```
1854_at      MYBL2  3.34    8.53  20.2 1.81e-08  7.63e-05  9.64
38116_at  KIAA0101  3.76    9.51  16.9 8.12e-08  2.51e-04  8.48
38065_at     HMGB2  2.99    9.10  16.2 1.12e-07  2.51e-04  8.21
39755_at      XBP1  1.77   12.13  15.8 1.36e-07  2.51e-04  8.05
1592_at      TOP2A  2.30    8.31  15.8 1.39e-07  2.51e-04  8.03
41400_at       TK1  2.24   10.04  15.3 1.81e-07  2.75e-04  7.81
33730_at    GPRC5A -2.04    8.57 -15.1 1.96e-07  2.75e-04  7.74
1651_at      UBE2C  2.97   10.50  14.8 2.39e-07  3.02e-04  7.57
38414_at     CDC20  2.02    9.46  14.6 2.66e-07  3.05e-04  7.48
1943_at      CCNA2  2.19    7.60  14.0 3.72e-07  3.69e-04  7.18
40117_at      MCM6  2.28    9.68  14.0 3.80e-07  3.69e-04  7.17
40533_at     BIRC5  1.64    8.47  13.5 4.94e-07  4.45e-04  6.93
39642_at    ELOVL2  1.61    7.88  13.0 6.71e-07  5.18e-04  6.65
34851_at     AURKA  1.96    9.96  12.8 7.51e-07  5.18e-04  6.55
1824_s_at     PCNA  1.64    9.24  12.8 7.95e-07  5.18e-04  6.50
35995_at     ZWINT  2.76    8.87  12.7 8.32e-07  5.18e-04  6.46
893_at       UBE2S  1.54   10.95  12.7 8.43e-07  5.18e-04  6.45
40079_at    GPRC5A -2.41    8.23 -12.6 8.62e-07  5.18e-04  6.42

> sessionInfo()
R version 3.0.0 (2013-04-03)
Platform: i386-w64-mingw32/i386 (32-bit)

locale:
[1] LC_COLLATE=English_Australia.1252  LC_CTYPE=English_Australia.1252
[3] LC_MONETARY=English_Australia.1252 LC_NUMERIC=C
[5] LC_TIME=English_Australia.1252

attached base packages:
[1] parallel  stats     graphics  grDevices utils     datasets  methods   base

other attached packages:
 [1] hgu95av2.db_2.9.0    org.Hs.eg.db_2.9.0   RSQLite_0.11.3
 [4] DBI_0.2-6            annotate_1.38.0      hgu95av2cdf_2.12.0
 [7] AnnotationDbi_1.22.2 affy_1.38.1          Biobase_2.20.0
[10] BiocGenerics_0.6.0   limma_3.17.7         BiocInstaller_1.10.0

loaded via a namespace (and not attached):
[1] affyio_1.28.0         IRanges_1.18.0        preprocessCore_1.22.0
[4] stats4_3.0.0          tools_3.0.0           XML_3.96-1.1
[7] xtable_1.7-1          zlibbioc_1.6.0
```

17.3 Comparing Mammary Progenitor Cell Populations with Illumina BeadChips

17.3.1 Introduction

This case study examines the expression profiles of adult mammary stem cells and of progenitor and mature mammary lumina cells. The data was first published by Lim et al [18], who used the expression profiles to show that limina progenitor cells are the likely cell of origina for basal-like breast cancer.

The data files used in this case study can be downloaded from `http://bioinf.wehi.edu.au/marray/IlluminaCaseStudy`. The expression data files are provided as gzip files and will need to be uncompressed before they can be used for a limma analysis. To run the analysis described here,

download and unzip the data files and set the working directory of R to the folder containing the files.

