-1.47	2.8	1.6602e-05	No	N
25	7076	TIMP1	TIMP metallopeptidase inhibitor 1	1.32	2.5	1.6886e-05	No	N
26	2549	GAB1	GRB2-associated binding protein 1	-1.24	2.4	1.7897e-05	No	N
27	3257	HPS1	Hermansky-Pudlak syndrome 1	1.05	2.1	1.8493e-05	Yes	Y
28	2239	GPC4	glypican 4	-0.93	1.9	1.8834e-05	No	N
29	83468	GLT8D2	glycosyltransferase 8 domain containing 2	-1.68	3.2	1.9156e-05	No	N
30	7850	IL1R2	interleukin 1 receptor, type II	2.34	5.0	2.0284e-05	No	N
31	4139	MARK1	MAP/microtubule affinity-regulating kinase 1	-0.96	1.9	2.0820e-05	No	N
32	2180	ACSL1	acyl-CoA synthetase long-chain family member 1	1.63	3.1	2.0858e-05	No	N

Table 12 – Continued on next page

Table 12 – continued from previous page

33	4069	LYZ	lysozyme	3.90	14.9	2.2290e-05	No	N
34	7357	UGCG	UDP-glucose ceramide glucosyltransferase	1.35	2.5	2.4131e-05	No	N
35	1287	COL4A5	collagen, type IV, alpha 5	-1.04	2.1	2.6186e-05	No	N
36	3718	JAK3	Janus kinase 3	1.10	2.1	2.6820e-05	No	N
37	4318	MMP9	matrix metallopeptidase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)	1.58	3.0	2.8587e-05	No	N
38	151742	PPM1L	protein phosphatase, Mg2+/Mn2+ dependent, 1L	-1.72	3.3	2.8727e-05	No	N
39	2729	GCLC	glutamate-cysteine ligase, catalytic subunit	2.06	4.2	2.9372e-05	Yes	Y
40	51306	FAM13B	family with sequence similarity 13, member B	-1.13	2.2	2.9782e-05	No	N
41	5196	PF4	platelet factor 4	1.90	3.7	3.0030e-05	No	N
42	55082	ARGLU1	arginine and glutamate rich 1	-1.41	2.7	3.0045e-05	No	N
43	6347	CCL2	chemokine (C-C motif) ligand 2	3.31	9.9	3.1674e-05	No	N
44	56829	ZC3HAV1	zinc finger CCCH-type, antiviral 1	1.59	3.0	3.1761e-05	Yes	Y
45	3671	ISLR	immunoglobulin superfamily containing leucine-rich repeat	-1.51	2.9	3.2474e-05	No	N
46	130271	PLEKHH2	pleckstrin homology domain containing, family H (with MyTH4 domain) member 2	-1.56	3.0	3.5350e-05	No	N
47	5473	PPBP	pro-platelet basic protein (chemokine (C-X-C motif) ligand 7)	4.84	28.7	3.5480e-05	No	N
48	2258	FGF13	fibroblast growth factor 13	-1.09	2.1	3.6003e-05	No	N
49	10395	DLC1	DLC1 Rho GTPase activating protein	-1.27	2.4	3.7680e-05	No	N
50	150381	PRR34-AS1	PRR34 antisense RNA 1	-1.12	2.2	3.9134e-05	Yes	N
51	80896	NPL	N-acetylneuraminate pyruvate lyase (dihydrodipicolinate synthase)	1.09	2.1	4.0651e-05	Yes	N
52	2857	GPR34	G protein-coupled receptor 34	-1.21	2.3	4.2178e-05	No	N
53	57126	CD177	CD177 molecule	1.05	2.1	4.4156e-05	No	N
54	5926	ARID4A	AT rich interactive domain 4A (RBP1-like)	-1.30	2.5	4.4708e-05	Yes	N
55	199675	MCEMP1	mast cell-expressed membrane protein 1	1.91	3.7	4.8061e-05	No	N
56	118924	FRA10AC1	fragile site, folic acid type, rare, fra(10)(q23.3) or fra(10)(q24.2) candidate 1	-1.02	2.0	4.8782e-05	Yes	N
57	54665	RSBN1	round spermatid basic protein 1	-1.48	2.8	5.0038e-05	Yes	N
58	55250	ELP2	elongator acetyltransferase complex subunit 2	-1.02	2.0	5.3422e-05	Yes	N
59	6583	SLC22A4	solute carrier family 22 (organic cation/zwitterion transporter), member 4	2.01	4.0	5.6387e-05	Yes	N
60	3046	HBE1	hemoglobin, epsilon 1	4.36	20.6	5.7674e-05	Yes	Y
61	5156	PDGFRA	platelet-derived growth factor receptor, alpha polypeptide	1.63	3.1	5.8574e-05	No	N
62	598	BCL2L1	BCL2-like 1	2.28	4.9	5.9007e-05	Yes	N
63	84952	CGNL1	cingulin-like 1	-1.07	2.1	5.9071e-05	No	N
64	154043	CNKSR3	CNKSR family member 3	-1.40	2.6	5.9344e-05	No	N
65	79415	C17orf62	chromosome 17 open reading frame 62	0.93	1.9	6.2682e-05	No	N

Table 12 – Continued on next page

Table 12 – continued from previous page

66	151987	PPP4R2	protein phosphatase 4, regulatory subunit 2	1.03	2.0	6.3269e-05	No	N
67	1825	DSC3	desmocollin 3	-1.36	2.6	6.4843e-05	No	N
68	9547	CXCL14	chemokine (C-X-C motif) ligand 14	-2.05	4.1	6.9885e-05	No	N
69	2235	FECH	ferrochelatase	3.22	9.3	7.2018e-05	Yes	Y
70	119504	ANAPC16	anaphase promoting complex subunit 16	-0.94	1.9	7.2154e-05	No	N
71	80333	KCNIP4	Kv channel interacting protein 4	-0.93	1.9	7.2333e-05	No	N
72	602	BCL3	B-cell CLL/lymphoma 3	1.09	2.1	7.3660e-05	No	N
73	2876	GPX1	glutathione peroxidase 1	1.22	2.3	7.3767e-05	Yes	N
74	6374	CXCL5	chemokine (C-X-C motif) ligand 5	2.93	7.6	7.4658e-05	No	N
75	121536	AEBP2	AE binding protein 2	-1.00	2.0	7.6443e-05	No	N
76	56952	PRTFDC1	phosphoribosyl transferase domain containing 1	-1.74	3.3	7.7466e-05	No	N
77	100128252	ZNF667-AS1	ZNF667 antisense RNA 1 (head to head)	-1.10	2.1	7.7973e-05	Yes	N
78	56947	MFF	mitochondrial fission factor	-1.02	2.0	7.8197e-05	Yes	N
79	84281	C2orf88	chromosome 2 open reading frame 88	1.50	2.8	7.8640e-05	Yes	Y
80	5348	FXYD1	FXYD domain containing ion transport regulator 1	-1.13	2.2	8.4555e-05	No	N
81	12	SERPINA3	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3	1.19	2.3	8.5122e-05	No	N
82	5118	PCOLCE	procollagen C-endopeptidase enhancer	-1.24	2.4	9.2349e-05	Yes	Y
83	51386	EIF3L	eukaryotic translation initiation factor 3, subunit L	-0.86	1.8	9.4464e-05	Yes	N
84	4815	NINJ2	ninjurin 2	1.87	3.6	9.7903e-05	Yes	N
85	54477	PLEKHA5	pleckstrin homology domain containing, family A member 5	-1.09	2.1	9.8958e-05	No	N
86	64431	ACTR6	ARP6 actin-related protein 6 homolog (yeast)	-0.96	1.9	1.0233e-04	Yes	N
87	55783	CMTR2	cap methyltransferase 2	-0.77	1.7	1.0259e-04	Yes	N
88	115330	GPR146	G protein-coupled receptor 146	1.45	2.7	1.0303e-04	Yes	Y
89	9180	OSMR	oncostatin M receptor	1.09	2.1	1.0338e-04	No	N
90	4594	MUT	methylmalonyl CoA mutase	-0.83	1.8	1.0531e-04	No	N
91	11099	PTPN21	protein tyrosine phosphatase, non-receptor type 21	-0.85	1.8	1.0946e-04	No	N
92	7750	ZMYM2	zinc finger, MYM-type 2	-1.16	2.2	1.1603e-04	Yes	Y
93	23379	ICE1	interactor of little elongator complex ELL subunit 1	-0.81	1.8	1.1612e-04	Yes	N
94	2332	FMR1	fragile X mental retardation 1	-1.00	2.0	1.1909e-04	No	N
95	100288911	LOC100288911	uncharacterized LOC100288911	-0.97	2.0	1.1935e-04	No	N
96	57608	KIAA1462	KIAA1462	-1.13	2.2	1.2029e-04	Yes	N
97	195828	ZNF367	zinc finger protein 367	-1.35	2.6	1.2396e-04	No	N
98	148203	ZNF738	zinc finger protein 738	-1.01	2.0	1.2452e-04	Yes	N
99	7001	PRDX2	peroxiredoxin 2	2.13	4.4	1.2503e-04	Yes	Y
100	143684	FAM76B	family with sequence similarity 76, member B	-0.95	1.9	1.3008e-04	Yes	N
101	51759	C9orf78	chromosome 9 open reading frame 78	1.62	3.1	1.3208e-04	Yes	Y
102	81578	COL21A1	collagen, type XXI, alpha 1	-1.36	2.6	1.3290e-04	No	N
103	23376	UFL1	UFM1-specific ligase 1	-0.88	1.8	1.3333e-04	Yes	N
104	3480	IGF1R	insulin-like growth factor 1 receptor	-1.14	2.2	1.3629e-04	Yes	Y

Table 12 – Continued on next page

Table 12 – continued from previous page

105	27102	EIF2AK1	eukaryotic translation initiation factor 2-alpha kinase 1	2.06	4.2	1.3903e-04	Yes	Ye
106	80232	WDR26	WD repeat domain 26	2.47	5.5	1.4085e-04	Yes	Ye
107	27252	KLHL20	kelch-like family member 20	-0.82	1.8	1.4096e-04	Yes	N
108	4239	MFAP4	microfibrillar-associated protein 4	-1.41	2.7	1.4175e-04	No	N
109	79884	MAP9	microtubule-associated protein 9	-0.94	1.9	1.5144e-04	No	N
110	200014	CC2D1B	coiled-coil and C2 domain containing 1B	0.94	1.9	1.5193e-04	Yes	Ye
111	26127	FGFR1OP2	FGFR1 oncogene partner 2	2.59	6.0	1.5546e-04	Yes	N
112	4603	MYBL1	v-myb avian myeloblastosis viral oncogene homolog-like 1	-1.01	2.0	1.5672e-04	Yes	N
113	7247	TSN	translin	-0.98	2.0	1.5895e-04	Yes	N
114	5265	SERPINA1	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1	1.88	3.7	1.6094e-04	No	N
115	9474	ATG5	autophagy related 5	-0.78	1.7	1.6205e-04	No	N
116	80817	CEP44	centrosomal protein 44kDa	-1.09	2.1	1.6331e-04	No	N
117	79834	PEAK1	pseudopodium-enriched atypical kinase 1	-0.85	1.8	1.6477e-04	Yes	N
118	3741	KCNA5	potassium voltage-gated channel, shaker-related subfamily, member 5	-1.52	2.9	1.6654e-04	No	N
119	4542	MYO1F	myosin IF	1.15	2.2	1.6738e-04	No	N
120	3298	HSF2	heat shock transcription factor 2	-0.86	1.8	1.6975e-04	Yes	N
121	51274	KLF3	Kruppel-like factor 3 (basic)	1.73	3.3	1.7069e-04	Yes	Ye
122	6478	SIAH2	siah E3 ubiquitin protein ligase 2	1.63	3.1	1.7290e-04	Yes	N
123	669	BPGM	2,3-bisphosphoglycerate mutase	4.77	27.3	1.7331e-04	Yes	Ye
124	6146	RPL22	ribosomal protein L22	-0.81	1.8	1.7416e-04	Yes	N
125	57631	LRCH2	leucine-rich repeats and calponin homology (CH) domain containing 2	-1.51	2.8	1.7442e-04	No	N
126	84901	NFATC2IP	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2 interacting protein	-0.63	1.5	1.7493e-04	Yes	N
127	338707	B4GALNT4	beta-1,4-N-acetyl-galactosaminyl transferase 4	0.69	1.6	1.7558e-04	Yes	Ye
128	10135	NAMPT	nicotinamide phosphoribosyltransferase	2.37	5.2	1.7851e-04	No	N
129	9235	IL32	interleukin 32	1.38	2.6	1.7890e-04	No	N
130	100507303	SNHG19	small nucleolar RNA host gene 19 (non-protein coding)	-1.16	2.2	1.8654e-04	No	N
131	10661	KLF1	Kruppel-like factor 1 (erythroid)	1.86	3.6	1.9437e-04	Yes	Ye
132	81624	DIAPH3	diaphanous-related formin 3	-0.95	1.9	1.9518e-04	No	N
133	26033	ATRNL1	attractin-like 1	-1.51	2.8	1.9794e-04	No	N
134	1657	DMXL1	Dmx-like 1	-0.75	1.7	1.9878e-04	Yes	N
135	1122	CHML	choroideremia-like (Rab escort protein 2)	-1.31	2.5	1.9999e-04	No	N
136	5747	PTK2	protein tyrosine kinase 2	-0.88	1.8	2.0444e-04	No	N
137	7130	TNFAIP6	tumor necrosis factor, alpha-induced protein 6	1.97	3.9	2.0702e-04	No	N
138	5806	PTX3	pentraxin 3, long	4.08	16.9	2.1541e-04	No	N
139	629	CFB	complement factor B	1.97	3.9	2.1541e-04	No	N
140	983	CDK1	cyclin-dependent kinase 1	-0.70	1.6	2.1546e-04	Yes	Ye
141	51327	AHSP	alpha hemoglobin stabilizing protein	4.25	19.1	2.1783e-04	Yes	Ye

Table 12 – Continued on next page

Table 12 – continued from previous page

142	7504	XK	X-linked Kx blood group	4.09	17.0	2.1953e-04	Yes	Y
143	433	ASGR2	asialoglycoprotein receptor 2	0.84	1.8	2.2346e-04	No	N
144	3290	HSD11B1	hydroxysteroid (11-beta) dehydrogenase 1	1.24	2.4	2.3104e-04	No	N
145	85313	PPIL4	peptidylprolyl isomerase (cyclophilin)-like 4	-0.59	1.5	2.3419e-04	Yes	Y
146	79074	C2orf49	chromosome 2 open reading frame 49	-0.86	1.8	2.3473e-04	Yes	N
147	657	BMPR1A	bone morphogenetic protein receptor, type IA	-1.05	2.1	2.3648e-04	No	N
148	11027	LILRA2	leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 2	0.93	1.9	2.3653e-04	No	N
149	2995	GYPC	glycophorin C (Gerbich blood group)	1.87	3.7	2.4025e-04	Yes	N
150	7133	TNFRSF1B	tumor necrosis factor receptor superfamily, member 1B	1.10	2.1	2.4053e-04	No	N
151	1719	DHFR	dihydrofolate reductase	-0.71	1.6	2.4152e-04	Yes	N
152	23306	TMEM194A	transmembrane protein 194A	-0.98	2.0	2.4232e-04	Yes	N
153	196047	EMX2OS	EMX2 opposite strand/antisense RNA	-0.98	2.0	2.4258e-04	No	N
154	81027	TUBB1	tubulin, beta 1 class VI	2.11	4.3	2.4714e-04	Yes	N
155	9055	PRC1	protein regulator of cytokinesis 1	-1.22	2.3	2.4883e-04	No	N
156	2018	EMX2	empty spiracles homeobox 2	-1.18	2.3	2.5046e-04	No	N
157	5789	PTPRD	protein tyrosine phosphatase, receptor type, D	-1.35	2.5	2.5515e-04	No	N
158	25853	DCAF12	DDB1 and CUL4 associated factor 12	3.50	11.3	2.5591e-04	Yes	Y
159	10318	TNIP1	TNFAIP3 interacting protein 1	1.10	2.1	2.5768e-04	No	N
160	54463	FAM134B	family with sequence similarity 134, member B	-0.85	1.8	2.6001e-04	Yes	Y
161	1052	CEBPD	CCAAT/enhancer binding protein (C/EBP), delta	1.34	2.5	2.6038e-04	No	N
162	9706	ULK2	unc-51 like autophagy activating kinase 2	-0.69	1.6	2.6213e-04	No	N
163	3590	IL11RA	interleukin 11 receptor, alpha	-1.02	2.0	2.6556e-04	No	N
164	6210	RPS15A	ribosomal protein S15a	-0.88	1.8	2.6677e-04	No	N
165	1305	COL13A1	collagen, type XIII, alpha 1	-1.32	2.5	2.7259e-04	No	N
166	25878	MXRA5	matrix-remodelling associated 5	-2.35	5.1	2.7442e-04	Yes	Y
167	11163	NUDT4	nudix (nucleoside diphosphate linked moiety X)-type motif 4	2.62	6.2	2.7635e-04	Yes	Y
168	23514	SPIDR	scaffolding protein involved in DNA repair	1.47	2.8	2.8072e-04	No	N
169	27109	ATP5S	ATP synthase, H+ transporting, mitochondrial Fo complex, subunit s (factor B)	-1.00	2.0	2.8107e-04	No	N
170	100507311	LOC100507311	uncharacterized LOC100507311	-1.28	2.4	2.8366e-04	No	N
171	84932	RAB2B	RAB2B, member RAS oncogene family	1.53	2.9	2.9007e-04	Yes	Y
172	84188	FAR1	fatty acyl CoA reductase 1	-0.91	1.9	2.9142e-04	Yes	Y
173	2268	FGR	FGR proto-oncogene, Src family tyrosine kinase	1.28	2.4	2.9645e-04	No	N
174	80315	CPEB4	cytoplasmic polyadenylation element binding protein 4	2.02	4.1	3.0077e-04	No	N
175	6286	S100P	S100 calcium binding protein P	1.35	2.5	3.0332e-04	No	N
176	10486	CAP2	CAP, adenylate cyclase-associated protein, 2 (yeast)	-0.88	1.8	3.0665e-04	No	N

Table 12 – Continued on next page

Table 12 – continued from previous page

177	11098	PRSS23	protease, serine, 23	-0.93	1.9	3.0830e-04	Yes	N
178	347902	AMIGO2	adhesion molecule with Ig-like domain 2	-1.70	3.3	3.1764e-04	No	N
179	5305	PIP4K2A	phosphatidylinositol-5-phosphate 4-kinase, type II, alpha	2.02	4.1	3.1989e-04	Yes	N
180	57538	ALPK3	alpha-kinase 3	-1.01	2.0	3.2138e-04	Yes	Y
181	23213	SULF1	sulfatase 1	1.26	2.4	3.2524e-04	No	N
182	79974	CPED1	cadherin-like and PC-esterase domain containing 1	-1.03	2.0	3.2798e-04	No	N
183	9849	ZNF518A	zinc finger protein 518A	-0.90	1.9	3.2985e-04	Yes	N
184	140733	MACROD2	MACRO domain containing 2	-0.86	1.8	3.3136e-04	No	N
185	286205	SCAI	suppressor of cancer cell invasion	-1.36	2.6	3.3250e-04	No	N
186	5743	PTGS2	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	1.73	3.3	3.3267e-04	No	N
187	716	C1S	complement component 1, s subcomponent	1.01	2.0	3.3622e-04	No	N
188	5927	KDM5A	lysine (K)-specific demethylase 5A	-0.90	1.9	3.3667e-04	Yes	Y
189	9021	SOCS3	suppressor of cytokine signaling 3	1.67	3.2	3.3740e-04	No	N
190	55120	FANCL	Fanconi anemia, complementation group L	-1.25	2.4	3.4315e-04	No	N
191	286827	TRIM59	tripartite motif containing 59	-0.82	1.8	3.4318e-04	Yes	N
192	81620	CDT1	chromatin licensing and DNA replication factor 1	0.88	1.8	3.4456e-04	Yes	Y
193	284	ANGPT1	angiopoietin 1	-1.04	2.1	3.4845e-04	No	N
194	64795	RMND5A	required for meiotic nuclear division 5 homolog A (S. cerevisiae)	2.45	5.5	3.5520e-04	Yes	Y
195	781	CACNA2D1	calcium channel, voltage-dependent, alpha 2/delta subunit 1	-0.95	1.9	3.5544e-04	No	N
196	9306	SOCS6	suppressor of cytokine signaling 6	-0.88	1.8	3.6194e-04	Yes	N
197	1021	CDK6	cyclin-dependent kinase 6	-0.90	1.9	3.6296e-04	No	N
198	6815	STYX	serine/threonine/tyrosine interacting protein	-0.85	1.8	3.6379e-04	No	N
199	9852	EPM2AIP1	EPM2A (laforin) interacting protein 1	-1.10	2.1	3.6421e-04	Yes	N
200	9945	GFPT2	glutamine-fructose-6-phosphate transaminase 2	0.73	1.7	3.6746e-04	No	N
201	51575	ESF1	ESF1, nucleolar pre-rRNA processing protein, homolog (S. cerevisiae)	-0.95	1.9	3.6779e-04	Yes	Y
202	360023	ZBTB41	zinc finger and BTB domain containing 41	-0.97	2.0	3.6920e-04	Yes	N
203	387647	PTCHD3P1	patched domain containing 3 pseudogene 1	-0.73	1.7	3.7156e-04	Yes	N
204	84164	ASCC2	activating signal cointegrator 1 complex subunit 2	1.40	2.6	3.7229e-04	Yes	Y
205	324	APC	adenomatous polyposis coli	-0.93	1.9	3.7702e-04	No	N
206	55363	HEMGN	hemogen	4.16	17.9	3.7829e-04	Yes	Y
207	22929	SEPHS1	selenophosphate synthetase 1	-0.79	1.7	3.7841e-04	No	N
208	8622	PDE8B	phosphodiesterase 8B	-0.99	2.0	3.8307e-04	No	N
209	9535	GMFG	glia maturation factor, gamma	1.29	2.4	3.8716e-04	No	N
210	5494	PPM1A	protein phosphatase, Mg2+/Mn2+ dependent, 1A	1.34	2.5	3.8891e-04	Yes	Y

Table 12 – Continued on next page

Table 12 – continued from previous page

211	645	BLVRB	biliverdin reductase B (flavin reductase (NADPH))	2.91	7.5	3.9199e-04	Yes	Y
212	49854	ZBTB21	zinc finger and BTB domain containing 21	-0.96	1.9	3.9320e-04	Yes	Y
213	57570	TRMT5	tRNA methyltransferase 5	-0.66	1.6	3.9961e-04	Yes	N
214	397	ARHGDIB	Rho GDP dissociation inhibitor (GDI) beta	1.15	2.2	4.0923e-04	No	N
215	374354	NHLRC2	NHL repeat containing 2	-0.63	1.5	4.1586e-04	No	N
216	8175	SF3A2	splicing factor 3a, subunit 2, 66kDa	-0.77	1.7	4.1649e-04	No	N
217	166815	TIGD2	tigger transposable element derived 2	-0.63	1.6	4.2106e-04	Yes	N
218	57037	ANKMY2	ankyrin repeat and MYND domain containing 2	-0.86	1.8	4.2411e-04	No	N
219	85461	TANC1	tetratricopeptide repeat, ankyrin repeat and coiled-coil containing 1	-0.85	1.8	4.2840e-04	Yes	Y
220	4354	MPP1	membrane protein, palmitoylated 1, 55kDa	3.04	8.2	4.3698e-04	Yes	Y
221	57494	RIMKLB	ribosomal modification protein rimK-like family member B	-1.07	2.1	4.4145e-04	No	N
222	8031	NCOA4	nuclear receptor coactivator 4	2.58	6.0	4.4693e-04	Yes	Y
223	10632	ATP5L	ATP synthase, H+ transporting, mitochondrial Fo complex, subunit G	-0.89	1.9	4.4977e-04	No	N
224	9829	DNAJC6	DnaJ (Hsp40) homolog, subfamily C, member 6	1.86	3.6	4.5768e-04	Yes	Y
225	58191	CXCL16	chemokine (C-X-C motif) ligand 16	1.25	2.4	4.5967e-04	Yes	N
226	3045	HBD	hemoglobin, delta	1.98	4.0	4.6247e-04	Yes	Y
227	8780	RIOK3	RIO kinase 3	2.64	6.2	4.6346e-04	Yes	Y
228	715	C1R	complement component 1, r subcomponent	1.20	2.3	4.6446e-04	No	N
229	140739	UBE2F	ubiquitin-conjugating enzyme E2F (putative)	0.88	1.8	4.6701e-04	Yes	Y
230	6622	SNCA	synuclein, alpha (non A4 component of amyloid precursor)	5.43	43.2	4.7673e-04	Yes	Y
231	23089	PEG10	paternally expressed 10	-0.77	1.7	4.7984e-04	No	N
232	8760	CDS2	CDP-diacylglycerol synthase (phosphatidate cytidylyltransferase) 2	-0.77	1.7	4.8043e-04	Yes	N
233	55074	OXR1	oxidation resistance 1	-0.83	1.8	4.8274e-04	Yes	N
234	25893	TRIM58	tripartite motif containing 58	3.42	10.7	4.8505e-04	Yes	Y
235	63899	NSUN3	NOP2/Sun domain family, member 3	1.53	2.9	4.9077e-04	Yes	Y
236	6303	SAT1	spermidine/spermine N1-acetyltransferase 1	1.08	2.1	4.9152e-04	No	N
237	23506	GLTSCR1L	GLTSCR1-like	-0.82	1.8	4.9344e-04	Yes	N
238	590	BCHE	butyrylcholinesterase	-1.49	2.8	4.9912e-04	No	N
239	100130581	LINC00910	long intergenic non-protein coding RNA 910	1.66	3.1	5.0447e-04	Yes	Y
240	6398	SECTM1	secreted and transmembrane 1	0.68	1.6	5.0490e-04	No	N
241	25884	CHRDL2	chordin-like 2	1.56	2.9	5.0500e-04	No	N
242	23762	OSBP2	oxysterol binding protein 2	1.80	3.5	5.0624e-04	Yes	Y
243	112476	PRRT2	proline-rich transmembrane protein 2	-1.12	2.2	5.1077e-04	No	N
244	4978	OPCML	opioid binding protein/cell adhesion molecule-like	-1.52	2.9	5.1648e-04	No	N
245	55632	G2E3	G2/M-phase specific E3 ubiquitin protein ligase	-0.68	1.6	5.1880e-04	Yes	N

Table 12 – Continued on next page

Table 12 – continued from previous page

246	212	ALAS2	aminolevulinate, delta-, synthase 2	5.02	32.6	5.2418e-04	Yes	Y
247	55544	RBM38	RNA binding motif protein 38	2.33	5.0	5.2563e-04	Yes	Y
248	57326	PBXIP1	pre-B-cell leukemia homeobox interacting protein 1	-0.90	1.9	5.3210e-04	No	N
249	27303	RBMS3	RNA binding motif, single stranded interacting protein 3	-1.04	2.1	5.3285e-04	No	N
250	2035	EPB41	erythrocyte membrane protein band 4.1	3.59	12.0	5.3581e-04	Yes	Y
251	10010	TANK	TRAF family member-associated NFKB activator	0.90	1.9	5.4048e-04	Yes	N
252	412	STS	steroid sulfatase (microsomal), isozyme S	-1.06	2.1	5.4657e-04	No	N
253	116151	FAM210B	family with sequence similarity 210, member B	2.12	4.4	5.4679e-04	Yes	Y
254	8515	ITGA10	integrin, alpha 10	-1.02	2.0	5.5357e-04	No	N
255	6843	VAMP1	vesicle-associated membrane protein 1 (synaptobrevin 1)	-1.00	2.0	5.5468e-04	Yes	N
256	81792	ADAMTS12	ADAM metallopeptidase with thrombospondin type 1 motif, 12	-0.93	1.9	5.5545e-04	Yes	N
257	3899	AFF3	AF4/FMR2 family, member 3	-1.17	2.3	5.5628e-04	No	N
258	6565	SLC15A2	solute carrier family 15 (oligopeptide transporter), member 2	-1.06	2.1	5.5665e-04	No	N
259	8818	DPM2	dolichyl-phosphate mannosyltransferase polypeptide 2, regulatory subunit	0.97	2.0	5.5863e-04	Yes	N
260	4900	NRGN	neurogranin (protein kinase C substrate, RC3)	1.12	2.2	5.6083e-04	Yes	N
261	6872	TAF1	TAF1 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 250kDa	-0.66	1.6	5.6325e-04	Yes	N
262	23582	CCNDBP1	cyclin D-type binding-protein 1	1.98	3.9	5.6327e-04	Yes	Y
263	79666	PLEKHF2	pleckstrin homology domain containing, family F (with FYVE domain) member 2	-0.77	1.7	5.6965e-04	Yes	N
264	3043	HBB	hemoglobin, beta	3.33	10.1	5.8187e-04	Yes	Y
265	6813	STXBP2	syntaxin binding protein 2	1.37	2.6	5.8356e-04	No	N
266	10457	GPNMB	glycoprotein (transmembrane) nmb	-1.19	2.3	5.8650e-04	Yes	N
267	51390	AIG1	androgen-induced 1	-1.11	2.2	5.8915e-04	No	N
268	2038	EPB42	erythrocyte membrane protein band 4.2	4.02	16.2	5.9060e-04	Yes	Y
269	23411	SIRT1	sirtuin 1	-0.78	1.7	5.9356e-04	Yes	N
270	27090	ST6GALNAC4	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 4	0.91	1.9	6.0340e-04	Yes	Y
271	50943	FOXP3	forkhead box P3	0.64	1.6	6.0913e-04	Yes	Y
272	2752	GLUL	glutamate-ammonia ligase	2.26	4.8	6.1231e-04	No	N
273	284406	ZFP82	ZFP82 zinc finger protein	-0.79	1.7	6.1308e-04	Yes	N
274	51596	CUTA	cutA divalent cation tolerance homolog (E. coli)	-0.80	1.7	6.1361e-04	Yes	N
275	64174	DPEP2	dipeptidase 2	-0.94	1.9	6.1689e-04	No	N
276	23076	RRP1B	ribosomal RNA processing 1B	-0.71	1.6	6.1722e-04	Yes	Y
277	202018	TAPT1	transmembrane anterior posterior transformation 1	-0.72	1.6	6.2030e-04	Yes	N

Table 12 – Continued on next page

Table 12 – continued from previous page

278	57082	CASC5	cancer susceptibility candidate 5	-0.74	1.7	6.2212e-04	Yes	Y
279	665	BNIP3L	BCL2/adenovirus E1B 19kDa interacting protein 3-like	2.44	5.4	6.2700e-04	Yes	N
280	3044	HBBP1	hemoglobin, beta pseudogene 1	1.91	3.8	6.3088e-04	Yes	Y
281	5784	PTPN14	protein tyrosine phosphatase, non-receptor type 14	-1.32	2.5	6.3100e-04	No	N
282	5008	OSM	oncostatin M	0.80	1.7	6.3899e-04	No	N
283	100190939	TPT1-AS1	TPT1 antisense RNA 1	-1.18	2.3	6.4169e-04	Yes	Y
284	948	CD36	CD36 molecule (thrombospondin receptor)	1.93	3.8	6.4382e-04	No	N
285	286042	FAM86B3P	family with sequence similarity 86, member A pseudogene	1.91	3.8	6.4417e-04	Yes	Y
286	9263	STK17A	serine/threonine kinase 17a	1.19	2.3	6.4418e-04	Yes	Y
287	23236	PLCB1	phospholipase C, beta 1 (phosphoinositide-specific)	-1.28	2.4	6.5178e-04	No	N
288	54855	FAM46C	family with sequence similarity 46, member C	4.88	29.5	6.5696e-04	Yes	Y
289	51029	DESI2	desumoylating isopeptidase 2	-1.20	2.3	6.6680e-04	Yes	N
290	80853	KDM7A	lysine (K)-specific demethylase 7A	1.67	3.2	6.7034e-04	No	N
291	388743	CAPN8	calpain 8	0.71	1.6	6.9090e-04	Yes	Y
292	23435	TARDBP	TAR DNA binding protein	-0.97	2.0	6.9099e-04	Yes	N
293	27013	CNPPD1	cyclin Pas1/PHO80 domain containing 1	1.11	2.2	6.9708e-04	Yes	N
294	2801	GOLGA2	golgin A2	-1.35	2.5	7.0006e-04	Yes	Y
295	55437	STRADB	STE20-related kinase adaptor beta	3.26	9.6	7.0059e-04	Yes	Y
296	4093	SMAD9	SMAD family member 9	-1.24	2.4	7.0067e-04	No	N
297	220972	MARCH8	membrane-associated ring finger (C3HC4) 8, E3 ubiquitin protein ligase	2.15	4.4	7.0228e-04	Yes	Y
298	6563	SLC14A1	solute carrier family 14 (urea transporter), member 1 (Kidd blood group)	2.24	4.7	7.0262e-04	Yes	Y
299	64770	CCDC14	coiled-coil domain containing 14	-0.64	1.6	7.1191e-04	Yes	N
300	10675	CSPG5	chondroitin sulfate proteoglycan 5 (neuroglycan C)	-0.88	1.8	7.2079e-04	No	N
301	340481	ZDHHC21	zinc finger, DHHC-type containing 21	-0.82	1.8	7.2170e-04	Yes	Y
302	51203	NUSAP1	nucleolar and spindle associated protein 1	-1.11	2.2	7.2254e-04	Yes	Y
303	55088	CCDC186	coiled-coil domain containing 186	-0.74	1.7	7.2289e-04	Yes	N
304	8404	SPARCL1	SPARC-like 1 (hevin)	-1.17	2.3	7.2315e-04	No	N
305	6001	RGS10	regulator of G-protein signaling 10	2.20	4.6	7.2405e-04	Yes	N
306	84376	HOOK3	hook microtubule-tethering protein 3	-0.76	1.7	7.2542e-04	Yes	N
307	55719	FAM178A	family with sequence similarity 178, member A	-0.92	1.9	7.3976e-04	Yes	N
308	84935	MEDAG	mesenteric estrogen-dependent adipogenesis	1.05	2.1	7.4487e-04	No	N
309	6006	RHCE	Rh blood group, CcEe antigens	1.17	2.2	7.6147e-04	Yes	Y
310	51094	ADIPOR1	adiponectin receptor 1	2.41	5.3	7.6235e-04	Yes	Y
311	55607	PPP1R9A	protein phosphatase 1, regulatory subunit 9A	-1.29	2.4	7.6863e-04	No	N
312	54210	TREM1	triggering receptor expressed on myeloid cells 1	0.87	1.8	7.7221e-04	No	N
313	23608	MKRN1	makorin ring finger protein 1	2.56	5.9	7.7575e-04	Yes	Y
314	4926	NUMA1	nuclear mitotic apparatus protein 1	-2.38	5.2	7.7616e-04	Yes	Y

Table 12 – Continued on next page

Table 12 – continued from previous page

315	2977	GUCY1A2	guanylate cyclase 1, soluble, alpha 2	-1.23	2.3	7.8425e-04	No	N
316	539	ATP5O	ATP synthase, H+ transporting, mitochondrial F1 complex, O subunit	-0.82	1.8	7.8750e-04	Yes	N
317	9994	CASP8AP2	caspase 8 associated protein 2	-0.85	1.8	7.8887e-04	Yes	N
318	10923	SUB1	SUB1 homolog (S. cerevisiae)	-0.77	1.7	8.0010e-04	Yes	N
319	383	ARG1	arginase 1	0.77	1.7	8.0078e-04	Yes	N
320	3576	CXCL8	chemokine (C-X-C motif) ligand 8	2.72	6.6	8.0496e-04	No	N
321	4649	MYO9A	myosin IXA	-0.83	1.8	8.0726e-04	Yes	N
322	9958	USP15	ubiquitin specific peptidase 15	1.50	2.8	8.1234e-04	Yes	N
323	2993	GYPA	glycophorin A (MNS blood group)	3.96	15.6	8.1546e-04	Yes	Y
324	3042	HBM	hemoglobin, mu	4.51	22.8	8.1865e-04	Yes	Y
325	9921	RNF10	ring finger protein 10	1.26	2.4	8.2499e-04	Yes	Y
326	2994	GYPB	glycophorin B (MNS blood group)	3.37	10.4	8.2603e-04	Yes	Y
327	19	ABCA1	ATP-binding cassette, sub-family A (ABC1), member 1	1.04	2.1	8.4459e-04	No	N
328	389114	ZNF662	zinc finger protein 662	-0.82	1.8	8.4671e-04	Yes	Y
329	79630	C1orf54	chromosome 1 open reading frame 54	-0.71	1.6	8.5018e-04	Yes	N
330	10766	TOB2	transducer of ERBB2, 2	-1.00	2.0	8.5056e-04	Yes	Y
331	54997	TESC	tescalcin	2.36	5.1	8.5294e-04	Yes	N
332	57531	HACE1	HECT domain and ankyrin repeat containing E3 ubiquitin protein ligase 1	-0.93	1.9	8.5572e-04	No	N
333	4257	MGST1	microsomal glutathione S-transferase 1	1.13	2.2	8.5951e-04	No	N
334	481	ATP1B1	ATPase, Na+/K+ transporting, beta 1 polypeptide	-1.04	2.1	8.6017e-04	Yes	Y
335	23516	SLC39A14	solute carrier family 39 (zinc transporter), member 14	0.85	1.8	8.6207e-04	No	N
336	11151	CORO1A	coronin, actin binding protein, 1A	1.32	2.5	8.7134e-04	No	N
337	8991	SELENBP1	selenium binding protein 1	2.49	5.6	8.7791e-04	Yes	N
338	3275	PRMT2	protein arginine methyltransferase 2	-1.09	2.1	9.0093e-04	No	N
339	84418	CYSTM1	cysteine-rich transmembrane module containing 1	1.29	2.4	9.2150e-04	Yes	N
340	3049	HBQ1	hemoglobin, theta 1	3.11	8.7	9.2190e-04	Yes	Y
341	26122	EPC2	enhancer of polycomb homolog 2 (Drosophila)	-0.86	1.8	9.2859e-04	Yes	N
342	25793	FBXO7	F-box protein 7	2.09	4.2	9.3451e-04	Yes	Y
343	80173	IFT74	intraflagellar transport 74	-0.89	1.8	9.3765e-04	Yes	N
344	5738	PTGFRN	prostaglandin F2 receptor inhibitor	-1.13	2.2	9.3912e-04	Yes	N
345	64130	LIN7B	lin-7 homolog B (C. elegans)	-0.75	1.7	9.4255e-04	Yes	N
346	55022	PID1	phosphotyrosine interaction domain containing 1	1.27	2.4	9.4622e-04	No	N
347	4066	LYL1	lymphoblastic leukemia associated hematopoiesis regulator 1	1.34	2.5	9.4799e-04	Yes	Y
348	5880	RAC2	ras-related C3 botulinum toxin substrate 2 (rho family, small GTP binding protein Rac2)	1.12	2.2	9.5641e-04	No	N
349	240	ALOX5	arachidonate 5-lipoxygenase	1.94	3.8	9.5747e-04	No	N
350	22934	RPIA	ribose 5-phosphate isomerase A	1.11	2.2	9.6453e-04	Yes	Y
351	11101	ATE1	arginyltransferase 1	-0.81	1.8	9.7734e-04	Yes	N
352	6888	TALDO1	transaldolase 1	1.07	2.1	9.7783e-04	Yes	Y
353	285382	C3orf70	chromosome 3 open reading frame 70	-0.71	1.6	9.7909e-04	Yes	N
354	83941	TM2D1	TM2 domain containing 1	-0.84	1.8	9.8100e-04	Yes	N

Table 12 – Continued on next page

Table 12 – continued from previous page

355	8667	EIF3H	eukaryotic translation initiation factor 3, subunit H	-1.09	2.1	9.9160e-04	No	N
356	23216	TBC1D1	TBC1 (tre-2/USP6, BUB2, cdc16) domain family, member 1	-0.76	1.7	1.0007e-03	Yes	N
357	55084	SOBP	sine oculis binding protein homolog (Drosophila)	-1.13	2.2	1.0013e-03	Yes	N
358	93134	ZNF561	zinc finger protein 561	-0.81	1.8	1.0053e-03	No	N
359	57205	ATP10D	ATPase, class V, type 10D	-0.84	1.8	1.0077e-03	Yes	Y
360	22943	DKK1	dickkopf WNT signaling pathway inhibitor 1	1.26	2.4	1.0099e-03	Yes	Y
361	124491	TMEM170A	transmembrane protein 170A	-0.77	1.7	1.0124e-03	Yes	Y
362	6521	SLC4A1	solute carrier family 4 (anion exchanger), member 1 (Diego blood group)	3.48	11.2	1.0160e-03	Yes	Y
363	160897	GPR180	G protein-coupled receptor 180	-0.97	2.0	1.0218e-03	Yes	N
364	5054	SERPINE1	serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1	1.16	2.2	1.0287e-03	No	N
365	1116	CHI3L1	chitinase 3-like 1 (cartilage glycoprotein-39)	2.36	5.1	1.0330e-03	No	N
366	3033	HADH	hydroxyacyl-CoA dehydrogenase	-0.91	1.9	1.0331e-03	No	N
367	23142	DCUN1D4	DCN1, defective in cullin neddylation 1, domain containing 4	-0.71	1.6	1.0343e-03	No	N
368	2766	GMPR	guanosine monophosphate reductase	1.85	3.6	1.0344e-03	Yes	N
369	5918	RARRES1	retinoic acid receptor responder (tazarotene induced) 1	1.25	2.4	1.0377e-03	No	N
370	54757	FAM20A	family with sequence similarity 20, member A	0.70	1.6	1.0422e-03	Yes	N
371	286	ANK1	ankyrin 1, erythrocytic	1.37	2.6	1.0441e-03	Yes	Y
372	2621	GAS6	growth arrest-specific 6	-1.00	2.0	1.0474e-03	Yes	N
373	5329	PLAUR	plasminogen activator, urokinase receptor	0.85	1.8	1.0572e-03	No	N
374	29994	BAZ2B	bromodomain adjacent to zinc finger domain, 2B	-0.90	1.9	1.0651e-03	Yes	N
375	79144	PPDPF	pancreatic progenitor cell differentiation and proliferation factor	-0.73	1.7	1.0736e-03	Yes	N
376	23658	LSM5	LSM5 homolog, U6 small nuclear RNA associated (S. cerevisiae)	-1.13	2.2	1.0742e-03	No	N
377	51196	PLCE1	phospholipase C, epsilon 1	-1.57	3.0	1.0767e-03	No	N
378	55704	CCDC88A	coiled-coil domain containing 88A	-0.72	1.6	1.0776e-03	Yes	N
379	4778	NFE2	nuclear factor, erythroid 2	3.73	13.2	1.0816e-03	No	N
380	3070	HELLS	helicase, lymphoid-specific	-0.61	1.5	1.0953e-03	Yes	N
381	55228	PNMAL1	paraneoplastic Ma antigen family-like 1	-1.68	3.2	1.1099e-03	No	N
382	1032	CDKN2D	cyclin-dependent kinase inhibitor 2D (p19, inhibits CDK4)	1.34	2.5	1.1099e-03	Yes	N
383	57095	PITHD1	PITH (C-terminal proteasome-interacting domain of thioredoxin-like) domain containing 1	1.78	3.4	1.1161e-03	Yes	Y
384	9667	SAFB2	scaffold attachment factor B2	-0.71	1.6	1.1168e-03	Yes	N
385	643911	CRNDE	colorectal neoplasia differentially expressed (non-protein coding)	-1.04	2.1	1.1180e-03	Yes	Y
386	4837	NNMT	nicotinamide N-methyltransferase	1.09	2.1	1.1184e-03	No	N

Table 12 – Continued on next page

Table 12 – continued from previous page

387	150378	FLJ30901	uncharacterized protein FLJ30901	1.01	2.0	1.1257e-03	Yes	Ye
388	23493	HEY2	hes-related family bHLH transcription factor with YRPW motif 2	-0.90	1.9	1.1348e-03	No	N
389	51629	SLC25A39	solute carrier family 25, member 39	3.21	9.3	1.1373e-03	Yes	Ye
390	58494	JAM2	junctional adhesion molecule 2	-0.92	1.9	1.1552e-03	No	N
391	10818	FRS2	fibroblast growth factor receptor substrate 2	-0.92	1.9	1.1714e-03	Yes	Ye
392	55432	YOD1	YOD1 deubiquitinase	3.40	10.5	1.1723e-03	Yes	Ye
393	57575	PCDH10	protocadherin 10	0.97	2.0	1.1728e-03	No	N
394	2769	GNA15	guanine nucleotide binding protein (G protein), alpha 15 (Gq class)	1.10	2.1	1.1751e-03	No	N
395	2987	GUK1	guanylate kinase 1	1.00	2.0	1.1765e-03	Yes	Ye
396	1303	COL12A1	collagen, type XII, alpha 1	-0.70	1.6	1.1803e-03	Yes	Ye
397	79822	ARHGAP28	Rho GTPase activating protein 28	-1.07	2.1	1.1919e-03	No	N
398	51122	COMMD2	COMM domain containing 2	-0.61	1.5	1.1921e-03	Yes	N
399	80256	FAM214B	family with sequence similarity 214, member B	0.92	1.9	1.1997e-03	Yes	N
400	80144	FRAS1	Fraser extracellular matrix complex subunit 1	-1.22	2.3	1.2028e-03	No	N
401	2039	DMTN	dematin actin binding protein	2.16	4.5	1.2063e-03	Yes	Ye
402	5537	PPP6C	protein phosphatase 6, catalytic subunit	-0.76	1.7	1.2106e-03	Yes	N
403	55599	RNPC3	RNA-binding region (RNP1, RRM) containing 3	-0.80	1.7	1.2129e-03	Yes	N
404	26098	EDRF1	erythroid differentiation regulatory factor 1	-0.66	1.6	1.2195e-03	Yes	N
405	64393	ZMAT3	zinc finger, matrin-type 3	-1.47	2.8	1.2273e-03	Yes	Ye
406	56906	THAP10	THAP domain containing 10	-0.70	1.6	1.2428e-03	Yes	N
407	395	ARHGAP6	Rho GTPase activating protein 6	-1.15	2.2	1.2672e-03	No	N
408	6678	SPARC	secreted protein, acidic, cysteine-rich (osteonectin)	-0.68	1.6	1.2813e-03	Yes	N
409	28984	RGCC	regulator of cell cycle	2.78	6.9	1.2892e-03	Yes	Ye
410	4725	NDUFS5	NADH dehydrogenase (ubiquinone) Fe-S protein 5, 15kDa (NADH-coenzyme Q reductase)	-0.75	1.7	1.2918e-03	Yes	Ye
411	6670	SP3	Sp3 transcription factor	-0.93	1.9	1.2956e-03	Yes	N
412	22990	PCNX	pecanex homolog (Drosophila)	-1.26	2.4	1.3046e-03	Yes	Ye
413	1870	E2F2	E2F transcription factor 2	1.52	2.9	1.3057e-03	Yes	Ye
414	27346	TMEM97	transmembrane protein 97	-1.09	2.1	1.3103e-03	No	N
415	5971	RELB	v-rel avian reticuloendotheliosis viral oncogene homolog B	0.87	1.8	1.3156e-03	No	N
416	1633	DCK	deoxycytidine kinase	2.03	4.1	1.3325e-03	Yes	N
417	64847	SPATA20	spermatogenesis associated 20	-0.90	1.9	1.3448e-03	Yes	N
418	54386	TERF2IP	telomeric repeat binding factor 2, interacting protein	1.24	2.4	1.3507e-03	Yes	Ye
419	51633	OTUD6B	OTU domain containing 6B	-0.94	1.9	1.3508e-03	Yes	N
420	7275	TUB	tubby bipartite transcription factor	-1.27	2.4	1.3814e-03	No	N
421	3674	ITGA2B	integrin, alpha 2b (platelet glycoprotein IIb of IIb/IIIa complex, antigen CD41)	0.93	1.9	1.3856e-03	Yes	Ye
422	79768	KATNBL1	katanin p80 subunit B-like 1	-2.16	4.5	1.3862e-03	Yes	Ye
423	5552	SRGN	serglycin	1.40	2.6	1.3906e-03	No	N
424	7337	UBE3A	ubiquitin protein ligase E3A	-0.80	1.7	1.3923e-03	Yes	N

Table 12 – Continued on next page

Table 12 – continued from previous page

425	149076	ZNF362	zinc finger protein 362	-0.76	1.7	1.4102e-03	Yes	N
426	10974	ADIRF	adipogenesis regulatory factor	-1.53	2.9	1.4189e-03	No	N
427	81689	ISCA1	iron-sulfur cluster assembly 1	0.75	1.7	1.4243e-03	Yes	N
428	23635	SSBP2	single-stranded DNA binding protein 2	-0.84	1.8	1.4278e-03	No	N
429	5341	PLEK	pleckstrin	0.93	1.9	1.4285e-03	No	N
430	493856	CISD2	CDGSH iron sulfur domain 2	2.19	4.5	1.4459e-03	Yes	Y
431	79071	ELOVL6	ELOVL fatty acid elongase 6	1.46	2.8	1.4529e-03	Yes	Y
432	6397	SEC14L1	SEC14-like 1 (S. cerevisiae)	1.67	3.2	1.4594e-03	Yes	Y
433	2532	ACKR1	atypical chemokine receptor 1 (Duffy blood group)	1.03	2.0	1.4636e-03	No	N
434	286097	MICU3	mitochondrial calcium uptake family, member 3	-0.73	1.7	1.5481e-03	No	N
435	114876	OSBPL1A	oxysterol binding protein-like 1A	-1.03	2.0	1.5505e-03	No	N
436	3575	IL7R	interleukin 7 receptor	1.08	2.1	1.5605e-03	Yes	N
437	23526	HMHA1	histocompatibility (minor) HA-1	0.85	1.8	1.5619e-03	No	N
438	2357	FPR1	formyl peptide receptor 1	1.12	2.2	1.5788e-03	No	N
439	10189	ALYREF	Aly/REF export factor	-2.03	4.1	1.5886e-03	Yes	Y
440	5730	PTGDS	prostaglandin D2 synthase 21kDa (brain)	-0.98	2.0	1.5934e-03	Yes	N
441	5886	RAD23A	RAD23 homolog A (S. cerevisiae)	0.88	1.8	1.5974e-03	Yes	Y
442	283298	OLFML1	olfactomedin-like 1	-1.14	2.2	1.6104e-03	No	N
443	55717	WDR11	WD repeat domain 11	-1.17	2.3	1.6158e-03	Yes	Y
444	9048	ARTN	artemin	0.78	1.7	1.6171e-03	Yes	Y
445	9173	IL1RL1	interleukin 1 receptor-like 1	3.00	8.0	1.6293e-03	No	N
446	8740	TNFSF14	tumor necrosis factor (ligand) superfamily, member 14	0.92	1.9	1.6357e-03	Yes	N
447	978	CDA	cytidine deaminase	0.74	1.7	1.6498e-03	Yes	N
448	5594	MAPK1	mitogen-activated protein kinase 1	1.00	2.0	1.6598e-03	Yes	Y
449	51133	KCTD3	potassium channel tetramerization domain containing 3	-0.76	1.7	1.6663e-03	No	N
450	26230	TIAM2	T-cell lymphoma invasion and metastasis 2	-0.90	1.9	1.6668e-03	Yes	N
451	157638	FAM84B	family with sequence similarity 84, member B	-0.89	1.8	1.6789e-03	No	N
452	26586	CKAP2	cytoskeleton associated protein 2	-0.72	1.7	1.6904e-03	Yes	Y
453	123283	TARSL2	threonyl-tRNA synthetase-like 2	-0.83	1.8	1.6920e-03	Yes	N
454	53354	PANK1	pantothenate kinase 1	-0.86	1.8	1.6923e-03	No	N
455	2040	STOM	stomatin	1.01	2.0	1.7001e-03	Yes	N
456	6774	STAT3	signal transducer and activator of transcription 3 (acute-phase response factor)	0.85	1.8	1.7007e-03	No	N
457	23641	LDOC1	leucine zipper, down-regulated in cancer 1	-0.87	1.8	1.7139e-03	No	N
458	3554	IL1R1	interleukin 1 receptor, type I	1.06	2.1	1.7152e-03	No	N
459	79956	ERMP1	endoplasmic reticulum metallopeptidase 1	-0.91	1.9	1.7182e-03	No	N
460	6372	CXCL6	chemokine (C-X-C motif) ligand 6	2.79	6.9	1.7275e-03	No	N
461	400916	CHCHD10	coiled-coil-helix-coiled-coil-helix domain containing 10	-0.87	1.8	1.7329e-03	No	N
462	51218	GLRX5	glutaredoxin 5	1.56	2.9	1.7699e-03	Yes	Y
463	4946	OAZ1	ornithine decarboxylase antizyme 1	1.48	2.8	1.7710e-03	Yes	Y
464	1890	TYMP	thymidine phosphorylase	0.81	1.8	1.7877e-03	No	N

Table 12 – Continued on next page

Table 12 – continued from previous page

465	54800	KLHL24	kelch-like family member 24	-1.08	2.1	1.8017e-03	No	N
466	8531	YBX3	Y box binding protein 3	1.12	2.2	1.8089e-03	Yes	N
467	56987	BBX	bobby sox homolog (Drosophila)	-1.27	2.4	1.8226e-03	Yes	Y
468	83719	YPEL3	yippee-like 3 (Drosophila)	1.56	2.9	1.8241e-03	Yes	Y
469	2926	GRSF1	G-rich RNA sequence binding factor 1	-0.80	1.7	1.8349e-03	No	N
470	79180	EFHD2	EF-hand domain family, member D2	0.78	1.7	1.8364e-03	No	N
471	415116	PIM3	Pim-3 proto-oncogene, serine/threonine kinase	0.59	1.5	1.8483e-03	No	N
472	7037	TFRC	transferrin receptor	2.27	4.8	1.8493e-03	No	N
473	4116	MAGOH	mago-nashi homolog, proliferation-associated (Drosophila)	-1.05	2.1	1.8502e-03	Yes	N
474	143686	SESN3	sestrin 3	2.58	6.0	1.8619e-03	Yes	Y
475	4660	PPP1R12B	protein phosphatase 1, regulatory subunit 12B	-1.56	2.9	1.8677e-03	Yes	Y
476	51393	TRPV2	transient receptor potential cation channel, subfamily V, member 2	1.05	2.1	1.8773e-03	Yes	N
477	9079	LDB2	LIM domain binding 2	-0.98	2.0	1.8799e-03	No	N
478	83696	TRAPPC9	trafficking protein particle complex 9	1.16	2.2	1.8818e-03	Yes	Y
479	57684	ZBTB26	zinc finger and BTB domain containing 26	-0.67	1.6	1.8819e-03	Yes	N
480	100873954	SNRK-AS1	SNRK antisense RNA 1	-0.64	1.6	1.8960e-03	Yes	N
481	219348	PLAC9	placenta-specific 9	-1.03	2.0	1.8962e-03	Yes	Y
482	91151	TIGD7	tigger transposable element derived 7	-0.80	1.7	1.9070e-03	Yes	N
483	54602	NDFIP2	Nedd4 family interacting protein 2	-0.65	1.6	1.9078e-03	Yes	Y
484	79667	KLF3-AS1	KLF3 antisense RNA 1	-0.93	1.9	1.9290e-03	Yes	N
485	57703	CWC22	CWC22 spliceosome-associated protein	-0.66	1.6	1.9292e-03	Yes	N
486	79701	OGFOD3	2-oxoglutarate and iron-dependent oxygenase domain containing 3	-1.06	2.1	1.9308e-03	Yes	Y
487	4792	NFKBIA	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha	0.83	1.8	1.9368e-03	No	N
488	1475	CSTA	cystatin A (stefin A)	1.55	2.9	1.9387e-03	No	N
489	3848	KRT1	keratin 1	3.30	9.8	1.9419e-03	Yes	Y
490	997	CDC34	cell division cycle 34	1.04	2.1	1.9478e-03	Yes	Y
491	1879	EBF1	early B-cell factor 1	-0.87	1.8	1.9616e-03	Yes	Y
492	701	BUB1B	BUB1 mitotic checkpoint serine/threonine kinase B	-0.66	1.6	1.9806e-03	Yes	Y
493	10181	RBM5	RNA binding motif protein 5	-1.14	2.2	1.9841e-03	Yes	Y
494	4791	NFKB2	nuclear factor of kappa light polypeptide gene enhancer in B-cells 2 (p49/p100)	1.25	2.4	2.0048e-03	No	N
495	6535	SLC6A8	solute carrier family 6 (neurotransmitter transporter), member 8	1.49	2.8	2.0366e-03	Yes	N
496	64098	PARVG	parvin, gamma	0.71	1.6	2.0638e-03	No	N
497	9584	RBM39	RNA binding motif protein 39	-0.92	1.9	2.0753e-03	Yes	Y
498	5586	PKN2	protein kinase N2	-0.84	1.8	2.0872e-03	Yes	N
499	3385	ICAM3	intercellular adhesion molecule 3	0.99	2.0	2.1034e-03	No	N
500	10900	RUNDC3A	RUN domain containing 3A	2.12	4.3	2.1048e-03	Yes	Y
501	9837	GINS1	GINS complex subunit 1 (Psf1 homolog)	-0.92	1.9	2.1214e-03	Yes	N
502	162466	PHOSPHO1	phosphatase, orphan 1	1.43	2.7	2.1289e-03	Yes	Y
503	682	BSG	basigin (Ok blood group)	1.36	2.6	2.1338e-03	Yes	Y

Table 12 – Continued on next page

Table 12 – continued from previous page

504	64426	SUDS3	suppressor of defective silencing 3 homolog (S. cerevisiae)	-0.70	1.6	2.1348e-03	No	N
505	25909	AHCTF1	AT hook containing transcription factor 1	-0.66	1.6	2.1465e-03	Yes	N
506	83852	SETDB2	SET domain, bifurcated 2	-0.71	1.6	2.1534e-03	Yes	N
507	283991	UBALD2	UBA-like domain containing 2	0.82	1.8	2.1814e-03	Yes	N
508	112770	C1orf85	chromosome 1 open reading frame 85	-0.92	1.9	2.1854e-03	No	N
509	2534	FYN	FYN proto-oncogene, Src family tyrosine kinase	-1.00	2.0	2.1938e-03	Yes	N
510	51312	SLC25A37	solute carrier family 25 (mitochondrial iron transporter), member 37	3.45	10.9	2.2079e-03	No	N
511	79710	MORC4	MORC family CW-type zinc finger 4	-0.82	1.8	2.2195e-03	No	N
512	83543	AIF1L	allograft inflammatory factor 1-like	-1.23	2.4	2.2245e-03	No	N
513	8993	PGLYRP1	peptidoglycan recognition protein 1	0.78	1.7	2.2333e-03	Yes	N
514	57685	CACHD1	cache domain containing 1	-0.87	1.8	2.2340e-03	Yes	Y
515	1669	DEFA4	defensin, alpha 4, corticostatin	0.80	1.7	2.2472e-03	Yes	Y
516	57541	ZNF398	zinc finger protein 398	-0.71	1.6	2.2607e-03	Yes	N
517	8548	BLZF1	basic leucine zipper nuclear factor 1	-0.63	1.5	2.2684e-03	Yes	N
518	57545	CC2D2A	coiled-coil and C2 domain containing 2A	-0.80	1.7	2.2727e-03	No	N
519	1908	EDN3	endothelin 3	-1.72	3.3	2.2898e-03	No	N
520	3643	INSR	insulin receptor	-0.72	1.6	2.3250e-03	Yes	N
521	57188	ADAMTSL3	ADAMTS-like 3	-0.88	1.8	2.3275e-03	No	N
522	79682	CENPU	centromere protein U	-0.86	1.8	2.3320e-03	Yes	N
523	64333	ARHGAP9	Rho GTPase activating protein 9	0.91	1.9	2.3389e-03	No	N
524	1629	DBT	dihydrolipoamide branched chain transacylase E2	-0.72	1.6	2.3587e-03	Yes	N
525	84131	CEP78	centrosomal protein 78kDa	-0.97	2.0	2.3629e-03	No	N
526	8924	HERC2	HECT and RLD domain containing E3 ubiquitin protein ligase 2	-0.75	1.7	2.3796e-03	Yes	N
527	4067	LYN	LYN proto-oncogene, Src family tyrosine kinase	1.40	2.6	2.3962e-03	No	N
528	170691	ADAMTS17	ADAM metallopeptidase with thrombospondin type 1 motif, 17	-1.16	2.2	2.4047e-03	No	N
529	2261	FGFR3	fibroblast growth factor receptor 3	-1.12	2.2	2.4071e-03	No	N
530	53335	BCL11A	B-cell CLL/lymphoma 11A (zinc finger protein)	-0.98	2.0	2.4196e-03	Yes	N
531	7009	TMBIM6	transmembrane BAX inhibitor motif containing 6	0.77	1.7	2.4271e-03	Yes	N
532	527	ATP6V0C	ATPase, H+ transporting, lysosomal 16kDa, V0 subunit c	1.29	2.4	2.4436e-03	Yes	Y
533	4050	LTB	lymphotoxin beta (TNF superfamily, member 3)	1.03	2.0	2.4515e-03	No	N
534	9411	ARHGAP29	Rho GTPase activating protein 29	-0.92	1.9	2.4673e-03	Yes	N
535	1381	CRABP1	cellular retinoic acid binding protein 1	-0.96	1.9	2.4829e-03	Yes	Y
536	7512	XPNPEP2	X-prolyl aminopeptidase (aminopeptidase P) 2, membrane-bound	1.24	2.4	2.4935e-03	Yes	N
537	56994	CHPT1	choline phosphotransferase 1	1.47	2.8	2.5168e-03	Yes	Y
538	83442	SH3BGRL3	SH3 domain binding glutamate-rich protein like 3	0.76	1.7	2.5293e-03	No	N
539	92	ACVR2A	activin A receptor, type IIA	-1.07	2.1	2.5365e-03	Yes	Y

Table 12 – Continued on next page

Table 12 – continued from previous page

540	5045	FURIN	furin (paired basic amino acid cleaving enzyme)	0.71	1.6	2.5388e-03	Yes	N
541	29761	USP25	ubiquitin specific peptidase 25	-1.12	2.2	2.5562e-03	No	N
542	54518	APBB1IP	amyloid beta (A4) precursor protein-binding, family B, member 1 interacting protein	0.78	1.7	2.5578e-03	No	N
543	28639	TRBC1	T cell receptor beta constant 1	1.26	2.4	2.5584e-03	Yes	Y
544	23189	KANK1	KN motif and ankyrin repeat domains 1	-0.94	1.9	2.5859e-03	No	N
545	7552	ZNF711	zinc finger protein 711	-0.90	1.9	2.5892e-03	No	N
546	10966	RAB40B	RAB40B, member RAS oncogene family	-0.95	1.9	2.5893e-03	No	N
547	4353	MPO	myeloperoxidase	0.74	1.7	2.6159e-03	Yes	N
548	6355	CCL8	chemokine (C-C motif) ligand 8	0.91	1.9	2.6177e-03	No	N
549	54462	CCSER2	coiled-coil serine-rich protein 2	-0.70	1.6	2.6358e-03	Yes	N
550	3176	HNMT	histamine N-methyltransferase	-0.97	2.0	2.6380e-03	No	N
551	28968	SLC6A16	solute carrier family 6, member 16	-0.86	1.8	2.6657e-03	Yes	Y
552	3050	HBZ	hemoglobin, zeta	4.12	17.4	2.6724e-03	Yes	Y
553	7072	TIA1	TIA1 cytotoxic granule-associated RNA binding protein	-0.91	1.9	2.6739e-03	Yes	N
554	7351	UCP2	uncoupling protein 2 (mitochondrial, proton carrier)	2.67	6.4	2.6772e-03	No	N
555	79930	DOK3	docking protein 3	0.96	1.9	2.6795e-03	No	N
556	6556	SLC11A1	solute carrier family 11 (proton-coupled divalent metal ion transporter), member 1	1.03	2.0	2.6930e-03	No	N
557	10929	SRSF8	serine/arginine-rich splicing factor 8	0.84	1.8	2.6983e-03	Yes	Y
558	54819	ZCCHC10	zinc finger, CCHC domain containing 10	-0.76	1.7	2.7064e-03	Yes	N
559	9843	HEPH	hephaestin	-0.90	1.9	2.7162e-03	Yes	N
560	55727	BTBD7	BTB (POZ) domain containing 7	-0.90	1.9	2.7224e-03	Yes	N
561	26046	LTN1	listerin E3 ubiquitin protein ligase 1	-1.27	2.4	2.7273e-03	Yes	Y
562	91300	R3HDM4	R3H domain containing 4	2.49	5.6	2.7515e-03	Yes	Y
563	84168	ANTXR1	anthrax toxin receptor 1	-0.65	1.6	2.7533e-03	Yes	Y
564	115123	MARCH3	membrane-associated ring finger (C3HC4) 3, E3 ubiquitin protein ligase	0.86	1.8	2.7619e-03	No	N
565	643837	LINC01128	long intergenic non-protein coding RNA 1128	-0.85	1.8	2.7687e-03	Yes	N
566	26037	SIPA1L1	signal-induced proliferation-associated 1 like 1	-0.84	1.8	2.8203e-03	No	N

3.9 Session Information

```
> toLatex(sessionInfo())
```

- R version 3.1.3 (2015-03-09), `x86_64-unknown-linux-gnu`
- Locale: `LC_CTYPE=en_US.UTF8`, `LC_NUMERIC=C`, `LC_TIME=en_US.UTF8`, `LC_COLLATE=en_US.UTF8`, `LC_MONETARY=en_US.UTF8`, `LC_MESSAGES=en_US.UTF8`, `LC_PAPER=en_US.UTF8`, `LC_NAME=en_US.UTF8`, `LC_ADDRESS=en_US.UTF8`, `LC_TELEPHONE=en_US.UTF8`, `LC_MEASUREMENT=en_US.UTF8`, `LC_IDENTIFICATION=en_US.UTF8`
- Base packages: base, datasets, graphics, grDevices, grid, methods, parallel, stats, stats4, utils

- Other packages: annotate 1.44.0, AnnotationDbi 1.28.2, Biobase 2.26.0, BiocGenerics 0.12.1, BiocParallel 1.0.3, DBI 0.3.1, Formula 1.2-1, frma 1.18.0, genefilter 1.48.1, geneplotter 1.44.0, GenomeInfoDb 1.2.5, ggplot2 1.0.1, hgu133plus2.db 3.0.0, Hmisc 3.16-0, IRanges 2.0.1, lattice 0.20-31, limma 3.22.7, mgcv 1.8-6, nlme 3.1-120, org.Hs.eg.db 3.0.0, plotrix 3.5-11, RColorBrewer 1.1-2, rJava 0.9-6, RSQLite 1.0.0, S4Vectors 0.4.0, survival 2.38-1, sva 3.12.0, xlsx 0.5.7, xlsxjars 0.6.1, XML 3.98-1.1, xtable 1.7-4
- Loaded via a namespace (and not attached): acepack 1.3-3.3, affxparser 1.38.0, affy 1.44.0, affyio 1.34.0, base64enc 0.1-2, BatchJobs 1.6, BBmisc 1.9, BiocInstaller 1.16.4, Biostrings 2.34.1, bit 1.1-12, brew 1.0-6, checkmate 1.5.2, cluster 2.0.1, codetools 0.2-11, colorspace 1.2-6, digest 0.6.8, fail 1.2, ff 2.2-13, foreach 1.4.2, foreign 0.8-63, GenomicRanges 1.18.4, gridExtra 0.9.1, gtable 0.1.2, iterators 1.0.7, latticeExtra 0.6-26, MASS 7.3-40, Matrix 1.2-0, munsell 0.4.2, nnet 7.3-9, oligo 1.30.0, oligoClasses 1.28.0, plyr 1.8.2, preprocessCore 1.28.0, proto 0.3-10, Rcpp 0.11.6, reshape2 1.4.1, rpart 4.1-9, scales 0.2.4, sendmailR 1.2-1, splines 3.1.3, stringr 0.6.2, tools 3.1.3, XVector 0.6.0, zlibbioc 1.12.0

4 Functional analysis of differentially expressed genes

4.1 Enrichment of DE genes in Gene Ontology (GO) terms

We perform a functional enrichment analysis with Gene Ontology (GO) among the enes forming the list of 1097 DE genes. We do it, however, separately for up-regulated and down-regulated genes because we expect that these two categories of DE genes are involved in very different molecular processes under the FIR affected and nonaffected infants.

For the purpose of filtering by variability (IQR) during this analysis and plotting a heatmap of gene expression profiles for the DE genes we are going to transform gene expression values to remove biological and non-biological variation unrelated to FIR.

```
> cleanExps <- removeBatchEffect(exprs(frmaEsetFilt), batch=frmaEsetFilt$Batch,
                                 covariates=sv$sv,
                                 design=model.matrix(~ factor(frmaEsetFilt$FIR)))
```

Throughout the two GO enrichment analyses, we will use the following *p*-value cutoff for the conditional hypergeometric test which removes overlapping genes in parent GO terms from child GO terms meeting this *p*-value cutoff, prior to their test (10).

```
> goAnalysisCondPvalueCutoff <- 0.01
```

Due to the large number of GO gene sets that may meet this cutoff we will follow a conservative approach and correct the *p*-values by FDR, selecting those with FDR $< 10\%$ and further restrict the selected gene sets to those with OR> 1.5. Prior to the multiple testing correction we perform a non-specific filtering step consisting of discarding gene sets with less than 5 and more than 300 genes annotated to them from the gene universe. We also filter out those gene sets whose average IQR across its genes, is below the 90% of a simulated distribution of average IQR values calculated from random gene sets sampled uniformly at random with matching gene set size. The enriched GO terms for up-regulated genes can be found in Table 13.

The gene universe is formed by all the genes either with up or down fold changes among those meeting the functional non-specific filter (20155 genes), and depending on whether the enrichment analysis is for up or down regulated genes.

```
> maskUp <- ttAll$logFC > 0
> DEgenesUp <- intersect(ttAll[maskUp, "ID"], DEgenes)
> length(DEgenesUp)
[1] 592
> gUniverse <- ttAll[maskUp, "ID"]
> length(gUniverse)
```

```
[1] 10868
> GOparams <- new("GOHyperGParams", geneIds=DEgenesUp, universeGeneIds=gUniverse,
                  annotation="org.Hs.eg.db", ontology="BP",
                  pvalueCutoff=goAnalysisCondPvalueCutoff, conditional=TRUE,
                  testDirection="over")
> cache(hgOverGOBPdeGenesUp <- hyperGTest2(GOparams, minSizeCutoff=5,
                                           maxSizeCutoff=300, orCutoff=1.5),
        dir=cacheDir_FE, prefix=cachePrefix_FE)
> hgOverGOBPdeGenesUp
Gene to GO BP Conditional test for over-representation
5790 GO BP ids tested (716 have p < 0.01)
Selected gene set size: 562
    Gene universe size: 7703
    Annotation package: org.Hs.eg
> IQRs <- apply(cleanExps[gUniverse, ], 1, IQR)
> cache(reportOverGOBPdeGenesUp <- GOreport(hgOverGOBPdeGenesUp, minSize=5, maxSize=300,
                                            minCount=5, minOddsRatio=1.5,
                                            featureFilterValues=IQRs,
                                            ff.R=1000, ff.prob=0.90, maxReported=Inf,
                                            p.value=0.1, p.adjust.method="fdr"),
        dir=cacheDir_FE, prefix=cachePrefix_FE)
> dim(reportOverGOBPdeGenesUp)
[1] 542  10
```

Table 13. Gene Ontology (GO) enrichment (biological process) of upregulated genes called differentially expressed between FIR and nonFIR affected infants at FDR $< 1\%$ and with a minimum 1.5-fold change. Column "GeneSyms" contains the genes enriching the corresponding GO term.

Term	OddsRatio	ExpCount	Count	Size	GeneSyms
neutrophil extravasation	Inf	0.36	5	5	ADAM8, AMICA1, PIK3CD, PIK3CG, TREM1
leukocyte migration involved in inflammatory response	77.05	0.51	6	7	ADAM8, ITGAM, PPBP, S100A8, S100A9, SELE
leukocyte aggregation	51.55	0.73	8	10	SEMA4D, HAS2, IL1B, RAC2, S100A8, S100A9, STK10, CD44
neutrophil activation	45.50	0.65	7	9	CD300A, CXCL8, CXCR2, PRKCD, PREX1, CCL5, CXCR4
cellular response to zinc ion	30.01	0.73	7	10	MT1E, MT1F, MT1G, MT1H, MT1M, MT1X, MT2A
response to peptidoglycan	22.50	0.80	7	11	IRAK3, IL6, IRAK1, MYD88, NOD2, TLR2, C5AR1
mononuclear cell migration	21.36	0.58	5	8	LGALS3, CCL2, CCL5, THBS1, C3AR1
negative regulation of interleukin-12 production	19.25	0.73	6	10	IRAK3, JAK3, ARRB2, NOD2, THBS1, TLR2
respiratory burst	17.26	1.02	8	14	CD24, CD52, CYBA, CYBB, CD55, NCF2, RAC2, SLC11A1
positive regulation of monocyte chemotaxis	16.02	0.66	5	9	CCR1, AIF1, SERPINE1, CCL5, CCR2

Table 13 – Continued on next page

Table 13 – continued from previous page

neutrophil chemotaxis	16.00	3.06	23	42	AMICA1, CSF3R, FCER1G, CXCL3, IL1B, CXCL8, CXCR2, ITGAM, ITGB2, LGALS3, PIK3CD, PIK3CG, TREM1, PREX1, S100A8, S100A9, S100A12, CCL2, NOD2, SPP1, SYK, C3AR1, C5AR1
positive regulation of acute inflammatory response	15.82	1.46	11	20	ADAM8, FCER1G, ALOX5AP, IL1B, IL6, IL6ST, OSM, PIK3CG, PTGS2, C3, OSMR
positive regulation of smooth muscle cell migration	15.40	0.80	6	11	LPAR1, AIF1, ITGA2, RETN, CCL5, NRP1
defense response to protozoan	15.40	0.80	6	11	IL4R, IL6, BCL3, CLEC7A, SLC11A1, CD37
granulocyte migration	14.53	3.48	25	48	AMICA1, CSF3R, FCER1G, CXCL3, IL1B, CXCL8, CXCR2, ITGAM, ITGB2, PIK3CD, PIK3CG, TREM1, PREX1, RAC2, S100A8, S100A9, S100A12, CCL2, NOD2, SPP1, SYK, THBS1, C3AR1, C5AR1, VEGFA
positive regulation vascular endothelial growth factor production	14.36	1.39	10	19	HPSE, CYP1B1, SULF1, HIF1A, IL1A, IL1B, IL6ST, PTGS2, C3, C3AR1
negative regulation of peptidyl-tyrosine phosphorylation	12.90	1.31	9	18	SEMA4D, CD300A, ERRFI1, PRKCD, PTPN2, PTPN6, SAMSN1, SOCS1, MVP
macrophage differentiation	12.89	0.73	5	10	CSF1R, ITGAM, MMP9, SPI1, VEGFA
phagocytosis, engulfment	12.88	1.17	8	16	CD300A, ABCA1, AIF1, FCER1G, ITGA2, BIN2, THBS1, CD36
leukocyte chemotaxis	12.85	2.62	18	37	ADAM8, CORO1A, AMICA1, CSF3R, FCER1G, CXCL3, IL1B, ITGAM, ITGB2, PF4, PIK3CD, TREM1, PREX1, S100A8, S100A9, SPP1, SYK, C5AR1
ovulation	12.81	0.73	5	10	IL4R, MMP19, PLAT, PTGS2, ADAMTS1
cellular response to cadmium ion	12.81	0.73	5	10	MT1E, MT1F, MT1G, MT1H, MT1X
dendritic cell chemotaxis	11.00	0.95	6	13	CCR1, CXCR2, PIK3CG, CCL5, CCR2, CXCR4
eosinophil migration	10.67	0.80	5	11	ADAM8, CD300A, LGALS3, PTGER4, CCL5
monocyte chemotaxis	10.08	1.16	7	16	FLT1, GREM1, IL6, LGALS3, LYN, S100A12, CCL2

Table 13 – Continued on next page

Table 13 – continued from previous page

negative regulation of cytokine-mediated signaling pathway	9.99	1.17	7	16	IRAK3, IL1RN, IL6ST, PTPN2, PTPRC, ROBO1, CCL5
negative regulation of insulin receptor signaling pathway	9.99	1.17	7	16	IL1B, PID1, PRKCD, PTPN2, PTPRE, SOCS1, SOCS3
production of molecular mediator involved in inflammatory response	9.67	1.53	9	21	FCER1G, IL17RA, ALOX5, ALOX5AP, IL4R, LYN, SERPINE1, SLC7A2, SYK
decidualization	9.62	1.02	6	14	JUNB, LIF, PTGS2, SPP1, STC1, VDR
regulation of lipopolysaccharide-mediated signaling pathway	9.15	0.88	5	12	TRIB1, CD55, LY96, LTF, TNFAIP3
chemokine metabolic process	9.15	0.88	5	12	IL1B, IL6, MYD88, TREM1, WNT5A
positive regulation of leukocyte apoptotic process	9.15	0.88	5	12	ADAM8, JAK3, LYN, CCL5, WNT5A
muscle cell migration	9.14	2.12	12	29	TRIB1, LPAR1, AIF1, IGFBP3, ITGA2, SERPINE1, PLAT, PLAU, RETN, CCL5, SORL1, NRP1
chronic inflammatory response	8.99	1.24	7	17	IL1RN, S100A8, S100A9, CCL5, THBS1, VCAM1, PTGES
embryo implantation	8.89	1.97	11	27	FBLN1, IL1B, LIF, MMP9, PRLR, PTGS2, SPP1, STC1, TGFBR2, TIMP1, VMP1
positive regulation of phagocytosis	8.61	1.82	10	25	DOCK2, FCER1G, PYCARD, ITGA2, PTX3, CCL2, NOD2, SLC11A1, C3, CD36
JAK-STAT cascade involved in growth hormone signaling pathway	8.58	1.46	8	20	JAK3, LYN, PRLR, PTPN6, STAT3, SOCS1, IRS2, SOCS3
myeloid leukocyte activation	8.54	2.72	15	39	AIF1, IL4R, CXCL8, CXCR2, MT1G, PRKCD, PREX1, CCL5, SLC7A2, SLC11A1, THBS1, WNT5A, CXCR4, ZC3H12A, IL1RL1
lipopolysaccharide-mediated signaling pathway	8.35	1.67	9	23	HCK, IL1B, IRAK1, LYN, MYD88, NFKBIA, CCL2, CCL5, TLR2
negative regulation of chemotaxis	8.18	1.31	7	18	AIF1, GREM1, PTPN2, ROBO1, CCL2, WNT5A, NRP1
positive regulation of T cell migration	8.18	1.31	7	18	ADAM8, AIF1, PYCARD, CCL5, CCL20, CCR2, WNT5A

Table 13 – Continued on next page

Table 13 – continued from previous page

positive regulation of release of cytochrome c from mitochondria	8.00	0.95	5	13	PPIF, PYCARD, ARRB2, BID, BNIP3
regulation of regulated secretory pathway	7.92	1.53	8	21	CD300A, FCER1G, FGR, LYN, STXBP2, SYK, CCR2, PRAM1
mast cell activation involved in immune response	7.73	1.75	9	24	CD300A, FCER1G, FGR, LYN, PIK3CD, PIK3CG, STXBP2, SYK, LAT2
cell chemotaxis	7.73	2.88	15	42	LPAR1, CXCL1, CXCL2, RHOG, ARRB2, BIN2, PDGFRA, PPBP, CCL8, CCL20, CXCL6, CXCL5, VCAM1, NRP1, DOCK4
substrate-dependent cell migration	7.70	1.17	6	16	ADAM8, ITGA2, PTPRC, ROBO1, NRP2, NRP1
negative regulation of toll-like receptor signaling pathway	7.70	1.17	6	16	IRAK3, PIK3AP1, LYN, ARRB2, NOD2, TNFAIP3
regulation of mast cell degranulation	7.70	1.17	6	16	CD300A, FCER1G, FGR, LYN, STXBP2, SYK
regulation of interferon-gamma-mediated signaling pathway	7.70	1.17	6	16	IFNGR1, IFNGR2, PTPN2, PTPN6, SOCS1, SOCS3
regulation of lymphocyte migration	7.59	1.97	10	27	ADAM8, AIF1, PYCARD, CCL2, CCL4, CCL5, CCL20, STK10, CCR2, WNT5A
positive regulation of inflammatory response	7.41	2.80	14	39	TNIP1, CTSS, ITGA2, SERPINE1, PLA2G2A, PTGER4, S100A8, S100A9, S100A12, TGM2, TLR2, CCR2, WNT5A, IL1RL1
mast cell mediated immunity	7.25	1.82	9	25	CD300A, FCER1G, FGR, LYN, PIK3CD, PIK3CG, STXBP2, SYK, LAT2
positive regulation of tumor necrosis factor production	7.25	1.82	9	25	FCER1G, PYCARD, MYD88, PF4, SASH3, NOD2, TLR2, CD14, CD36
leukocyte degranulation	7.22	2.03	10	28	CD300A, FCER1G, FGR, HCK, LYN, PIK3CD, PIK3CG, RAB27A, CCR2, LAT2
negative regulation of adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	7.11	1.02	5	14	PTPN6, PTPRC, C10orf54, NOD2, IL1RL1