17.3.2 The target RNA samples

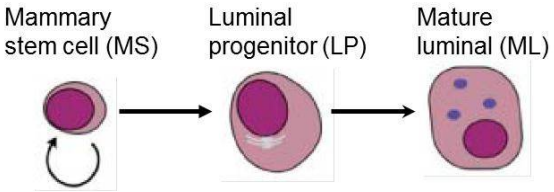
The data consists of three files:

```
> dir()
[1] "control probe profile.txt"
[2] "probe profile.txt"
[3] "Targets.txt"
```

First read the targets file describing the target RNA samples.

```
> library(limma)
> targets <- readTargets()
> targets
   Donor Age CellType
1      1  39       MS
2      1  39   Stroma
3      1  39       ML
4      1  39       LP
5      2  57       MS
6      2  57   Stroma
7      2  57       ML
8      2  57       LP
9      3  21       MS
10     3  21   Stroma
11     3  21       ML
12     3  21       LP
```

Breast tissue was obtained from three healthy human donors who were undergoing reduction mammoplasties. Epithelial cells were sorted into three subpopulations enriched for mammary stem cells (MS), luminal progenitor cells (LP) and mature luminal cells (ML) [18]. The MS, LP and ML cells representative a lineage of luminal cells use to construct the ducts used to transport milk in the breast [44]:



Stromal cells were also profiles as a comparison group. There were therefore four cell populations from each person. RNA was extracted from freshly sorted cells, making twelve RNA samples in total.

17.3.3 The expression profiles

The RNA samples were hybridized to two Illumina HumanWG-6 version 3 BeadChips. Each BeadChip is able to accommodate six samples. The BeadChips images were scanned and summarized using BeadStudio. Un-normalized summary probe profiles were exported from from BeadStudio to tab-delimited text files.

Separate files were written for regular probes and for control probes. The file `probe profile.txt` contains the expression profiles for regular probes, designed to interrogate the expression levels of genes. `control probe profile.txt` contains the profiles of control probes, including negative control probes. Note that BeadStudio by default writes the profiles for all the samples to the same two files.

We read in the expression profiles for both regular and control probes, telling `read.ilm` that we wish to read the detection p-values as well as the expression values:

```
> x <- read.ilmn(files="probe profile.txt",ctrlfiles="control probe profile.txt",
+ other.columns="Detection")
Reading file probeprofile.txt ... ...
Reading file controlprobeprofile.txt ... ...
```

This reads a `EListRaw` object. There are about 750 negative probes and about 49,000 regular probes:

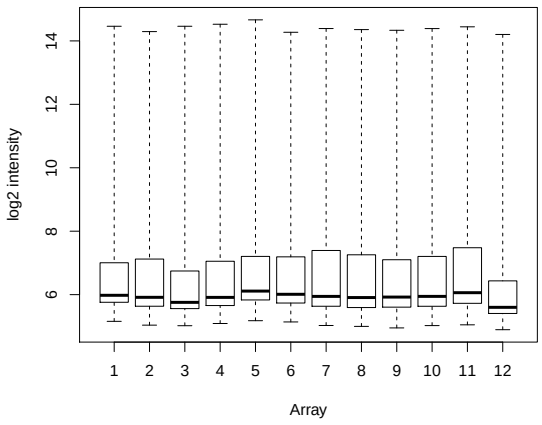
```
> table(x$genes$Status)
            BIOTIN            CY3_HYB       HOUSEKEEPING           LABELING
                 2                  6                  7                  2
LOW_STRINGENCY_HYB           NEGATIVE            regular
                 8                759              48803
```

The component `E` contains the expression value for each probe

```
> options(digits=3)
> head(x$E)
                1    2    3    4    5    6    7    8    9   10   11   12
ILMN_1762337 52.3 46.1 54.0 47.7 54.8 47.4 67.4 47.9 40.5 44.7 80.6 42.5
ILMN_2055271 69.9 73.9 58.6 72.4 77.1 82.1 69.1 81.3 79.1 82.5 87.0 60.9
ILMN_1736007 57.5 53.7 53.4 49.4 58.6 59.9 56.4 51.6 58.7 51.7 58.4 43.9
ILMN_2383229 53.6 57.5 48.2 48.2 61.8 64.5 52.7 43.5 65.5 49.8 53.9 39.3
ILMN_1806310 58.1 55.1 50.5 60.0 64.2 58.4 58.0 52.3 56.6 55.6 65.3 46.4
ILMN_1779670 64.5 61.2 52.8 61.9 67.9 59.7 68.2 63.1 65.1 65.7 69.6 52.0
```