Table 13 – Continued on next page

Table 13 – continued from previous page

response to ischemia	7.11	1.02	5	14	PPIF, RCAN1, PANX1, MAP3K5, CPEB4
T cell chemotaxis	7.11	1.02	5	14	PIK3CD, PIK3CG, CCL5, CCR2, WNT5A
positive regulation of lipid storage	7.11	1.02	5	14	HILPDA, NFKB1, C3, ZC3H12A, CD36
response to hyperoxia	7.11	1.02	5	14	CDKN1A, PDPN, NCF2, BNIP3, SOD2
positive regulation of smooth muscle cell proliferation	7.11	2.26	11	31	HBEGF, AIF1, IL6, IRAK1, ITGA2, MYD88, RETN, PTGS2, CCL5, TGFBR2, TGM2
positive regulation of epithelial cell migration	7.06	1.23	6	17	CTSH, HBEGF, HIF1A, ITGA2, MMP9, TGFBR2
superoxide anion generation	6.99	1.24	6	17	CYBA, CYBB, NCF2, PRKCD, SOD2, SYK
fibrinolysis	6.92	1.46	7	20	FAP, SERPINE1, SERPINB2, PLAT, PLAU, PLAUR, THBS1
positive regulation of nitric oxide biosynthetic process	6.92	1.46	7	20	ICAM1, IL1B, IL6, PTGS2, PTX3, SOD2, TLR2
regulation of mast cell activation	6.86	1.68	8	23	CD300A, FCER1G, FGR, LYN, PLSCR1, PTPRE, STXBP2, SYK
acute-phase response	6.47	1.74	8	24	CEBPB, SERPINA3, HP, IL6, SERPINA1, PLSCR1, STAT3, CD163
cellular response to biotic stimulus	6.46	6.93	31	95	TRIB1, LILRB2, CSF3, CD55, ABCA1, LY96, PYCARD, HCK, ICAM1, IL1B, IL6, CXCL8, IRAK1, LTF, LYN, MYD88, NFKB1, NFKBIA, SERPINE1, CCL2, CCL5, NOD2, SYK, TLR2, TNFAIP3, TNFRSF1B, WNT5A, ZC3H12A, CD14, CD86, CD36
regulation of cell adhesion mediated by integrin	6.44	1.97	9	27	CYP1B1, LYN, SERPINE1, PIK3CG, PLAU, PTPN6, CCL5, SYK, FERMT3
cellular response to growth hormone stimulus	6.43	1.75	8	24	JAK3, LYN, PRLR, PTPN6, STAT3, SOCS1, IRS2, SOCS3
positive regulation of tissue remodeling	6.41	1.31	6	18	ADAM8, B4GALT1, SPP1, SYK, TFRC, VDR
regulation of antigen processing and presentation	6.40	1.09	5	15	LILRB2, PYCARD, NOD2, SLC11A1, THBS1
positive regulation of blood coagulation	6.40	1.09	5	15	HPSE, F3, SERPINE1, THBS1, CD36

Table 13 – Continued on next page

Table 13 – continued from previous page

monocyte differentiation	6.40	1.09	5	15	CSF1R, IFI16, MT1G, MYC, VEGFA
negative regulation of bone mineralization	6.40	1.09	5	15	CCR1, GREM1, HIF1A, NFE2, SRGN
vascular endothelial growth factor signaling pathway	6.40	1.09	5	15	FLT1, PDGFRA, VEGFA, NRP2, NRP1
regulation of membrane protein ectodomain proteolysis	6.40	1.09	5	15	ADAM8, IL1B, TIMP1, TIMP3, TNFRSF1B
multicellular organismal homeostasis	6.26	2.22	10	31	AMPD2, AMPD3, METRNL, IL1A, IL1B, IL1RN, SLC11A2, PTGS2, SLC11A1, STAT3
positive regulation of leukocyte migration	6.22	4.07	18	56	CCR1, AIF1, PYCARD, IL6, CXCL8, ITGA2, LGALS3, SERPINE1, CCL2, CCL4, CCL5, CCL20, SELE, THBS1, TLR2, C3AR1, VEGFA, WNT5A
regulation of leukocyte chemotaxis	6.13	3.65	16	50	CCR1, AIF1, GREM1, IL6, CXCL8, LYN, SERPINE1, CCL2, CCL4, CCL5, NOD2, THBS1, C3AR1, CCR2, VEGFA, WNT5A
collagen fibril organization	6.10	2.04	9	28	COL1A1, COL3A1, COL5A2, COL11A1, CYP1B1, DPT, GREM1, LOX, ADAMTS2
cellular response to lipopolysaccharide	6.04	3.43	15	48	LILRB2, CSF3, ABCA1, PYCARD, ICAM1, IL6, CXCL8, NFKB1, SERPINE1, TNFRSF1B, WNT5A, ZC3H12A, CD14, CD86, CD36
brown fat cell differentiation	5.99	1.61	7	22	CEBPB, LRG1, METRNL, ERO1L, PLAC8, PTGS2, BNIP3
response to transition metal nanoparticle	5.99	3.94	17	54	ICAM1, IL1A, AQP9, MT1E, MT1F, MT1G, MT1H, MT1M, MT1X, MT2A, NFKB1, SLC11A2, PTGS2, S100A8, S100A9, SOD2, VCAM1
regulation of type 2 immune response	5.92	1.39	6	19	IL4R, IL6, BCL6, NOD2, CCR2, CD86
response to vitamin D	5.92	1.39	6	19	TNC, PTGS2, SPP1, STC1, VDR, TRIM25
positive regulation of interleukin-6 production	5.86	2.33	10	32	LILRB2, FCER1G, PYCARD, IL1B, IL6, MYD88, NOD2, TLR2, WNT5A, CD36
regulation of fibroblast growth factor receptor signaling pathway	5.82	1.17	5	16	SULF1, FAM20C, THBS1, WNT5A, RUNX2

Table 13 – Continued on next page

Table 13 – continued from previous page

regulation of alpha-beta T cell proliferation	5.82	1.17	5	16	CD55, IRF1, C10orf54, SYK, CCR2
detection of external biotic stimulus	5.82	1.17	5	16	LY96, NLRC4, NOD2, TLR2, CD1D
tumor necrosis factor superfamily cytokine production	5.78	4.52	19	62	ADAM8, IRAK3, FCER1G, PYCARD, LTF, ARRB2, MYD88, PF4, SASH3, BCL3, NOD2, THBS1, TLR2, TNFAIP3, CCR2, ZC3H12A, CD14, CD86, CD36
cellular response to interleukin-1	5.75	1.89	8	26	PYCARD, HAS2, HIF1A, ICAM1, CXCL8, NFKB1, CCL2, CCL5
integrin-mediated signaling pathway	5.73	4.30	18	59	COL3A1, FCER1G, FGR, HCK, ITGA2, ITGAM, ITGAX, ITGB2, ITGB8, NEDD9, PLEK, CEACAM1, SYK, TIMP1, TYROBP, FERMT3, PRAM1, ADAMTS1
positive regulation of T cell proliferation	5.73	4.30	18	59	CD24, LILRB2, TNFSF13B, CORO1A, CD55, AIF1, PYCARD, IL1B, IL6, IL6ST, SASH3, PTPRC, CCL5, SYK, CCR2, VCAM1, CD1D, CD86
positive regulation of cytokine secretion	5.65	3.12	13	43	FCN1, FGR, CLEC5A, PANX1, CLEC4E, IL1A, PTGER4, NOD2, SYK, TLR2, WNT5A, CD14, TNFSF15
regulation of lymphocyte proliferation	5.65	4.08	17	57	CD24, CORO1A, CD55, AIF1, PYCARD, IL1B, IL6, IL6ST, IRF1, LGALS3, MYD88, CCL5, SYK, CCR2, VCAM1, CD1D, CD86
negative regulation of B cell activation	5.62	1.68	7	23	CD300A, TNFRSF21, LYN, MNDA, BCL6, SAMSN1, TNFAIP3
regulation of lymphocyte apoptotic process	5.61	2.41	10	33	ADAM8, HIF1A, JAK3, LGALS3, LYN, BCL6, CCL5, WNT5A, DOCK8, IRS2
positive regulation of angiogenesis	5.57	4.14	17	57	CHI3L1, LRG1, ADM, CTSH, CYP1B1, F3, FLT1, HIF1A, IL1A, IL1B, SERPINE1, ANGPTL4, TGFBR2, THBS1, C3, C3AR1, WNT5A
I-kappaB kinase/NF-kappaB signaling	5.56	1.92	8	27	LY96, TLR8, PTGS2, BCL3, TLR2, IL18R1, TIFA, CD14
response to fungus	5.52	2.19	9	30	LTF, MYD88, PTX3, S100A8, S100A9, S100A12, SYK, TLR2, CD86

Table 13 – Continued on next page

Table 13 – continued from previous page

negative regulation of leukocyte apoptotic process	5.52	2.19	9	30	FCER1G, HCLS1, HIF1A, JAK3, BCL6, CCL5, NOD2, DOCK8, IRS2
interleukin-1 beta secretion	5.49	1.46	6	20	ABCA1, PANX1, PYCARD, NLRC4, NOD2, WNT5A
adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	5.47	1.70	7	24	IL4R, IL6, JAK3, NFKB2, RELB, TNFAIP3, IL18R1
cellular response to amino acid stimulus	5.41	1.97	8	27	CEBPB, COL1A1, COL3A1, COL5A2, CYBA, PDGFRA, CPEB4, SOCS1
negative regulation of lymphocyte mediated immunity	5.33	1.24	5	17	ARRB2, PTPN6, PTPRC, C10orf54, NOD2
regulation of T-helper 1 type immune response	5.33	1.24	5	17	IL4R, JAK3, SLC11A1, CCR2, IL1RL1
positive regulation of phosphatidylinositol 3-kinase activity	5.33	1.24	5	17	FGR, FLT1, LYN, PDGFRA, NOD2
positive regulation of CD4-positive, alpha-beta T cell activation	5.33	1.24	5	17	CD55, IL4R, IL6, SASH3, CD86
regulation of blood coagulation	5.30	3.26	13	45	PDPN, FAP, FCER1G, LYN, SERPINB2, PDGFRA, SERPINE2, PLAT, PLAU, PLAUR, PLEK, PRKCD, SYK
intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator	5.29	1.75	7	24	CDKN1A, PYCARD, IFI16, TAF9B, DDIT4, BCL3, CD44
negative regulation of immune response	5.20	4.09	16	56	OLFM4, IRAK3, COL3A1, CD55, IFI16, LYN, ARRB2, PTPN6, PTPRC, BCL6, SAMSN1, C10orf54, NOD2, TNFAIP3, CCR2, IL1RL1
response to interferon-gamma	5.18	6.20	24	85	IFITM2, AIF1, FCGR1B, GBP2, GCH1, HCK, ICAM1, IFNGR1, IFNGR2, IRF1, MT2A, PRKCD, PTPN2, PTPN6, CCL2, CCL5, SLC11A1, VCAM1, WNT5A, SOCS1, KYNU, SOCS3, CD86, CD44

Table 13 – Continued on next page

Table 13 – continued from previous page

regulation of bone remodeling	5.14	2.04	8	28	ADAM8, CSF1R, GREM1, TNFRSF11B, SPP1, SYK, TFRC, TNFAIP3
cytokine secretion	5.14	3.06	12	43	CHI3L1, ABCA1, LRRC32, TNFRSF21, IL6, LCP2, LYN, TLR8, TREM1, SRGN, NLRC4, S100A12
positive regulation of cytokine biosynthetic process	5.10	3.36	13	46	IL1A, IL1B, IRF1, LTB, MYD88, TLR8, BCL3, SYK, THBS1, TLR2, CCR2, WNT5A, CD86
myeloid cell differentiation	5.06	3.36	13	47	CSF1R, CSF3, SBNO2, SNX10, IFI16, JUNB, MMP9, MT1G, PIK3CD, RELB, TFRC, TGFBR2, CD86
alpha-beta T cell differentiation involved in immune response	5.04	2.33	9	32	IL4R, IL6, JAK3, PTGER4, RELB, BCL3, BCL6, IL18R1, CD86
phagosome maturation	5.04	2.33	9	32	TCIRG1, CORO1A, CYBA, CYBB, LTF, NCF2, NCF4, SLAMF8, SLC11A1
negative regulation of hemostasis	5.04	2.33	9	32	FAP, SERPINE1, SERPINB2, PDGFRA, SERPINE2, PLAT, PLAU, PLAUR, THBS1
leukocyte cell-cell adhesion	5.03	2.59	10	36	TNIP1, ICAM1, ITGAM, ITGB2, CCL5, SELE, SELPLG, SYK, VCAM1, FERMT3
positive regulation of protein kinase B signaling	5.00	3.14	12	43	ADAM8, HPSE, CHI3L1, CSF3, HBEGF, F3, HCLS1, IL6, ARRB2, PIK3CG, FAM110C, THBS1
regulation of response to cytokine stimulus	4.99	4.74	18	65	CD24, IRAK3, PYCARD, HIF1A, IFNAR2, IFNGR1, IFNGR2, IL1RN, IL6ST, IRAK1, PTPN2, PTPN6, PTPRC, ROBO1, CCL5, WNT5A, SOCS1, SOCS3
positive regulation of production of molecular mediator of immune response	4.99	1.82	7	25	FCER1G, SASH3, NOD2, TLR2, WNT5A, CD36, CD37
response to activity	4.99	1.82	7	25	METRNL, HIF1A, ITGA2, NCF2, CCL2, SOD2, SCO2

Table 13 – Continued on next page

Table 13 – continued from previous page

inflammatory response	4.99	12.63	47	191	PDPN, CHI3L1, CCR1, CSF1R, CYBA, CYBB, S1PR3, AIF1, KDM6B, STAB1, LY96, ACKR1, PYCARD, CXCL1, CXCL2, CXCL3, HIF1A, IFI16, CXCL8, CXCR2, ITGB2, LYZ, TLR8, PIK3CD, PTX3, NLRC4, PROK2, BDKRB2, CCL2, CCL4, CCL8, CCL20, CXCL6, NFKBIZ, CLEC7A, SLC11A1, SPP1, TNFAIP6, C3AR1, C5AR1, CXCR4, IL18RAP, PSTPIP1, NMI, CD14, THEMIS2, CD44
positive regulation of interleukin-1 beta production	4.92	1.31	5	18	PANX1, PYCARD, IFI16, NOD2, WNT5A
cellular metabolic compound salvage	4.92	1.31	5	18	TYMP, AMPD2, AMPD3, UPP1, PDXK
apoptotic cell clearance	4.92	1.31	5	18	CCL2, TGM2, THBS1, C3, CD36
regulation of actin cytoskeleton reorganization	4.92	1.31	5	18	CSF1R, CSF3, HCK, HCLS1, PDGFRA
regulation of endothelial cell apoptotic process	4.92	1.31	5	18	ICAM1, SERPINE1, ANGPTL4, THBS1, TNFAIP3
regulation of interleukin-10 production	4.90	2.12	8	29	FCER1G, TNFRSF21, PYCARD, JAK3, SASH3, BCL3, NOD2, TLR2
positive regulation of endocytosis	4.88	1.58	6	22	GREM1, ARRB2, SERPINE1, SELE, VEGFA, CD14
cellular extravasation	4.87	1.59	6	22	ITGAM, PTGER4, CCL2, SELE, SELPLG, VCAM1
regulation of alpha-beta T cell activation	4.86	2.39	9	33	IKZF1, CD300A, IL4R, IL6, JAK3, SASH3, BCL6, TGFBR2, CD86
cytokine production involved in immune response	4.86	2.39	9	33	CD55, FCER1G, TREM1, SASH3, NOD2, SLC11A1, TLR2, WNT5A, CD36
regulation of adaptive immune response	4.85	2.66	10	37	IL4R, IRF1, JAK3, PTPN6, PTPRC, BCL6, SAMSN1, C10orf54, TNFAIP3, IL1RL1
positive regulation of phosphatidylinositol 3-kinase signaling	4.83	2.41	9	33	SEMA4D, PIK3AP1, CSF3, FGR, FLT1, HCLS1, PDGFRA, PTPN6, CCL5
regulation of cell shape	4.81	4.60	17	63	SEMA4D, PDPN, CORO1A, FGD4, CSF1R, FGR, SH3KBP1, HCK, ICAM1, IL6, ITGB2, MYO10, CDC42SE1, CCL2, VEGFA, LST1, FGD3

Table 13 – Continued on next page

Table 13 – continued from previous page

protein import into nucleus, translocation	4.81	1.61	6	22	CDKN1A, IL6, NFKBIA, BCL3, BCL6, SLC11A1
patterning of blood vessels	4.81	1.61	6	22	FLT1, SPI1, TGFBR2, VEGFA, CXCR4, NRP1
positive regulation of mesenchymal cell proliferation	4.81	1.61	6	22	GAS1, MYC, TGFBR2, VEGFA, WNT5A, IRS2
negative regulation of I-kappaB kinase/NF-kappaB signaling	4.81	1.61	6	22	TNIP1, OLFM4, PYCARD, NFKBIA, TNFAIP3, IL1RL1
regulation of chemokine production	4.76	1.89	7	26	CSF1R, ACKR1, PYCARD, HIF1A, IL4R, TLR2, IL1RL1
positive regulation of alpha-beta T cell differentiation	4.73	1.90	7	26	IKZF1, IL4R, IL6, SASH3, SYK, TGFBR2, CD86
response to bacterium	4.71	11.19	40	160	CD24, TRIB1, LILRB2, CEBPB, IRAK3, ADM, CSF2RB, CSF3, CD55, ABCA1, LY96, GCH1, CXCL2, HCK, ICAM1, IL1B, IL1RN, CXCL8, IL10RA, IRAK1, LYN, MGST1, NCF2, NFKB1, NFKBIA, PTGER4, PTGS2, CCL2, CCL5, SELE, SOD2, TNFAIP3, TNFRSF1B, VCAM1, WNT5A, ZC3H12A, CD1D, CD14, CD86, PTGES
positive regulation of chemotaxis	4.71	4.67	17	64	CCR1, LPAR1, AIF1, F3, IL6, CXCL8, ITGA2, SERPINE1, CCL2, CCL4, CCL5, THBS1, C3AR1, CCR2, VEGFA, WNT5A, NRP1
toll-like receptor 4 signaling pathway	4.68	5.25	19	72	PIK3AP1, LY96, BIRC3, IRAK1, ITGAM, ITGB2, LTF, LYN, MYD88, NFKB1, NFKB2, NFKBIA, RPS6KA1, NOD2, TLR2, TNFAIP3, TRAF3, CD14, CD36
positive chemotaxis	4.67	2.19	8	30	CORO1A, F3, CXCL8, ITGA2, LGALS3, CCL5, VEGFA, NRP1
regulation of leukocyte proliferation	4.65	6.40	23	89	CD24, CDKN1A, TNFSF13B, CORO1A, CD55, AIF1, PYCARD, IL1B, IL6, IL6ST, IRF1, LGALS3, MYD88, SASH3, PTPRC, BCL6, CCL5, SYK, CCR2, VCAM1, IRS2, CD1D, CD86

Table 13 – Continued on next page

Table 13 – continued from previous page

negative regulation of NF-kappaB transcription factor activity	4.60	2.77	10	38	IRAK3, CYP1B1, PYCARD, IRAK1, ARRB2, NFKBIA, NOD2, TNFAIP3, TRAF3, ZC3H12A
positive regulation of NF-kappaB transcription factor activity	4.58	5.91	21	81	ADAM8, IRAK3, PYCARD, ICAM1, IL1B, IRAK1, LTF, MYD88, NFKB1, NFKB2, NFKBIA, NLRC4, S100A8, S100A9, S100A12, NOD2, TLR2, WNT5A, TRIM25, TRIM8, CFLAR
positive regulation of cellular component movement	4.57	13.94	48	191	ADAM8, SEMA4D, PDPN, CORO1A, CCR1, COL1A1, CSF1R, CTSH, HBEGF, LPAR1, AIF1, F3, FGR, FLT1, SRPX2, PYCARD, HAS2, HIF1A, ICAM1, IL6, CXCL8, ITGA2, LGALS3, LYN, MMP9, MYO1F, SERPINE1, PDGFRA, RETN, PTGS2, PTPRC, BCL6, CCL2, CCL4, CCL5, CCL20, SELE, FAM110C, TGFBR2, THBS1, TLR2, C3AR1, CCR2, VEGFA, WNT5A, IRS2, NRP2, NRP1
defense response to bacterium	4.57	6.20	22	86	FCER1G, STAB1, PYCARD, HP, LTF, LYZ, SERPINE1, PLAC8, CFP, TREM1, PPBP, PRKCD, SLAMF8, NLRC4, BCL3, S100A8, S100A9, S100A12, CCL20, CXCL6, SLC11A1, SYK
positive regulation of gliogenesis	4.52	1.68	6	23	IL6ST, LIF, LYN, SERPINE2, TLR2, CXCR4
regulation of tyrosine phosphorylation of Stat3 protein	4.52	1.68	6	23	CSF1R, IL6, IL6ST, LIF, PTPN2, SOCS1
response to interleukin-4	4.52	1.68	6	23	CORO1A, IL1RN, IL4R, JAK3, NFIL3, PTPN2
zymogen activation	4.51	5.11	18	70	CTSH, F3, PYCARD, IFI16, MAP3K5, MFI2, SLC11A2, SERPINE1, SERPINE2, PLAT, NLRC4, ROBO1, S100A8, S100A9, BID, THBS1, TNFRSF10B, TNFSF15
positive regulation of mononuclear cell proliferation	4.51	6.27	22	86	CD24, CDKN1A, LILRB2, TNFSF13B, CORO1A, CD55, AIF1, PYCARD, IL1B, IL6, IL6ST, MYD88, SASH3, PTPRC, BCL6, CCL5, SYK, CCR2, VCAM1, IRS2, CD1D, CD86

Table 13 – Continued on next page

Table 13 – continued from previous page

regulation of cellular response to insulin stimulus	4.49	1.97	7	27	IL1B, PID1, PRKCD, PTPN2, PTPRE, SOCS1, SOCS3
nucleotide-binding domain, leucine rich repeat containing receptor signaling pathway	4.45	2.55	9	35	PANX1, PYCARD, BIRC3, IRAK1, NFKBIA, NLRC4, NOD2, TNFAIP3, PSTPIP1
negative regulation of coagulation	4.45	2.55	9	35	FAP, SERPINE1, SERPINB2, PDGFRA, SERPINE2, PLAT, PLAU, PLAUR, THBS1
negative regulation of leukocyte proliferation	4.44	3.14	11	43	LILRB2, CD300A, LRRC32, GREM1, TNFRSF21, LYN, MNDA, PTPN6, C10orf54, TNFAIP3, LST1
regulation of innate immune response	4.43	13.94	47	191	ADAM8, TNIP1, IRAK3, PIK3AP1, CTSS, FCN1, FGR, LY96, SAMHD1, PYCARD, HCK, BIRC3, IFI16, IFNAR2, IFNGR1, IFNGR2, IRAK1, IRF1, ITGAM, ITGB2, LTF, LYN, ARRB2, MYO1F, MYD88, NFKB1, NFKB2, NFKBIA, TLR8, PLSCR1, PTPN2, PTPN6, NLRC4, RPS6KA1, CCL5, NOD2, CLEC7A, TLR2, TNFAIP3, TRAF3, WNT5A, SOCS1, SOCS3, CD1D, CD14, CD86, CD36
MyD88-dependent toll-like receptor signaling pathway	4.42	4.30	15	59	TNIP1, IRAK3, LY96, IRAK1, IRF1, MYD88, NFKB1, NFKB2, NFKBIA, TLR8, RPS6KA1, NOD2, TLR2, CD14, CD36
response to virus	4.36	4.93	17	69	IRAK3, CD55, ENO1, FGR, CLEC5A, ANPEP, ICAM1, IFNAR2, IFNGR1, IFNGR2, ITGA2, MYD88, BCL3, CCL4, CCL5, CCL8, CXCR4
regulation of vasculature development	4.35	8.24	28	113	CHI3L1, LRG1, ADM, CTSH, CYP1B1, F3, VASH1, STAB1, FLT1, SULF1, SRPX2, HIF1A, IL1A, IL1B, IL6, SERPINE1, ANGPTL4, PF4, PTGS2, CCL2, TGFBR2, THBS1, TNFAIP3, C3, C3AR1, CCR2, VEGFA, WNT5A
regulation of tumor necrosis factor production	4.34	2.30	8	32	IRAK3, LTF, ARRB2, BCL3, THBS1, TNFAIP3, CCR2, ZC3H12A

Table 13 – Continued on next page

Table 13 – continued from previous page

positive regulation of endothelial cell proliferation	4.29	2.63	9	36	LRG1, CYBA, F3, HIF1A, CCL2, VEGFA, WNT5A, NRP2, NRP1
sprouting angiogenesis	4.28	2.33	8	32	FLT1, GREM1, SRPX2, PTGS2, ROBO1, THBS1, VEGFA, NRP1
response to lipopolysaccharide	4.28	5.88	20	84	CEBPB, IRAK3, ADM, CSF2RB, GCH1, IL1RN, IL10RA, MGST1, NCF2, PTGER4, PTGS2, S100A8, S100A9, SELE, NOD2, SLC11A1, SOD2, C5AR1, VCAM1, PTGES
regulation of CD4-positive, alpha-beta T cell differentiation	4.27	1.75	6	24	IL4R, IL6, JAK3, SASH3, BCL6, CD86
foam cell differentiation	4.27	1.75	6	24	ABCA1, NFKB1, NFKBIA, PF4, PLA2G2A, CD36
temperature homeostasis	4.26	1.46	5	20	IL1A, IL1B, IL1RN, PTGS2, STAT3
B cell homeostasis	4.26	1.46	5	20	TNFSF13B, HIF1A, LYN, PIK3CD, TNFAIP3
positive regulation of fat cell differentiation	4.26	1.46	5	20	CEBPB, METRNL, PTGS2, ZC3H12A, MEDAG
negative regulation of carbohydrate metabolic process	4.26	1.46	5	20	IL6, PLEK, DDIT4, STAT3, IER3
nucleotide-binding oligomerization domain containing signaling pathway	4.26	1.46	5	20	BIRC3, IRAK1, NFKBIA, NOD2, TNFAIP3
pattern recognition receptor signaling pathway	4.25	8.39	28	115	TNIP1, IRAK3, PIK3AP1, CTSS, FCN1, LY96, BIRC3, IRAK1, IRF1, ITGAM, ITGB2, LTF, LYN, ARRB2, MYD88, NFKB1, NFKB2, NFKBIA, TLR8, RPS6KA1, NOD2, CLEC7A, TLR2, TNFAIP3, TRAF3, CD14, CD86, CD36

Table 13 – Continued on next page

Table 13 – continued from previous page

regulation of cell motility	4.23	17.01	55	238	ADAM8, SEMA4D, PDPN, CORO1A, CCR1, COL1A1, CSF1R, CTSH, HBEGF, LPAR1, F3, SPATA13, FGR, FLT1, SRPX2, PYCARD, HAS2, HIF1A, ICAM1, IL6, CXCL8, ITGA2, LAMA2, LGALS3, LYN, MMP9, MMP10, MYO1F, PDGFRA, SERPINE2, PLAU, RETN, PTGS2, PTPRC, RAC2, CCL4, CCL5, CCL20, SELE, NOD2, FAM110C, SGK1, SORL1, STK10, TGFBR2, TLR2, C3AR1, CCR2, VEGFA, WNT5A, CXCR4, SH3BGRL3, IRS2, NRP2, NRP1
positive regulation of defense response	4.23	9.62	32	137	IRAK3, PIK3AP1, FCN1, LY96, PYCARD, HCK, BIRC3, IFI16, IRAK1, IRF1, ITGAM, ITGB2, LTF, LYN, ARRB2, MYD88, NFKB1, NFKB2, NFKBIA, TLR8, PLSCR1, NLRC4, RPS6KA1, CCL5, NOD2, CLEC7A, TNFAIP3, TRAF3, CD1D, CD14, CD86, CD36
adaptive immune response	4.21	1.46	5	21	CTSS, TNFRSF21, PIK3CD, PIK3CG, SYK
homotypic cell-cell adhesion	4.20	2.66	9	37	LYN, PDGFRA, SERPINE2, PIK3CG, PLEK, PRKCD, PTPN6, SYK, FERMT3
cellular defense response	4.17	3.28	11	45	LILRB2, TCIRG1, CLEC5A, LY96, CXCR2, LSP1, MNDA, NCF2, C5AR1, CCR2, TYROBP
defense response to Gram-positive bacterium	4.15	2.99	10	41	FGR, IL6, MYO1F, MYD88, PLA2G2A, PLD1, NOD2, TLR2, C5AR1, CD36
cellular response to hydrogen peroxide	4.13	2.70	9	37	PPIF, CYP1B1, KDM6B, GPX3, HP, IL6, MAP3K5, BNIP3, TNFAIP3
organ regeneration	4.11	2.41	8	33	CDKN1A, ADM, LCP1, NNMT, CCL2, TGFBR2, C5AR1, UCP2
positive regulation of interferon-gamma production	4.11	2.41	8	33	PYCARD, IL1B, SASH3, BCL3, SLC11A1, CCR2, WNT5A, IL18R1
myeloid cell development	4.11	2.41	8	33	ITGAM, LTF, SLC11A2, FAM20C, PTPN6, BCL6, TYROBP, PTBP3
T cell proliferation	4.08	3.63	12	51	DOCK2, LRRC32, TNFRSF21, IRF1, ITGAM, LGALS3, PIK3CG, PTPN6, C10orf54, SLC11A1, CXCR4, DOCK8

Table 13 – Continued on next page

Table 13 – continued from previous page

heterotypic cell-cell adhesion	4.08	2.12	7	29	LILRB2, IL1B, IL1RN, ITGAX, ITGB2, CD1D, CD44
NIK/NF-kappaB signaling	4.08	2.12	7	29	CHI3L1, BIRC3, IRAK1, NFKB2, RELB, TNFRSF10B, TNFSF15
positive regulation of osteoblast differentiation	4.08	2.12	7	29	CEBPB, CTHRC1, IL6, IL6ST, LTF, FAM20C, RUNX2
mammary gland morphogenesis	4.08	2.12	7	29	CSF1R, B4GALT1, NR3C1, ROBO1, TGFBR2, VDR, WNT5A
regulation of response to wounding	4.07	4.84	16	71	PDPN, HPSE, HBEGF, F3, FAP, SERPINB2, PDGFRA, SERPINE2, PLAT, PLAU, PLAUR, PLEK, SYK, TGFBR2, THBS1, CD36
cellular response to mechanical stimulus	4.05	3.36	11	46	COL1A1, CYBA, IL1B, IRF1, MYD88, NFKB1, TLR8, PTGER4, PTGS2, BNIP3, TNFRSF10B
muscle cell proliferation	4.05	4.60	15	63	TRIB1, HBEGF, AIF1, IGFBP3, IL6, IRAK1, ITGA2, MYD88, RETN, PTGS2, CCL5, TGFBR2, TGM2, TNFAIP3, CFLAR
negative regulation of multi-organism process	4.05	4.60	15	63	TNIP1, IFITM2, IFI16, ISG20, LTF, MYD88, PLSCR1, PTX3, CCL4, CCL5, NOD2, TLR2, TRIM25, TRIM8, CD36
intrinsic apoptotic signaling pathway in response to oxidative stress	4.05	1.82	6	25	CYP1B1, HIF1A, MCL1, MAP3K5, PRKCD, SOD2
regulation of NF-kappaB import into nucleus	4.05	1.82	6	25	IL1B, NFKBIA, PTGS2, BCL3, TLR2, IL18R1
hydrogen peroxide metabolic process	4.05	1.82	6	25	CYBA, CYBB, GPX3, HP, RAC2, SOD2
positive regulation of transcription factor import into nucleus	4.05	1.82	6	25	CSF3, HCLS1, IL1B, PTGS2, TLR2, IL18R1
regulation of blood vessel endothelial cell migration	4.05	1.82	6	25	VASH1, SRPX2, PTGS2, THBS1, VEGFA, SH3BGRL3
negative regulation of viral genome replication	4.05	1.82	6	25	TNIP1, IFITM2, IFI16, ISG20, PLSCR1, CCL5
interferon-gamma-mediated signaling pathway	4.03	2.75	9	38	FCGR1B, GBP2, HCK, ICAM1, IRF1, MT2A, PRKCD, VCAM1, CD44

Table 13 – Continued on next page

Table 13 – continued from previous page

T cell activation involved in immune response	4.03	4.30	14	59	ICAM1, IL4R, IL6, JAK3, LCP1, LGALS3, PTGER4, RAB27A, RELB, BCL3, BCL6, SLC11A1, IL18R1, CD86
ruffle organization	4.00	1.53	5	21	CSF1R, AIF1, ICAM1, RHOG, PLEK
response to progesterone	4.00	1.53	5	21	JUNB, NCF2, CCL2, THBS1, TLR2
positive regulation of oxidoreductase activity	4.00	1.53	5	21	GCH1, HIF1A, IL1B, NOD2, VDR
negative regulation of signal transduction in absence of ligand	4.00	1.53	5	21	GAS1, IL1A, IL1B, MCL1, PF4
negative regulation of extrinsic apoptotic signaling pathway in absence of ligand	4.00	1.53	5	21	GAS1, IL1A, IL1B, MCL1, PF4
negative regulation of leukocyte activation	4.00	5.91	19	81	LILRB2, CD300A, LRRC32, TNFRSF21, IL4R, IRF1, JAK3, LGALS3, LYN, MNDA, PTPN6, BCL6, SAMSN1, C10orf54, TNFAIP3, CCR2, LST1, ZC3H12A, CD37
humoral immune response mediated by circulating immunoglobulin	3.95	2.48	8	34	CD55, C1RL, PTPN6, BCL3, NOD2, C1R, C1S, C3
membrane protein proteolysis	3.95	2.48	8	34	ADAM8, CTSH, IL1B, NFKB1, TIMP1, TIMP3, TNFRSF1B, SPPL2A
B cell receptor signaling pathway	3.95	2.48	8	34	CD300A, LYN, MNDA, PIK3CD, PTPN6, PTPRC, SYK, LAT2
toll-like receptor TLR1:TLR2 signaling pathway	3.94	3.43	11	47	LY96, IRAK1, MYD88, NFKB1, NFKB2, NFKBIA, RPS6KA1, NOD2, TLR2, CD14, CD36
toll-like receptor TLR6:TLR2 signaling pathway	3.94	3.43	11	47	LY96, IRAK1, MYD88, NFKB1, NFKB2, NFKBIA, RPS6KA1, NOD2, TLR2, CD14, CD36
regulation of cell adhesion	3.94	1.54	5	22	CYTH4, ICAM1, CXCL8, LAMA2, CYTIP
tyrosine phosphorylation of STAT protein	3.93	2.18	7	30	CSF1R, HCLS1, IL6, JAK3, LYN, CCL5, SOCS1
negative regulation of immune effector process	3.92	3.12	10	43	IRAK3, CD300A, CD55, LGALS3, ARRB2, PTPN6, PTPRC, C10orf54, NOD2, CCR2

Table 13 – Continued on next page

Table 13 – continued from previous page

cellular response to cytokine stimulus	3.91	2.47	8	37	CORO1A, DPYSL3, AIF1, HCLS1, MME, MT1X, NFIL3, CD86
positive regulation of endothelial cell migration	3.90	2.19	7	30	SRPX2, PTGS2, THBS1, VEGFA, WNT5A, NRP2, NRP1
positive regulation of reactive oxygen species metabolic process	3.90	2.19	7	30	CDKN1A, PID1, PRKCD, TGFBR2, THBS1, ZC3H12A, CD36
regulation of ERK1 and ERK2 cascade	3.90	1.87	6	26	TNIP1, LIF, LYN, PTPN2, PTPN6, SYK
cellular response to inorganic substance	3.89	4.74	15	65	PPIF, SPIDR, ALOX5AP, JUNB, MT1E, MT1F, MT1G, MT1H, MT1M, MT1X, MT2A, SLC11A2, BNIP3, WNT5A, CD86
positive regulation of locomotion	3.88	9.63	30	135	ADAM8, SEMA4D, PDPN, CORO1A, COL1A1, CSF1R, CTSH, HBEGF, FGR, FLT1, SRPX2, PYCARD, HAS2, HIF1A, ICAM1, LGALS3, LYN, MMP9, MYO1F, PDGFRA, RETN, PTGS2, PTPRC, CCL20, SELE, FAM110C, TGFBR2, TLR2, IRS2, NRP2
negative regulation of extrinsic apoptotic signaling pathway	3.87	3.79	12	52	GAS1, ICAM1, IL1A, IL1B, LGALS3, MCL1, SERPINE1, PF4, THBS1, TNFAIP3, NRP1, CFLAR
regulation of behavior	3.86	8.03	25	110	CCR1, LPAR1, AIF1, F3, GREM1, IL6, CXCL8, ITGA2, LYN, SERPINE1, PDGFRA, RETN, PTPN2, ROBO1, CCL2, CCL4, CCL5, NOD2, THBS1, C3AR1, CCR2, VEGFA, WNT5A, CXCR4, NRP1
T-helper cell differentiation	3.86	1.89	6	26	IL6, PTGER4, RELB, BCL3, IL18R1, CD86
toll-like receptor 2 signaling pathway	3.85	3.49	11	48	PIK3AP1, LY96, IRAK1, MYD88, NFKB1, NFKB2, NFKBIA, RPS6KA1, TLR2, CD14, CD36
platelet aggregation	3.79	1.60	5	22	PDGFRA, PIK3CG, PLEK, PTPN6, FERMT3
regulation of B cell proliferation	3.79	3.21	10	44	CDKN1A, TNFSF13B, CD300A, TNFRSF21, LYN, MNDA, SASH3, PTPRC, BCL6, IRS2
positive regulation of JAK-STAT cascade	3.79	3.21	10	44	CD300A, CSF1R, CYP1B1, HCLS1, IL6, IL6ST, LIF, LYN, PRLR, CCL5

Table 13 – Continued on next page

Table 13 – continued from previous page

cellular response to tumor necrosis factor	3.78	5.18	16	71	CHI3L1, CYBA, TNFRSF21, PYCARD, HAS2, ICAM1, CXCL8, SLC11A2, PID1, PTPN2, CCL2, CCL5, THBS1, TNFRSF1B, TRAF3, VCAM1
regulation of T cell activation	3.76	4.85	15	68	CD24, TNFSF13B, CORO1A, MAP3K8, AIF1, PYCARD, IL1B, IL6ST, LYN, PTPRC, CCL2, CCL5, NOD2, VCAM1, CD1D
response to cold	3.76	1.61	5	22	ADM, METRNL, IL6, NFKBIA, PLAC8
negative regulation of cell-substrate adhesion	3.76	1.61	5	22	COL1A1, MFI2, SERPINE1, BCL6, THBS1
negative regulation of tumor necrosis factor production	3.76	1.61	5	22	IRAK3, ARRB2, NOD2, TNFAIP3, ZC3H12A
mammary gland duct morphogenesis	3.76	1.61	5	22	CSF1R, NR3C1, ROBO1, VDR, WNT5A
regulation of interleukin-1 production	3.74	2.26	7	31	S1PR3, PANX1, PYCARD, IFI16, ARRB2, NOD2, WNT5A
filopodium assembly	3.74	2.26	7	31	FGD4, DPYSL3, SPATA13, ARHGAP26, MYO10, PRKCD, FGD3
platelet degranulation	3.73	4.23	13	58	FCER1G, LYN, SERPINE1, PF4, SERPINA1, PLEK, PPBP, SRGN, SYK, THBS1, TIMP1, VEGFA, CD36
antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent	3.73	2.92	9	40	CYBA, CYBB, FCGR1B, NCF2, NCF4, PSMB9, TAP1, TAPBP, CD36
negative regulation of lymphocyte proliferation	3.73	2.92	9	40	LILRB2, CD300A, LRRC32, TNFRSF21, LYN, MNDA, PTPN6, C10orf54, LST1
pyruvate metabolic process	3.73	3.57	11	49	ENO1, HIF1A, HK2, HK3, PFKFB3, PFKFB4, PGK1, DDIT4, STAT3, IER3, SLC16A3
negative regulation of phosphorylation	3.73	6.55	20	91	PPIF, TRIB1, CDKN1A, SEMA4D, GREM1, PYCARD, BIRC3, IGFBP3, IL6, ERRFI1, DDIT4, PID1, DNAJC3, PTPRC, BDKRB2, SAMSN1, SOCS1, SOCS3, GMFG, MVP

Table 13 – Continued on next page

Table 13 – continued from previous page

leukocyte mediated immunity	3.71	12.77	38	175	CTSC, CD300A, CTSH, CD55, FCER1G, FGR, ICAM1, IL1B, IL6, JAK3, LYN, ARRB2, MYO1F, MYD88, C1RL, TLR8, PIK3CD, PIK3CG, TREM1, SASH3, PRKCD, PTPN6, PTPRC, RAB27A, BCL3, BCL6, C10orf54, NOD2, SLC11A1, STXBP2, SYK, C1R, C1S, C3, CCR2, LAT2, PRAM1, TNFSF13
regulation of reactive oxygen species metabolic process	3.69	1.95	6	27	CYP1B1, HP, BIRC3, RAC2, SYK, IER3
activation of cysteine-type endopeptidase activity involved in apoptotic process	3.69	3.94	12	54	CTSH, F3, PYCARD, MAP3K5, SLC11A2, NLRC4, ROBO1, S100A8, S100A9, BID, TNFRSF10B, TNFSF15
cell activation	3.68	17.78	52	263	CD300A, AMICA1, DOCK2, SBNO2, IL17RA, LRRC32, TNFRSF21, ICAM1, CXCL8, CXCR2, IRF1, ITGAM, JAK3, LCP1, LGALS3, MNDA, MT1G, MYO1F, NFKB2, TLR8, PIK3CD, PRLR, PTGER4, PREX1, PTPN2, PTPRE, RAB27A, RELB, BCL3, S100A12, SELPLG, SAMSN1, C10orf54, CLEC7A, SLC7A2, SLC11A1, SPI1, STXBP2, TLR2, TNFAIP3, TYROBP, LAT2, CXCR4, LST1, ZC3H12A, DOCK8, PRAM1, RUNX2, IL18R1, SKAP2, CD37, CD48
complement activation	3.67	2.63	8	36	CD55, FCN1, C1RL, CFP, CFB, C1R, C1S, C3
regulation of interleukin-8 production	3.67	2.63	8	36	FCN1, PYCARD, IL1B, SERPINE1, TLR8, BCL3, TLR2, WNT5A
regulation of oxidoreductase activity	3.67	2.63	8	36	GCH1, HIF1A, HP, IL1B, NFKB1, SLAMF8, NOD2, VDR
regulation of protein secretion	3.66	8.03	24	110	ADAM8, CSF1R, FCN1, FGR, CLEC5A, PANX1, LRRC32, CLEC4E, TNFRSF21, PYCARD, IL1A, IL4R, IL6, LYN, TLR8, SRGN, PTGER4, NOD2, SYK, TLR2, WNT5A, IL1RL1, CD14, TNFSF15

Table 13 – Continued on next page

Table 13 – continued from previous page

embryonic cranial skeleton morphogenesis	3.66	1.97	6	27	RDH10, GAS1, CHST11, PDGFRA, TGFBR2, RUNX2
alpha-beta T cell activation	3.63	1.65	5	23	DOCK2, PTGER4, RELB, BCL3, IL18R1
regulation of humoral immune response	3.59	2.33	7	32	CD55, CFP, PTPN6, CFB, NOD2, C3, CD37
response to fatty acid	3.59	2.33	7	32	ACSL1, TNC, PID1, PTGS2, TLR2, UCP2, DGAT2
response to cytokine	3.58	12.32	36	181	CCR1, COL3A1, CSF1R, CSF2RB, CSF3, CSF3R, DPYSL3, ACSL1, IL17RA, ACKR1, HCLS1, IL1A, IL1B, IL1R1, IL6, CXCR2, IL10RA, JUNB, MCL1, MME, MT1X, MYD88, NFKB1, PF4, PLSCR1, PRLR, SELPLG, STAT3, TIMP1, CCR2, IL1R2, CXCR4, IL18R1, IL1RL1, OSMR, PTGES
positive regulation of B cell activation	3.58	3.36	10	46	CDKN1A, TNFSF13B, IL6, SASH3, PTPRC, BCL6, NOD2, SYK, IRS2, TNFSF13
complement activation, classical pathway	3.55	1.68	5	23	CD55, C1RL, C1R, C1S, C3
developmental programmed cell death	3.55	1.68	5	23	B4GALT1, IL1A, IL1B, SPI1, VDR
regulation of interferon-beta production	3.55	1.68	5	23	PYCARD, IRF1, TLR8, TLR2, TRAF3
regulation of lipid kinase activity	3.55	1.68	5	23	FGR, FLT1, LYN, PDGFRA, NOD2
regulation of cellular response to oxidative stress	3.55	1.68	5	23	HIF1A, HP, MCL1, SOD2, CD36
negative regulation of T cell activation	3.54	3.72	11	51	LILRB2, CD300A, LRRC32, TNFRSF21, IL4R, IRF1, JAK3, LGALS3, PTPN6, BCL6, C10orf54
negative regulation of angiogenesis	3.54	2.70	8	37	VASH1, STAB1, SULF1, SERPINE1, PF4, CCL2, THBS1, CCR2
regulation of extrinsic apoptotic signaling pathway in absence of ligand	3.54	2.70	8	37	GAS1, IL1A, IL1B, JAK3, MCL1, PF4, TNFRSF10B, CFLAR
leukocyte migration	3.53	5.40	16	80	COL1A1, SIRPA, B4GALT1, HCK, ITGAX, MMP9, CD177, PTPN6, RAC2, SELL, SHC1, STK10, DOCK8, SLC16A3, CD44, CD48

Table 13 – Continued on next page

Table 13 – continued from previous page

regulation of inflammatory response	3.53	7.87	23	112	PIK3AP1, CD55, SBNO2, IL17RA, METRNL, HCK, BIRC3, IL1R1, LYN, MYD88, NFKB1, CFP, PRKCD, PTPN2, BCL6, CFB, CCL5, SELE, NOD2, SLC7A2, TNFAIP3, TNFRSF1B, IER3
chemokine-mediated signaling pathway	3.51	2.03	6	28	CCR1, ACKR1, CXCR2, CCL2, CCR2, CXCR4
regulation of immunoglobulin mediated immune response	3.49	2.04	6	28	FCER1G, PTPN6, BCL6, NOD2, C3, TNFSF13
positive regulation of B cell proliferation	3.49	2.04	6	28	CDKN1A, TNFSF13B, SASH3, PTPRC, BCL6, IRS2
regulation of extrinsic apoptotic signaling pathway via death domain receptors	3.49	2.04	6	28	ICAM1, LGALS3, SERPINE1, THBS1, TNFAIP3, CFLAR
lymphocyte homeostasis	3.45	2.41	7	33	TNFSF13B, CORO1A, HIF1A, JAK3, LYN, PIK3CD, TNFAIP3
regulation of cytokine secretion	3.43	2.06	6	29	LRRC32, TNFRSF21, IL6, LYN, TLR8, SRGN
regulation of cell-cell adhesion	3.43	4.52	13	62	ADAM8, HAS2, IL1B, IL1RN, LYN, MYO10, SERPINE2, PRKCD, CCL5, SYK, WNT5A, FERMT3, CD44
response to heat	3.42	2.77	8	38	HSPA6, IL1A, IL6, LYN, OSM, CCL2, THBS1, CD14
positive regulation of cell activation	3.41	14.37	40	197	CD24, ADAM8, CDKN1A, LILRB2, IKZF1, PDPN, TNFSF13B, CORO1A, MAP3K8, CD55, AIF1, FCER1G, FGR, PYCARD, IL1B, IL4R, IL6, IL6ST, LYN, MYD88, PLEK, SASH3, PTPN6, PTPRC, BCL6, RPS6KA1, CCL2, CCL5, NOD2, SYK, TGFBR2, THBS1, CCR2, VCAM1, WNT5A, IRS2, TNFSF13, CD1D, IL1RL1, CD86
Fc receptor signaling pathway	3.37	2.45	7	34	PPAPDC1A, FCGR1B, FCGR2A, FGR, HCK, MYO10, PRKCD
regulation of production of molecular mediator of immune response	3.37	2.10	6	29	IRAK3, IL6, JAK3, BCL6, C10orf54, TNFSF13

Table 13 – Continued on next page

Table 13 – continued from previous page

regulation of response to stress	3.36	11.44	32	173	IRAK3, DOCK2, FCN1, FGR, LY96, SAMHD1, IFI16, IFNAR2, IFNGR1, IFNGR2, IRAK1, IRF1, ITGAM, ITGB2, LTF, ARRB2, MYO1F, NFKB2, NFKBIA, TLR8, PLSCR1, PTPN6, NLRC4, RPS6KA1, CLEC7A, TRAF3, SOCS1, SOCS3, CD1D, CD14, CD86, CD37
cellular transition metal ion homeostasis	3.36	4.60	13	63	TCIRG1, SLC39A14, HIF1A, LTF, MFI2, MT2A, MYC, SLC11A2, S100A8, S100A9, SLC11A1, TFRC, SCO2
regulation of B cell mediated immunity	3.34	2.12	6	29	FCER1G, PTPN6, BCL6, NOD2, C3, TNFSF13
regulation of glial cell differentiation	3.34	2.12	6	29	TNFRSF21, IL6ST, LIF, SERPINE2, TLR2, CXCR4
positive regulation of tyrosine phosphorylation of STAT protein	3.32	2.48	7	34	CSF1R, HCLS1, IL6, IL6ST, LIF, LYN, CCL5
modification of morphology or physiology of other organism involved in symbiotic interaction	3.32	2.48	7	34	TNIP1, TLR8, TREM1, CCL4, CCL5, NOD2, TLR2
interaction with host	3.31	5.00	14	69	TCIRG1, TNIP1, IFITM2, CYBA, CYBB, NCF2, NCF4, PTX3, SLC11A1, TAP1, TRIM25, CXCR4, TRIM8, VAMP8
interleukin-2 production	3.31	2.85	8	39	IL1A, IL1B, SASH3, NOD2, SLC11A1, TNFAIP3, CCR2, CD86
epithelium migration	3.30	7.59	21	104	CTSH, CYP1B1, HBEGF, FAP, VASH1, GREM1, SRPX2, HIF1A, ITGA2, MMP9, PTGS2, ROBO1, S100P, STC1, TGFBR2, THBS1, VEGFA, WNT5A, SH3BGRL3, NRP2, NRP1
positive regulation of peptidyl-tyrosine phosphorylation	3.30	7.22	20	99	CD24, SEMA4D, CSF1R, CSF3, GREM1, HCLS1, ICAM1, IL6, IL6ST, LIF, LYN, ARRB2, OSM, CCL5, NOD2, SYK, VEGFA, NRP1, CD36, CD44
T cell mediated immunity	3.30	3.94	11	54	CTSC, CTSH, CD55, ICAM1, IL1B, SASH3, PTPRC, RAB27A, C10orf54, NOD2, SLC11A1

Table 13 – Continued on next page

Table 13 – continued from previous page

iron ion homeostasis	3.29	4.30	12	59	TCIRG1, HIF1A, LCN2, LTF, MFI2, MYC, SLC11A2, SLC25A37, SLC11A1, SOD2, TFRC, STEAP4
positive regulation of cell migration	3.29	5.37	15	77	SEMA4D, PDPN, CORO1A, COL1A1, CSF1R, F3, FGR, FLT1, HAS2, LYN, MYO1F, PDGFRA, PTPRC, FAM110C, IRS2
extracellular matrix organization	3.28	9.85	27	138	COMP, FBLN1, SULF1, FMOD, EGFL6, B4GALT1, ERO1L, HAS2, HAS3, TNC, ICAM1, ICAM3, ITGA2, ITGAM, ITGAX, ITGB2, ITGB8, LAMA2, LGALS3, NFKB2, TNFRSF11B, SERPINE1, PDGFRA, PLOD2, BCL3, THBS1, VCAM1
regulation of locomotion	3.26	11.74	32	171	TRIB1, IFITM2, CD300A, COL3A1, CYP1B1, DPYSL3, SPATA13, VASH1, SULF1, GREM1, IGFBP3, IL1RN, LAMA2, MMP10, SERPINE2, PLAU, PTGER4, PTPN2, PTX3, RAC2, ROBO1, NOD2, SGK1, SORL1, STAT3, STC1, STK10, TIMP1, TRIM25, CXCR4, TRIM8, SH3BGRL3
platelet activation	3.23	11.89	32	163	PDPN, COL1A1, COL3A1, FCER1G, GNA15, IL6, ITPR2, RHOG, LCP2, LYN, ARRB2, SERPINE1, PDGFRA, PF4, SERPINA1, SERPINE2, PIK3CG, PLEK, PLSCR1, APBB1IP, PPBP, SRGN, PRKCD, PTPN6, RAC2, SHC1, SYK, THBS1, TIMP1, VEGFA, FERMT3, CD36
glycolytic process	3.22	2.54	7	35	ENO1, HIF1A, HK2, HK3, PFKFB3, PFKFB4, PGK1
regulation of cell migration	3.21	9.59	26	140	TRIB1, CD300A, COL3A1, CYP1B1, DPYSL3, SPATA13, VASH1, SULF1, GREM1, IGFBP3, IL1RN, LAMA2, MMP10, SERPINE2, PLAU, PTGER4, RAC2, NOD2, SGK1, SORL1, STAT3, STC1, STK10, TIMP1, CXCR4, SH3BGRL3

Table 13 – Continued on next page

Table 13 – continued from previous page

positive regulation of adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains	3.21	3.28	9	45	TNFSF13B, FCER1G, IL1B, SASH3, NOD2, SLC11A1, C3, CCR2, TNFSF13
osteoclast differentiation	3.21	3.28	9	45	CCR1, CSF1R, SBNO2, SNX10, JUNB, LTF, FAM20C, TFRC, TYROBP
mononuclear cell proliferation	3.21	6.24	17	88	CD300A, DOCK2, LRRC32, TNFRSF21, IRF1, ITGAM, LGALS3, LYN, MNDA, PIK3CG, PRKCD, PTPN6, C10orf54, SLC11A1, CXCR4, LST1, DOCK8
regulation of interleukin-2 production	3.20	2.55	7	35	IL1A, IL1B, SASH3, NOD2, TNFAIP3, CCR2, CD86
response to metal ion	3.20	10.43	28	143	PPIF, ALOX5AP, ICAM1, IL1A, IL6, AQP9, JUNB, MT1E, MT1F, MT1G, MT1H, MT1M, MT1X, MT2A, NFKB1, SLC11A2, TNFRSF11B, PTGS2, S100A8, S100A9, BNIP3, SOD2, THBS1, VCAM1, WNT5A, CD14, CD86, PTGES
cellular response to drug	3.20	2.19	6	30	RNF149, IL1B, MT2A, MYC, SLAMF8, CD69
regulation of type I interferon-mediated signaling pathway	3.20	2.19	6	30	IFNAR2, PTPN2, PTPN6, WNT5A, SOCS1, SOCS3
positive regulation of protein transport	3.19	9.70	26	133	ADAM8, CSF1R, CSF3, FCN1, FGR, CLEC5A, PANX1, CLEC4E, PYCARD, HCLS1, IL1A, IL1B, IL4R, IL6, PTGER4, PTGS2, NOD2, SORL1, SYK, TLR2, WNT5A, IL18R1, RAB29, IL1RL1, CD14, TNFSF15
response to oxygen levels	3.18	12.04	32	165	CD24, ADAM8, CDKN1A, PDPN, ADM, LPAR1, HIF1A, ICAM1, IRAK1, ITGA2, ITPR2, NCF2, SLC11A2, ANGPTL4, PLAT, PLAU, PLOD2, DDIT4, AGTRAP, BACH1, PTGS2, CCL2, BNIP3, SOD2, STC1, THBS1, TLR2, UCP2, VCAM1, VEGFA, CXCR4, CPEB4

Table 13 – Continued on next page

Table 13 – continued from previous page

TRIF-dependent toll-like receptor signaling pathway	3.14	3.72	10	51	LY96, BIRC3, IRAK1, NFKB1, NFKB2, NFKBIA, RPS6KA1, NOD2, TRAF3, CD14
negative regulation of protein processing	3.12	3.36	9	46	CSTA, CD55, GAS1, SERPINE1, SERPINE2, PLAT, THBS1, TIMP1, TIMP3
acute inflammatory response	3.12	2.97	8	42	CD55, F3, B4GALT1, CFP, S100A8, CFB, VCAM1, PTGES
negative regulation of cell motility	3.12	6.79	18	93	TRIB1, CD300A, COL3A1, CYP1B1, DPYSL3, AIF1, VASH1, SULF1, GREM1, IGFBP3, IL1RN, SERPINE1, PTGER4, CCL2, STAT3, STC1, THBS1, TIMP1
response to temperature stimulus	3.10	2.24	6	31	ADM, METRNL, ERO1L, ARRB2, NFKBIA, PLAC8
single organism cell adhesion	3.07	17.22	44	236	CD24, ADAM8, LILRB2, TNIP1, SEMA4D, OLFM4, PDPN, COL11A1, CSTA, CYP1B1, SRPX2, HAS2, ICAM1, ICAM3, IL1B, IL1RN, ITGAM, ITGAX, ITGB2, LYN, MFI2, MYO10, PDGFRA, SERPINE2, PIK3CG, PLEK, PRKCD, PREX1, PTPN6, RAC2, S100A8, S100A9, CCL5, SELE, SELPLG, SHC1, STK10, SYK, VCAM1, WNT5A, VMP1, FERMT3, CD1D, CD44
cytokine-mediated signaling pathway	3.07	10.35	27	148	CD24, CCR1, CSF1R, CSF2RB, CSF3, CSF3R, ACSL1, IL17RA, ACKR1, PYCARD, HIF1A, IL1A, IL1B, CXCR2, IL10RA, MYD88, PF4, PRLR, CCL2, TNFRSF1B, TRAF3, CCR2, IL1R2, CXCR4, IL18R1, IL1RL1, OSMR
cellular carbohydrate catabolic process	3.07	2.26	6	31	GK, HIF1A, DDIT4, STAT3, IER3, MGAM
platelet-derived growth factor receptor signaling pathway	3.07	2.26	6	31	F3, PDGFRA, PLAT, PTPN2, VEGFA, NRP1
negative regulation of cell adhesion	3.07	5.33	14	73	SEMA4D, COL1A1, CYP1B1, TNC, IL1RN, ARHGDIB, MFI2, MYO1F, SERPINE1, SERPINE2, PRKCD, PTPRC, BCL6, THBS1
regulation of peptidyl-serine phosphorylation	3.06	3.79	10	52	CSF3, HCLS1, IL6, LIF, OSM, DDIT4, BDKRB2, VEGFA, WNT5A, CD44

Table 13 – Continued on next page

Table 13 – continued from previous page

toll-like receptor 9 signaling pathway	3.04	3.43	9	47	PIK3AP1, IRAK1, MYD88, NFKB1, NFKB2, NFKBIA, TLR8, RPS6KA1, NOD2
response to mechanical stimulus	3.04	4.97	13	69	CHI3L1, COL3A1, COL11A1, RCAN1, TNC, IL6, ITGA2, JUNB, SERPINE2, RETN, CCL2, TGFBR2, THBS1
negative regulation of phosphorus metabolic process	3.03	13.72	35	188	PPIF, TRIB1, CDKN1A, TNIP1, SEMA4D, IRAK3, CD300A, GREM1, RNF149, PYCARD, BIRC3, IGFBP3, IL1B, IL6, LIF, LYN, MYC, PLEK, P2RY13, ERRFI1, DDIT4, PID1, PRKCD, DNAJC3, PTPN2, PTPN6, PTPRC, BDKRB2, SAMSN1, SORL1, CCR2, SOCS1, SOCS3, GMFG, MVP
regulation of type I interferon production	3.03	4.60	12	63	PYCARD, IFI16, IRF1, MYD88, NFKB1, NFKB2, NFKBIA, TLR8, TLR2, TNFAIP3, TRAF3, TRIM25
negative regulation of intrinsic apoptotic signaling pathway	3.02	3.06	8	42	PPIF, HIF1A, MCL1, TAF9B, BCL2A1, BDKRB2, SOD2, CD44
B cell mediated immunity	3.02	5.40	14	74	CD55, FCER1G, MYD88, C1RL, TLR8, PRKCD, PTPN6, BCL3, BCL6, NOD2, C1R, C1S, C3, TNFSF13
toll-like receptor 3 signaling pathway	3.01	4.23	11	58	LY96, BIRC3, IRAK1, NFKB1, NFKB2, NFKBIA, RPS6KA1, NOD2, TNFAIP3, TRAF3, CD14
negative regulation of protein kinase activity	3.00	7.00	18	96	TRIB1, CDKN1A, IRAK3, CD300A, PYCARD, IL1B, IL6, LYN, ERRFI1, PRKCD, DNAJC3, PTPN6, PTPRC, SORL1, SOCS1, SOCS3, GMFG, MVP
type I interferon signaling pathway	2.99	3.87	10	53	IFITM2, GBP2, IFNAR2, IRF1, ISG20, PTPN2, PTPN6, WNT5A, SOCS1, SOCS3
unsaturated fatty acid biosynthetic process	2.99	2.70	7	37	ALOX5, ALOX5AP, FADS3, ELOVL2, PTGS2, SYK, PTGES
negative regulation of MAPK cascade	2.97	4.67	12	64	TNIP1, IRAK3, CD300A, RNF149, IL1B, LIF, LYN, MYC, PRKCD, PTPN2, PTPN6, SORL1
positive regulation of immune effector process	2.97	5.06	13	70	CD300A, FGR, IL1B, IL4R, IL6, LYN, CCL2, STXBP2, SYK, C3, CCR2, TNFSF13, CD86

Table 13 – Continued on next page

Table 13 – continued from previous page

cytokine metabolic process	2.97	5.07	13	70	CEBPB, IL1A, IRF1, LTB, NFKB1, TLR8, BCL3, SYK, THBS1, TLR2, CCR2, CD86, TNFSF15
positive regulation of cysteine-type endopeptidase activity	2.97	5.47	14	75	CTSH, F3, PYCARD, IFI16, MAP3K5, MYC, SLC11A2, NLRC4, ROBO1, S100A8, S100A9, BID, TNFRSF10B, TNFSF15
Fc-gamma receptor signaling pathway involved in phagocytosis	2.96	3.50	9	48	PPAPDC1A, FCGR2A, FGR, HCK, ITPR2, LYN, MYO10, PRKCD, SYK
regulation of vesicle-mediated transport	2.96	12.33	31	169	CD300A, DOCK2, FCER1G, FGR, GAS1, GREM1, B4GALT1, PYCARD, HCK, ITGA2, LGALS3, LYN, ARRB2, SERPINE1, PTX3, RAB27A, CCL2, SELE, NOD2, SLC11A1, SORL1, STXBP2, SYK, C3, CCR2, VEGFA, PRAM1, VAMP8, NRP1, CD14, CD36
negative regulation of hormone secretion	2.95	2.33	6	32	IL1B, IL6, LIF, PIM3, OSM, UCP2
regulation of antigen receptor-mediated signaling pathway	2.95	2.33	6	32	CD300A, LGALS3, LYN, PTPN2, PTPN6, PTPRC
cellular iron ion homeostasis	2.93	3.14	8	43	TCIRG1, HIF1A, LTF, MFI2, MYC, SLC11A2, SLC11A1, TFRC
positive regulation of phospholipase C activity	2.93	3.14	8	43	LPAR1, FLT1, GNA15, ITPR2, PDGFRA, PRKCD, SELE, C5AR1
fatty acid transport	2.93	3.14	8	43	PLIN2, ACSL1, IL1B, BDKRB2, SYK, THBS1, IRS2, CD36
peptidyl-tyrosine modification	2.93	14.96	37	205	CD24, SEMA4D, CAMKK2, CD300A, CSF1R, CSF3, FGR, FLT1, GREM1, HCK, HCLS1, ICAM1, IL6, IL6ST, ITGB2, JAK3, LIF, LYN, ARRB2, OSM, PDGFRA, ERRFI1, PRKCD, PRLR, PTPN2, PTPN6, CCL5, SAMSN1, NOD2, SHC1, SYK, VEGFA, SOCS1, NRP1, CD36, CD44, MVP
response to type I interferon	2.92	3.94	10	54	IFITM2, GBP2, IFNAR2, IRF1, ISG20, PTPN2, PTPN6, WNT5A, SOCS1, SOCS3

Table 13 – Continued on next page

Table 13 – continued from previous page

regulation of leukocyte activation	2.90	12.50	31	177	CD24, ADAM8, CDKN1A, IKZF1, TNFSF13B, CORO1A, MAP3K8, CD55, AIF1, PYCARD, IL1B, IL6, IL6ST, MYD88, SASH3, PTPN2, PTPRC, CCL2, CCL5, NOD2, SLC7A2, TGFBR2, THBS1, VCAM1, WNT5A, PRAM1, IRS2, TNFSF13, CD1D, IL1RL1, CD86
leukocyte activation involved in immune response	2.90	5.95	15	84	ICAM1, IL4R, IL6, JAK3, LCP1, LGALS3, PTGER4, RAB27A, RELB, BCL3, BCL6, SLC11A1, TNFSF13, IL18R1, CD86
immune response-activating cell surface receptor signaling pathway	2.90	11.31	28	155	CD300A, PPAPDC1A, FCER1G, FCGR2A, FCN1, FGR, TNFRSF21, HCK, ITPR2, LCP2, LGALS3, LYN, MNDA, MYO10, NFKB1, NFKBIA, PIK3CD, PLSCR1, PRKCD, PTPN2, PTPN6, PTPRC, SYK, TLR2, C3AR1, C5AR1, LAT2, THEMIS2
myeloid cell homeostasis	2.89	4.38	11	60	FCER1G, HCLS1, HIF1A, IL6, LYN, SLC11A2, PTPN2, BCL6, SPI1, VEGFA, PTBP3
erythrocyte differentiation	2.89	3.57	9	49	HCLS1, HIF1A, LYN, SLC11A2, PTPN2, BCL6, SPI1, VEGFA, PTBP3
endothelial cell migration	2.88	3.18	8	44	CYP1B1, FAP, VASH1, GREM1, ROBO1, S100P, STC1, SH3BGRL3
activation of immune response	2.87	8.85	22	128	CD300A, PPAPDC1A, FCER1G, FCGR2A, FGR, TNFRSF21, ITPR2, LCP2, LGALS3, MNDA, MYO10, PIK3CD, PLSCR1, PRKCD, PTPN2, PTPN6, PTPRC, SYK, C3AR1, C5AR1, LAT2, THEMIS2

Table 13 – Continued on next page

Table 13 – continued from previous page

coagulation	2.87	21.67	52	300	PDPN, HPSE, AMICA1, COL1A1, COL3A1, SIRPA, F3, FCER1G, FGR, MAFF, GNA15, IL6, IRF1, ITGA2, ITGAM, ITGAX, ITGB2, ITPR2, RHOG, LCP2, LYN, ARRB2, NFE2, PF4, SERPINA1, PIK3CG, PLEK, PLSCR1, TREM1, APBB1IP, PPBP, SRGN, PRKCD, CD177, PTPN6, RAB27A, RAC2, SELE, SELL, SELPLG, SHC1, SYK, TIMP1, VEGFA, TFPI2, DOCK8, FERMT3, PAPSS2, SLC16A3, CD36, CD44, CD48
calcium-mediated signaling	2.86	4.82	12	66	CAMKK2, RCAN1, LMCD1, CXCL8, PLEK, MCTP2, CCL4, SELE, SYK, LAT2, CXCR4, MCTP1
positive regulation of peptidyl-serine phosphorylation	2.85	3.21	8	44	CSF3, HCLS1, IL6, LIF, OSM, VEGFA, WNT5A, CD44
regulation of phospholipase C activity	2.85	3.21	8	44	LPAR1, FLT1, GNA15, ITPR2, PDGFRA, PRKCD, SELE, C5AR1
response to nutrient	2.85	7.30	18	100	COL1A1, CYBB, ABCA1, ACSL1, TNC, ITGA2, TNFRSF11B, PTGS2, CCL2, NOD2, SOD2, SPP1, STC1, TGFBR2, VCAM1, VDR, TRIM25, KYNU
response to hypoxia	2.83	11.09	27	152	CD24, ADAM8, ADM, HIF1A, ICAM1, IRAK1, ITGA2, ITPR2, SLC11A2, ANGPTL4, PLAT, PLAU, PLOD2, DDIT4, AGTRAP, BACH1, PTGS2, CCL2, BNIP3, SOD2, STC1, THBS1, TLR2, UCP2, VCAM1, VEGFA, CXCR4
reactive oxygen species metabolic process	2.82	3.23	8	45	CYP1B1, BIRC3, NCF2, DDIT4, PREX1, BNIP3, SYK, IER3
regulation of cytokine production	2.80	9.04	22	132	CD24, ADAM8, LILRB2, CEBPB, FGR, CLEC5A, IL17RA, LRRC32, CLEC4E, HILPDA, BIRC3, LTB, LYN, SRGN, PTGER4, BCL6, C10orf54, SLC11A1, SYK, IL18R1, CD86, TNFSF15

Table 13 – Continued on next page

Table 13 – continued from previous page

regulation of immune response	2.79	15.92	38	240	CD24, CDKN1A, LILRB2, CD300A, AMICA1, CTSH, HBEGF, PPAPDC1A, FCGR1B, FCGR2A, FCGR2B, TNFRSF21, ICAM1, ICAM3, ITPR2, LCP2, LGALS3, MNDA, MYO10, C1RL, PDGFRA, PIK3CD, PRKCD, SELL, SHC1, STXBP2, SYK, C1R, C1S, C3AR1, C5AR1, TYROBP, VCAM1, LAT2, PRAM1, SPPL2A, IRS2, THEMIS2
extrinsic apoptotic signaling pathway via death domain receptors	2.77	3.28	8	45	ICAM1, LGALS3, SERPINE1, BID, THBS1, TNFAIP3, TNFRSF10B, CFLAR
odontogenesis of dentin-containing tooth	2.77	3.28	8	45	FST, ADM, CSF3R, TNC, TNFRSF11B, PDGFRA, FAM20C, RUNX2
carbohydrate biosynthetic process	2.75	7.08	17	97	ENO1, B4GALT1, NR3C1, HAS2, HAS3, IL6, NFKB1, CHST11, CHST15, PGD, PGK1, PLEK, PTPN2, UGP2, B3GNT5, IRS2, GFPT2
cellular glucose homeostasis	2.75	3.72	9	51	CYBA, HIF1A, HK2, HK3, ICAM1, PIM3, UCP2, VCAM1, IRS2
cellular carbohydrate metabolic process	2.75	10.07	24	138	B4GALT1, GK, NR3C1, HAS2, HAS3, HIF1A, HK2, HK3, IGFBP3, IL6, IL6ST, PFKFB3, PFKFB4, PGD, PLEK, DDIT4, PTPN2, STAT3, UGP2, B3GNT5, DGAT2, IRS2, IER3, MGAM
leukocyte differentiation	2.74	13.98	33	194	CSF1R, CSF3, DOCK2, SBNO2, SNX10, IFI16, IRF1, ITGAM, JAK3, JUNB, LTF, LYN, MMP9, MT1G, MYC, PIK3CD, FAM20C, PTGER4, PREX1, PTPN2, PTPN6, PTPRC, RELB, BCL3, BCL6, SPI1, TFRC, TYROBP, VCAM1, VEGFA, RUNX2, IL18R1, CD1D
collagen catabolic process	2.73	4.16	10	57	COL1A1, COL3A1, COL5A2, COL11A1, CTSS, FAP, MMP9, MMP10, MMP19, ADAMTS2
positive regulation of peptidase activity	2.73	6.27	15	86	CTSH, F3, FBLN1, PYCARD, IFI16, MAP3K5, MYC, SLC11A2, NLRC4, ROBO1, S100A8, S100A9, BID, TNFRSF10B, TNFSF15

Table 13 – Continued on next page

Table 13 – continued from previous page

negative regulation of response to wounding	2.72	4.60	11	63	METRNL, NFKB1, SERPINE1, PRKCD, PTGER4, PTPN2, NOD2, SPP1, TNFAIP3, TNFRSF1B, IER3
endothelium development	2.72	2.92	7	40	S1PR3, KDM6B, GJA4, ICAM1, MYD88, STC1, NRP1
negative regulation of ion transport	2.72	2.92	7	40	LILRB2, ICAM1, IL1RN, SERPINE2, PTGS2, THBS1, IRS2
mammary gland epithelium development	2.72	2.92	7	40	CEBPB, CSF1R, NR3C1, HIF1A, ROBO1, VDR, WNT5A
lymphocyte activation	2.72	14.08	33	201	CD300A, AMICA1, LRRC32, TNFRSF21, ICAM1, IRF1, ITGAM, JAK3, LCP1, MNDA, PIK3CD, PIK3CG, PRKCD, PRLR, PTGER4, PREX1, PTPN2, RAB27A, RELB, BCL3, SAMSN1, C10orf54, CLEC7A, SLC11A1, SPI1, TNFAIP3, LAT2, CXCR4, LST1, RUNX2, IL18R1, SKAP2, CD48
positive regulation of lymphocyte activation	2.71	9.30	22	129	CD24, ADAM8, LILRB2, IKZF1, CORO1A, MAP3K8, CD55, AIF1, PYCARD, IL1B, IL4R, IL6ST, LYN, MYD88, PTPN6, CCL2, CCL5, TGFBR2, CCR2, VCAM1, CD1D, CD86
regulation of protein maturation	2.71	7.59	18	104	ADAM8, TRIB1, CTSC, CSTA, CD55, GAS1, IL1B, MFI2, SERPINE1, CFP, SERPINE2, PLAT, CFB, THBS1, TIMP1, TIMP3, TNFRSF1B, C3
bone development	2.70	5.47	13	75	SEMA4D, COL1A1, COMP, SULF1, GREM1, HAS2, LTF, PTGER4, PTPRC, STC1, THBS1, RUNX2, PAPSS2
immunoglobulin production	2.70	3.36	8	46	TNFSF13B, IL6, SASH3, PTPRC, BCL6, NOD2, TNFSF13, CD37
regulation of fat cell differentiation	2.70	3.36	8	46	CEBPB, METRNL, IL6, PTGS2, SOD2, WNT5A, ZC3H12A, MEDAG
multicellular organismal macromolecule metabolic process	2.70	5.91	14	81	COL1A1, COL3A1, COL5A2, COL11A1, CTSS, FAP, HIF1A, IL6, ITGA2, MMP9, MMP10, MMP19, RETN, ADAMTS2
positive regulation of phospholipase activity	2.69	3.79	9	52	LPAR1, FLT1, GNA15, ITPR2, PDGFRA, PRKCD, CCL5, SELE, C5AR1

Table 13 – Continued on next page

Table 13 – continued from previous page

regulation of sequence-specific DNA binding transcription factor activity	2.69	13.79	32	189	ADAM8, TRIB1, IRAK3, CYP1B1, PYCARD, HCK, ICAM1, IL1B, IL6, IRAK1, LTF, ARRB2, MYD88, NFKB1, NFKB2, NFKBIA, NLRC4, S100A8, S100A9, S100A12, NOD2, SGK1, TLR2, TNFAIP3, TRAF3, VEGFA, WNT5A, TRIM25, ZC3H12A, TRIM8, BHLHE40, CFLAR
positive regulation of cell adhesion	2.68	7.66	18	105	ADAM8, OLFM4, EGFL6, HAS2, IL1B, ITGA2, MYO10, PREX1, PTPN6, CCL5, SPP1, SYK, TGM2, THBS1, VEGFA, WNT5A, CD36, CD44
negative regulation of growth	2.68	8.97	21	123	CDKN1A, ENO1, GAS1, HIF1A, MT1E, MT1F, MT1G, MT1H, MT1M, MT1X, MT2A, MYD88, PLAC8, SERPINE2, BCL6, NOD2, SPP1, TLR2, WNT5A, NRP1, CD36
regulation of endopeptidase activity	2.67	5.96	14	83	CTSH, F3, PYCARD, BIRC3, MAP3K5, MYC, SLC11A2, NLRC4, ROBO1, S100A8, S100A9, BID, TNFRSF10B, TNFSF15
T cell differentiation	2.66	4.25	10	59	DOCK2, PIK3CD, PTGER4, PREX1, PTPRC, RELB, BCL3, RUNX2, IL18R1, CD1D
positive regulation of protein serine/threonine kinase activity	2.66	11.24	26	154	CD24, ADAM8, CHI3L1, CSF1R, LPAR1, ACSL1, FLT1, FPR1, IL1B, IL1RN, IRAK1, LTF, MAP3K5, PIK3CG, PROK2, S100A12, NOD2, SHC1, SYK, THBS1, C5AR1, VEGFA, WNT5A, CXCR4, TNFRSF10B, TNFSF15
innate immune response	2.66	19.77	45	294	CDKN1A, CORO1A, CSF1R, CYBA, CYBB, CD55, HBEGF, PPAPDC1A, FCER1G, FCGR2A, CLEC5A, PANX1, ITPR2, JAK3, LCN2, LCP2, LGALS3, MAP3K5, MYO10, NCF2, CLEC4A, C1RL, PDGFRA, CFP, PIK3CD, PIK3CG, TREM1, PTX3, RAB27A, S100A8, S100A9, S100A12, CFB, SHC1, SYK, C1R, C1S, C3, CCR2, TYROBP, LAT2, TRIM25, TRIM8, IRS2, PSTPIP1

Table 13 – Continued on next page

Table 13 – continued from previous page

regulation of epithelial cell proliferation	2.65	9.92	23	136	LRG1, AMICA1, CYBA, F3, VASH1, SULF1, GAS1, B4GALT1, GLUL, HIF1A, IL6, MYC, PLA2G2A, ROBO1, CCL2, NOD2, THBS1, C5AR1, VDR, VEGFA, WNT5A, NRP2, NRP1
regulation of establishment of protein localization	2.65	18.02	41	247	ADAM8, CDKN1A, CSF1R, CSF3, FCN1, FGR, SPIDR, CLEC5A, PANX1, LRRC32, CLEC4E, TNFRSF21, PYCARD, HCLS1, IL1A, IL1B, IL4R, IL6, RHOG, LCP1, LYN, NFKBIA, NFKBIE, TLR8, PID1, SRGN, PTGER4, PTGS2, BCL3, BID, NOD2, SORL1, SYK, TLR2, WNT5A, IL18R1, RAB29, IL1RL1, CD14, CD36, TNFSF15
female gonad development	2.63	3.43	8	47	FST, ICAM1, ARRB2, MMP19, PDGFRA, RETN, VEGFA, ADAMTS1
myelination	2.63	3.43	8	47	LPAR1, EGR2, TNFRSF21, LAMA2, MBP, TLR2, CXCR4, CNTNAP1
signal transduction in absence of ligand	2.63	3.87	9	53	GAS1, IL1A, IL1B, JAK3, MCL1, PF4, BCL2A1, TNFRSF10B, CFLAR
unsaturated fatty acid metabolic process	2.62	4.74	11	65	PDPN, CYP1B1, ACSL1, ALOX5, ALOX5AP, FADS3, ELOVL2, PTGS2, SYK, TLR2, PTGES
cellular response to decreased oxygen levels	2.62	4.74	11	65	ADAM8, HIF1A, ICAM1, IRAK1, SLC11A2, BACH1, PTGS2, BNIP3, STC1, VEGFA, CPEB4
intrinsic apoptotic signaling pathway	2.62	6.49	15	90	CD24, CDKN1A, PYCARD, ERO1L, IFI16, LYN, DDIT4, BCL3, S100A8, S100A9, BID, BNIP3, TNFRSF1B, TNFRSF10B, IER3
response to monosaccharide	2.62	6.06	14	83	CYBA, GLUL, HIF1A, ICAM1, ITGA2, PIM3, NCF2, PTGS2, SOD2, TGFBR2, THBS1, UCP2, VCAM1, IRS2
regulation of T cell differentiation	2.62	5.62	13	77	ADAM8, LILRB2, IKZF1, IL4R, IL6, IRF1, JAK3, SASH3, PTPN2, BCL6, SYK, TGFBR2, CD86
positive regulation of ERK1 and ERK2 cascade	2.62	5.62	13	77	CHI3L1, CCR1, CSF1R, PYCARD, ICAM1, IL6, ARRB2, PDGFRA, NOD2, C5AR1, VEGFA, NRP1, CD44

Table 13 – Continued on next page

Table 13 – continued from previous page

regulation of MAP kinase activity	2.60	11.89	27	163	CD24, ADAM8, TRIB1, IRAK3, CD300A, LPAR1, FLT1, FPR1, IL1B, IL1RN, IRAK1, LYN, MAP3K5, PIK3CG, PRKCD, PTPN6, PROK2, S100A12, NOD2, SHC1, SORL1, SYK, THBS1, C5AR1, VEGFA, WNT5A, CXCR4
tissue remodeling	2.60	4.78	11	66	CTHRC1, CSF1R, GREM1, HIF1A, IL6, LIF, TNFRSF11B, RAC2, TGM2, TNFAIP3, CCR2
positive regulation of epithelial cell proliferation	2.60	3.46	8	48	AMICA1, GAS1, B4GALT1, GLUL, IL6, MYC, NOD2, C5AR1
vasculature development	2.59	15.55	35	221	ADAM8, PDPN, HPSE, COL1A1, COL3A1, TYMP, ELK3, GREM1, B4GALT1, GJA4, ANPEP, HAS2, CXCL8, ITGB8, JUNB, LIF, LOX, MMP19, NFE2, PDGFRA, PIK3CG, ERRFI1, PROK2, ROBO1, SAT1, CEACAM1, SHC1, SPI1, SYK, TNFAIP2, CXCR4, ZC3H12A, NRP2, NRP1, SOCS3
activation of MAPK activity	2.58	6.13	14	84	ADAM8, LPAR1, FPR1, IL1B, IRAK1, MAP3K5, PROK2, NOD2, SHC1, SYK, THBS1, C5AR1, WNT5A, CXCR4
regulation of protein import into nucleus	2.57	5.25	12	72	CDKN1A, CSF3, HCLS1, IL1B, IL6, NFKBIA, NFKBIE, PTGS2, BCL3, TLR2, IL18R1, CD36
regulation of apoptotic signaling pathway	2.57	14.74	33	202	PPIF, CTSC, CTSH, GAS1, PYCARD, HIF1A, ICAM1, IL1A, IL1B, JAK3, LGALS3, ARRB2, MCL1, OSM, G0S2, SERPINE1, TAF9B, PF4, PRKCD, BCL2A1, BDKRB2, S100A8, S100A9, BID, BNIP3, SOD2, THBS1, TNFAIP3, TNFRSF10B, NRP1, CFLAR, IER3, CD44
transition metal ion transport	2.57	4.38	10	60	TCIRG1, SLC39A14, LTF, MFI2, SLC11A2, SLC25A37, SLC11A1, TFRC, STEAP4, SCO2
coenzyme biosynthetic process	2.56	3.50	8	48	NAMPT, MTHFD2, ACSL1, GCH1, ELOVL2, NADK, PDXK, KYNU