The intensities vary from about 5 to 14 on the $\log_2$ scale:

```
> boxplot(log2(x$E),range=0,ylab="log2 intensity")
```



17.3.4 How many probes are truly expressed?

The detection values contain p-values for testing whether each probe is more intense than the negative control probes. Small values are evidence that the probe corresponds to a truly expressed gene:

```
> head(x$other$Detection)
                  1     2      3       4      5     6      7      8      9     10     11    12
ILMN_1762337 0.5585 0.675 0.1370 0.60139 0.5776 0.782 0.0503 0.4781 0.9082 0.7145 0.0000 0.460
ILMN_2055271 0.0306 0.000 0.0493 0.00278 0.0364 0.000 0.0391 0.0000 0.0000 0.0000 0.0000 0.000
ILMN_1736007 0.2772 0.292 0.1534 0.48611 0.4112 0.220 0.2318 0.3145 0.1554 0.3774 0.2539 0.360
ILMN_2383229 0.4735 0.187 0.3658 0.56389 0.2951 0.124 0.3408 0.7447 0.0537 0.4680 0.3986 0.747
ILMN_1806310 0.2618 0.248 0.2589 0.12778 0.2196 0.264 0.1955 0.2920 0.1963 0.2382 0.1220 0.203
ILMN_1779670 0.0850 0.113 0.1644 0.10417 0.1469 0.224 0.0461 0.0691 0.0621 0.0655 0.0709 0.058
```

We can go further than this and estimate the overall proportion of the regular probes that correspond to expressed transcript, using the method of Shi et al [35].

```
> pe <- propexpr(x)
> dim(pe) <- c(4,3)
> dimnames(pe) <- list(CellType=c("MS","Stroma","ML","LP"),Donor=c(1,2,3))
> pe
        Donor
CellType     1     2     3
  MS     0.557 0.504 0.529
  Stroma 0.518 0.514 0.514
  ML     0.549 0.495 0.535
  LP     0.555 0.518 0.517
```

The proportion of probes that are expressed varies from 50–56%. The average is 52.5%.

17.3.5 Normalization and filtering

Background correction and normalize:

```
> y <- neqc(x)
```

The `neqc` functions performs normexp background correction using negative controls, then quantile normalizes and finally $\log_2$ transforms [36]. It also automatically removes the control probes, leaving only the regular probes in `y`:

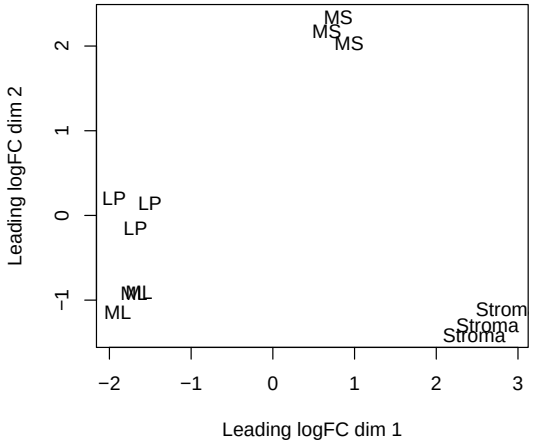
```
> dim(y)
[1] 48803    12
```

Filter out probes that are not expressed. We keep probes that are expressed in at least three arrays according to a detection p-values of 5%:

```
> expressed <- rowSums(y$other$Detection < 0.05) >= 3
> y <- y[expressed,]
> dim(y)
[1] 24691    12
```

A multi-dimensional scaling plot shows that the cell types are cell separated:

```
> plotMDS(y,labels=targets$CellType)
```



17.3.6 Within-patient correlations

The study involves multiple cell types from the same patient. Arrays from the same donor are not independent, so we need to estimate the within-donor correlation:

```
> ct <- factor(targets$CellType)
> design <- model.matrix(~0+ct)
> colnames(design) <- levels(ct)
> dupcor <- duplicateCorrelation(y,design,block=targets$Donor)
> dupcor$consensus.correlation
[1] 0.134
```

As expected, the within-donor correlation is small but positive.

17.3.7 Differential expression between cell types

Now we look for differentially expressed genes. We make all possible pairwise comparisons between the epithelial cell types, allowing for the correlation within donors:

```
> fit <- lmFit(y,design,block=targets$Donor,correlation=dupcor$consensus.correlation)
> contrasts <- makeContrasts(ML-MS, LP-MS, ML-LP, levels=design)
> fit2 <- contrasts.fit(fit, contrasts)
> fit2 <- eBayes(fit2, trend=TRUE)
> summary(decideTests(fit2, method="global"))
       ML - MS LP - MS ML - LP
Down      3074    2836    1631
NotSig   18088   18707   21307
Up        3529    3148    1753
```

Top ten differentially expressed probes between ML and MS:

```
> topTable(fit2, coef=1)
                SYMBOL logFC AveExpr     t  P.Value adj.P.Val    B
ILMN_1766707     IL17B -4.19    5.94 -29.0 2.51e-12  5.19e-08 18.1
ILMN_1706051      PLD5 -4.00    5.67 -27.8 4.20e-12  5.19e-08 17.7
ILMN_1656369    C8orf4  5.24   11.26  25.1 1.36e-11  1.06e-07 16.7
```

```
ILMN_2413323       GRP -6.60    6.82 -24.3 2.01e-11  1.06e-07 16.4
ILMN_1787750     CD200 -5.68    6.30 -23.9 2.43e-11  1.06e-07 16.2
ILMN_1669819 LOC402569  2.52    5.50  23.8 2.57e-11  1.06e-07 16.2
ILMN_1777998  ARHGAP25 -4.78    6.26 -23.1 3.50e-11  1.15e-07 15.9
ILMN_1701933      SNCA -5.26    6.27 -22.8 4.19e-11  1.15e-07 15.7
ILMN_1777199       GRP -5.47    6.36 -22.8 4.23e-11  1.15e-07 15.7
ILMN_1708303   CYP4F22  4.09    5.94  22.5 4.76e-11  1.15e-07 15.6
```

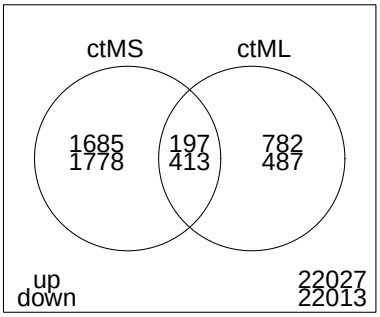
17.3.8 Signature genes for luminal progenitor cells

Now we find genes uniquely expressed in LP cells, as compared to MS and ML. We refit the linear model, making LP the reference cell type:

```
> ct <- relevel(ct, ref="LP")
> design <- model.matrix(~ct)
> fit <- lmFit(y,design,block=targets$Donor,correlation=dupcor$consensus.correlation)
> fit2 <- fit[,c("ctMS","ctML")]
> fit2 <- eBayes(fit2, trend=TRUE)
```

The we find all those genes that are up-regulated in LP vs both MS and ML, using a 2-fold-change and 5% FDR:

```
> results <- decideTests(fit2, lfc=1)
> vennDiagram(results, include=c("up","down"))
```



There are 197 positive signature genes and 413 negative. To see the top positive signature genes with their fold-changes:

```
> LP.sig <- rowSums(results>0)==2
> topTable(fit2[LP.sig,])
               SYMBOL ctMS ctML AveExpr     F  P.Value adj.P.Val
ILMN_1716370     TNS4 4.87 1.47    6.81 156.3 3.31e-09  4.81e-07
ILMN_1810054     CNN1 6.41 1.42    8.47 146.0 4.88e-09  4.81e-07
ILMN_1713744 C14orf132 3.85 1.74    7.34 103.5 3.37e-08  2.21e-06
ILMN_1720513   SETBP1 4.17 1.27    8.27  92.5 6.28e-08  3.09e-06
ILMN_1767662    LASS6 2.60 1.98   10.01  85.2 9.88e-08  3.46e-06
ILMN_1681984  GALNT10 3.54 1.23    8.09  84.2 1.06e-07  3.46e-06
ILMN_1731374      CPE 4.23 3.01    9.12  81.9 1.23e-07  3.46e-06
ILMN_1789639     FMOD 3.57 2.24    8.21  75.7 1.89e-07  4.65e-06
ILMN_1743933    TSHZ3 3.94 1.22    8.93  72.6 2.38e-07  5.21e-06
ILMN_1721876    TIMP2 3.57 1.38   10.10  69.0 3.13e-07  6.17e-06
```

17.4 Time Course Effects of Corn Oil on Rat Thymus with Agilent 4x44K Arrays

17.4.1 Introduction

This case study analyses a time-course experiment using single-channel Agilent Whole Rat Genome Microarray 4x44K v3 arrays.