Table 13 – Continued on next page

Table 13 – continued from previous page

modification of morphology or physiology of other organism	2.56	3.50	8	48	TNIP1, TLR8, TREM1, S100A12, CCL4, CCL5, NOD2, TLR2
positive regulation of calcium ion transport	2.56	3.50	8	48	CCR1, MCHR1, LGALS3, ARRB2, CCL2, CCL4, CCL5, STC1
regulation of cysteine-type endopeptidase activity involved in apoptotic process	2.56	7.95	18	109	CTSH, F3, PYCARD, IL6, MAP3K5, MYC, SLC11A2, NLRC4, ROBO1, RPS6KA1, S100A8, S100A9, BID, THBS1, VEGFA, TNFRSF10B, CD44, TNFSF15
angiogenesis	2.56	7.04	16	101	ADAM8, TYMP, ELK3, ANPEP, IL6, CXCL8, MMP19, PIK3CG, PROK2, SAT1, CEACAM1, SHC1, SYK, TNFAIP2, ZC3H12A, NRP2
positive regulation of I-kappaB kinase/NF-kappaB signaling	2.55	7.08	16	97	LPAR1, BIRC3, IL1B, IRAK1, LTF, MYD88, S100A4, S100A12, SECTM1, NOD2, TGM2, TRIM25, TRIM8, TNFRSF10B, CFLAR, CD36
maintenance of location	2.55	9.34	21	128	PLIN2, ABCA1, FBN1, HILPDA, ERO1L, IL1B, IL6, NFKB1, NFKBIA, NFKBIE, SRGN, PTPN2, PTPRC, BCL3, S100A8, S100A9, SORL1, C3, ZC3H12A, DGAT2, CD36
defense response to virus	2.55	9.34	21	128	IFITM2, DOCK2, SAMHD1, PYCARD, HCK, BIRC3, IFI16, IL6, IRF1, ISG20, TLR8, PLSCR1, DDIT4, DNAJC3, PTPRC, BNIP3, TNFAIP3, TRAF3, TRIM25, CD86, CD37
cell-type specific apoptotic process	2.55	15.30	34	211	ADAM8, CEBPB, FAP, FCER1G, CLEC5A, NR3C1, HCLS1, HIF1A, ICAM1, IL6, JAK3, LGALS3, LYN, ARRB2, MCL1, MYC, NCF2, SERPINE1, ANGPTL4, PIK3CG, BCL6, CCL5, NOD2, BNIP3, SOD2, TGFBR2, THBS1, TNFAIP3, C5AR1, WNT5A, CPEB4, DOCK8, IRS2, NRP1

Table 13 – Continued on next page

Table 13 – continued from previous page

regulation of cellular component movement	2.55	12.04	27	176	TRIB1, CD300A, COL3A1, CYP1B1, DPYSL3, SPATA13, VASH1, SULF1, GREM1, B4GALT1, IGFBP3, IL1RN, LAMA2, MMP10, SERPINE2, PLAU, PTGER4, RAC2, NOD2, SGK1, SORL1, STAT3, STC1, STK10, TIMP1, CXCR4, SH3BGRL3
protein maturation	2.54	8.45	19	118	CTSH, F3, ANPEP, PYCARD, ERO1L, IFI16, MAP3K5, MME, SLC11A2, SRGN, NLRC4, ROBO1, S100A8, S100A9, BID, SORL1, TNFRSF10B, ADAMTS2, TNFSF15
response to purine-containing compound	2.54	5.76	13	79	COL1A1, EGR2, PANX1, IL1B, IL6, AQP9, ITPR2, JUNB, MMP19, PLAT, PTGS2, SELL, STC1
positive regulation of cytosolic calcium ion concentration	2.53	9.85	22	135	CD24, CD52, CCR1, ADM, CD55, LPAR1, S1PR3, GNA15, MCHR1, ERO1L, LYN, PDGFRA, PIK3CG, PTPN6, PTPRC, PROK2, BDKRB2, TGM2, C3AR1, C5AR1, CCR2, CXCR4
negative regulation of defense response	2.53	5.33	12	73	IRAK3, METRNL, IFI16, ARRB2, NFKB1, PRKCD, PTGER4, PTPN2, NOD2, TNFAIP3, TNFRSF1B, IER3
regulation of ossification	2.52	7.15	16	98	SEMA4D, CEBPB, CTHRC1, CCR1, EGR2, GREM1, HIF1A, IL6, IL6ST, LTF, NFE2, SRGN, FAM20C, PTGER4, WNT5A, RUNX2
leukocyte mediated cytotoxicity	2.51	4.01	9	55	CTSC, CTSH, ICAM1, ARRB2, TREM1, PTPN6, PTPRC, RAB27A, STXBP2
development of primary female sexual characteristics	2.50	3.57	8	49	FST, ICAM1, ARRB2, MMP19, PDGFRA, RETN, VEGFA, ADAMTS1
response to organophosphorus	2.49	5.40	12	74	COL1A1, EGR2, PANX1, IL1B, AQP9, ITPR2, JUNB, MMP19, PLAT, PTGS2, SELL, STC1
blood coagulation	2.49	9.97	22	143	AMICA1, SIRPA, FGR, MAFF, IRF1, ITGA2, ITGAM, ITGAX, ITGB2, NFE2, TREM1, CD177, RAB27A, SELE, SELL, SELPLG, TFPI2, DOCK8, PAPSS2, SLC16A3, CD44, CD48

Table 13 – Continued on next page

Table 13 – continued from previous page

extracellular matrix disassembly	2.48	6.79	15	93	ADAM8, COL1A1, COL3A1, COL5A2, COL11A1, CTSS, FAP, FBN1, MFI2, MMP9, MMP10, MMP19, SPP1, TIMP1, CD44
protein localization to nucleus	2.48	6.79	15	93	CDKN1A, COL1A1, CSF3, HCLS1, IL1B, IL6, LIF, NFKBIA, NFKBIE, PTGS2, BCL3, TLR2, RASSF5, IL18R1, CD36
negative regulation of protein modification process	2.48	13.79	30	189	TRIB1, CDKN1A, SEMA4D, IRAK3, CD300A, GREM1, PYCARD, IGFBP3, IL1B, IL6, LYN, ARRB2, ERRFI1, DDIT4, PID1, PRKCD, DNAJC3, PSMB9, PTPN2, PTPN6, PTPRC, BDKRB2, SAMSN1, SORL1, SPI1, TNFAIP3, SOCS1, SOCS3, GMFG, MVP
ameboidal cell migration	2.48	10.51	23	144	CTSH, CYP1B1, HBEGF, FAP, VASH1, GREM1, SRPX2, HIF1A, ITGA2, MMP9, PTGS2, ROBO1, S100P, STC1, TGFBR2, THBS1, TIMP1, VEGFA, WNT5A, CXCR4, SH3BGRL3, NRP2, NRP1
divalent inorganic cation homeostasis	2.46	14.81	32	203	CD24, CD52, CCR1, ADM, CD55, LPAR1, S1PR3, SLC39A14, GNA15, MCHR1, ERO1L, LYN, PDGFRA, PIK3CG, CCDC109B, PTPN6, PTPRC, PROK2, BDKRB2, S100A8, S100A9, CCL2, CCL5, CCL8, SLC11A1, STC1, TGM2, C3AR1, C5AR1, CCR2, VDR, CXCR4
immune effector process	2.46	4.97	11	74	PPAPDC1A, FCGR2A, SBNO2, ITPR2, LCP1, MYO10, PTGER4, PTX3, RELB, TYROBP, IL18R1
regulation of actin polymerization or depolymerization	2.45	5.47	12	75	ARPC1B, CORO1A, CSF3, PYCARD, HCK, HCLS1, ICAM1, PLEK, PRKCD, PREX1, CAPG, SH3BGRL3
biomineral tissue development	2.45	5.47	12	75	CCR1, COL1A1, SBNO2, GREM1, HIF1A, LTF, ASGR2, NFE2, SRGN, FAM20C, PTGS2, SPP1
fat cell differentiation	2.45	4.55	10	63	EGR2, IL6, RETN, SOD2, WNT5A, STEAP4, ZC3H12A, MEDAG, SOCS1, CREB5

Table 13 – Continued on next page

Table 13 – continued from previous page

phosphatidylinositol-mediated signaling	2.44	7.81	17	107	CDKN1A, SEMA4D, PIK3AP1, CSF1R, CSF3, HBEGF, FGR, FLT1, HCLS1, PDGFRA, SERPINE2, PIK3CD, PIK3CG, PTPN6, CCL5, IRS2, CD86
morphogenesis of a branching epithelium	2.44	8.76	19	120	ADM, CTSH, RDH10, FLT1, SULF1, B4GALT1, TNC, IL6, MYC, SPI1, TGFBR2, TGM2, VDR, VEGFA, WNT5A, CXCR4, NRP1, SOCS3, CD44
positive regulation of cell death	2.43	16.42	35	225	ADAM8, CDKN1A, FGD4, ADM, CYP1B1, LPAR1, ALDH1A3, PHLDA1, B4GALT1, NR3C1, PYCARD, HP, IGFBP3, IL6, JAK3, LYN, ARRB2, MAP3K5, MMP9, MNDA, NCF2, ARHGEF3, DDIT4, PTGS2, NLRC4, BCL6, CCL5, BID, BNIP3, TGM2, THBS1, UCP2, VDR, WNT5A, FGD3
metal ion homeostasis	2.43	21.30	45	292	CD24, TCIRG1, CD52, CCR1, ADM, CD55, LPAR1, S1PR3, SLC39A14, GNA15, MCHR1, ERO1L, HIF1A, LCN2, LTF, LYN, MFI2, MT2A, MYC, SLC11A2, SLC25A37, PDGFRA, PIK3CG, CCDC109B, PTPN6, PTPRC, PROK2, BDKRB2, S100A8, S100A9, CCL2, CCL5, CCL8, SLC11A1, SOD2, STC1, TFRC, TGM2, C3AR1, C5AR1, CCR2, VDR, CXCR4, STEAP4, SCO2
cellular calcium ion homeostasis	2.43	13.06	28	179	CD24, CD52, CCR1, ADM, CD55, LPAR1, S1PR3, GNA15, MCHR1, ERO1L, LYN, PDGFRA, PIK3CG, CCDC109B, PTPN6, PTPRC, PROK2, BDKRB2, CCL2, CCL5, CCL8, STC1, TGM2, C3AR1, C5AR1, CCR2, VDR, CXCR4
positive regulation of Rho GTPase activity	2.43	5.98	13	82	SEMA4D, FGD4, DOCK2, SPATA13, ARHGAP26, ICAM1, RHOG, ARHGEF3, ERRFI1, PREX1, SH3BGRL3, FGD3, DOCK4
cellular response to hypoxia	2.42	4.60	10	63	ADAM8, HIF1A, ICAM1, IRAK1, SLC11A2, BACH1, PTGS2, BNIP3, STC1, VEGFA

Table 13 – Continued on next page

Table 13 – continued from previous page

tissue homeostasis	2.42	7.88	17	108	ADAM8, CORO1A, SERPINA3, CSF1R, CTSH, ITGA2, LTF, LYZ, TNFRSF11B, SASH3, RAC2, NOD2, SPP1, SYK, TFRC, TNFAIP3, VEGFA
regulation of Rho protein signal transduction	2.42	7.88	17	108	SEMA4D, FGD4, COL3A1, DOCK2, ABCA1, LPAR1, SPATA13, ARHGAP26, ICAM1, RHOG, ARHGEF3, ERRFI1, PREX1, BCL6, SH3BGRL3, FGD3, DOCK4
cellular cation homeostasis	2.42	19.92	42	273	CD24, TCIRG1, CD52, CCR1, ADM, CD55, LPAR1, S1PR3, SLC39A14, GNA15, MCHR1, ERO1L, HIF1A, LTF, LYN, MFI2, MT2A, MYC, SLC11A2, PDGFRA, PIK3CG, CCDC109B, SLAMF8, PTPN6, PTPRC, PROK2, BDKRB2, S100A8, S100A9, CCL2, CCL5, CCL8, SLC11A1, STC1, TFRC, TGM2, C3AR1, C5AR1, CCR2, VDR, CXCR4, SCO2
regulation of viral process	2.41	5.54	12	76	TNIP1, IFITM2, IFI16, CXCL8, ISG20, PLSCR1, PTX3, CCL4, CCL5, TAP1, TRIM25, TRIM8
female sex differentiation	2.40	4.16	9	57	FST, ICAM1, ARRB2, MMP19, PDGFRA, RETN, VEGFA, WNT5A, ADAMTS1
negative regulation of inflammatory response	2.40	4.16	9	57	METRNL, NFKB1, PRKCD, PTGER4, PTPN2, NOD2, TNFAIP3, TNFRSF1B, IER3
cofactor biosynthetic process	2.40	4.16	9	57	NAMPT, MTHFD2, ACSL1, GCH1, SLC11A2, ELOVL2, NADK, PDXK, KYNU
positive regulation of lipase activity	2.40	4.16	9	57	LPAR1, FLT1, GNA15, ITPR2, PDGFRA, PRKCD, CCL5, SELE, C5AR1
negative regulation of protein serine/threonine kinase activity	2.40	4.16	9	57	CDKN1A, IRAK3, CD300A, PYCARD, IL1B, LYN, PRKCD, PTPN6, SORL1
negative regulation of multicellular organismal process	2.40	14.64	31	203	CD24, IRAK3, CCR1, ADM, CYBA, GREM1, TNFRSF21, PYCARD, HIF1A, IL6, JAK3, LTF, ARRB2, NFE2, NFKB1, TNFRSF11B, PLAC8, PIK3CG, SRGN, PTGS2, BCL6, C10orf54, NOD2, SLC11A1, TGFBR2, TLR2, TNFAIP3, TRAF3, TRIM25, ZC3H12A, IL1RL1

Table 13 – Continued on next page

Table 13 – continued from previous page

regulation of leukocyte mediated immunity	2.39	5.11	11	71	ICAM1, IL1B, JAK3, ARRB2, SASH3, PTPN6, PTPRC, BCL6, C10orf54, NOD2, TNFSF13
regulation of apoptotic process	2.38	18.63	39	274	CDKN1A, SEMA4D, CEBPB, COMP, EGR2, FAP, FCER1G, CLEC5A, HCK, HCLS1, BIRC3, IL1RN, IL6ST, IRAK1, LTF, PIM3, MYD88, NFKB1, NFKBIA, CHST11, SERPINB2, PLAC8, PIK3CG, PRLR, PSMB9, BCL3, PROK2, CCL2, NOD2, SGK1, TIMP1, TRAF3, C5AR1, UCP2, CPEB4, DOCK8, IRS2, TNFRSF10D, SOCS3
actin filament polymerization	2.36	5.18	11	71	ARPC1B, CORO1A, CSF3, AIF1, PYCARD, HCK, HCLS1, ICAM1, PRKCD, PREX1, CAPG
apoptotic mitochondrial changes	2.36	4.23	9	58	PPIF, PYCARD, HK2, ARRB2, MCL1, BID, BNIP3, SOD2, IER3
regulation of phospholipase activity	2.36	4.23	9	58	LPAR1, FLT1, GNA15, ITPR2, PDGFRA, PRKCD, CCL5, SELE, C5AR1
regulation of response to DNA damage stimulus	2.36	4.23	9	58	PPP4R2, SPIDR, MCL1, MYC, NPAS2, TAF9B, PRKCD, IER3, CD44
regulation of immune effector process	2.35	4.22	9	60	DOCK2, HCK, BIRC3, ICAM1, CFP, CFB, TNFAIP3, TRAF3, PRAM1
response to inorganic substance	2.35	4.23	9	60	COL1A1, CYP1B1, KDM6B, GPX3, HP, MAP3K5, TNFAIP3, UCP2, CD36
cytokine production	2.34	11.93	25	177	CD24, LILRB2, CEBPB, HPSE, CYP1B1, SULF1, IL17RA, HILPDA, IL6ST, IRF1, LTB, NFKB1, NFKB2, NFKBIA, SERPINE1, PIK3CD, PIK3CG, PRKCD, PTGS2, CCL2, C3, TRAF3, C3AR1, TRIM25, IL18R1
response to ethanol	2.33	4.74	10	65	TNC, ICAM1, ITPR2, S100A8, S100A9, CCL2, SOD2, STAT3, VCAM1, CD14
multi-multicellular organism process	2.32	10.58	22	145	ADM, CD55, FBLN1, MAFF, B4GALT1, IL1B, IL1RN, JUNB, LIF, MMP9, PAPPA, SERPINE2, PRLR, PTGS2, CCL2, SPP1, STC1, TGFBR2, TIMP1, UCP2, VDR, VMP1

Table 13 – Continued on next page

Table 13 – continued from previous page

positive regulation of MAPK cascade	2.32	10.58	22	147	CD24, ADAM8, LPAR1, FLT1, FPR1, IGFBP3, IL1B, IL1RN, IRAK1, LIF, MAP3K5, MYD88, OSM, PIK3CG, PROK2, S100A12, SHC1, SYK, THBS1, WNT5A, CXCR4, CD36
embryonic skeletal system development	2.32	5.25	11	72	COL1A1, COL11A1, RDH10, SULF1, GAS1, CHST11, PDGFRA, KIAA1217, TGFBR2, WNT5A, RUNX2
aging	2.32	9.63	20	132	CDKN1A, CTSC, COL3A1, ADM, VASH1, ICAM1, IL6, MME, NCF2, SERPINE1, PRKCD, RETN, CCL2, SOD2, TGFBR2, TIMP1, TNFRSF1B, UCP2, VCAM1, CD86
negative regulation of transferase activity	2.30	9.70	20	133	TRIB1, CDKN1A, IRAK3, CD300A, PYCARD, IL1B, IL6, LYN, ERRFI1, PRKCD, DNAJC3, PSMB9, PTPN6, PTPRC, SORL1, SOCS1, IRS2, SOCS3, GMFG, MVP
placenta development	2.29	6.27	13	86	CEBPB, ADM, HIF1A, ITGB8, JUNB, LIF, NFE2, PTGS2, SPP1, STC1, VCAM1, VDR, SOCS3
response to ionizing radiation	2.29	4.82	10	66	CDKN1A, CYBA, SPIDR, ICAM1, IFI16, MYC, CCL2, NABP1, SOD2, VCAM1
odontogenesis	2.29	4.82	10	66	FST, COL1A1, ADM, CSF3R, GAS1, TNC, TNFRSF11B, PDGFRA, FAM20C, RUNX2
cellular response to fibroblast growth factor stimulus	2.28	8.76	18	120	CDKN1A, HBEGF, SULF1, CXCL8, ITPR2, PDGFRA, PIK3CD, PRKCD, FAM20C, CCL2, CCL5, SHC1, THBS1, WNT5A, RUNX2, IRS2, CD86, CD44
cartilage development	2.28	7.30	15	100	CHI3L1, COL1A1, COL11A1, COMP, SULF1, CHRDL2, HIF1A, ITGB8, CHST11, STC1, TGFBR2, THBS1, WNT5A, RUNX2, CD44
response to hexose	2.27	5.84	12	80	CYBA, GLUL, HIF1A, ICAM1, PIM3, NCF2, PTGS2, TGFBR2, THBS1, UCP2, VCAM1, IRS2
regulation of cell-substrate adhesion	2.26	6.35	13	87	OLFM4, COL1A1, EGFL6, MFI2, SERPINE1, PLAU, PREX1, RAC2, BCL6, SPP1, THBS1, VEGFA, CD36

Table 13 – Continued on next page

Table 13 – continued from previous page

regulation of myeloid cell differentiation	2.26	6.35	13	87	IKZF1, CCR1, HCLS1, HIF1A, LIF, LTF, LYN, MYC, NFKBIA, PF4, PTPN2, SPI1, TYROBP
positive regulation of leukocyte differentiation	2.26	6.35	13	87	ADAM8, LILRB2, IKZF1, CCR1, HCLS1, IL4R, IL6, LIF, PF4, SASH3, SYK, TGFBR2, CD86
regulation of actin filament polymerization	2.25	4.89	10	67	ARPC1B, CORO1A, CSF3, PYCARD, HCK, HCLS1, ICAM1, PRKCD, PREX1, CAPG
negative regulation of cell proliferation	2.25	20.07	40	279	TRIB1, CDKN1A, ADM, CYP1B1, DPT, AIF1, FAP, VASH1, SULF1, GAS1, B4GALT1, CXCL1, IGFBP3, IL1A, IL1B, IL6, CXCL8, IRF1, LIF, MYC, OSM, SERPINE2, PLA2G2A, PTGS2, PTPN2, RARRES1, BCL6, ROBO1, SOD2, STAT3, TGFBR2, THBS1, TLR2, VDR, WNT5A, RASSF5, SKAP2, CD37, ADAMTS1, PTGES
Fc-epsilon receptor signaling pathway	2.25	7.37	15	101	CDKN1A, HBEGF, FCER1G, ITPR2, LCP2, LYN, NFKB1, NFKBIA, PDGFRA, PIK3CD, SHC1, SYK, LAT2, IRS2, CD86
positive regulation of apoptotic signaling pathway	2.24	5.91	12	81	CTSC, CTSH, PYCARD, JAK3, MCL1, OSM, G0S2, PRKCD, S100A8, S100A9, BID, THBS1
response to reactive oxygen species	2.23	6.42	13	88	PPIF, COL1A1, CYP1B1, KDM6B, GPX3, HP, IL6, MAP3K5, BNIP3, SOD2, TNFAIP3, UCP2, CD36
negative regulation of secretion	2.23	6.42	13	88	CD300A, LRRC32, TNFRSF21, IL1B, IL1RN, IL6, LIF, PIM3, OSM, SRGN, PTGER4, CCR2, UCP2
response to drug	2.22	16.12	32	221	CDKN1A, PLIN2, CYBA, CYBB, ABCA1, ACSL1, RNF149, ICAM1, IL1B, IL1RN, IL6, ITGA2, JUNB, LOX, LYN, MGST1, MT2A, MYC, NCF2, NNMT, TNFRSF11B, SLAMF8, PTGS2, CCL2, SLC22A3, SOD2, STAT3, TGFBR2, THBS1, TLR2, CD86, CD69

Table 13 – Continued on next page

Table 13 – continued from previous page

response to extracellular stimulus	2.22	15.61	31	214	CDKN1A, COL1A1, ADM, CYBA, CYBB, ABCA1, ACSL1, GLUL, TNC, ICAM1, IFI16, IL6, ITGA2, LYN, TNFRSF11B, PTGS2, CCL2, NOD2, SLC22A3, SOD2, SPP1, STC1, TGFBR2, UCP2, UPP1, VCAM1, VDR, TRIM25, CPEB4, VAMP8, KYNU
positive regulation of secretion	2.22	13.06	26	179	ADAM8, CSF1R, FCER1G, FCN1, FGR, CLEC5A, PANX1, CLEC4E, GLUL, PYCARD, HIF1A, IL1A, IL1B, IL4R, IL6, RETN, PTGER4, RAB27A, NOD2, SYK, TLR2, WNT5A, IRS2, IL1RL1, CD14, TNFSF15
carbohydrate catabolic process	2.20	6.49	13	89	ENO1, GK, HIF1A, HK2, HK3, PFKFB3, PFKFB4, PGD, PGK1, DDIT4, STAT3, IER3, MGAM
protein targeting to nucleus	2.20	8.03	16	110	CDKN1A, CSF3, HCLS1, IL1B, IL6, JAK3, NFKBIA, NFKBIE, PTGS2, BCL3, BCL6, SLC11A1, STAT3, TLR2, IL18R1, CD36
single-organism nuclear import	2.20	8.03	16	110	CDKN1A, CSF3, HCLS1, IL1B, IL6, JAK3, NFKBIA, NFKBIE, PTGS2, BCL3, BCL6, SLC11A1, STAT3, TLR2, IL18R1, CD36
positive regulation of kinase activity	2.20	21.01	41	288	CD24, ADAM8, CDKN1A, CHI3L1, CD300A, MAP3K8, CSF1R, LPAR1, ACSL1, FGR, FLT1, FPR1, GREM1, IL1B, IL1RN, CXCL8, IRAK1, LTF, LYN, MAP3K5, PDGFRA, PIK3CG, DNAJC3, PRLR, PTPRC, PROK2, S100A12, CCL2, CCL5, NOD2, SHC1, SLC11A1, SYK, TGFBR2, THBS1, C5AR1, VEGFA, WNT5A, CXCR4, TNFRSF10B, TNFSF15
cellular response to oxidative stress	2.19	5.00	10	69	PPIF, KDM6B, GPX3, HP, IL6, MGST1, SLC11A2, BNIP3, TNFAIP3, CD36
regulation of neuron death	2.17	8.61	17	118	ADAM8, CEBPB, CSF3, NR3C1, HIF1A, IL6, MCL1, MAP3K5, NCF2, DDIT4, CCL2, CCL5, SOD2, SORL1, C5AR1, CPEB4, NRP1

Table 13 – Continued on next page

Table 13 – continued from previous page

activation of protein kinase activity	2.17	5.54	11	77	CHI3L1, CD300A, MAP3K8, GREM1, PRLR, CCL5, SLC11A1, TGFBR2, VEGFA, TNFRSF10B, TNFSF15
response to glucose	2.17	5.54	11	76	CYBA, GLUL, HIF1A, ICAM1, PIM3, NCF2, TGFBR2, THBS1, UCP2, VCAM1, IRS2
negative regulation of neuron death	2.17	5.54	11	76	ADAM8, CEBPB, CSF3, HIF1A, IL6, CCL2, SOD2, SORL1, C5AR1, CPEB4, NRP1
single-organism carbohydrate catabolic process	2.17	6.06	12	83	ENO1, GK, HIF1A, HK2, HK3, PFKFB3, PFKFB4, PGD, PGK1, DDIT4, STAT3, IER3
B cell activation	2.16	5.58	11	78	IKZF1, JAK3, PIK3CD, PRKCD, PTPN2, PTPN6, BCL3, VCAM1, LAT2, SKAP2, CD86
negative regulation of endopeptidase activity	2.13	9.27	18	127	SERPINA3, CSTA, SERPINB1, IFI16, IL6, SERPINE1, SERPINB2, SERPINA1, SERPINE2, RPS6KA1, SORL1, THBS1, TIMP1, TIMP3, C3, VEGFA, TFPI2, CD44
ossification	2.12	6.17	12	86	COL1A1, SBNO2, CLEC5A, CHRDL2, TNC, IGFBP3, JUNB, MMP9, ASGR2, PTGS2, SPP1, STC1
negative regulation of transport	2.12	14.08	27	193	PPIF, LILRB2, CD300A, LRRC32, TNFRSF21, ICAM1, IL1B, IL1RN, IL6, LGALS3, LIF, PIM3, MMP9, NFKB1, NFKBIA, NFKBIE, OSM, SERPINE2, PID1, SRGN, PTGER4, PTGS2, THBS1, CCR2, UCP2, IRS2, CD36
response to corticosteroid	2.11	6.20	12	85	CDKN1A, COL1A1, ADM, NR3C1, IL1RN, IL6, JUNB, PAPPA, PLAT, PTGS2, CCL2, STC1
regulation of calcium ion transport	2.09	6.79	13	93	LILRB2, CCR1, MCHR1, ICAM1, LGALS3, LYN, ARRB2, PTGS2, PTPN6, CCL2, CCL4, CCL5, STC1
negative regulation of response to external stimulus	2.08	6.28	12	87	TRIB1, METRNL, LTF, NFKB1, SERPINE1, PRKCD, PTGER4, NOD2, SPP1, TNFAIP3, TNFRSF1B, IER3

Table 13 – Continued on next page

positive regulation of apoptotic process	2.07	13.83	26	191	FGD4, ADM, CYP1B1, LPAR1, ALDH1A3, PHLDA1, B4GALT1, NR3C1, PYCARD, IGFBP3, IL6, ARRB2, MAP3K5, MMP9, MNDA, NCF2, ARHGEF3, PTGS2, NLRC4, BCL6, BID, BNIP3, TGM2, THBS1, VDR, FGD3
endocytosis	2.06	19.40	36	272	CORO1A, CD300A, MARCH3, ADM, ABCA1, PPAPDC1A, AIF1, FCGR2A, FCN1, FGR, STAB1, SH3KBP1, HCK, CXCL8, CXCR2, ITPR2, LGALS3, LYN, ASGR2, MYO10, BIN2, PIK3CG, PRKCD, APOBR, CDC42SE1, CLEC7A, SORL1, SYK, TFRC, TGFBR2, TGM2, THBS1, LOXL3, VAMP8, PSTPIP1, CD163
positive regulation of cytokine production	2.05	7.95	15	113	ADAM8, CSF1R, IL17RA, PANX1, HILPDA, BIRC3, IFI16, IL4R, IRF1, NFKB1, NFKB2, NFKBIA, SERPINE1, THBS1, IL1RL1
response to nitrogen compound	2.05	12.31	23	173	CDKN1A, IRAK3, ADM, CD55, PANX1, IL1RN, IL6, IRAK1, ITGA2, JUNB, MGST1, MMP19, MYD88, NFKBIA, NNMT, PLAT, RETN, CCL2, SELL, TIMP1, C5AR1, VCAM1, KYNU
regulation of proteolysis	2.04	19.56	36	270	TRIB1, CTSC, SERPINA3, CSTA, CTSH, SERPINB1, F3, FBLN1, PYCARD, BIRC3, IFI16, IL6, MAP3K5, MYC, SLC11A2, SERPINE1, SERPINB2, SERPINA1, SERPINE2, PLAT, PLAUR, NLRC4, ROBO1, RPS6KA1, S100A8, S100A9, BID, SORL1, THBS1, C3, TRAF3, VEGFA, TFPI2, TNFRSF10B, CD44, TNFSF15
developmental growth involved in morphogenesis	2.03	6.42	12	88	SEMA4D, COMP, RDH10, TNC, ROBO1, SPP1, STC1, THBS1, VEGFA, WNT5A, NRP2, NRP1
antigen processing and presentation	2.02	8.62	16	119	CTSH, CTSS, CYBA, CYBB, FCER1G, FCGR1B, ICAM1, NCF2, NCF4, PSMB9, RAB27A, RELB, TAP1, TAPBP, CD1D, CD36

Table 13 – Continued on next page

Table 13 – continued from previous page

cellular response to growth factor stimulus	1.99	17.65	32	246	LRG1, FGD4, COL1A1, COL3A1, FBN1, FLT1, FMOD, GAS1, HAS2, HIF1A, IL1B, IRAK1, JUNB, ARRB2, MT1G, MYC, MYD88, NFKB1, NFKBIA, CHST11, SERPINE1, ARHGEF3, DDIT4, RPS6KA1, STAT3, TGFBR2, TNFRSF1B, VCAM1, VEGFA, NRP2, NRP1, FGD3
response to acid chemical	1.99	9.84	18	136	CEBPB, COL1A1, COL3A1, COL5A2, CTSH, CYBA, ABCA1, ICAM1, IL6, ITGA2, LYN, PDGFRA, CCL2, SOD2, WNT5A, CPEB4, SOCS1, PTGES
response to organic cyclic compound	1.98	10.44	19	147	IRAK3, PLIN2, CYBA, CYP1B1, ABCA1, ACSL1, SPIDR, ICAM1, ITGA2, LYN, NFKB1, NFKBIA, TAF9B, CCL5, NOD2, VCAM1, KYNU, CD36, PTGES
protein import	1.96	8.83	16	121	CDKN1A, CSF3, HCLS1, IL1B, IL6, JAK3, NFKBIA, NFKBIE, PTGS2, BCL3, BCL6, SLC11A1, STAT3, TLR2, IL18R1, CD36
protein autophosphorylation	1.94	8.90	16	122	CAMKK2, IRAK3, CSF1R, FGR, FLT1, GREM1, HCK, IRAK1, LYN, PIM3, PDGFRA, ERRFI1, STK10, SYK, VEGFA, MVP
regulation of kinase activity	1.94	8.36	15	119	TRIB1, CAMKK2, IRAK3, PYCARD, IL6, ERRFI1, PRKCD, SLAMF8, PTPN6, SORL1, SOCS1, IRS2, SOCS3, GMFG, MVP
generation of precursor metabolites and energy	1.93	12.40	22	170	PPIF, ENO1, MCHR1, HIF1A, HK2, HK3, IL6ST, ITPR2, MYC, PFKFB3, PFKFB4, PGK1, DDIT4, PID1, SOD2, STAT3, UCP2, UGP2, IRS2, IER3, GFPT2, SCO2
lung development	1.89	7.95	14	109	PDPN, CHI3L1, CTSH, RDH10, HSD11B1, TNC, LIF, LOX, PDGFRA, ERRFI1, TGFBR2, VEGFA, WNT5A, ADAMTS2
stem cell development	1.89	9.12	16	125	COL1A1, RDH10, GREM1, HAS2, HIF1A, FAM101B, LIF, PLA2G2A, PTPRC, S100A4, SPI1, STAT3, WNT5A, LOXL3, NRP2, NRP1

Table 13 – Continued on next page

Table 13 – continued from previous page

receptor-mediated endocytosis	1.87	9.78	17	134	ADM, STAB1, GREM1, CXCL8, CXCR2, ARRB2, ASGR2, SERPINE1, APOBR, SELE, SORL1, TFRC, TGFBR2, VEGFA, LOXL3, CD163, CD36
regulation of cellular component size	1.84	9.92	17	136	ARPC1B, SEMA4D, CORO1A, CSF3, PYCARD, HCK, HCLS1, ICAM1, PLEK, PRKCD, PREX1, SPP1, VEGFA, WNT5A, CAPG, SH3BGRL3, NRP1
divalent metal ion transport	1.83	14.74	25	202	LILRB2, CORO1A, CCR1, SLC39A14, PANX1, GJA4, MCHR1, ERO1L, ICAM1, ITPR2, LGALS3, LYN, ARRB2, SLC11A2, CCDC109B, PTGS2, PTPN6, PTPRC, CCL2, CCL4, CCL5, CCL8, SLC11A1, STC1, VDR
exocytosis	1.82	9.41	16	131	CD300A, CCR1, FGR, B4GALT1, HCK, MYO1F, PIK3CD, PIK3CG, RAB27A, CCL5, CCL8, STXBP2, TNFAIP2, CCR2, LAT2, PRAM1
response to oxidative stress	1.82	12.40	21	170	PPIF, COL1A1, CYP1B1, RCAN1, KDM6B, GPX3, HIF1A, HP, IL6, MCL1, MAP3K5, MGST1, NFKB1, SLC11A2, PRKCD, PTGS2, BNIP3, SOD2, TNFAIP3, UCP2, CD36
regulation of protein kinase activity	1.82	20.88	35	292	CD24, ADAM8, CAMKK2, CHI3L1, MAP3K8, CSF1R, LPAR1, ACSL1, FGR, FLT1, FPR1, GREM1, IL1RN, CXCL8, IRAK1, LTF, MAP3K5, PIK3CG, PRLR, PROK2, S100A12, CCL2, CCL5, NOD2, SHC1, SLC11A1, SYK, TGFBR2, THBS1, C5AR1, VEGFA, WNT5A, CXCR4, TNFRSF10B, TNFSF15
reproductive system development	1.81	17.29	29	237	FST, CEBPB, ADM, RDH10, SULF1, GJB2, HIF1A, TNC, ICAM1, ITGB8, JUNB, LIF, ARRB2, MGST1, MMP19, NFE2, PDGFRA, SERPINE2, RETN, PTGS2, SPP1, STC1, VCAM1, VDR, VEGFA, WNT5A, SOCS3, ADAMTS1, CD44

Table 13 – Continued on next page

Table 13 – continued from previous page

negative regulation of hydrolase activity	1.81	14.30	24	196	PPIF, SEMA4D, SERPINA3, CSTA, SERPINB1, IFI16, IL6, ARRB2, SERPINE1, SERPINB2, ANGPTL4, SERPINA1, SERPINE2, PTX3, RPS6KA1, SORL1, THBS1, TIMP1, TIMP3, C3, VEGFA, TFPI2, SPOCD1, CD44
regulation of DNA metabolic process	1.81	9.48	16	130	CDKN1A, PPP4R2, SPIDR, GREM1, IL6, CXCL8, ARRB2, MYC, NPAS2, PDGFRA, PID1, BCL6, SHC1, SPI1, TNFSF13, IER3
positive regulation of cell differentiation	1.81	16.74	28	235	CD24, SEMA4D, OLFM4, COL1A1, ADM, GREM1, METRNL, HIF1A, IGFBP3, JUNB, LYN, SERPINE2, ADAMTS9, PTGS2, PREX1, ROBO1, RPS6KA1, C10orf54, TCF4, TLR2, VDR, VEGFA, CXCR4, ZC3H12A, MEDAG, NRP1, SOCS3, BOC
cellular response to nitrogen compound	1.80	20.43	34	280	TCIRG1, CEBPB, COL1A1, COL3A1, COL5A2, CYBA, EGR2, SPIDR, ICAM1, IL1B, AQP9, ITPR2, JAK3, LYN, MAP3K5, NFKB1, PDGFRA, PID1, PRKCD, PRLR, PTGS2, PTPN2, PTPN6, PTPRE, NOD2, SHC1, STAT3, STC1, TLR2, UCP2, CPEB4, SOCS1, IRS2, SOCS3
transmembrane receptor protein serine/threonine kinase signaling pathway	1.80	11.31	19	155	FST, LRG1, COL3A1, FBN1, SULF1, FMOD, GREM1, IRAK1, JUNB, ARRB2, MYC, CHST11, SERPINE1, CCL2, C10orf54, TGFBR2, THBS1, WNT5A, RUNX2
tube morphogenesis	1.80	13.13	22	180	PDPN, CTHRC1, ADM, CSF1R, CTSH, RDH10, FLT1, B4GALT1, NR3C1, HIF1A, TNC, MYC, ROBO1, SPI1, TGFBR2, TGM2, VDR, VEGFA, WNT5A, CXCR4, NRP1, CD44

Table 13 – Continued on next page

Table 13 – continued from previous page

symbiosis, encompassing mutualism through parasitism	1.78	21.33	35	299	DOCK2, FBLN1, FCGR2B, CLEC5A, ANPEP, PILRA, HCK, ICAM1, IFI16, IL6ST, CXCL8, ISG20, ITGA2, LYN, MAP3K5, MYD88, NFKBIA, PDGFRA, PLSCR1, TREM1, PSMB9, CCL2, CCL4, CCL5, BNIP3, STAT3, SYK, TFRC, CCR2, VCAM1, CFLAR, ATG16L2, CD1D, CD86, CD36
interspecies interaction between organisms	1.78	21.33	35	299	DOCK2, FBLN1, FCGR2B, CLEC5A, ANPEP, PILRA, HCK, ICAM1, IFI16, IL6ST, CXCL8, ISG20, ITGA2, LYN, MAP3K5, MYD88, NFKBIA, PDGFRA, PLSCR1, TREM1, PSMB9, CCL2, CCL4, CCL5, BNIP3, STAT3, SYK, TFRC, CCR2, VCAM1, CFLAR, ATG16L2, CD1D, CD86, CD36
response to steroid hormone	1.77	15.76	26	216	CD24, CDKN1A, COL1A1, ADM, RCAN1, NR3C1, IL1RN, IL4R, IL6, JUNB, LOX, NCF2, TNFRSF11B, PAPPA, PLAT, PTGS2, CCL2, BID, SPP1, STAT3, STC1, TGFBR2, THBS1, TLR2, VDR, TRIM25
regulation of cellular response to stress	1.76	15.90	26	218	PPP4R2, VASH1, SPIDR, PYCARD, HIF1A, HP, IL1B, IL1RN, LYN, MCL1, MAP3K5, MYC, MYD88, NPAS2, TAF9B, PRKCD, BDKRB2, NOD2, SOD2, SPP1, SYK, VEGFA, WNT5A, IER3, CD36, CD44
response to peptide hormone	1.75	16.56	27	227	TCIRG1, COL1A1, ADM, CD55, EGR2, IL1B, IL6, JAK3, JUNB, LYN, NFKB1, PLAT, PID1, PRKCD, PRLR, RETN, PTPN2, PTPN6, PTPRE, SHC1, STAT3, TIMP1, TLR2, UCP2, SOCS1, IRS2, SOCS3
stem cell differentiation	1.75	11.60	19	159	COL1A1, RDH10, GREM1, HAS2, HIF1A, FAM101B, LIF, PDGFRA, PLA2G2A, PTPRC, S100A4, C10orf54, SPI1, STAT3, WNT5A, LOXL3, RUNX2, NRP2, NRP1
lipid localization	1.73	12.99	21	178	PLIN2, ABCA1, ACSL1, HILPDA, IL1B, IL6, NFKB1, NFKBIA, PRKCD, APOBR, RETN, PTPN2, BDKRB2, SORL1, SYK, THBS1, C3, ZC3H12A, DGAT2, IRS2, CD36

Table 13 – Continued on next page

Table 13 – continued from previous page

epithelial tube morphogenesis	1.71	12.48	20	171	CTHRC1, ADM, CSF1R, CTSH, RDH10, FLT1, B4GALT1, NR3C1, HIF1A, TNC, MYC, ROBO1, SPI1, TGFBR2, VDR, VEGFA, WNT5A, CXCR4, NRP1, CD44
actin cytoskeleton organization	1.70	20.87	33	286	ARPC1B, CORO1A, FGD4, CSF1R, CSF3, DOCK2, DPYSL3, AIF1, ARHGAP26, PYCARD, CXCL1, HCK, HCLS1, ICAM1, FAM101B, RHOG, LCP1, ARHGDIB, MYO1F, NEDD9, PDGFRA, PLEK, PRKCD, PTGER4, PREX1, RAC2, BCL6, PARVG, SHC1, CAPG, SH3BGRL3, CFLAR, FGD3
negative regulation of cellular component organization	1.67	16.63	26	228	PPIF, CORO1A, IRAK3, CD300A, DPYSL3, LPAR1, FAP, LGALS3, LIF, MBP, MFI2, NFE2, OSM, PID1, PRKCD, PTGER4, BNIP3, SORL1, SPI1, SPP1, THBS1, WNT5A, CAPG, NRP1, CFLAR, IER3
calcium ion transport	1.65	14.15	22	194	LILRB2, CORO1A, CCR1, PANX1, GJA4, MCHR1, ERO1L, ICAM1, ITPR2, LGALS3, LYN, ARRB2, CCDC109B, PTGS2, PTPN6, PTPRC, CCL2, CCL4, CCL5, CCL8, STC1, VDR
tissue morphogenesis	1.62	21.01	32	288	FST, CTHRC1, COL11A1, ADM, CSF1R, CTSH, RDH10, HBEGF, ALDH1A3, KDM6B, FLT1, SULF1, B4GALT1, NR3C1, HIF1A, TNC, IL6, ITGA2, LIF, MYC, ROBO1, SPI1, STC1, TGFBR2, TGM2, VDR, VEGFA, WNT5A, CXCR4, NRP1, SOCS3, CD44
positive regulation of protein phosphorylation	1.62	21.03	32	295	ADAM8, CDKN1A, CHI3L1, CD300A, MAP3K8, LPAR1, ACSL1, FLT1, FPR1, IL1B, IL1RN, CXCL8, IRAK1, LTF, MAP3K5, PIK3CG, DNAJC3, PRLR, PTPRC, PROK2, S100A12, CCL2, SHC1, SLC11A1, TGFBR2, THBS1, C3, C5AR1, WNT5A, CXCR4, TNFRSF10B, TNFSF15

Table 13 – Continued on next page

Table 13 – continued from previous page

viral process	1.61	18.46	28	256	DOCK2, FBLN1, FCGR2B, CLEC5A, ANPEP, PILRA, HCK, ICAM1, IL6ST, ITGA2, LYN, MAP3K5, NFKBIA, PDGFRA, PSMB9, CCL2, BNIP3, STAT3, SYK, TFRC, CCR2, VCAM1, CXCR4, VAMP8, CFLAR, ATG16L2, CD1D, CD86
chemotaxis	1.60	16.49	25	247	CD24, PLXNC1, SEMA4D, COL3A1, COL5A2, DOCK2, DPYSL3, TYMP, EGR2, FPR1, GAS1, LAMA2, LSP1, MYO10, PLAU, PLAUR, PLD1, PTPRC, RAC2, PROK2, RPS6KA1, TYROBP, CNTNAP1, NRP2, BOC

Next, we carry out the GO enrichment analysis for the significantly down-regulated genes. Results can be found in Table 14.

```
> maskDown <- ttAll$logFC < 0
> DEgenesDown <- intersect(ttAll[maskDown, "ID"], DEgenes)
> length(DEgenesDown)
[1] 505
> gUniverse <- ttAll[maskDown, "ID"]
> length(gUniverse)
[1] 9287
> GOparams <- new("GOHyperGParams", geneIds=DEgenesDown, universeGeneIds=gUniverse,
                  annotation="org.Hs.eg.db", ontology="BP",
                  pvalueCutoff=goAnalysisCondPvalueCutoff, conditional=TRUE,
                  testDirection="over")
> cache(hgOverGOBPdeGenesDown <- hyperGTest2(GOparams, minSizeCutoff=5, maxSizeCutoff=300, orCuto
        dir=cacheDir_FE, prefix=cachePrefix_FE)
> hgOverGOBPdeGenesDown
Gene to GO BP Conditional test for over-representation
4046 GO BP ids tested (176 have p < 0.01)
Selected gene set size: 425
    Gene universe size: 6983
    Annotation package: org.Hs.eg
> IQRs <- apply(cleanExps[gUniverse, ], 1, IQR)
> cache(reportOverGOBPdeGenesDown <- GOreport(hgOverGOBPdeGenesDown, minSize=5, maxSize=300,
                                              minCount=5, minOddsRatio=1.5, featureFilterValues=
                                              ff.R=1000, ff.prob=0.90, maxReported=Inf,
                                              p.value=0.1, p.adjust.method="fdr"),
        dir=cacheDir_FE, prefix=cachePrefix_FE)
> dim(reportOverGOBPdeGenesDown)
[1] 35 10
```

Table 14. Gene Ontology (GO) enrichment (biological process) of downregulated genes called differentially expressed between FIR and nonFIR affected infants at FDR $< 1\%$ and with a minimum 1.5-fold change. Column "GeneSyms" contains the genes enriching the corresponding GO term.

Term	OddsRatio	ExpCount	Count	Size	GeneSyms
positive regulation of JNK cascade	11.72	0.85	6	14	GADD45G, DUSP19, ANKRD6, PLCB1, FZD7, DIXDC1
positive regulation of stress-activated protein kinase signaling cascade	6.69	1.22	6	20	GADD45G, DUSP19, ANKRD6, PLCB1, FZD7, DIXDC1
regulation of smooth muscle contraction	6.25	1.28	6	21	GUCY1A3, ATP1A2, P2RX1, PLCE1, PRKG1, CAV1
neural crest cell development	6.25	1.28	6	21	CITED2, EDN3, FOXC1, ISL1, PITX2, BMPR1A
astrocyte differentiation	5.51	1.40	6	23	CDK6, FGFR3, BIN1, NTRK3, PLP1, SOX6
cell fate specification	4.56	1.89	7	31	HOXA13, HOXC10, ISL1, NTRK3, SOX6, BMPR1A, FZD7
response to alkaloid	4.56	1.89	7	31	GSTM2, HNMT, ATP1A2, BCHE, RYR2, RYR3, CASQ2
calcium-mediated signaling	4.28	2.56	9	42	DMD, DMPK, LPHN1, GSTM2, PLCE1, PPP1R9A, RYR2, CASQ2, RGN
negative regulation of neurogenesis	4.17	2.31	8	38	FGF13, FGFR3, SPG20, NTRK3, PCM1, PTK2, BMPR1A, ULK2
negative regulation of cell adhesion	4.04	2.37	8	39	KANK1, ANGPT1, ARHGAP6, PIK3R1, PRKG1, PTK2, JAM2, FZD7
negative regulation of canonical Wnt signaling pathway	4.03	2.98	10	49	PRICKLE1, AMER2, ANKRD6, FRZB, DKK3, APC, ISL1, WNT4, TCF7L2, CAV1
embryonic limb morphogenesis	4.03	3.59	12	59	FGF9, HOXA10, HOXA13, HOXC10, AFF3, LRP5, MYH3, PBX1, PITX2, PTCH1, BMPR1A, FRAS1
regulation of blood vessel size	3.57	3.29	10	54	EDN3, FOXC1, GUCY1A3, KCNA5, KCNJ8, ATP1A2, P2RX1, PRKG1, CAV1, ATG5
appendage morphogenesis	3.57	3.96	12	65	FGF9, HOXA10, HOXA13, HOXC10, AFF3, LRP5, MYH3, PBX1, PITX2, PTCH1, BMPR1A, FRAS1
limb development	3.48	4.38	13	72	FGF9, HOXA10, HOXA13, HOXC10, AFF3, LRP5, MYH3, PBX1, PITX2, PTCH1, BMPR1A, NR2F2, FRAS1

Table 14 – Continued on next page

Table 14 – continued from previous page

heart contraction	3.25	4.99	14	82	DMD, DMPK, EDN3, GPD1L, HEY2, GSTM2, KCNA5, ATP1A2, FXYD1, PRKCA, RYR2, CASQ2, CAV1, ATG5
cell-matrix adhesion	3.20	3.96	11	65	CDK6, DLC1, COL13A1, ECM2, ARHGAP6, PIK3R1, WNT4, EPDR1, PTK2, AJUBA, ITGA10
regulation of Wnt signaling pathway	2.98	6.15	16	101	PRICKLE1, AMER2, FGF9, FGFR3, ANKRD6, KANK1, FRZB, DKK3, APC, ISL1, LRP5, WNT4, TCF7L2, TLE1, DIXDC1, CAV1
regulation of metal ion transport	2.63	6.39	15	105	CNTN1, CNKSR3, DMD, DMPK, GPD1L, GSTM2, KCNA5, P2RX1, FXYD1, RYR2, CACNA2D1, KCNAB1, CAMK2G, CASQ2, CAV1
regulation of anatomical structure size	2.60	8.22	19	135	DSTN, PLEKHH2, JMY, EDN3, FGF13, FOXC1, SPG20, KANK1, GUCY1A3, KCNA5, KCNJ8, ATP1A2, NTRK3, P2RX1, PFN2, PRKG1, CAV1, ATG5, ULK2
regulation of lipid metabolic process	2.59	6.03	14	99	FABP3, FGFR3, DKK3, DNAJC15, PRKAG2, PDK2, PDK4, PIK3R1, WNT4, PTK2, TCF7L2, CAV1, PDE8B, NR1D2
circulatory system process	2.54	9.74	22	160	DMD, DMPK, EDN3, FOXC1, GPD1L, HEY2, ANGPT1, GSTM2, GUCY1A3, KCNA5, KCNJ8, LRP5, ATP1A2, P2RX1, FXYD1, PRKCA, PRKG1, RYR2, OR51E2, CASQ2, CAV1, ATG5
stem cell development	2.54	7.49	17	123	CITED2, EDN3, FGFR3, FOXC1, HEY2, APC, ISL1, LRP5, PCM1, PITX2, WNT4, BMPR1A, TCF7L2, LDLRAD4, TET1, FZD7, LDB2
negative regulation of cell differentiation	2.38	12.72	27	209	CDK6, PRICKLE1, FGF13, FGFR3, SPG20, KANK1, HEY2, FRZB, APC, ISL1, LRP5, NTRK3, PBX1, PCM1, PIK3R1, WNT4, PTCH1, PTK2, BMPR1A, TCF7L2, LDLRAD4, TET1, FZD7, DIXDC1, CAV1, LDB2, ULK2

Table 14 – Continued on next page

Table 14 – continued from previous page

forebrain development	2.33	9.98	21	164	CDK6, DLC1, CKB, EMX2, FGF9, FGF13, FGFR3, BTBD3, PLCB1, ISL1, PCM1, PITX2, PLP1, WNT4, PRKG1, SUDS3, BMPR1A, TCF7L2, NR2F2, FAT4, DIXDC1
sulfur compound metabolic process	2.31	9.07	19	149	CSPG5, GPC4, MAMDC2, PHGDH, ANGPT1, GSTM1, GSTM2, ACACB, IDH1, MUT, OGN, PCYOX1, PDK2, PDK4, PRELP, MICAL1, TCF7L2, SUCLG2, DSEL
regulation of cell projection organization	2.28	8.70	18	143	CHN1, CNTN1, DMD, FGF13, SPG20, KANK1, NEGR1, SIPA1L1, APC, NTRK3, PFN2, PTK2, PTPRD, RAP1A, BHLHB9, DIXDC1, ATG5, ULK2
muscle system process	2.22	10.41	21	171	CRYAB, DMD, DMPK, EDN3, HEY2, GSTM2, GUCY1A3, MYH3, ATP1A2, P2RX1, PLCE1, FXYD1, PRKCA, PRKG1, RYR2, RYR3, SSPN, CAMK2G, CASQ2, CAV1, MYOM1
regulation of cell motility	2.15	11.75	23	193	CDK6, CITED2, DLC1, EDN3, KANK1, PLCB1, GAB1, ANGPT1, SCAI, APC, NTRK3, PFN2, PIK3R1, PITX2, WNT4, PRKCA, PTK2, HACE1, NR2F2, LDLRAD4, MMP28, AJUBA, NEXN
sensory organ development	2.10	11.99	23	197	CITED2, LRIG3, COL8A2, CRYAB, FGF9, FGFR3, FOXC1, HEY2, TSPAN12, FRZB, HOXA13, APC, LRP5, NTRK3, PITX2, MDM1, TUB, FAT4, DIXDC1, PROM1, KCNQ4, ATG5, CXCL14
regulation of nervous system development	1.98	13.75	25	226	CHN1, CNTN1, DMD, FGF13, FGFR3, LPHN1, SPG20, KANK1, HEY2, NEGR1, SIPA1L1, BIN1, ISL1, NTRK3, PBX1, PCM1, PTK2, PTPRD, RAP1A, BMPR1A, TCF7L2, SYNDIG1, BHLHB9, DIXDC1, ULK2

Table 14 – Continued on next page

Table 14 – continued from previous page

embryonic morphogenesis	1.97	14.36	26	236	CITED2, DLC1, LRIG3, PRICKLE1, FGF9, FOXC1, FRZB, GRSF1, HOXA10, HOXA13, HOXB7, HOXC10, AFF3, LRP5, MYH3, PBX1, PITX2, WNT4, PTCH1, RYR2, BMPR1A, TCF7L2, TXNRD1, FRAS1, FZD7, KCNQ4
signal transduction by phosphorylation	1.95	14.49	26	238	GADD45G, CRYAB, DUSP19, PPM1L, CNKSR3, EDN3, FGF9, FGF13, FGFR3, ANKRD6, PLCB1, GAB1, ANGPT1, APC, NTRK3, PLCE1, TNFRSF19, PRKCA, RAP1A, RGS4, BMPR1A, STYX, FZD7, AJUBA, DIXDC1, CAV1
synaptic transmission	1.90	16.62	29	273	CSPG5, ADCY5, PRIMA1, DMPK, EPB41L1, LPHN1, PLCB1, SIPA1L1, GNAI1, APBA1, KCNA5, KCND3, KCNJ8, MAOA, ATP1A2, NFATC4, P2RX1, PFN2, PLCL1, PLP1, PRKCA, PTK2, BCHE, RAP1A, KCNAB1, SYNDIG1, BHLHB9, CAMK2G, KCNQ4
regulation of nucleotide metabolic process	1.85	18.20	31	299	DLC1, CAP2, CHML, CHN1, PLCB1, SIPA1L1, ARHGEF26, BIN1, DNAJC15, GUCY1A2, GUCY1A3, ARHGAP6, NTRK3, PLCE1, FAM13B, PFN2, WNT4, PRKCA, PRKG1, PTK2, ARHGAP21, RAP1A, RGS4, RYR2, ARHGAP28, GARNL3, AJUBA, RGS9, RGN, RASAL2, RABGAP1L

Next, we are going to build an integrated display of differential expression, functional enrichment and phenotypic data. For this purpose, we build first incidence matrices of genes by GO terms using the GSEABase package.

```
> library(GSEABase)
> gscUp <- do.call("GeneSetCollection",
                   mapply(function(goID, geneIDs)
                          GeneSet(EntrezIdentifier("org.Hs.eg.db"), geneIds=geneIDs,
                                  setName=goID),
                          as.character(reportOverGOBPdeGenesUp$GOBPID),
                          strsplit(gsub(" ", "",
                                        as.character(reportOverGOBPdeGenesUp$Genes)), ","),
                          USE.NAMES=FALSE))
> gscUp
```

```
GeneSetCollection
  names: GO:0072672, GO:0002523, ..., GO:0006935 (542 total)
  unique identifiers: 101, 120425, ..., 10154 (467 total)
  types in collection:
    geneIdType: EntrezIdentifier (1 total)
    collectionType: NullCollection (1 total)
> gscDown <- do.call("GeneSetCollection",
                     mapply(function(goID, geneIDs)
                            GeneSet(EntrezIdentifier("org.Hs.eg.db"), geneIds=geneIDs,
                                    setName=goID),
                            as.character(reportOverGOBPdeGenesDown$GOBPID),
                            strsplit(gsub(" ", "",
                                          as.character(reportOverGOBPdeGenesDown$Genes)), ","),
                            USE.NAMES=FALSE))
> gscDown
GeneSetCollection
  names: GO:0046330, GO:0070304, ..., GO:0006140 (35 total)
  unique identifiers: 10912, 142679, ..., 9910 (166 total)
  types in collection:
    geneIdType: EntrezIdentifier (1 total)
    collectionType: NullCollection (1 total)
> imatUp <- t(incidence(gscUp))
> dim(imatUp)
[1] 467 542
> imatDown <- t(incidence(gscDown))
> dim(imatDown)
[1] 166  35
```

Perform a hierarchical clustering of genes and samples.

```
> cleanExpsDE <- cleanExps[DEgenes, ]
> sampleClustering <- hclust(as.dist(1-cor(cleanExpsDE, method="spearman")))
> geneClustering <- hclust(as.dist(1-cor(t(cleanExpsDE), method="pearson")))
> ord <- colnames(cleanExpsDE)[order.dendrogram(as.dendrogram(sampleClustering))]
> ord <- pData(frmaEsetFilt)[ord, "FIR"]
> names(ord) <- colnames(cleanExpsDE)[order.dendrogram(as.dendrogram(sampleClustering))]
```

To facilitate the display of functional enrichment information we pull GO term descriptions and abbreviate some of their words, such as *positive* to *pos.* or *regulation* to *reg.*, and remove prepositions *of* and *to*.

```
> goTermsUp <- reportOverGOBPdeGenesUp$Term
> names(goTermsUp) <- reportOverGOBPdeGenesUp$GOBPID
> goTermsDown <- reportOverGOBPdeGenesDown$Term
> names(goTermsDown) <- reportOverGOBPdeGenesDown$GOBPID
> goTerms <- c(goTermsUp, goTermsDown[is.na(match(names(goTermsDown),
                                                   names(goTermsUp)))])
> shortenGOterms <- function(gt) {
    gt <- gsub("inflammatory", "inflam.", gt)
    gt <- gsub("positive", "pos.", gt)
    gt <- gsub("negative", "neg.", gt)
    gt <- gsub("molecular", "mol.", gt)
    gt <- gsub("mediator", "med.", gt)
    gt <- gsub("response", "resp.", gt)
    gt <- gsub("regulation", "reg.", gt)
    gt <- gsub("activation", "activ.", gt)
```

```
    gt <- gsub("production", "prod.", gt)
    gt <- gsub("developmental", "dev.", gt)
    gt <- gsub("development", "dev.", gt)
    gt <- gsub("differentiation", "diff.", gt)
    gt <- gsub("cytokine-mediated", "cytok.", gt)
    gt <- gsub("lipopolysaccharide-mediated", "lipopolysacc.", gt)
    gt <- gsub("stress-activated", "stress-act.", gt)
    gt <- gsub("peptidyl-tyrosine", "pep.-tyr.", gt)
    gt <- gsub("signaling", "sig.", gt)
    gt <- gsub("pathway", "pwy.", gt)
    gt <- gsub("involved ", "", gt)
    gt <- gsub(" of ", " ", gt)
    gt <- gsub(" to ", " ", gt)
    gt <- gsub(" in ", " ", gt)

    gt
  }
> goTerms <- shortenGOterms(goTerms)
> funAnnotations <- list(up=imatUp, down=imatDown) ## genes x GO terms
> ## put back the original ranking order
> funAnnotations[["up"]] <- funAnnotations[["up"]][, match(colnames(funAnnotations[["up"]]),
                                                           reportOverGOBPdeGenesUp$GOBPID)]
> funAnnotations[["down"]] <- funAnnotations[["down"]][, match(colnames(funAnnotations[["down"]]
                                                               reportOverGOBPdeGenesDown$GOBPID)
```

We cannot include in a single figure the names of all of the 1097 DE genes. For this reason, we select the top 60 DE genes with lowest p-value.

```
> selGenes <- ttDEgenes$ID[1:60]
> selGenesUp <- selGenes[selGenes %in% ttDEgenes$ID[ttDEgenes$logFC > 0]]
> selGenesDown <- selGenes[selGenes %in% ttDEgenes$ID[ttDEgenes$logFC < 0]]
> selGenesText <- getSYMBOL(selGenes, "org.Hs.eg.db")
> stopifnot(all(!is.na(selGenesText)))
```

Finally, we build an incidence matrix of the phenotypic data to facilitate their visualization across samples. We group phenotypic variables into those measuring labor-related features and those measuring infant-related features.

```
> phendata <- rbind(as.numeric(frmaEsetFilt$Elbw == "yes"),
                    as.numeric(frmaEsetFilt$Bpd == "yes"),
                    as.numeric(frmaEsetFilt$VentGT7days == "yes"),
                    as.numeric(frmaEsetFilt$O2GE28days == "yes"),
                    as.numeric(frmaEsetFilt$Ivh == "yes"),
                    as.numeric(frmaEsetFilt$NecEnt == "yes"),
                    as.numeric(frmaEsetFilt$Rop == "yes"),
                    as.numeric(frmaEsetFilt$Pda == "yes"),
                    as.numeric(frmaEsetFilt$Sepsis == "yes"),
                    as.numeric(frmaEsetFilt$Ptl == "yes"),
                    as.numeric(frmaEsetFilt$Pih == "yes"),
                    as.numeric(frmaEsetFilt$Cs == "yes"),
                    as.numeric(frmaEsetFilt$Acs == "yes"),
                    as.numeric(frmaEsetFilt$Plate == "yes"),
                    as.numeric(frmaEsetFilt$Cord == "yes"),
                    as.numeric(frmaEsetFilt$Funisitis == "yes"),
                    as.numeric(frmaEsetFilt$Wmd == "yes"),
                    as.numeric(frmaEsetFilt$Cp == "yes"),
                    as.numeric(frmaEsetFilt$Mdi == "yes"),
```

```
                              as.numeric(frmaEsetFilt$Pdi == "yes")
                              )
> rownames(phendata) <- c("Elbw", "Bpd", "Ventilation", "Oxygen",
                          "Ivh", "NecEnt", "Rop", "Pda", "Sepsis",
                          "Ptl", "Pih", "Cs", "Acs", "Plate",
                          "Cord", "Funisitis", "Wmd", "Cp", "Mdi", "Pdi")
> colnames(phendata) <- colnames(frmaEsetFilt)
> phengroups <- list(Infant=c("Elbw", "Bpd", "Ventilation", "Oxygen", "NecEnt",
                              "Rop", "Pda", "Sepsis", "Pdi", "Mdi", "Cp", "Wmd", "Ivh"),
                     Labor=c("Ptl", "Pih", "Cs", "Acs", "Plate", "Cord", "Funisitis"))
```

The resulting graphical representation is shown in Figure 14. Its left-hand side contains a heatmap of gene expression values after adjusting for unknown sources of variability with surrogate variables. On top of the heatmap we find a hierarchical clustering of the samples which shows that they cluster according to the FIR condition of the displayed DE genes. On the bottom of the heatmap we find an incidence matrix of different phenotypic variables were a black cell indicates the presence of that phenotypic in the corresponding samples. On the left of the heatmap there is a hierarchical clustering of the genes with color annotations for those with evidence of a regulatory role, concretely, transcription factors (violet) and RNA-binding proteins (green).

The right-hand side of the figure contains two incidence matrices stacked on top of each other, of genes (rows) by GO terms (columns). One of them shows GO terms enriched with up-regulated genes (bottom) and the other shows GO terms enriched with down-regulated genes (top). The horizontal lines of the grid layed over the two matrices represent a resolution of 25 genes per cell. Dots in these matrices correspond to genes, where red indicate up-regulation and blue down-regulation. To ease their display, their size is bigger (i.e., have lower resolution) than the scale of the vertical axis which correponds to the clustered 1097 DE genes. On the right of these two matrices we find the top-60 DE genes with lowest *p*-value.

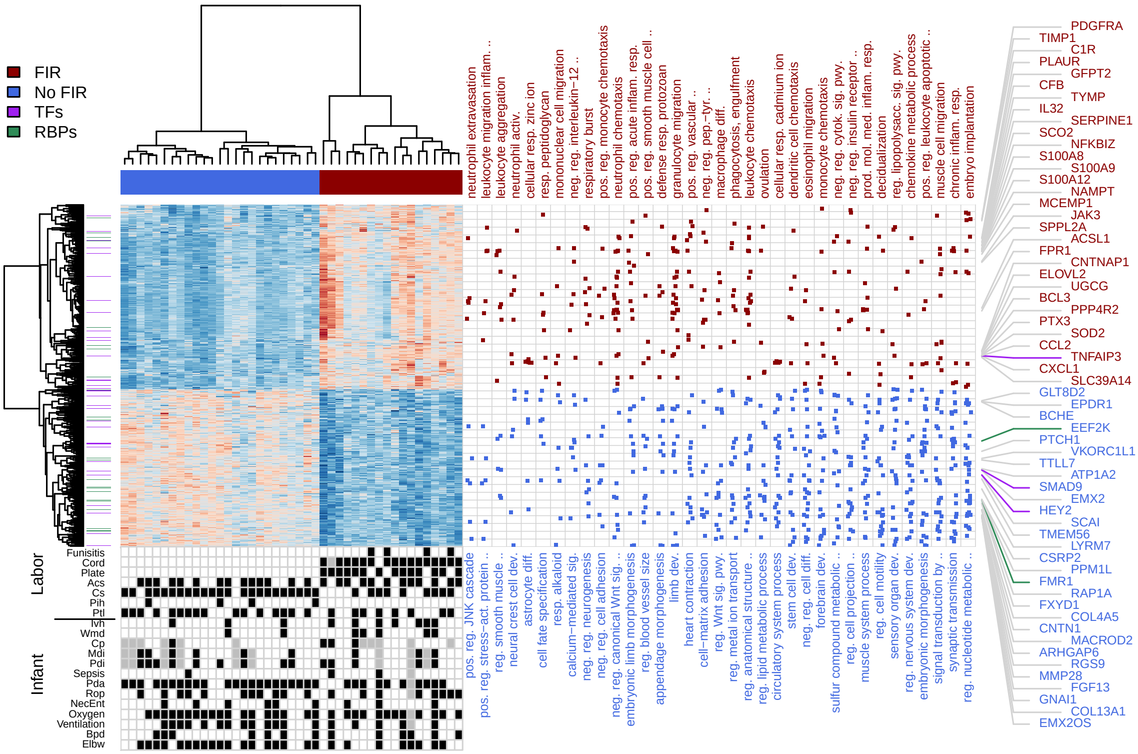

Figure 14. Integrated graphical representation of differential expression in between FIR and non-FIR infants, functional enrichment and phenotypic data.

We also perform a GO enrichment analysis among up and downregulated genes found by comparing infants who were not administered with antenatal glucocorticoids (ACS) and showed the combination of FIR, IVH and WMD phenotype against those who did not have any of these three phenotypes.

```
> DEgenesFDR <- ttAllIQRfilt[ttAllIQRfilt$adj.P.Val < FDRcutoff, "ID"]
> maskUp <- ttAllIVHWMDnoACS$logFC > 0
> DEgenesUpIVHWMDnoACS <- intersect(ttAllIVHWMDnoACS[maskUp, "ID"], DEgenesIVHWMDnoACS)
> length(DEgenesUpIVHWMDnoACS)
[1] 256
> gUniverse <- ttAllIVHWMDnoACS[maskUp, "ID"]
> length(gUniverse)
[1] 10386
> GOparams <- new("GOHyperGParams", geneIds=DEgenesUpIVHWMDnoACS, universeGeneIds=gUniverse,
                  annotation="org.Hs.eg.db", ontology="BP",
                  pvalueCutoff=goAnalysisCondPvalueCutoff, conditional=TRUE,
                  testDirection="over")
> cache(hgOverGOBPdeGenesUpIVHWMDnoACS <- hyperGTest2(GOparams, minSizeCutoff=5,
                                                      maxSizeCutoff=300, orCutoff=1.5),
        dir=cacheDir_FE, prefix=cachePrefix_FE)
> hgOverGOBPdeGenesUpIVHWMDnoACS
Gene to GO BP Conditional test for over-representation
3866 GO BP ids tested (285 have p < 0.01)
Selected gene set size: 235
    Gene universe size: 7316
    Annotation package: org.Hs.eg
> IQRs <- apply(exprs(cleanEsetIVHWMDnoACS)[gUniverse, ], 1, IQR)
> cache(reportOverGOBPdeGenesUpIVHWMDnoACS <- GOreport(hgOverGOBPdeGenesUpIVHWMDnoACS, minSize=5
                                                       minCount=5, minOddsRatio=1.5,
                                                       featureFilterValues=IQRs,
                                                       ff.R=1000, ff.prob=0.90, maxReported=Inf,
                                                       p.value=0.1, p.adjust.method="fdr",
                                                       highlightGenes=setdiff(DEgenesUpIVHWMDnoAC
                                                       highlightStr="\\textbf{%s}"),
        dir=cacheDir_FE, prefix=cachePrefix_FE)
> dim(reportOverGOBPdeGenesUpIVHWMDnoACS)
[1] 70 10
```

We can find the results of the GO enrichment analysis for up-regulated genes in Table 15.

Table 15. Gene Ontology (GO) enrichment (biological process) of upregulated genes called differentially expressed (DE) at FDR <1% and with a minimum 1.5-fold change between FIR/IVH/WMD and non-FIR/nonIVH/nonWMD affected infants. Column "GeneSyms" contains the genes enriching the corresponding GO term, where those highlighted in boldface are not DE between FIR and nonFIR affected infants at FDR <1%, and therefore, their upregulation is specific to the joint FIR/IVH/WMD condition.

Term	OddsRatio	ExpCount	Count	Size	GeneSyms
oxygen transport	37.08	0.35	6	11	**HBM**, **HBB**, **HBD**, **HBE1**, **HBQ1**, **HBZ**
response to superoxide	19.22	0.42	5	13	**MPO**, SOD2, **PRDX2**, UCP2, CD36
JAK-STAT cascade involved in growth hormone signaling pathway	13.97	0.51	5	16	JAK3, LYN, **MAPK1**, STAT3, SOCS3

Table 15 – Continued on next page

Table 15 – continued from previous page

hydrogen peroxide catabolic process	12.81	0.55	5	17	**GPX1**, **HBB**, **MPO**, **SNCA**, **PRDX2**
defense response to fungus	10.97	0.61	5	19	**DEFA4**, **MPO**, S100A8, S100A9, S100A12
cellular response to growth hormone stimulus	10.97	0.61	5	19	JAK3, LYN, **MAPK1**, STAT3, SOCS3
platelet aggregation	10.84	0.87	7	27	**DMTN**, **HBB**, **ITGA2B**, LYN, PDGFRA, PLEK, FERMT3
platelet degranulation	9.35	1.54	11	48	**DMTN**, **ITGA2B**, LYN, SERPINE1, PF4, SERPINA1, PLEK, PPBP, SRGN, TIMP1, CD36
glutathione metabolic process	8.08	0.77	5	24	PNKD, **GCLC**, **GPX1**, MGST1, SOD2
regulation of NF-kappaB import into nucleus	8.08	0.77	5	24	NFKBIA, **PPM1A**, PTGS2, BCL3, **TNFSF14**
regulation of mitochondrial membrane potential	8.08	0.77	5	24	**GCLC**, PID1, **BCL2L1**, SOD2, UCP2
cell chemotaxis	6.87	1.23	7	40	CXCL1, CXCL2, PDGFRA, PPBP, CCL8, CXCL6, CXCL5
T-helper 1 type immune response	6.39	0.93	5	29	JAK3, RELB, BCL3, SLC11A1, IL1RL1
leukocyte homeostasis	6.37	1.12	6	35	CORO1A, JAK3, LYN, **FOXP3**, TNFAIP3, **TNFSF14**
neutrophil chemotaxis	6.18	1.35	7	42	CXCL8, **MPP1**, TREM1, S100A8, S100A9, S100A12, CCL2
erythrocyte homeostasis	6.09	1.77	9	55	**KLF1**, **EPB42**, **DMTN**, **ALAS2**, **FECH**, **ANK1**, **HBZ**, LYN, **BPGM**
positive regulation of inflammatory response	5.72	1.86	9	58	TNIP1, OSM, SERPINE1, PTGS2, S100A8, S100A9, S100A12, IL1RL1, OSMR
complement activation	5.68	1.03	5	32	**RGCC**, **KRT1**, CFB, C1R, C1S
tetrapyrrole metabolic process	5.60	1.25	6	39	**ALAS2**, **FECH**, **EIF2AK1**, **ANK1**, **SLC25A39**, **BLVRB**
lipopolysaccharide-mediated signaling pathway	5.48	1.06	5	33	LYN, NFKBIA, **MAPK1**, CCL2, TNFAIP3
cellular response to lipopolysaccharide	5.37	1.29	6	41	ABCA1, CXCL8, **ARG1**, SERPINE1, TNFRSF1B, CD36
iron ion transport	5.29	1.09	5	34	SLC39A14, SLC25A37, **ATP6V0C**, SLC11A1, TFRC
response to amino acid	5.27	1.54	7	48	**GPX1**, **ARG1**, LYN, PDGFRA, **BCL2L1**, CCL2, CPEB4

Table 15 – Continued on next page

Table 15 – continued from previous page

leukocyte chemotaxis	5.25	1.32	6	42	CORO1A, PF4, TREM1, **CXCL16**, S100A8, S100A9
regulation of hemostasis	4.96	1.86	8	58	**DMTN**, **KRT1**, LYN, SERPINE1, PDGFRA, PLAUR, PLEK, CD36
granulocyte migration	4.86	1.90	8	59	CXCL8, **MPP1**, TREM1, RAC2, S100A8, S100A9, S100A12, CCL2
iron ion homeostasis	4.76	1.93	8	60	**EPB42**, **ALAS2**, **FECH**, SLC25A37, **ATP6V0C**, SLC11A1, SOD2, TFRC
cellular response to biotic stimulus	4.71	2.70	11	84	ABCA1, CXCL8, **ARG1**, LYN, NFKBIA, SERPINE1, **MAPK1**, CCL2, TNFAIP3, TNFRSF1B, CD36
regulation of coagulation	4.67	1.96	8	61	**DMTN**, **KRT1**, LYN, SERPINE1, PDGFRA, PLAUR, PLEK, CD36
homeostasis of number of cells	4.54	2.26	9	72	**KLF1**, **EPB42**, **DMTN**, **ALAS2**, **FECH**, **ANK1**, **HBZ**, **IL7R**, **BPGM**
response to hydrogen peroxide	4.50	2.02	8	63	**GPX1**, **HBB**, **ARG1**, **MPO**, **SNCA**, SOD2, **PRDX2**, TNFAIP3
defense response to bacterium	4.40	3.95	15	123	**DEFA4**, FGR, LYZ, MYO1F, SERPINE1, TREM1, PPBP, BCL3, S100A8, S100A9, S100A12, CXCL6, SLC11A1, **PGLYRP1**, CD36
regulation of wound healing	4.37	2.34	9	73	**DMTN**, **KRT1**, LYN, SERPINE1, PDGFRA, PLAUR, PLEK, TNFAIP3, CD36
cytokine secretion	4.27	2.67	10	83	CHI3L1, ABCA1, FGR, **RGCC**, LYN, **FOXP3**, TREM1, SRGN, S100A12, IL1RL1
response to bacterium	4.23	4.37	16	143	ABCA1, **GPX1**, CXCL2, CXCL8, **ARG1**, LYN, MGST1, **MPO**, NFKBIA, **MAPK1**, PTGS2, CCL2, **SNCA**, SOD2, TNFAIP3, TNFRSF1B
regulation of protein stability	4.15	1.90	7	59	**FBXO7**, ASGR2, **AHSP**, **TESC**, **MAPK1**, **SNCA**, **DPM2**
platelet activation	3.83	3.23	11	103	GNA15, SERPINE1, PF4, SERPINA1, APBB1IP, PPBP, SRGN, **MAPK1**, RAC2, TIMP1, CD36
negative regulation of response to wounding	3.78	2.06	7	64	**GPX1**, **KRT1**, SERPINE1, **FOXP3**, TNFAIP3, TNFRSF1B, **PGLYRP1**

Table 15 – Continued on next page

Table 15 – continued from previous page

response to tumor necrosis factor	3.69	2.41	8	75	CHI3L1, CXCL8, PID1, PTGS2, **CXCL16**, CCL2, TNFRSF1B, **YBX3**
response to lipopolysaccharide	3.52	2.51	8	81	MGST1, **MPO**, PTGS2, S100A8, S100A9, SLC11A1, **SNCA**, SOD2
regulation of endocytosis	3.49	2.86	9	89	FGR, **DKK1**, SERPINE1, PTX3, CCL2, SLC11A1, **SNCA**, CD36, **DNAJC6**
protein import into nucleus	3.45	3.21	10	100	JAK3, NFKBIA, **PPM1A**, **MAPK1**, PTGS2, BCL3, SLC11A1, STAT3, **TNFSF14**, CD36
nuclear import	3.45	3.21	10	100	JAK3, NFKBIA, **PPM1A**, **MAPK1**, PTGS2, BCL3, SLC11A1, STAT3, **TNFSF14**, CD36
myeloid leukocyte activation	3.38	3.28	10	102	FGR, CXCL8, LYN, MYO1F, RELB, S100A12, SLC11A1, **SNCA**, STXBP2, IL1RL1
myeloid cell differentiation	3.38	5.33	16	166	**KLF1**, **EPB42**, **DMTN**, **ALAS2**, **FECH**, **ANK1**, **HBZ**, LYN, MMP9, NFKBIA, PF4, **PIP4K2A**, **TESC**, RELB, **BPGM**, TFRC
reactive oxygen species metabolic process	3.23	3.40	10	106	**GPX1**, **HBB**, **MPO**, PID1, RAC2, **SNCA**, SOD2, **PRDX2**, **STK17A**, CD36
cofactor metabolic process	3.20	4.14	12	129	NAMPT, **ALAS2**, ACSL1, **FECH**, **RPIA**, **EIF2AK1**, **ANK1**, **SLC25A39**, **BLVRB**, **SNCA**, **TALDO1**, **ELOVL6**
cellular response to peptide	3.19	4.88	14	152	**KLF1**, JAK3, **ARG1**, LYN, **ADIPOR1**, **KLF3**, **ATP6V0C**, **PPM1A**, PID1, **STRADB**, **MAPK1**, STAT3, UCP2, SOCS3
positive regulation of cell migration	3.17	5.27	15	164	CORO1A, **DMTN**, FGR, CXCL8, **ITGA2B**, LYN, MMP9, MYO1F, SERPINE1, PDGFRA, **MAPK1**, PTGS2, **CXCL16**, CCL2, **TNFSF14**
positive regulation of locomotion	3.14	5.69	16	177	CORO1A, **DMTN**, FGR, CXCL8, **ITGA2B**, LYN, MMP9, MYO1F, SERPINE1, PDGFRA, **MAPK1**, PTGS2, **CXCL16**, CCL2, **TNFSF14**, **ARTN**
positive regulation of cellular component movement	3.07	5.43	15	169	CORO1A, **DMTN**, FGR, CXCL8, **ITGA2B**, LYN, MMP9, MYO1F, SERPINE1, PDGFRA, **MAPK1**, PTGS2, **CXCL16**, CCL2, **TNFSF14**