The experiment concerns the effect of corn oil on gene expression in the thymus or rats. The data was submitted by Hong Weiguo to ArrayExpress as series E-GEOD-33005. The description of the experiment reads:

> "To investigate the effects of corn oil (CO), common drug vehicle, on the gene expression profiles in rat thymus with microarray technique. Female Wistar Rats were administered daily with normal saline (NS), CO 2, 5, 10 ml/kg for 14 days, respectively. Then, the thymus samples of rats were collected for microarray test and histopathology examination. The microarray data showed that 0, 40, 458 differentially expressed genes (DEGs) in 2, 5, 10 ml/kg CO group compared to NS group, respectively. The altered genes were associated with immune response, cellular response to organic cyclic substance, regulation of fatty acid beta-oxidation, et al. However, no obvious histopathologic change was observed in the three CO dosage groups. These data show that 10 ml/kg CO, that dosage has been determined as the vehicle in drug safety assessment, can cause obvious influence on gene expression in rat thymus. Our study suggest that the dosage of CO gavage as the vehicle for water-in-soluble agents in drug development should be no more than 5 ml/kg if agents' molecular effects in thymus want to be assessed. Gene expression in thymus from female Wistar rats daily administered with 2, 5, 10 ml/kg of corn oil or 10 ml/kg of saline by gavage for 14 consecutive days were measured using Agilent Rat Whole Genome 8×60K array."

17.4.2 Data availability

All files files were downloaded from `http://www.ebi.ac.uk/arrayexpress/experiments/E-GEOD-33005`. The data is also available as GEO series GSE33005.

Using R, the data can be downloaded to your working directory by:

```
> URL <- "https://www.ebi.ac.uk/arrayexpress/files/E-GEOD-33005"
> SDRF.file <- "E-GEOD-33005.sdrf.txt"
> Data.file <- "E-GEOD-33005.raw.1.zip"
> download.file(paste(URL,SDRF.file,sep="/"), SDRF.file)
> download.file(paste(URL,Data.file,sep="/"), Data.file)
> unzip(Data.file)
```

17.4.3 Reading the data

First we read the sample and data relationship format (SDRF) file. This is equivalent to what is known as the "targets file" in limma:

```
> SDRF <- read.delim("E-GEOD-33005.sdrf.txt",check.names=FALSE,stringsAsFactors=FALSE)
> SDRF[,c("Array Data File","Characteristics[treatment]")]
                                          Array Data File Characteristics[treatment]
1  GSM819076_US10283824_252828210181_S01_GE1_107_Sep09_1_4.txt          10 ml/kg corn oil
```

```
2  GSM819075_US10283824_252828210181_S01_GE1_107_Sep09_1_3.txt          10 ml/kg corn oil
3  GSM819074_US10283824_252828210181_S01_GE1_107_Sep09_1_2.txt           5 ml/kg corn oil
4  GSM819073_US10283824_252828210180_S01_GE1_107_Sep09_1_4.txt           2 ml/kg corn oil
5  GSM819072_US10283824_252828210180_S01_GE1_107_Sep09_1_3.txt           2 ml/kg corn oil
6  GSM819071_US10283824_252828210180_S01_GE1_107_Sep09_1_2.txt           10 ml/kg saline
7  GSM819070_US10283824_252828210180_S01_GE1_107_Sep09_1_1.txt           10 ml/kg saline
8  GSM819069_US10283824_252828210179_S01_GE1_107_Sep09_1_4.txt           10 ml/kg saline
9  GSM819068_US10283824_252828210179_S01_GE1_107_Sep09_1_3.txt          10 ml/kg corn oil
10 GSM819067_US10283824_252828210179_S01_GE1_107_Sep09_1_2.txt          10 ml/kg corn oil
11 GSM819066_US10283824_252828210179_S01_GE1_107_Sep09_1_1.txt          10 ml/kg corn oil
12 GSM819065_US10283824_252828210178_S01_GE1_107_Sep09_1_4.txt           5 ml/kg corn oil
13 GSM819064_US10283824_252828210178_S01_GE1_107_Sep09_1_3.txt           5 ml/kg corn oil
14 GSM819063_US10283824_252828210178_S01_GE1_107_Sep09_1_2.txt           5 ml/kg corn oil
15 GSM819062_US10283824_252828210178_S01_GE1_107_Sep09_1_1.txt           2 ml/kg corn oil
16 GSM819061_US10283824_252828210177_S01_GE1_107_Sep09_1_4.txt           2 ml/kg corn oil
17 GSM819060_US10283824_252828210177_S01_GE1_107_Sep09_1_3.txt           2 ml/kg corn oil
18 GSM819059_US10283824_252828210177_S01_GE1_107_Sep09_1_2.txt           10 ml/kg saline
19 GSM819058_US10283824_252828210177_S01_GE1_107_Sep09_1_1.txt           10 ml/kg saline
```