Table 15 – Continued on next page

Table 15 – continued from previous page

wound healing	3.01	9.08	24	292	FGR, **FGFR1OP2**, GNA15, **HBB**, **HBD**, **HBE1**, **HPS1**, **ITGA2B**, NFE2, **NINJ2**, PF4, SERPINA1, TREM1, APBB1IP, PPBP, SRGN, **MAPK1**, CD177, RAC2, S100A8, SLC11A1, **BSG**, TIMP1, FERMT3
negative regulation of endopeptidase activity	2.97	4.05	11	126	**CDKN2D**, SERPINA3, CSTA, **GPX1**, **FURIN**, SERPINE1, SERPINA1, **SIAH2**, **SNCA**, TIMP1, **TNFSF14**
coagulation	2.91	7.70	20	246	FGR, GNA15, **HBB**, **HBD**, **HBE1**, **HPS1**, **ITGA2B**, NFE2, PF4, SERPINA1, TREM1, APBB1IP, PPBP, SRGN, **MAPK1**, CD177, RAC2, **BSG**, TIMP1, FERMT3
hemostasis	2.84	7.85	20	251	FGR, GNA15, **HBB**, **HBD**, **HBE1**, **HPS1**, **ITGA2B**, NFE2, PF4, SERPINA1, TREM1, APBB1IP, PPBP, SRGN, **MAPK1**, CD177, RAC2, **BSG**, TIMP1, FERMT3
multicellular organismal homeostasis	2.83	4.63	12	144	CORO1A, SERPINA3, AMPD3, **GPX1**, **KRT1**, LYZ, PTGS2, RAC2, SLC11A1, STAT3, TFRC, TNFAIP3
regulation of apoptotic signaling pathway	2.82	5.85	15	182	**CDKN2D**, **GCLC**, **GPX1**, JAK3, OSM, SERPINE1, PF4, **STRADB**, **BCL2L1**, S100A8, S100A9, **SIAH2**, SOD2, TNFAIP3, **YBX3**
nuclear transport	2.79	4.69	12	146	**NUDT4**, JAK3, NFKBIA, **PPM1A**, **STRADB**, **MAPK1**, PTGS2, BCL3, SLC11A1, STAT3, **TNFSF14**, CD36
cellular response to oxygen-containing compound	2.73	6.45	16	215	**DMTN**, **FECH**, SPIDR, **GPX1**, **HBB**, PIM3, MGST1, **MPO**, PDGFRA, **TESC**, PTGS2, **BCL2L1**, **SNCA**, SOD2, **PRDX2**, CPEB4
response to oxidative stress	2.68	4.86	12	153	**SESN3**, **GCLC**, **HBB**, **ARG1**, **KRT1**, **MPO**, PTGS2, **SNCA**, SOD2, **PRDX2**, TNFAIP3, UCP2

Table 15 – Continued on next page

Table 15 – continued from previous page

regulation of cytokine production	2.67	9.60	23	299	FGR, SULF1, ACKR1, **RGCC**, JAK3, LTB, LYN, NFKB2, NFKBIA, **FURIN**, SERPINE1, **FOXP3**, PF4, SRGN, **ZC3HAV1**, PTGS2, BCL3, CCL2, SLC11A1, TNFAIP3, **PGLYRP1**, IL1RL1, CD36
protein targeting	2.64	6.20	15	193	**DMTN**, **FBXO7**, **ANK1**, JAK3, NFKBIA, **PPM1A**, **TESC**, **MAPK1**, PTGS2, BCL3, SLC11A1, STAT3, **BSG**, **TNFSF14**, CD36
negative regulation of cellular protein metabolic process	2.64	7.07	17	220	**CDKN2D**, CSTA, **DMTN**, **DKK1**, **FBXO7**, **EIF2AK1**, **GCLC**, LYN, **FURIN**, **FOXP3**, **TESC**, PID1, **SNCA**, TIMP1, TNFAIP3, SOCS3, GMFG
single organismal cell-cell adhesion	2.60	6.30	15	196	TNIP1, CSTA, **DMTN**, **RGCC**, **HBB**, ICAM3, **ITGA2B**, LYN, **NINJ2**, PDGFRA, PLEK, RAC2, S100A8, S100A9, FERMT3
leukocyte mediated immunity	2.59	5.01	12	156	FGR, **IL7R**, JAK3, LYN, MYO1F, **FOXP3**, TREM1, BCL3, SLC11A1, STXBP2, C1R, C1S
negative regulation of transport	2.57	6.36	15	198	**DMTN**, **RGCC**, PIM3, MMP9, NFKBIA, OSM, **FOXP3**, **PPM1A**, PID1, SRGN, PTGS2, **SNCA**, UCP2, **TNFSF14**, CD36
cell-type specific apoptotic process	2.49	7.00	16	218	**E2F2**, **GCLC**, **GPX1**, **RGCC**, JAK3, LYN, SERPINE1, **BCL2L1**, CCL2, **SNCA**, SOD2, **PRDX2**, TNFAIP3, CPEB4, **STK17A**, **CDC34**
negative regulation of phosphate metabolic process	2.49	5.65	13	176	TNIP1, **CDKN2D**, **DMTN**, **DKK1**, **FBXO7**, LYN, PLEK, **TESC**, PID1, **SNCA**, SOCS3, GMFG, **CDA**
regulation of protein transport	2.40	6.30	14	196	**DMTN**, FGR, **RGCC**, LYN, NFKBIA, **OAZ1**, **FOXP3**, **PPM1A**, SRGN, PTGS2, BCL3, **TNFSF14**, IL1RL1, CD36
response to cytokine	2.32	8.44	18	270	CORO1A, ACSL1, ACKR1, IL1R1, **IL7R**, JAK3, **ARG1**, **ADIPOR1**, PF4, **MAPK1**, **BCL2L1**, SLC11A1, **SNCA**, TIMP1, IL1R2, SOCS3, IL1RL1, OSMR

Table 15 – Continued on next page

Table 15 – continued from previous page

Next, we carry out the GO enrichment analysis for the significantly down-regulated genes between FIR/IVH/WMD infants and those without any of those phenotypes and who were not administered with ACS.

```
> maskDown <- ttAllIVHWMDnoACS$logFC < 0
> DEgenesDownIVHWMDnoACS <- intersect(ttAllIVHWMDnoACS[maskDown, "ID"], DEgenesIVHWMDnoACS)
> length(DEgenesDownIVHWMDnoACS)
[1] 310
> gUniverse <- ttAllIVHWMDnoACS[maskDown, "ID"]
> length(gUniverse)
[1] 9769
> GOparams <- new("GOHyperGParams", geneIds=DEgenesDownIVHWMDnoACS, universeGeneIds=gUniverse,
                  annotation="org.Hs.eg.db", ontology="BP",
                  pvalueCutoff=goAnalysisCondPvalueCutoff, conditional=TRUE,
                  testDirection="over")
> cache(hgOverGOBPdeGenesDownIVHWMDnoACS <- hyperGTest2(GOparams, minSizeCutoff=5,
                                                        maxSizeCutoff=300, orCutoff=1.5),
        dir=cacheDir_FE, prefix=cachePrefix_FE)
> hgOverGOBPdeGenesDownIVHWMDnoACS
Gene to GO BP Conditional test for over-representation
3284 GO BP ids tested (58 have p < 0.01)
Selected gene set size: 260
    Gene universe size: 7370
    Annotation package: org.Hs.eg
> IQRs <- apply(exprs(cleanEsetIVHWMDnoACS)[gUniverse, ], 1, IQR)
> cache(reportOverGOBPdeGenesDownIVHWMDnoACSFDR10 <- GOreport(hgOverGOBPdeGenesDownIVHWMDnoACS,
                                                              minSize=5, maxSize=300,
                                                              minCount=5, minOddsRatio=1.5,
                                                              featureFilterValues=IQRs,
                                                              ff.R=1000, ff.prob=0.90, maxReporte
                                                              p.value=0.1, p.adjust.method="fdr"
                                                              highlightGenes=setdiff(DEgenesDown
                                                              highlightStr="\\textbf{%s}"),
        dir=cacheDir_FE, prefix=cachePrefix_FE)
> dim(reportOverGOBPdeGenesDownIVHWMDnoACSFDR10)
[1]  0 10
```

As we can see, no enriched GO terms can be found among down-regulated genes. Let's do the analysis again increasing the FDR cutoff up to 20%.

```
> cache(reportOverGOBPdeGenesDownIVHWMDnoACS <- GOreport(hgOverGOBPdeGenesDownIVHWMDnoACS,
                                                         minSize=5, maxSize=300,
                                                         minCount=5, minOddsRatio=1.5,
                                                         featureFilterValues=IQRs,
                                                         ff.R=1000, ff.prob=0.90, maxReported=Inf
                                                         p.value=0.2, p.adjust.method="fdr",
                                                         highlightGenes=setdiff(DEgenesDownIVHWM
                                                         highlightStr="\\textbf{%s}"),
        dir=cacheDir_FE, prefix=cachePrefix_FE)
> dim(reportOverGOBPdeGenesDownIVHWMDnoACS)
[1]  2 10
```

This time, 2 GO terms are significantly enriched at 20% FDR and its description can be found in Table 16.

Table 16. Gene Ontology (GO) enrichment (biological process) of downregulated genes called differentially expressed (DE) at FDR <1% and with a minimum 1.5-fold change between FIR/IVH/WMD and non-FIR/nonIVH/nonWMD affected infants. Column "GeneSyms" contains the genes enriching the corresponding GO term, where those highlighted in boldface are not DE between FIR and nonFIR affected infants at FDR <1%, and therefore, their downregulation is specific to the joint FIR/IVH/WMD condition.

Term	OddsRatio	ExpCount	Count	Size	GeneSyms
negative regulation of neurogenesis	5.93	1.41	7	40	FGF13, FGFR3, **BCL11A**, PTK2, BMPR1A, **HOOK3**, ULK2
negative regulation of cytoskeleton organization	5.44	1.52	7	43	DLC1, PLEKHH2, FGF13, KANK1, **CKAP2**, APC, ARHGAP6

4.2 Enrichment of DE genes in immune reponse pathways

We assess enrichment of DE genes in pathways involved in the immune response. For this purpose we use gene sets describing innate inflammatory pathways defined in (11), from where we excluded the apoptosis signaling gene set. We also include a gene set describing the adaptive inflammatory pathway of the Th17 cell response defined in (12).

```
> library(xlsx)
> cache(inflamGenesPwys <- read.xlsx("annotData/LozaChang07_WFINFLAMtableS1.xls",
                                     sheetIndex=1, startRow=2, stringsAsFactors=FALSE,
                                     check.names=FALSE),
        dir=cacheDir_FE, prefix=cachePrefix_FE)
> head(inflamGenesPwys)
  geneID chrom Trasciption_Start Trasciption_End
1  PTPRU  chr1          29383643        29473927
2  VCAM1  chr1         100897317       100916620
3   CD58  chr1         116769198       116825645
4    CD2  chr1         117009143       117023892
5   MUC1  chr1         151971671       151975732
6   DARC  chr1         155987582       155989361
  5kb upstream of trasciption_Start 2kb downstream of trasciption_End
1                          29378643                          29475927
2                         100892317                         100918620
3                         116767198                         116830645
4                         117004143                         117025892
5                         151969671                         151980732
6                         155982582                         155991361
                                       PROTEIN                  Primary Pathway
1 Protein tyrosine phosphatase receptor type U Adhesion-Extravasation-Migration
2                                        VCAM1 Adhesion-Extravasation-Migration
3                                         CD58 Adhesion-Extravasation-Migration
4                                          CD2 Adhesion-Extravasation-Migration
5                               CD227, mucin-1 Adhesion-Extravasation-Migration
6        Duffy blood group, chemokine receptor Adhesion-Extravasation-Migration
                     Secondary Pathway # of tSNPs in gene_CEU
1                       MAPK signaling                     19
2                                 <NA>                     12
3                  Leukocyte signaling                      4
4                  Leukocyte signaling                      6
5                                 <NA>                      1
```

```
6 G-Protein Coupled Receptor Signaling                      5
  # of tSNPs in gene_YRI
1                    39
2                    33
3                    17
4                     9
5                     3
6                     1
> inflamGenesPwys <- inflamGenesPwys[, c("geneID", "PROTEIN", "Primary Pathway",
                                      "Secondary Pathway")]
> head(inflamGenesPwys)
  geneID                                     PROTEIN
1  PTPRU Protein tyrosine phosphatase receptor type U
2  VCAM1                                       VCAM1
3   CD58                                        CD58
4    CD2                                         CD2
5   MUC1                              CD227, mucin-1
6   DARC        Duffy blood group, chemokine receptor
                 Primary Pathway                  Secondary Pathway
1 Adhesion-Extravasation-Migration                     MAPK signaling
2 Adhesion-Extravasation-Migration                               <NA>
3 Adhesion-Extravasation-Migration                Leukocyte signaling
4 Adhesion-Extravasation-Migration                Leukocyte signaling
5 Adhesion-Extravasation-Migration                               <NA>
6 Adhesion-Extravasation-Migration G-Protein Coupled Receptor Signaling
> dim(inflamGenesPwys)
[1] 1027    4
> unique(inflamGenesPwys$"Primary Pathway")
 [1] "Adhesion-Extravasation-Migration"
 [2] "Apoptosis Signaling"
 [3] "Apoptosis signaling"
 [4] "Calcium Signaling"
 [5] "Complement Cascase"
 [6] "Cytokine signaling"
 [7] "Eicosanoid Signaling"
 [8] "Glucocorticoid/PPAR signaling"
 [9] "G-Protein Coupled Receptor Signaling"
[10] "Innate pathogen detection"
[11] "Leukocyte signaling"
[12] "MAPK signaling"
[13] "Natural Killer Cell Signaling"
[14] "NF-kB signaling"
[15] "Phagocytosis-Ag presentation"
[16] "PI3K/AKT Signaling"
[17] "ROS/Glutathione/Cytotoxic granules"
[18] "TNF Superfamily Signaling"
> length(unique(inflamGenesPwys$"Primary Pathway"))
[1] 18
> unique(inflamGenesPwys$"Secondary Pathway")
 [1] "MAPK signaling"
 [2] NA
 [3] "Leukocyte signaling"
 [4] "G-Protein Coupled Receptor Signaling"
 [5] "PI3K/AKT Signaling"
```

```
 [6] "TNF Superfamily Signaling"
 [7] "NF-kB signaling"
 [8] "Glucocorticoid/PPAR signaling"
 [9] "Apoptosis Signaling"
[10] "Innate pathogen detection"
[11] "Glucocorticoid signaling"
[12] "Apoptosis signaling"
[13] "Natural Killer Cell Signaling"
[14] "Cytokine signaling"
> length(unique(inflamGenesPwys$"Secondary Pathway"))
[1] 14
> length(unique(inflamGenesPwys$geneID))
[1] 1027
```

We have to first fix words that differ by letter capitalization and a couple of features in the gene idenfiers stored in the `geneID` column, such as remove spaces and replace uppercase `ORF` by lowercase `orf`:

```
> inflamGenesPwys$"Primary Pathway" <- gsub("signaling", "Signaling",
                                      inflamGenesPwys$"Primary Pathway")
> inflamGenesPwys$"Secondary Pathway" <- gsub("signaling", "Signaling",
                                        inflamGenesPwys$"Secondary Pathway")
> inflamGenesPwys$geneID <- gsub(" ", "", inflamGenesPwys$geneID)
> inflamGenesPwys$geneID <- gsub("ORF", "orf", inflamGenesPwys$geneID)
```

We first try to map the identifiers in the `geneID` column to Entrez identifiers as HGNC official symbols:

```
> egIDsFromSymbols <- select(org.Hs.eg.db, columns="ENTREZID",
                             keys=inflamGenesPwys$geneID, keytype="SYMBOL")
> dim(egIDsFromSymbols)
[1] 1029    2
> missingIDs <- egIDsFromSymbols$SYMBOL[is.na(egIDsFromSymbols$ENTREZID)]
> length(missingIDs)
[1] 60
> head(missingIDs)
[1] "DARC"  "MENA"  "BLR1"  "MLCK"  "CCBP2" "VIL2"
```

There are 60 genes whose symbols cannot be mapped to Entrez gene identifiers. We are going to try map those using aliases for these gene symbols, as follows.

```
> egIDsFromAliases <- select(org.Hs.eg.db, columns="ENTREZID",
                             keys=missingIDs, keytype="ALIAS")
> dim(egIDsFromAliases)
[1] 68  2
> missingIDs <- egIDsFromAliases$ALIAS[is.na(egIDsFromAliases$ENTREZID)]
> length(missingIDs)
[1] 1
> head(missingIDs)
[1] "KIAA1271"
```

There is (are) still 1 gene(s) whose symbols cannot be mapped to Entrez gene identifiers. The `PROTEIN` column in the original file seems to contain the description of the function and also sometimes an alternative symbol for the gene or protein. We are going to try to map this to an alias.

```
> mt <- match(inflamGenesPwys$geneID,
              egIDsFromAliases$ALIAS[is.na(egIDsFromAliases$ENTREZID)])
> missingIDs2 <- inflamGenesPwys[!is.na(mt), c("geneID", "PROTEIN")]
> colnames(missingIDs2) <- c("geneID", "ALIAS")
> missingIDs2
```

```
      geneID  ALIAS
567 KIAA1271 CARDIF
> egIDsFromAliases2 <- select(org.Hs.eg.db, columns="ENTREZID",
                              keys=missingIDs2$ALIAS, keytype="ALIAS")
> egIDsFromAliases2 <- merge(missingIDs2, egIDsFromAliases2)[, c("geneID", "ENTREZID")]
> colnames(egIDsFromAliases2) <- c("ALIAS", "ENTREZID")
> egIDsFromAliases <- rbind(egIDsFromAliases[!is.na(egIDsFromAliases$ENTREZID), ],
                            egIDsFromAliases2)
> sum(is.na(egIDsFromAliases$ENTREZID))
[1] 0
```

So, there is finally 0 gene(s) whose symbols cannot be mapped to Entrez gene identifiers. We rebuild now the table of inflammatory genes and pathways anchoring them at Entrez gene identifiers.

```
> egIDsFromSymbols <- egIDsFromSymbols[!is.na(egIDsFromSymbols$ENTREZID), ]
> egIDsFromAliases <- egIDsFromAliases[!is.na(egIDsFromAliases$ENTREZID), ]
> inflamGenesPwys1 <- merge(inflamGenesPwys, egIDsFromSymbols, by.x="geneID", by.y="SYMBOL")
> inflamGenesPwys2 <- merge(inflamGenesPwys, egIDsFromAliases, by.x="geneID", by.y="ALIAS")
> maskNAs <- is.na(match(inflamGenesPwys$geneID,
                         c(inflamGenesPwys1$geneID, inflamGenesPwys2$geneID)))
> inflamGenesPwys <- rbind(inflamGenesPwys1, inflamGenesPwys2,
                           cbind(inflamGenesPwys[maskNAs, ],
                                 ENTREZID=rep(NA_character_, times=sum(maskNAs))))
> inflamGenesPwys <- unique(inflamGenesPwys[, c("ENTREZID", "Primary Pathway",
                                                "Secondary Pathway")])
> inflamGenesPwys <- data.frame(EntrezID=inflamGenesPwys$ENTREZID,
                                Symbol=select(org.Hs.eg.db, columns="SYMBOL",
                                              keys=inflamGenesPwys$ENTREZID,
                                              keytype="ENTREZID")$SYMBOL,
                                PrimaryPathway=inflamGenesPwys$"Primary Pathway",
                                SecondaryPathway=inflamGenesPwys$"Secondary Pathway",
                                stringsAsFactors=FALSE)
> dim(inflamGenesPwys)
[1] 1034    4
> head(inflamGenesPwys)
  EntrezID Symbol                   PrimaryPathway SecondaryPathway
1    53947 A4GALT             Apoptosis Signaling             <NA>
2       25   ABL1             Leukocyte Signaling             <NA>
3     1636    ACE              Cytokine Signaling             <NA>
4    22985  ACIN1             Apoptosis Signaling             <NA>
5     6868 ADAM17              Cytokine Signaling             <NA>
6      101  ADAM8 Adhesion-Extravasation-Migration             <NA>
```

To have a set of inflammatory pathways involved exclusively in the innate immune response we are going to exclude the “Apoptosis Signaling pathway” and stored it separately as a targeted gene set.

```
> apopGenes <- inflamGenesPwys[grep("Apoptosis", inflamGenesPwys$PrimaryPathway), "EntrezID"]
> head(apopGenes)
[1] "53947" "22985" "317"   "572"   "578"   "581"
> length(apopGenes)
[1] 69
> inflamGenesPwys <- inflamGenesPwys[-grep("Apoptosis", inflamGenesPwys$PrimaryPathway), ]
> dim(inflamGenesPwys)
[1] 965   4
> save(apopGenes, file=file.path("annotData", "apopGenes.RData"))
```

Now we import a gene set defining the Th17 cell response described in (12), which in this case corresponds to an adaptive inflammatory pathway.

```
> th17Genes <- read.xlsx(file.path("annotData", "CiofaniLittman12_th17GenesTableS1.xlsx"),
                         sheetIndex=1, stringsAsFactors=FALSE)
> head(th17Genes)
   Gene    Gene.type   Effect                         PMID
1   AHR           TF positive           18362915, 18362914
2  Batf           TF positive                     19578362
3  Bcl3           TF negative                     20622172
4  BCL6           TF negative 19628815, 19631565, 20212093
5 CCL20    chemokine positive           19050256, 18025126
6  Ccr6 cell surface positive           19050256, 18025126
                                                                 Evidence
1                                                              deficiency
2                                                              deficiency
3                                                              deficiency
4                                                              deficiency
5 ligand deficiency, recombinant protein, neutralizing antibody to receptor
6                                    deficiency, neutralizing antibody
> th17Genes$Gene <- gsub(" ", "", toupper(th17Genes$Gene))
> th17Genes <- unique(th17Genes$Gene)
> sym2eg <- select(org.Hs.eg.db, keys=th17Genes, columns="ENTREZID", keytype="SYMBOL")
> sym2eg <- split(sym2eg$ENTREZID, sym2eg$SYMBOL)
> sym2eg <- sapply(sym2eg,
                   function(x, tt) {
                     if (any(!is.na(x))) {
                       lfc <- abs(tt$logFC)[tt$ID %in% x]
                       if (length(lfc) > 0)
                         x <- tt$ID[tt$ID %in% x][which.max(lfc)]
                       else
                         x <- x[1]
                     }
                     x
                   }, ttAll)
> th17Genes <- sym2eg
> head(th17Genes)
    AHR    BATF    BCL3    BCL6   CCL20    CCR6
  "196" "10538"   "602"   "604"  "6364"  "1235"
> length(th17Genes)
[1] 73
> head(inflamGenesPwys)
  EntrezID Symbol                        PrimaryPathway SecondaryPathway
2       25   ABL1                  Leukocyte Signaling             <NA>
3     1636    ACE                   Cytokine Signaling             <NA>
5     6868 ADAM17                   Cytokine Signaling             <NA>
6      101  ADAM8     Adhesion-Extravasation-Migration             <NA>
7      107  ADCY1 G-Protein Coupled Receptor Signaling             <NA>
8      108  ADCY2 G-Protein Coupled Receptor Signaling             <NA>
> dim(inflamGenesPwys)
[1] 965   4
> inflamGenesPwys <- rbind(inflamGenesPwys,
                           data.frame(EntrezID=th17Genes, Symbol=names(th17Genes),
                                      PrimaryPathway=rep("Th17 cell response",
                                                         length(th17Genes)),
```

```
                                          SecondaryPathway=rep(NA_character_,
                                                           length(th17Genes))))
> dim(inflamGenesPwys)
[1] 1038    4
```

Store as a text file the set of innate and adaptive inflammatory pathways.

```
> write.table(inflamGenesPwys, file="annotData/innateInflamGenesPwys.txt",
              sep="\t", quote=FALSE, col.names=TRUE, row.names=FALSE)
```

Finally, we perform Fisher's exact tests to assess enrichment of these pathways by the lists of signifincantly up and down regulated genes.

```
> table(inflamGenesPwys$PrimaryPathway)
    Adhesion-Extravasation-Migration                    Calcium Signaling
                                 145                                   14
                  Complement Cascase                   Cytokine Signaling
                                  40                                  172
                 Eicosanoid Signaling        Glucocorticoid/PPAR Signaling
                                  39                                   21
G-Protein Coupled Receptor Signaling             Innate pathogen detection
                                  42                                   51
                  Leukocyte Signaling                       MAPK Signaling
                                 122                                  118
       Natural Killer Cell Signaling                       NF-kB Signaling
                                  32                                   33
         Phagocytosis-Ag presentation                   PI3K/AKT Signaling
                                  39                                   37
  ROS/Glutathione/Cytotoxic granules                   Th17 cell response
                                  22                                   73
            TNF Superfamily Signaling
                                  38
> genesByPwys <- split(as.character(inflamGenesPwys$EntrezID),
                       inflamGenesPwys$PrimaryPathway)
> iPwys <- do.call("GeneSetCollection",
                   mapply(function(geneSetID, geneIDs) {
                     GeneSet(EntrezIdentifier("org.Hs.eg.db"),
                       geneIds=geneIDs,
                       setName=geneSetID)
                   }, names(genesByPwys),
                   genesByPwys, USE.NAMES=FALSE))
> funEnrichment <- function(gsc, genes, universe, minsize=1) {
    stopifnot(all(genes %in% universe))
    ans <- lapply(gsc, function(geneSet, genes, universe) {
                         N <- length(universe)
                         m <- length(intersect(geneIds(geneSet), universe))
                         n <- length(genes)
                         k <- length(intersect(geneIds(geneSet), genes))
                         t <- array(c(k, n - k, m - k, N + k - n - m), dim = c(2, 2))
                         ans <- fisher.test(t, alternative="greater")
                         c(OddsRatio=as.vector(ans$estimate),
                           Pvalue=ans$p.value, Counts=k, Size=m)
                       }, genes, gUniverse)
    ans <- as.data.frame(do.call("rbind", ans))
    rownames(ans) <- names(gsc)
    ans <- cbind(ans, FDR=p.adjust(ans$Pvalue, method="fdr"))
```

```
    mask <- ans$Counts >= minsize
    ans <- ans[mask, ]
    ans[order(ans$OddsRatio, decreasing=TRUE), ]
  }
> DEgenesUp <- DEgenes[ttAllIQRfilt[DEgenes, "logFC"] > 0]
> gUniverse <- ttAll[ttAll$logFC > 0, "ID"]
> iPwysFEup <- funEnrichment(iPwys, DEgenesUp, gUniverse, minsize=5)
> iPwysFEup
                                   OddsRatio       Pvalue Counts Size
Complement Cascase                  8.387679 3.336885e-06     10   31
Cytokine Signaling                  8.057515 2.416246e-20     41  135
TNF Superfamily Signaling           7.536651 2.025035e-05      9   30
ROS/Glutathione/Cytotoxic granules  7.284081 1.686958e-03      5   17
Th17 cell response                  7.031938 1.123693e-08     17   60
Adhesion-Extravasation-Migration    6.847685 1.083382e-14     32  117
NF-kB Signaling                     6.815932 2.789349e-04      7   25
Innate pathogen detection           6.069035 3.219721e-05     10   39
Leukocyte Signaling                 5.647744 1.129649e-09     23   96
Eicosanoid Signaling                4.770669 3.428558e-03      6   28
Phagocytosis-Ag presentation        3.232718 2.802634e-02      5   32
MAPK Signaling                      2.412985 2.623691e-02      8   66
                                            FDR
Complement Cascase                 1.134541e-05
Cytokine Signaling                 4.107618e-19
TNF Superfamily Signaling          5.737599e-05
ROS/Glutathione/Cytotoxic granules 3.186477e-03
Th17 cell response                 4.775697e-08
Adhesion-Extravasation-Migration   9.208749e-14
NF-kB Signaling                    5.927366e-04
Innate pathogen detection          7.819322e-05
Leukocyte Signaling                6.401346e-09
Eicosanoid Signaling               5.828549e-03
Phagocytosis-Ag presentation       3.970398e-02
MAPK Signaling                     3.970398e-02
> DEgenesDown <- DEgenes[ttAllIQRfilt[DEgenes, "logFC"] < 0]
> gUniverse <- ttAll[ttAll$logFC < 0, "ID"]
> iPwysFEdown <- funEnrichment(iPwys, DEgenesDown, gUniverse, minsize=5)
> iPwysFEdown
[1] OddsRatio Pvalue    Counts    Size      FDR
<0 rows> (or 0-length row.names)
```

The previous enrichment analysis with the Fisher's exact test shows that there are no inflammatory gene sets with at least 5 significant down-regulated genes, and 12 are enriched for significant up-regulated genes at a FDR$<$ 10%.

In Figure 15 we have the expression values of the DE genes annotated to each of the enriched pathways, compared between FIR and non-FIR samples.

We are going to assess now whether all these inflammatory pathways together are enriched for DE genes. We consider first as DE all genes with significant expression changes at FDR$<$1%:

```
> maskDEpwys <- iPwysFEup$FDR < 0.1 & iPwysFEup$OddsRatio > 1.5
> genesByDEiPwys <- lapply(genesByPwys[rownames(iPwysFEup)[maskDEpwys]],
                           function(x, g) x[!is.na(match(x, g))], featureNames(frmaEsetFilt))
> sigGenes <- ttAllIQRfilt$ID[ttAllIQRfilt$adj.P.Val < FDRcutoff]
> N <- nrow(ttAll)
> m <- length(intersect(unique(unlist(genesByDEiPwys, use.names=FALSE)), ttAll$ID))
> n <- length(sigGenes)
```

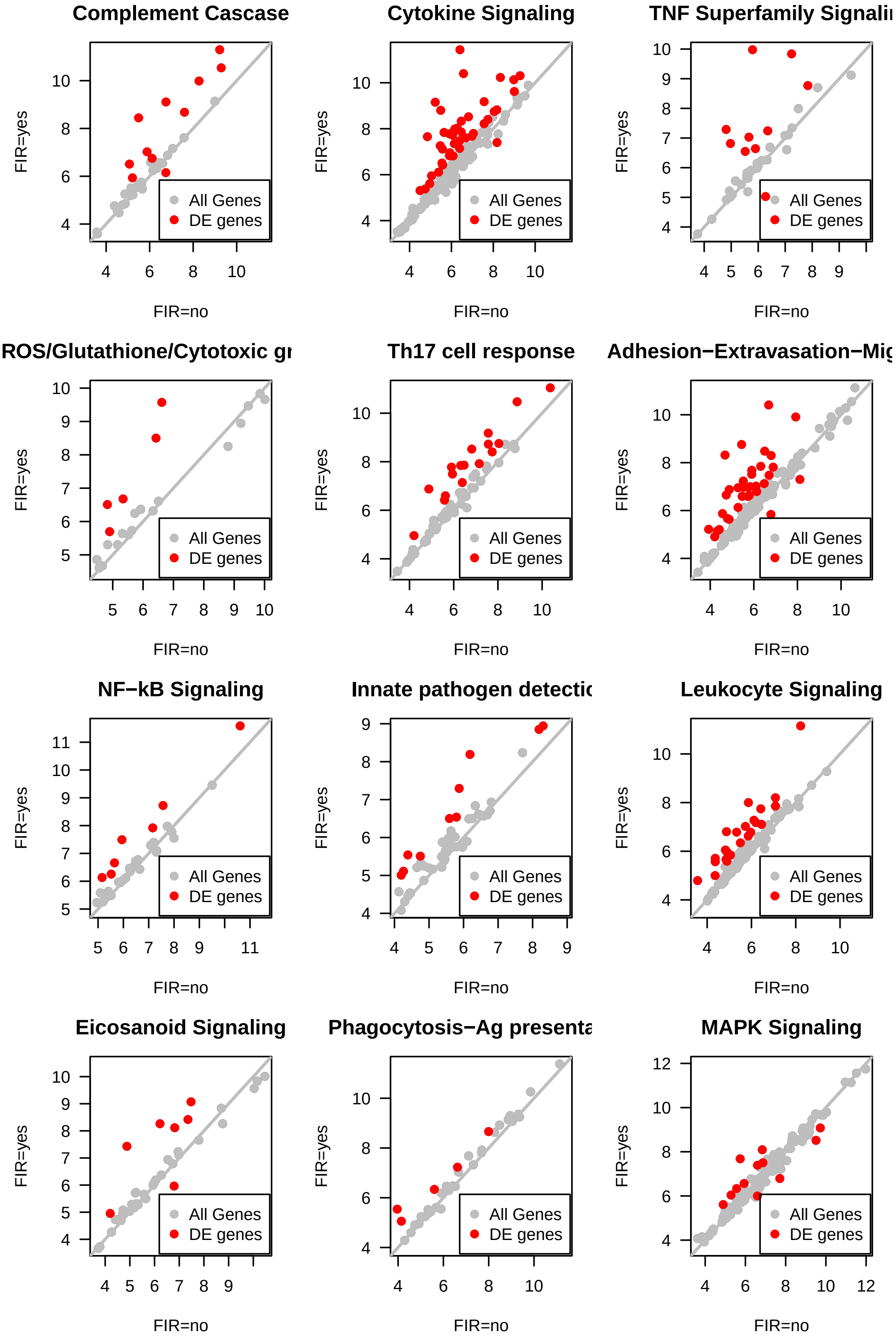

Figure 15. Expression values of genes belonging to inflammatory pathways significantly enriched by genes called DE between FIR and nonaffected-FIR infants at FDR<1% and minimum 1.5-fold change.

```
> k <- length(intersect(sigGenes, unique(unlist(genesByDEiPwys, use.names=FALSE))))
> t <- array(c(k, n - k, m - k, N + k - n - m), dim = c(2, 2),
             dimnames=list(InflamSenesc=c("in", "out"), DE=c("yes", "no")))
> t
            DE
InflamSenesc  yes    no
         in   308   514
         out 3588 15745
> fisher.test(t, alternative="greater")
        Fisher's Exact Test for Count Data

data:  t
p-value < 2.2e-16
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
 2.319274      Inf
sample estimates:
odds ratio
  2.629564
```

Thus, DE inflammatory pathways are enriched for DE genes at FDR<1%. Let's test now for up-regulated genes significantly DE at the same level:

```
> sigGenes <- ttAllIQRfilt$ID[ttAllIQRfilt$adj.P.Val < FDRcutoff & ttAllIQRfilt$logFC > 0]
> N <- nrow(ttAll[ttAll$logFC > 0, ])
> m <- length(intersect(unique(unlist(genesByDEiPwys, use.names=FALSE)),
                        ttAll$ID[ttAll$logFC > 0]))
> n <- length(sigGenes)
> k <- length(intersect(ttAllIQRfilt$ID[ttAllIQRfilt$logFC > 0],
                        intersect(sigGenes, unique(unlist(genesByDEiPwys, use.names=FALSE)))))
> t <- array(c(k, n - k, m - k, N + k - n - m), dim = c(2, 2),
             dimnames=list(InflamSenesc=c("in", "out"), DE=c("yes", "no")))
> t
            DE
InflamSenesc  yes   no
         in   265  369
         out 1214 9020
> fisher.test(t, alternative="greater")
        Fisher's Exact Test for Count Data

data:  t
p-value < 2.2e-16
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
 4.611766      Inf
sample estimates:
odds ratio
  5.334608
```

Up-regulated genes at FDR<1% enrich significantly DE inflammatory pathways with a larger odds ratio than all up and down regulated genes. Finally, let's test for up-regulated DE genes at FDR<1% and with a minimum 1.5-fold change:

```
> sigGenes <- ttAllIQRfilt$ID[ttAllIQRfilt$adj.P.Val < FDRcutoff &
                              ttAllIQRfilt$logFC > log2(fcCutoff)]
> N <- nrow(ttAll[ttAll$logFC > 0, ])
```

```
> m <- length(intersect(unique(unlist(genesByDEiPwys, use.names=FALSE)), ttAll$ID[ttAll$logFC >
> n <- length(sigGenes)
> k <- length(intersect(ttAllIQRfilt$ID[ttAllIQRfilt$logFC > 0],
                        intersect(sigGenes, unique(unlist(genesByDEiPwys, use.names=FALSE)))))
> t <- array(c(k, n - k, m - k, N + k - n - m), dim = c(2, 2),
             dimnames=list(Inflammatory=c("in", "out"), DE=c("yes", "no")))
> t
            DE
Inflammatory yes   no
         in  159  475
         out 433 9801
> fisher.test(t, alternative="greater")
        Fisher's Exact Test for Count Data

data:  t
p-value < 2.2e-16
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
 6.342848      Inf
sample estimates:
odds ratio
  7.574603
```

Inflammatory pathways show again a significant enrichment for genes called DE at FDR<1% and with a minimum 1.5-fold change with a larger odds ratio than the other two subsets of DE genes.

We cross-classify the set of significantly up-regulated DE genes into the different inflammatory pathways they enrich. This cross-classification is shown in Table 17.

4.3 Enrichment of DE genes in apoptosis and senescence pathways

We import gene sets associated to different related senescence pathways defined in the PathCards database (13). The gene sets defining these pathways were downloaded from (14).

```
> if (!file.exists(file.path("annotData", "sncPwys.RData"))) {
    sncPwys <- vector(mode="list", length=4)
    for (i in 1:4)
      sncPwys[[i]] <- read.xlsx(file.path("annotData", "senescenceGeneSetCompletePathCards.xls")
                              sheetIndex=i, stringsAsFactors=FALSE, header=FALSE)
    sncPwys <- lapply(sncPwys, unlist, use.names=FALSE)
    sncPwys <- lapply(sncPwys, function(x) { x <- x[!is.na(x)] ; x <- unique(gsub(" ", "", x)) ;
    names(sncPwys) <- c("CellularResponseStress", "CellularSenescence", "OxidativeStressInduced"

    sncPwys <- lapply(sncPwys, function(x) {
                        x <- select(org.Hs.eg.db, keys=x, columns="ENTREZID", keytype="SYMBOL")
                        missingIDs <- x$SYMBOL[is.na(x$ENTREZID)]
                        entrezIDs <- x$ENTREZID[!is.na(x$ENTREZID)]
                        if (length(grep("ENSG", missingIDs)) > 0) {
                          ensgmissingIDs <- missingIDs[grep("ENSG", missingIDs)]
                          missingIDs <- missingIDs[-grep("ENSG", missingIDs)]
                          y <- select(org.Hs.eg.db, keys=ensgmissingIDs, columns="ENTREZID", key
                          entrezIDs <- c(entrezIDs, y$ENTREZID[!is.na(y$ENTREZID)])
                          missingIDs <- c(missingIDs, y$ENSEMBL[is.na(y$ENTREZID)])
                        }
                        if (length(missingIDs) > 0) {
                          y <- select(org.Hs.eg.db, keys=missingIDs, columns="ENTREZID", keytype
                          entrezIDs <- c(entrezIDs, y$ENTREZID)
```

Table 17. Significantly upregulated genes called differentially expressed at FDR $< 1\%$ with a minimum 1.5-fold change cross-classified into inflammatory pathways selected by enrichment analysis at FDR $< 10\%$ and OR > 1.5.