We are interested in the treatment column:

```
> Treatment <- SDRF[,"Characteristics[treatment]"]
```

We set the saline control to be the first level of the Treatment factor.

```
> levels <- c("10 ml/kg saline","2 ml/kg corn oil","5 ml/kg corn oil","10 ml/kg corn oil")
> Treatment <- factor(Treatment, levels=levels)
```

Next read the intensity data:

```
> x <- read.maimages(SDRF[,"Array Data File"],
+           source="agilent", green.only=TRUE, other.columns="gIsWellAboveBG")
Read GSM819076_US10283824_252828210181_S01_GE1_107_Sep09_1_4.txt
Read GSM819075_US10283824_252828210181_S01_GE1_107_Sep09_1_3.txt
Read GSM819074_US10283824_252828210181_S01_GE1_107_Sep09_1_2.txt
Read GSM819073_US10283824_252828210180_S01_GE1_107_Sep09_1_4.txt
Read GSM819072_US10283824_252828210180_S01_GE1_107_Sep09_1_3.txt
Read GSM819071_US10283824_252828210180_S01_GE1_107_Sep09_1_2.txt
Read GSM819070_US10283824_252828210180_S01_GE1_107_Sep09_1_1.txt
Read GSM819069_US10283824_252828210179_S01_GE1_107_Sep09_1_4.txt
Read GSM819068_US10283824_252828210179_S01_GE1_107_Sep09_1_3.txt
Read GSM819067_US10283824_252828210179_S01_GE1_107_Sep09_1_2.txt
Read GSM819066_US10283824_252828210179_S01_GE1_107_Sep09_1_1.txt
Read GSM819065_US10283824_252828210178_S01_GE1_107_Sep09_1_4.txt
Read GSM819064_US10283824_252828210178_S01_GE1_107_Sep09_1_3.txt
Read GSM819063_US10283824_252828210178_S01_GE1_107_Sep09_1_2.txt
Read GSM819062_US10283824_252828210178_S01_GE1_107_Sep09_1_1.txt
Read GSM819061_US10283824_252828210177_S01_GE1_107_Sep09_1_4.txt
Read GSM819060_US10283824_252828210177_S01_GE1_107_Sep09_1_3.txt
Read GSM819059_US10283824_252828210177_S01_GE1_107_Sep09_1_2.txt
Read GSM819058_US10283824_252828210177_S01_GE1_107_Sep09_1_1.txt
```

Note that we have read in the extra column `gIsWellAboveBG`, which records whether the intensity of each spot is considered above the background level for that array. This column will help us later with probe filtering.

The data has 44,254 probes and 19 arrays:

```
> dim(x)
[1] 44254    19
```

17.4.4 Gene annotation

We can use the annotation package for this type of Agilent array, `RnAgilentDesign028282.db`, to get gene symbols and Entrez Gene Ids from the probe Ids:

```
> library(RnAgilentDesign028282.db)
> x$genes$EntrezID <- mapIds(RnAgilentDesign028282.db, x$genes$ProbeName,
+                            keytype="PROBEID", column="ENTREZID")
> x$genes$Symbol <- mapIds(RnAgilentDesign028282.db, x$genes$ProbeName,
+                            keytype="PROBEID", column="SYMBOL")
> x$genes[201:205,]
    Row Col ControlType    ProbeName SystematicName EntrezID Symbol
201   3  34           0 A_42_P642757      NM_031235    81918  Pard3
202   3  35           0 A_64_P066694   A_64_P066694     <NA>   <NA>
203   3  36           0 A_42_P699201      XM_227798   295538 Depdc1
204   3  37           0 A_42_P591944      NM_172093    94164   Hbg1
205   3  38           0 A_44_P115293   NM_001108294   360611  Copz2
```

17.4.5 Background correction and normalize

We use normexp background correction followed by quantile normalization:

```
> y <- backgroundCorrect(x, method="normexp")
Array 1 corrected
Array 2 corrected
Array 3 corrected
Array 4 corrected
Array 5 corrected
Array 6 corrected
Array 7 corrected
Array 8 corrected
Array 9 corrected
Array 10 corrected
Array 11 corrected
Array 12 corrected
Array 13 corrected
Array 14 corrected
Array 15 corrected
Array 16 corrected
Array 17 corrected
Array 18 corrected
Array 19 corrected
> y <- normalizeBetweenArrays(y, method="quantile")
```

17.4.6 Gene filtering

We will filter out control probes as indicated by the `ControlType` column:

```
> Control <- y$genes$ControlType==1L
```

We will also filter out probes with no Entrez Gene Id or Symbol

```
> NoSymbol <- is.na(y$genes$Symbol)
```

Finally, we will filter probes that don't appear to be expressed. We keep probes that are above background on at least four arrays (because there are four replicates of each treatment):

```
> IsExpr <- rowSums(y$other$gIsWellAboveBG > 0) >= 4
```

Now we select the probes to keep in a new data object `yfilt`:

```
> yfilt <- y[!Control & !NoSymbol & IsExpr, ]
> dim(yfilt)
[1] 27191    19
```

To be tidy, we remove annotation columns we no longer need:

```
> yfilt$genes <- yfilt$genes[,c("ProbeName","Symbol","EntrezID")]
> head(yfilt$genes)
      ProbeName  Symbol EntrezID
13 A_64_P002176   Wdfy3   305164
14 A_42_P664913 Ankrd37   361149
15  A_43_P13320    Ifng    25712
18  A_43_P11804     Ptn    24924
19 A_44_P808710     Rd3   684158
20 A_64_P142111  Gxylt1   300173
```

17.4.7 Differential expression

Now we can find genes differentially expressed for the corn oil treatments compared to the saline control:

```
> design <- model.matrix(~Treatment)
> fit <- lmFit(yfilt,design)
> fit <- eBayes(fit,trend=TRUE,robust=TRUE)
> summary(decideTests(fit[,-1]))
       Treatment2 ml/kg corn oil Treatment5 ml/kg corn oil Treatment10 ml/kg corn oil
Down                           0                       0                        756
NotSig                     27191                   27191                      24898
Up                             0                       0                       1537
```