Pathway	OR	P.value	Counts	Size	Genes
Complement Cascase	8.40	3.3e-06	10	31	C1R, C1S, C3, C3AR1, C5AR1, CD55, CFB, CFP, SERPINA1, SERPINE1
Cytokine Signaling	8.10	2.4e-20	41	135	CEBPB, CSF1R, CSF2RB, CSF3, CSF3R, CXCL8, CXCR2, IFNAR2, IFNGR1, IFNGR2, IL10RA, IL17RA, IL18R1, IL18RAP, IL1A, IL1B, IL1R1, IL1R2, IL1RL1, IL1RN, IL32, IL4R, IL6, IL6ST, IRF1, JAK3, LIF, NFIL3, NMI, OSM, OSMR, PDGFRA, PTPN2, S100A12, S100A8, S100A9, SOCS1, SOCS3, STAT3, TGFBR2, VEGFA
TNF Superfamily Signaling	7.50	2.0e-05	9	30	PTX3, TNFAIP3, TNFAIP6, TNFRSF11B, TNFRSF1B, TNFSF13B, TNFSF15, TNIP1, TRAF3
ROS/Glutathione/Cytotoxic granules	7.30	1.7e-03	5	17	ANPEP, CYBB, GPX3, NCF2, SOD2
Th17 cell response	7.00	1.1e-08	17	60	BCL3, BCL6, CCL20, HIF1A, IFNGR1, IL1R1, ITGB2, JAK3, LIF, NFKB1, NFKBIZ, PTGER4, SOCS1, SOCS3, STAT3, TGFBR2, VDR
Adhesion-Extravasation-Migration	6.80	1.1e-14	32	117	ACKR1, ADAM8, CCL2, CCL20, CCL4, CCL5, CCL8, CCR1, CCR2, CD36, CD48, CEACAM1, CXCL1, CXCL2, CXCL3, CXCL5, ICAM1, ICAM3, ITGA2, ITGAM, ITGAX, ITGB2, MMP10, MMP19, MMP9, PLAUR, PPBP, RASSF5, SELE, SELL, SELPLG, VCAM1
NF-kB Signaling	6.80	2.8e-04	7	25	BCL3, BCL6, NFKB1, NFKB2, NFKBIA, NFKBIE, RELB
Innate pathogen detection	6.10	3.2e-05	10	39	CD14, IRAK1, IRAK3, LY96, MYD88, NLRC4, NOD2, PYCARD, TLR2, TLR8
Leukocyte Signaling	5.60	1.1e-09	23	96	CD37, CD44, CD52, CD53, CD86, FCER1G, FCGR2A, FCGR2B, LCP2, LILRA2, LILRA3, LILRA6, LILRB2, NAMPT, PIK3AP1, PTPRC, SECTM1, SEMA4D, SIGLEC10, SIRPA, SYK, TCIRG1, TREM1
Eicosanoid Signaling	4.80	3.4e-03	6	28	ALOX5, ALOX5AP, PLA2G2A, PTGER4, PTGES, PTGS2
Phagocytosis-Ag presentation	3.20	2.8e-02	5	32	CD1D, CTSS, PSMB9, TAP1, TAPBP
MAPK Signaling	2.40	2.6e-02	8	66	IFI16, LYN, MAP3K5, MYC, PRKCD, RAC2, RPS6KA1, SHC1

```
                        }
                        unique(entrezIDs)
                      })
    save(sncPwys, file=file.path("annotData", "sncPwys.RData"))
  } else
    load(file.path("annotData", "sncPwys.RData"))
```

These pathways have the following numbers of Entrez genes annotated to them:

```
> sapply(sncPwys, length)
CellularResponseStress     CellularSenescence OxidativeStressInduced 
                   244                    197                    127 
                  SASP 
                   112 
```

We build a `GeneSetCollection` object with these pathways and the apoptosis gene set, and assess the null hypothesis that none of them is enriched with at least 5 up or down regulated DE genes and a minimum OR>1.5 at FDR$<10\%$:

```
> sncApoPwys <- sncPwys
> sncApoPwys$Apoptosis <- apopGenes
> sncApoPwys <- do.call("GeneSetCollection",
                   mapply(function(geneSetID, geneIDs) {
                     GeneSet(EntrezIdentifier("org.Hs.eg.db"),
                       geneIds=geneIDs,
                       setName=geneSetID)
                   }, names(sncApoPwys),
                   sncApoPwys, USE.NAMES=FALSE))
> DEgenesUp <- DEgenes[ttAllIQRfilt[DEgenes, "logFC"] > 0]
> gUniverse <- ttAll[ttAll$logFC > 0, "ID"]
> sncApoPwysFEup <- funEnrichment(sncApoPwys, DEgenesUp, gUniverse, minsize=5)
> sncApoPwysFEup
                       OddsRatio       Pvalue Counts Size          FDR
SASP                    4.677228 0.0008447174      8   38 0.0021117934
Apoptosis               4.250972 0.0014290837      8   41 0.0023818062
CellularResponseStress  3.271414 0.0001989556     15   96 0.0009947782
CellularSenescence      2.830182 0.0053956450     10   72 0.0067445562
> DEgenesDown <- DEgenes[ttAllIQRfilt[DEgenes, "logFC"] < 0]
> gUniverse <- ttAll[ttAll$logFC < 0, "ID"]
> sncApoPwysFEdown <- funEnrichment(sncApoPwys, DEgenesDown, gUniverse, minsize=5)
> sncApoPwysFEdown
                       OddsRatio   Pvalue Counts Size       FDR
CellularResponseStress  1.135733 0.443395      6   98 0.7485706
```

We can see that 4 and 0 gene sets are enriched with 5 or more up and down regulated DE genes, respectively, at FDR$<10\%$ and with a minimum OR>1.5.

In Figure 16 we have the expression values of the DE genes annotated to each of the enriched pathways, compared between FIR and non-FIR samples.

We are going to assess now whether these senescence and apoptosis pathways together are enriched for DE genes. We consider first as DE all genes with significant expression changes at FDR<1%:

```
> maskDEpwys <- sncApoPwysFEup$FDR < 0.1 & sncApoPwysFEup$OddsRatio > 1.5
> genesByDEsncApoPwys <- lapply(geneIds(sncApoPwys)[rownames(sncApoPwysFEup)[maskDEpwys]],
                                function(x, g) x[!is.na(match(x, g))], featureNames(frmaEsetFilt
> sigGenes <- ttAllIQRfilt$ID[ttAllIQRfilt$adj.P.Val < FDRcutoff]
> N <- nrow(ttAll)
```

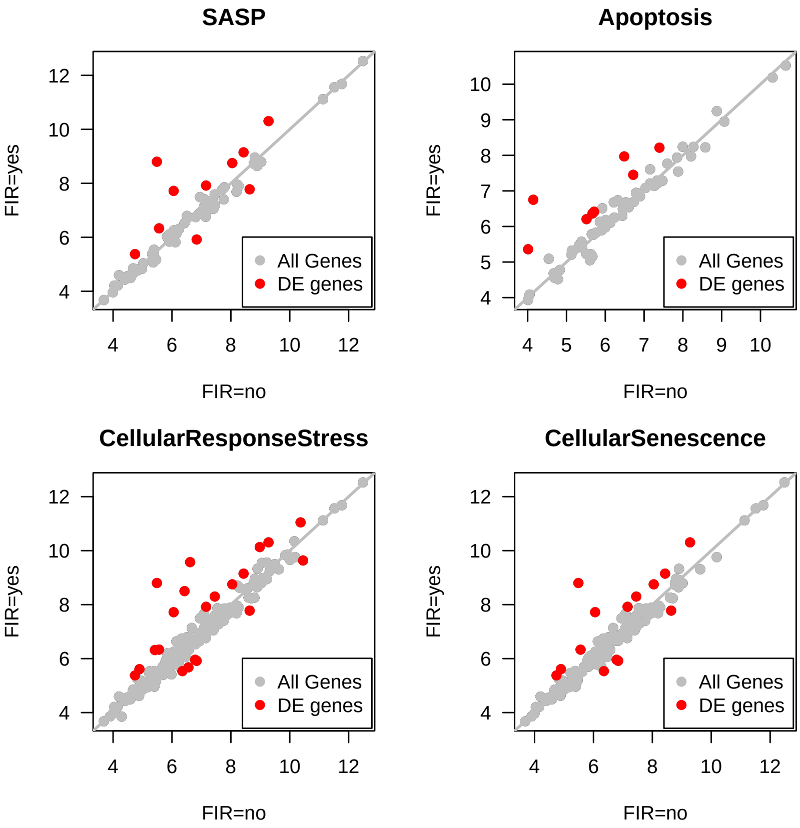

Figure 16. Expression values of genes belonging to apoptosis and senescence pathways significantly enriched by genes called DE between FIR and nonaffected-FIR infants at FDR<1% and minimum 1.5-fold change.

```
> m <- length(intersect(unique(unlist(genesByDEsncApoPwys, use.names=FALSE)), ttAll$ID))
> n <- length(sigGenes)
> k <- length(intersect(sigGenes, unique(unlist(genesByDEsncApoPwys, use.names=FALSE))))
> t <- array(c(k, n - k, m - k, N + k - n - m), dim = c(2, 2),
             dimnames=list(InflamSenesc=c("in", "out"), DE=c("yes", "no")))
> t
            DE
InflamSenesc  yes    no
         in    84   176
         out 3812 16083
> fisher.test(t, alternative="greater")
        Fisher's Exact Test for Count Data

data:  t
p-value = 4.003e-07
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
 1.598015      Inf
sample estimates:
odds ratio
  2.013625
```

Thus, DE senescence and apoptosis pathways are enriched for DE genes at FDR<1%. Let's test now for up-regulated genes significantly DE at the same level:

```
> sigGenes <- ttAllIQRfilt$ID[ttAllIQRfilt$adj.P.Val < FDRcutoff & ttAllIQRfilt$logFC > 0]
> N <- nrow(ttAll[ttAll$logFC > 0, ])
> m <- length(intersect(unique(unlist(genesByDEsncApoPwys, use.names=FALSE)),
                        ttAll$ID[ttAll$logFC > 0]))
> n <- length(sigGenes)
> k <- length(intersect(ttAllIQRfilt$ID[ttAllIQRfilt$logFC > 0],
                        intersect(sigGenes,
                                  unique(unlist(genesByDEsncApoPwys, use.names=FALSE)))))
> t <- array(c(k, n - k, m - k, N + k - n - m), dim = c(2, 2),
             dimnames=list(InflamSenesc=c("in", "out"), DE=c("yes", "no")))
> t
            DE
InflamSenesc  yes   no
         in    43   92
         out 1436 9297
> fisher.test(t, alternative="greater")
        Fisher's Exact Test for Count Data

data:  t
p-value = 3.393e-08
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
 2.176907      Inf
sample estimates:
odds ratio
  3.025565
```

Up-regulated genes at FDR<1% enrich significantly DE senescence and apoptosis pathways with a larger odds ratio than all up and down regulated genes. Finally, let's test for up-regulated DE genes at FDR<1% and with a minimum 1.5-fold change:

```
> sigGenes <- ttAllIQRfilt$ID[ttAllIQRfilt$adj.P.Val < FDRcutoff &
                              ttAllIQRfilt$logFC > log2(fcCutoff)]
> N <- nrow(ttAll[ttAll$logFC > 0, ])
> m <- length(intersect(unique(unlist(genesByDEsncApoPwys, use.names=FALSE)),
                        ttAll$ID[ttAll$logFC > 0]))
> n <- length(sigGenes)
> k <- length(intersect(ttAllIQRfilt$ID[ttAllIQRfilt$logFC > 0],
                        intersect(sigGenes,
                                  unique(unlist(genesByDEsncApoPwys, use.names=FALSE)))))
> t <- array(c(k, n - k, m - k, N + k - n - m), dim = c(2, 2),
             dimnames=list(Inflammatory=c("in", "out"), DE=c("yes", "no")))
> t
            DE
Inflammatory yes    no
         in   23   112
         out 569 10164
> fisher.test(t, alternative="greater")
        Fisher's Exact Test for Count Data

data:  t
p-value = 9.324e-07
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
 2.40281     Inf
```

```
sample estimates:
odds ratio
  3.667361
```

Inflammatory pathways show again a significant enrichment for genes called DE at FDR<1% and with a minimum 1.5-fold change with a larger odds ratio than the other two subsets of DE genes.

We cross-classify the set of significantly up-regulated DE genes into the the apoptosis and senescence pathways they are enriching. This cross-classification is shown in Table 18.

Table 18. Significantly upregulated genes called differentially expressed at FDR $< 1\%$ with a minimum 1.5-fold change cross-classified into apoptosis and senescence pathways selected by enrichment analysis at FDR $< 10\%$ and OR > 1.5.

Pathway	OR	P.value	Counts	Size	Genes
SASP	4.70	8.4e-04	8	38	CDKN1A, CEBPB, CXCL8, IL1A, IL6, NFKB1, RPS6KA1, STAT3
Apoptosis	4.30	1.4e-03	8	41	BCL2A1, BID, BIRC3, CFLAR, MCL1, TNFRSF10B, TNFRSF10D, TNFRSF21
CellularResponseStress	3.30	2.0e-04	15	96	CDKN1A, CEBPB, CXCL8, ERO1L, GPX3, HIF1A, IL1A, IL6, KDM6B, MAP3K5, NFKB1, RPS6KA1, SOD2, STAT3, VEGFA
CellularSenescence	2.80	5.4e-03	10	72	CDKN1A, CEBPB, CXCL8, IL1A, IL6, KDM6B, MAP3K5, NFKB1, RPS6KA1, STAT3

4.4 Enrichment of DE genes among targets of FMRP (FMR1 gene)

We assess enrichemnt of DE genes among targets of the fragile X syndrom protein (FMRP), encoded by the FMR1 gene. These targets were obtained from the publication by (15).

```
> library(xlsx)
> if (!file.exists("annotData/fmrpTargets.RData")) {
    ## in the Excel file from TableS2 in Darnel et al. (2011)
    ## gene symbol is the human gene symbol
    fmrpTargetsDf <- read.xlsx("annotData/DarnellDarnell11_FMRPtargetsTableS2.xls",
                               sheetIndex=1, startRow=2, stringsAsFactors=FALSE,
                               check.names=FALSE)
    fmrpTargets <- select(org.Hs.eg.db, keys=fmrpTargetsDf$`Gene Symbol`,
                          columns="ENTREZID", keytype="SYMBOL")
    missingIDs <- fmrpTargets$SYMBOL[is.na(fmrpTargets$ENTREZID)]
    missingIDs
    alias2eg <- select(org.Hs.eg.db, keys=missingIDs, columns="ENTREZID", keytype="ALIAS")
    alias2eg <- split(alias2eg$ENTREZID, alias2eg$ALIAS)
    alias2eg <- sapply(alias2eg,
                       function(x, tt) {
                         if (any(!is.na(x))) {
                           lfc <- abs(tt$logFC)[tt$ID %in% x]
                           if (length(lfc) > 0)
                             x <- tt$ID[tt$ID %in% x][which.max(lfc)]
                           else
                             x <- x[1]
                         }
                         x
```

```
                       }, ttAll)
    stopifnot(nrow(fmrpTargets) == length(unique(fmrpTargets$SYMBOL))) ## QC
    fmrpTargets[match(names(alias2eg), fmrpTargets$SYMBOL), "ENTREZID"] <- alias2eg
    fmrpTargets <- fmrpTargets$ENTREZID[!is.na(fmrpTargets$ENTREZID)]
    length(fmrpTargets)
    save(fmrpTargets, file="annotData/fmrpTargets.RData")
  } else
    load("annotData/fmrpTargets.RData")
> N <- nrow(ttAll)
> m <- length(intersect(fmrpTargets, ttAll$ID))
> mask <- ttAllIQRfilt$adj.P.Val < FDRcutoff
> n <- sum(mask)
> k <- length(intersect(ttAllIQRfilt$ID,
                        intersect(fmrpTargets, ttAllIQRfilt$ID[mask])))
> t <- array(c(k, n - k, m - k, N + k - n - m), dim = c(2, 2),
             dimnames=list(FMRPtargets=c("in", "out"), DE=c("yes", "no")))
> t
           DE
FMRPtargets  yes    no
        in   214   609
        out 3682 15650
> fisher.test(t, alternative="greater")
        Fisher's Exact Test for Count Data

data:  t
p-value = 1.086e-06
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
 1.30055     Inf
sample estimates:
odds ratio
   1.49354
```

4.5 Enrichment of DE genes among targets of the APC gene

We assess enrichemnt of DE genes among targets of the adenomatous polyposis coli (APC) gene. These targets were obtained from the publication by (16).

```
> if (!file.exists("annotData/apcTargets.RData")) {
    ## in the Excel file from TableS1 in Preitner et al. (2014)
    ## gene symbol is the human gene symbol
    apcTargetsDf <- read.xlsx("annotData/Preitner14_APCtargetsTableS1.xlsx",
                              sheetIndex=1, startRow=1, stringsAsFactors=FALSE,
                              check.names=FALSE)
    apcTargets <- select(org.Hs.eg.db, keys=toupper(apcTargetsDf$`Gene Symbol`),
                         columns="ENTREZID", keytype="SYMBOL")
    apcTargets <- split(apcTargets$ENTREZID, apcTargets$SYMBOL)
    apcTargets <- sapply(apcTargets,
                       function(x, tt) {
                         if (any(!is.na(x))) {
                           lfc <- abs(tt$logFC)[tt$ID %in% x]
                           if (length(lfc) > 0)
                             x <- tt$ID[tt$ID %in% x][which.max(lfc)]
                           else
                             x <- x[1]
```

```
                      }
                      x
                    }, ttAll)
    missingIDs <- names(apcTargets)[is.na(apcTargets)]
    missingIDs
    alias2eg <- select(org.Hs.eg.db, keys=missingIDs, columns="ENTREZID", keytype="ALIAS")
    alias2eg <- split(alias2eg$ENTREZID, alias2eg$ALIAS)
    alias2eg <- sapply(alias2eg,
                    function(x, tt) {
                      if (any(!is.na(x))) {
                        lfc <- abs(tt$logFC)[tt$ID %in% x]
                        if (length(lfc) > 0)
                          x <- tt$ID[tt$ID %in% x][which.max(lfc)]
                        else
                          x <- x[1]
                      }
                      x
                    }, ttAll)
    stopifnot(length(apcTargets) == length(unique(names(apcTargets)))) ## QC
    apcTargets[match(names(alias2eg), names(apcTargets))] <- alias2eg
    apcTargets <- apcTargets[!is.na(apcTargets)]
    length(apcTargets)
    save(apcTargets, file="annotData/apcTargets.RData")
  } else
    load("annotData/apcTargets.RData")
> N <- nrow(ttAll)
> m <- length(intersect(apcTargets, ttAll$ID))
> mask <- ttAllIQRfilt$adj.P.Val < FDRcutoff
> n <- sum(mask)
> k <- length(intersect(ttAllIQRfilt$ID,
                        intersect(apcTargets, ttAllIQRfilt$ID[mask])))
> t <- array(c(k, n - k, m - k, N + k - n - m), dim = c(2, 2),
             dimnames=list(APCtargets=c("in", "out"), DE=c("yes", "no")))
> t
          DE
APCtargets  yes    no
       in    62   181
       out 3834 16078
> fisher.test(t, alternative="greater")
        Fisher's Exact Test for Count Data

data:  t
p-value = 0.01047
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
 1.10933     Inf
sample estimates:
odds ratio
  1.436428
```

4.6 Enrichment of DE genes in human phenotype ontology (HPO) terms

We assess enrichment of DE genes in gene sets from the human phenotype ontology (17).

```
> load("annotData/human-phenotype-ontology.RData")
> gUniverse <- ttAll[ttAll$logFC > 0, "ID"]
> OBOparams <- new("OBOHyperGParams", geneIds=DEgenesUp, universeGeneIds=gUniverse,
                   pvalueCutoff=goAnalysisCondPvalueCutoff, conditional=TRUE,
                   testDirection="over", datPkg=OBOCollectionDatPkg(oboHPO, gscHPO))
> cache(hgOverHPOdeGenesUp <- hyperGTest2(OBOparams, minSizeCutoff=5, maxSizeCutoff=300,
                                          orCutoff=1.5),
        dir=cacheDir_FE, prefix=cachePrefix_FE)
> hgOverHPOdeGenesUp
Gene to OBO Conditional test for over-representation
5842 OBO ids tested (82 have p < 0.01)
Selected gene set size: 118
    Gene universe size: 1610
    Annotation package: OBO
> IQRs <- apply(cleanExps[gUniverse, ], 1, IQR)
> cache(reportOverHPOdeGenesUp <- GOreport(hgOverHPOdeGenesUp, minSize=5, maxSize=300,
                                     minCount=5, minOddsRatio=1.5, featureFilterValues=IQRs,
                                     ff.R=1000, ff.prob=0.90, maxReported=Inf,
                                     p.value=0.1, p.adjust.method="fdr"),
        dir=cacheDir_FE, prefix=cachePrefix_FE)
> dim(reportOverHPOdeGenesUp)
[1] 39 10
```

Table 19. Human Phenotype Ontology (HPO) enrichment of upregulated genes called differentially expressed at FDR $< 1\%$ and with a minimum 1.5-fold change between FIR and nonFIR affected infants. Column "GeneSyms" contains the genes enriching the corresponding HPO term.

Term	OddsRatio	ExpCount	Count	Size	GeneSyms
Recurrent staphylococcal infections	Inf	0.37	5	5	NCF2, ITGB2, STAT3, CYBA, CYBB
Liver abscess	Inf	0.37	5	5	NCF2, NCF4, CTSC, CYBA, CYBB
Recurrent bacterial skin infections	23.46	0.81	7	11	NCF2, CTSC, PSTPIP1, CYBA, CYBB, GJB2, IL10RA
Gingivitis	13.49	1.17	8	16	STAT3, CYBA, CYBB, NCF2, NCF4, ITGB2, CTSC, COL3A1
Aortic dissection	13.16	0.73	5	10	FBN1, TGFBR2, COL5A2, COL1A1, COL3A1
Chronic obstructive pulmonary disease	10.96	0.81	5	11	CYBA, CYBB, SERPINA1, NCF2, NCF4
Recurrent skin infections	9.39	1.61	9	22	NCF2, IL17RA, PSTPIP1, CYBA, CYBB, GJB2, CLEC7A, CTSC, IL10RA
Premature loss of primary teeth	8.21	0.95	5	13	VDR, CTSC, SNX10, COL3A1, TCIRG1
Mediastinal lymphadenopathy	8.21	0.95	5	13	CYBA, CYBB, NCF2, NCF4, BIRC3
Abnormality of the rectum	7.78	1.39	7	19	TGFBR2, PDGFRA, NCF2, ITGB2, CYBA, CYBB, GREM1
Abnormality of the thoracic cavity	7.29	1.03	5	14	CYBA, CYBB, NCF2, NCF4, BIRC3
Ectopia lentis	7.21	1.25	6	17	FBN1, COL11A1, COL5A2, TAP1, TAPBP, COL1A1

Table 19 – Continued on next page

Table 19 – continued from previous page

Premature loss of teeth	7.17	1.47	7	20	VDR, SNX10, COL3A1, TCIRG1, TNFRSF11B, RUNX2, CTSC
Urticaria	6.66	1.54	7	21	MYD88, PDGFRA, NCF2, NCF4, CYBA, CYBB, GJB2
Emphysema	6.61	1.32	6	18	FBN1, TAP1, SERPINA1, TAPBP, NFKB2, PRKCD
Cellulitis	6.61	1.32	6	18	STAT3, CYBA, CYBB, PIK3CD, NCF2, COL1A1
Recurrent fungal infections	6.21	1.61	7	22	NCF2, STAT3, IL17RA, CYBA, CYBB, DOCK8, CLEC7A
Abnormality of macrophages	5.96	1.17	5	16	STXBP2, NOD2, CYBA, CYBB, NCF2
Lymphadenopathy	5.53	4.76	18	65	MYD88, TGFBR2, ABCA1, SNX10, CTHRC1, TCIRG1, NOD2, PSTPIP1, CYBA, RAB27A, NFKB2, PRKCD, PIK3CD, NCF2, NCF4, PTPRC, BIRC3, CYBB
Osteomyelitis	5.16	1.83	7	25	IFNGR1, NCF2, TCIRG1, IL1RN, STAT3, CYBA, CYBB
Sinusitis	3.98	4.25	13	58	MYD88, SNX10, TCIRG1, CYBA, DOCK8, NFKB2, PRKCD, NCF2, NCF4, CYBB, RUNX2, TAP1, TAPBP
Skin ulcer	3.60	5.35	15	73	MYD88, NOD2, PSTPIP1, GJB2, NCF2, NCF4, CYBB, TAP1, TAPBP, CLEC7A, CTSC, IL17RA, CYBA, DOCK8, STAT3
Abnormal bleeding	3.17	9.45	23	129	TGFBR2, FERMT3, NOD2, PRLR, TCIRG1, ITGA2, SERPINE1, DOCK8, PRKCD, COL1A1, IL10RA, MYD88, SNX10, ADAMTS2, PDGFRA, COL5A2, COL3A1, GREM1, CD36, RAB27A, PLAU, NFKB2, FCGR2C
Subcutaneous hemorrhage	3.16	5.50	14	75	MYD88, FERMT3, SNX10, NOD2, ADAMTS2, COL5A2, COL3A1, TCIRG1, ITGA2, RAB27A, PLAU, NFKB2, PRKCD, COL1A1
Abnormality of oral mucosa	3.10	5.57	14	76	MYD88, IL1RN, GJB2, ADAMTS2, NCF2, NCF4, COL3A1, CYBB, CTSC, ITGB2, CYBA, FAM20C, WNT5A, STAT3
Arthralgia	3.07	4.76	12	65	IL1RN, NOD2, PSTPIP1, NFKB2, PRKCD, COMP, COL11A1, PTPN2, COL3A1, C1R, NLRC4, TYROBP

Table 19 – Continued on next page

Table 19 – continued from previous page

Abnormality of temperature regulation	3.03	7.92	19	108	STXBP2, MYD88, SNX10, NOD2, PSTPIP1, NCF2, NCF4, BIRC3, CYBB, CLEC7A, TCIRG1, IL17RA, CYBA, RAB27A, NFKBIA, COL1A1, GCH1, STAT3, NLRC4
Otitis media	2.86	5.50	13	75	FBN1, NCF2, NCF4, CYBB, RUNX2, TAP1, TAPBP, CYBA, DOCK8, NFKB2, PRKCD, PTPRC, STAT3
Abnormality of humoral immunity	2.82	6.01	14	82	MYD88, CFP, PSTPIP1, JAK3, CXCR4, DOCK8, NFKB2, PRKCD, COL1A1, PIK3CD, C1R, STAT3, C1S, C3
Recurrent infections	2.63	14.26	29	204	FBN1, TGFBR2, FERMT3, GJB2, JAK3, CTSC, ITGB2, TCIRG1, PRKCD, COL1A1, WNT5A, C1R, C3, IL10RA, SNX10, PSTPIP1, COL5A2, NCF4, BIRC3, CORO1A, RUNX2, TAP1, TAPBP, CXCR4, RAB27A, NFKBIA, NFKB2, PIK3CD, PTPRC
Recurrent respiratory infections	2.62	12.02	25	164	FBN1, CYBB, JAK3, CLEC7A, CTSC, TCIRG1, DOCK8, PRKCD, WNT5A, C1R, SNX10, NCF2, NCF4, BIRC3, CORO1A, RUNX2, TAP1, TAPBP, CXCR4, IL17RA, CYBA, NFKBIA, NFKB2, PIK3CD, STAT3
Arthritis	2.55	6.01	13	82	FBN1, NOD2, PSTPIP1, GJB2, COMP, COL11A1, COL5A2, COL3A1, CTSC, COL1A1, PTPN2, C1R, STAT3
Abnormality of skin physiology	2.42	9.97	20	136	NOD2, GJB2, CYBB, CLEC7A, CTSC, GLUL, DOCK8, IL10RA, LYZ, PAPSS2, SAT1, PSTPIP1, PDGFRA, NCF2, NCF4, OSMR, IL17RA, CYBA, PTPRC, STAT3
Abnormality of lymphocytes	2.41	6.82	14	93	STXBP2, MYD88, PSTPIP1, CORO1A, JAK3, CXCR4, RAB27A, DOCK8, NFKB2, PRKCD, COL1A1, PIK3CD, PTPRC, STAT3

Table 19 – Continued on next page

Table 19 – continued from previous page

Abnormality of cellular immune system	2.20	14.37	26	196	FERMT3, NOD2, CYBB, JAK3, BCL3, ITGB2, DOCK8, PRKCD, COL1A1, STXBP2, MYD88, ARHGAP26, PSTPIP1, PDGFRA, NCF2, NCF4, CORO1A, TYROBP, CXCR4, CYBA, RAB27A, NFKB2, RAC2, PIK3CD, PTPRC, STAT3
Splenomegaly	2.19	10.19	19	139	FERMT3, ABCA1, NOD2, CYBB, TCIRG1, PRKCD, MYD88, LYZ, SNX10, IL1RN, PDGFRA, NCF2, NCF4, CYBA, RAB27A, NFKB2, PIK3CD, PTPRC, NLRC4
Abnormality of the immune system	2.12	10.76	19	170	FBN1, IL17RA, FCGR2C, GJB2, CLEC7A, C1R, C1S, C3, RUNX2, TAP1, TAPBP, NFKBIA, CTSC, WNT5A, CFP, COL5A2, ITGAM, IL10RA, PTPN2
Abnormality of the lymphatic system	1.97	15.61	26	213	TGFBR2, FERMT3, ABCA1, NOD2, CYBB, JAK3, TCIRG1, PRKCD, STXBP2, MYD88, LYZ, SNX10, CTHRC1, IL1RN, PSTPIP1, PDGFRA, NCF2, NCF4, BIRC3, CYBA, RAB27A, NFKB2, PIK3CD, PTPRC, CSF3R, NLRC4
Abnormality of blood and blood-forming tissues	1.86	18.86	29	294	ABCA1, PGK1, PGD, PSTPIP1, IFNGR1, BIRC3, CORO1A, TYROBP, PLAT, CYBA, PIK3CD, CSF3R, SLC11A2, CYBB, JAK3, BCL3, ITGB2, B4GALT1, STXBP2, MYC, ARHGAP26, NCF2, NCF4, CXCR4, SLC19A2, RAC2, PTPRC, STAT3, NLRC4

```
> gUniverse <- ttAll[ttAll$logFC < 0, "ID"]
> OBOparams <- new("OBOHyperGParams", geneIds=DEgenesDown, universeGeneIds=gUniverse,
                   pvalueCutoff=goAnalysisCondPvalueCutoff, conditional=TRUE,
                   testDirection="over", datPkg=OBOCollectionDatPkg(oboHPO, gscHPO))
> cache(hgOverHPOdeGenesDown <- hyperGTest2(OBOparams, minSizeCutoff=5, maxSizeCutoff=300,
                                            orCutoff=1.5),
        dir=cacheDir_FE, prefix=cachePrefix_FE)
> hgOverHPOdeGenesDown
Gene to OBO Conditional test for over-representation
5338 OBO ids tested (25 have p < 0.01)
Selected gene set size: 104
    Gene universe size: 1385
    Annotation package: OBO
```

```
> IQRs <- apply(cleanExps[gUniverse, ], 1, IQR)
> cache(reportOverHPOdeGenesDown <- GOreport(hgOverHPOdeGenesDown, minSize=5, maxSize=300,
                                             minCount=5, minOddsRatio=1.5, featureFilterValues=I(
                                             ff.R=1000, ff.prob=0.90, maxReported=Inf,
                                             p.value=0.15, p.adjust.method="fdr"),
        dir=cacheDir_FE, prefix=cachePrefix_FE)
> dim(reportOverHPOdeGenesDown)
[1]  0 10
```

4.7 Enrichment of DE genes among umbilical cord mensenchymal stem cell markers

We assess enrichment of DE genes among mesenchymal stem cell markers of the umbilical cord tissue, using a signature from (18) of mesenchymal stem cells derived from umbilical cord. The genes forming this signature were downloaded from (19).

```
> if (!file.exists(file.path("annotData", "mscGenes.RData"))) {
    library(xlsx)

    ucMSC <- read.xlsx(file.path("annotData", "GeneSetMSCsLifeMpasc.xlsx"),
                       sheetIndex=1, stringsAsFactors=FALSE)

    table(ucMSC$EXPRESSION)
    ucMSCmarkers <- ucMSC$ID[ucMSC$EXPRESSION %in% "Marker"]
    length(ucMSCmarkers)
    ucMSCnonmarkers <- ucMSC$ID[ucMSC$EXPRESSION %in% "NonMarker"]
    length(ucMSCnonmarkers)
    ucMSCexpressed <- ucMSC$ID[ucMSC$EXPRESSION %in% "Expressed"]
    length(ucMSCexpressed)

    mscGenes <- ucMSCmarkers
    save(mscGenes, file=file.path("annotData", "mscGenes.RData"))
  } else
    load(file.path("annotData", "mscGenes.RData"))
```

First, let's assess enrichment with the set of genes called DE at 1% FDR:

```
> DEgenesFDR <- ttAllIQRfilt$ID[ttAllIQRfilt$adj.P.Val <= FDRcutoff]
> N <- nrow(ttAll)
> m <- length(intersect(mscGenes, ttAll$ID))
> n <- length(DEgenesFDR)
> k <- length(intersect(ttAllIQRfilt$ID, intersect(mscGenes, DEgenesFDR)))
> t <- array(c(k, n - k, m - k, N + k - n - m), dim = c(2, 2),
             dimnames=list(ucMSCgenes=c("in", "out"), DE=c("yes", "no")))
> t
          DE
ucMSCgenes  yes    no
       in    10     9
       out 3886 16250
> fisher.test(t, alternative="greater")
        Fisher's Exact Test for Count Data

data:  t
p-value = 0.001192
alternative hypothesis: true odds ratio is greater than 1
```

```
95 percent confidence interval:
 1.967201      Inf
sample estimates:
odds ratio
  4.645821
```

Now, with the genes that are called DE at 1% FDR and have a minimum 1.5-fold change:

```
> N <- nrow(ttAll)
> m <- length(intersect(mscGenes, ttAll$ID))
> n <- length(DEgenes)
> k <- length(intersect(ttAllIQRfilt$ID, intersect(mscGenes, DEgenes)))
> t <- array(c(k, n - k, m - k, N + k - n - m), dim = c(2, 2),
             dimnames=list(ucMSCgenes=c("in", "out"), DE=c("yes", "no")))
> t
          DE
ucMSCgenes  yes    no
       in     5    14
       out 1092 19044
> fisher.test(t, alternative="greater")
        Fisher's Exact Test for Count Data

data:  t
p-value = 0.0029
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
 2.149357      Inf
sample estimates:
odds ratio
  6.227101
```

Now, examine the enrichment restricted to upregulated genes:

```
> N <- nrow(ttAll[ttAll$logFC > 0, ])
> m <- length(intersect(mscGenes, ttAll$ID[ttAll$logFC > 0]))
> n <- length(DEgenesUp)
> k <- length(intersect(ttAllIQRfilt$ID[ttAllIQRfilt$logFC > 0],
                        intersect(mscGenes, DEgenesUp)))
> t <- array(c(k, n - k, m - k, N + k - n - m), dim = c(2, 2),
             dimnames=list(ucMSCgenes=c("in", "out"), DE=c("yes", "no")))
> t
          DE
ucMSCgenes yes    no
       in    4     6
       out 588 10270
> fisher.test(t, alternative="greater")
        Fisher's Exact Test for Count Data

data:  t
p-value = 0.001406
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
 3.073139      Inf
sample estimates:
odds ratio
  11.63681
```

Finally, examine the enrichment restricted to downregulated genes:

```
> N <- nrow(ttAll[ttAll$logFC < 0, ])
> m <- length(intersect(mscGenes, ttAll$ID[ttAll$logFC < 0]))
> n <- length(DEgenesDown)
> k <- length(intersect(ttAllIQRfilt$ID[ttAllIQRfilt$logFC < 0],
                        intersect(mscGenes, DEgenesDown)))
> t <- array(c(k, n - k, m - k, N + k - n - m), dim = c(2, 2),
             dimnames=list(ucMSCgenes=c("in", "out"), DE=c("yes", "no")))
> t
          DE
ucMSCgenes yes   no
       in    1    8
       out 504 8774
> fisher.test(t, alternative="greater")
        Fisher's Exact Test for Count Data

data:  t
p-value = 0.3955
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
 0.09930754        Inf
sample estimates:
odds ratio
  2.175819
```

In this latter case, there is no significant enrichment.

4.8 Summary table for targeted gene sets

Here we summarize the targeted gene sets we employed before and their annotation sources. This information is hardcoded in the hidden code building this table which is shown in this document as Table 20.

4.9 Session Information

```
> toLatex(sessionInfo())
```

- R version 3.1.3 (2015-03-09), `x86_64-unknown-linux-gnu`
- Locale: `LC_CTYPE=en_US.UTF8`, `LC_NUMERIC=C`, `LC_TIME=en_US.UTF8`, `LC_COLLATE=en_US.UTF8`, `LC_MONETARY=en_US.UTF8`, `LC_MESSAGES=en_US.UTF8`, `LC_PAPER=en_US.UTF8`, `LC_NAME=en_US.UTF8`, `LC_ADDRESS=en_US.UTF8`, `LC_TELEPHONE=en_US.UTF8`, `LC_MEASUREMENT=en_US.UTF8`, `LC_IDENTIFICATION=en_US.UTF8`
- Base packages: base, datasets, graphics, grDevices, methods, parallel, stats, stats4, utils
- Other packages: affy 1.44.0, affyPLM 1.42.0, annotate 1.44.0, AnnotationDbi 1.28.2, Biobase 2.26.0, BiocGenerics 0.12.1, Category 2.32.0, DBI 0.3.1, gcrma 2.38.0, genefilter 1.48.1, GenomeInfoDb 1.2.5, GO.db 3.0.0, GOstats 2.32.0, graph 1.44.1, GSEABase 1.28.0, IRanges 2.0.1, limma 3.22.7, Matrix 1.2-0, mgcv 1.8-6, nlme 3.1-120, org.Hs.eg.db 3.0.0, preprocessCore 1.28.0, RColorBrewer 1.1-2, rJava 0.9-6, RSQLite 1.0.0, S4Vectors 0.4.0, sva 3.12.0, xlsx 0.5.7, xlsxjars 0.6.1, XML 3.98-1.1, xtable 1.7-4
- Loaded via a namespace (and not attached): affyio 1.34.0, AnnotationForge 1.8.2, BiocInstaller 1.16.4, Biostrings 2.34.1, grid 3.1.3, lattice 0.20-31, RBGL 1.42.0, splines 3.1.3, survival 2.38-1, tools 3.1.3, XVector 0.6.0, zlibbioc 1.12.0

Table 20. Description of targeted gene sets significantly enriched by genes called differentially expressed between FIR and nonFIR affected infants. The "Size" column specifies the number of genes in the gene set that map to the microarray expression data set analyzed in this article.

Gene Set	Size	Source
Inflammation (innate response)	800	Table S1 from Loza et al. Assembly of inflammation-related genes for pathway-focused genetic analysis. PLOS One, 2(10):e1035, 2007 [PMID:17940599]
Inflammation (Th17 - adaptive response)	72	Table S1 from ciofani et al. A validated regulatory network for Th17 cell specification. Cell, 151(2):289-303, 2012 [PMDI:23021777]
Senescence	194	Belinky et al. PathCards: multi-source consolidation of human biological pathways. Database, 2015:bav006, 2015 [PMID:25725062]
Apoptosis	68	Table S1 from Loza et al. Assembly of inflammation-related genes for pathway-focused genetic analysis. PLOS One, 2(10):e1035, 2007 [PMID:17940599]
FMRP targets	823	Table S2 from Darnell et al. FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. Cell, 146(2):247-261, 2011 [PMID:21784246]
APC targets	243	Table S1 from Preitner et al. APC is an RNA-binding protein, and its interactome provides a link to neural development and microtubule assembly. Cell, 158(2):368-382, 2014 [PMID:25036633]
UC mesenchymal stem cells	19	Edgar et al. LifeMap discovery: the embryonic development, stem cells, and regenerative medicine research portal. PLOS One, 8(7):e66629, 2013 [PMID:23874394]

References

[1] The R Foundation. The Comprehensive R Archive Network (CRAN). (`http://cran.r-project.org`).

[2] Gentleman RC, Carey VJ, Bates DM, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 2004; 5:R80.

[3] National Center for Biotechnology Information. Gene Expression Omnibus (GEO). (`http://www.ncbi.nlm.nih.gov/geo`).

[4] Brettschneider J, Collin F, Bolstad BM, Speed TP. Quality assessment for short oligonucleotide microarray data. Technometrics 2008; 50:241–264.

[5] McCall MN, Bolstad BM, Irizarry RA. Frozen robust multiarray analysis (fRMA). Biostatistics 2010; 11:242–253.

[6] Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 2004; 3:1–25.

[7] Miyazaki K, Furuhashi M, Ishikawa K, Tamakoshi K, Ikeda T, et al. The effects of antenatal corticosteroids therapy on very preterm infants after chorioamnionitis. Arch Gynecol Obstet 2014; 289:1185–1190.

[8] Bourgon R, Gentleman R, Huber W. Independent filtering increases detection power for high-throughput experiments. Proc Natl Acad Sci U S A, May (2010) 107:9546–9551.

[9] Madsen-Bouterse SA, Romero R, Tarca AL, et al. The transcriptome of the fetal inflammatory response syndrome. Am J Reprod Immunol 2010; 63:73–92.

[10] Alexa A, Rahnenführer J, Lengauer T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics 2006; 22:1600–1607.

[11] Loza MJ, McCall CE, Li L, Isaacs WB, Xu J, Chang BL. Assembly of inflammation-related genes for pathway-focused genetic analysis. PLoS One 2007; 2:e1035.

[12] Ciofani M, Madar A, Galan C, et al. A validated regulatory network for Th17 cell specification. Cell 2012; 151:289–303.

[13] Belinky F, Nativ N, Stelzer G, et al. Pathcards: multi-source consolidation of human biological pathways. Database (Oxford) 2015; 2015:bav006.

[14] LifeMap Sciences. PathCards Pathway Unification Database: Cellular Senescence. (`http://pathcards.genecards.org/card/cellular_senescence?term=senescence`). [Online; accessed 30-March-2015].

[15] Darnell JC, Van Driesche SJ, Zhang C, et al. FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. Cell 2011; 146:247–261.

[16] Preitner N, Quan J, Nowakowski DW, et al. APC is an RNA-binding protein, and its interactome provides a link to neural development and microtubule assembly. Cell 2014; 158:368–382.

[17] Köhler S, Doelken SC, Mungall CJ, et al. The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data. Nucleic Acids Res 2014; 42:D966–D974.

[18] Edgar R, Mazor Y, Rinon A, et al. LifeMap Discovery™: The embryonic development, stem cells, and regenerative medicine research portal. PLoS One 2013; 8:e66629.

[19] LifeMap Sciences. Embryonic Development & Stem Cell Compendium: Umbilical cord-derived mesenchymal stem cells.
(`http://discovery.lifemapsc.com/stem-cell-differentiation/in-vitro-cells/`
`mesenchymal-stem-cells-umbilical-cord-homo-sapiens-umbilical-cord-derived-mesenchymal-stem-cells`
`panel_GeneExpression`). [Online; accessed 16-March-2015].