It appears that only the 10 ml/kg treatment is different from the saline control. The top 20 genes for the 10 ml/kg treatment are as follows:

```
> topTable(fit,coef=4,n=20)
         ProbeName    Symbol  EntrezID logFC AveExpr     t  P.Value adj.P.Val    B
28763 A_44_P552452    RT1-Bb    309622  9.33    8.78  45.3 1.78e-21  4.85e-17 32.3
9680  A_44_P991565       Bad     64639  3.73    7.91  39.1 2.14e-20  2.90e-16 31.4
42069 A_64_P160096      Mei1    315162  3.84    7.67  32.6 7.94e-19  7.20e-15 29.4
12513 A_42_P638620      Lcn2    170496  3.83    8.09  32.1 1.07e-18  7.29e-15 29.2
6942  A_42_P667782   Fastkd2    301463 -1.70   12.44 -24.5 2.14e-16  1.16e-12 25.7
4631  A_64_P006625     RT1-A 100188935  2.65   13.17  24.1 2.74e-16  1.24e-12 25.5
40353 A_64_P149280     Vegfb     89811 -3.26    8.69 -21.2 3.49e-15  1.36e-11 23.6
655   A_42_P667782   Fastkd2    301463 -1.44   12.13 -20.5 6.55e-15  2.23e-11 23.1
23459 A_42_P667782   Fastkd2    301463 -1.64   12.53 -20.3 7.43e-15  2.25e-11 23.0
22161 A_64_P154811   Ccdc146    499980  1.79    6.97  20.1 9.39e-15  2.55e-11 22.8
37136 A_42_P667782   Fastkd2    301463 -1.78   12.32 -20.0 1.04e-14  2.57e-11 22.7
10116 A_42_P667782   Fastkd2    301463 -1.71   12.49 -19.4 1.77e-14  4.01e-11 22.3
36556 A_42_P667782   Fastkd2    301463 -1.81   12.36 -18.9 3.10e-14  6.48e-11 21.8
29421 A_42_P667782   Fastkd2    301463 -1.72   12.53 -18.3 5.71e-14  1.11e-10 21.3
23088 A_64_P163386 LOC691921    691921  1.49   11.22  18.0 7.41e-14  1.34e-10 21.1
35094 A_64_P054586     Usp9x    363445 -1.59    9.35 -17.8 9.75e-14  1.66e-10 20.8
13512 A_42_P667782   Fastkd2    301463 -1.56   12.64 -17.3 1.67e-13  2.68e-10 20.4
11619 A_64_P107239     A4gnt    685758  1.17    6.46  16.8 2.72e-13  4.11e-10 20.0
22476 A_42_P667782   Fastkd2    301463 -1.55   12.65 -16.5 3.93e-13  5.63e-10 19.6
371   A_64_P059545     Mlycd     85239 -2.43   13.28 -16.4 6.07e-13  8.25e-10 19.2
```

17.4.8 Gene ontology analysis

```
> g <- goana(fit, coef=4, species="Rn", geneid="EntrezID")
> topGO(g,n=20,truncate="50")
                                                    Term Ont    N  Up Down    P.Up P.Down
GO:0006691                     leukotriene metabolic process  BP   18   9    2 6.1e-07 0.1048
GO:0042605                           peptide antigen binding  MF   24  10    2 1.2e-06 0.1681
GO:0019370                  leukotriene biosynthetic process  BP   11   7    0 1.3e-06 1.0000
GO:0002376                               immune system process  BP 1690 158   56 1.4e-06 0.2988
GO:0019882               antigen processing and presentation  BP   77  18    6 1.7e-06 0.0312
GO:0006955                                   immune response  BP  947  99   37 1.8e-06 0.0803
GO:0007169 transmembrane receptor protein tyrosine kinase ...  BP  445  55   17 3.7e-06 0.2134
GO:0042611                             MHC protein complex  CC   19   8    1 1.3e-05 0.4484
GO:0009605                     response to external stimulus  BP 1840 163   64 2.2e-05 0.1619
GO:0048002 antigen processing and presentation of peptide ...  BP   45  12    4 2.2e-05 0.0491
GO:0002474 antigen processing and presentation of peptide ...  BP   26   9    2 2.4e-05 0.1904
GO:0001819         positive regulation of cytokine production  BP  333  42   11 3.2e-05 0.4503
GO:0042107                        cytokine metabolic process  BP   88  17    3 4.8e-05 0.5124
GO:0009607                      response to biotic stimulus  BP  775  79   28 5.0e-05 0.2157
GO:0051883 killing of cells in other organism involved in ...  BP   17   7    0 5.5e-05 1.0000
GO:0006952                                 defense response  BP 1031  99   42 6.1e-05 0.0381
GO:0006954                               inflammatory response  BP  515  57   23 6.5e-05 0.0481
GO:0007166            cell surface receptor signaling pathway  BP 1783 156   73 6.5e-05 0.0063
GO:0046456                     icosanoid biosynthetic process  BP   36  10    2 7.4e-05 0.3050
GO:0044364              disruption of cells of other organism  BP   43  11    0 7.4e-05 1.0000
```