Cocaine_Gene_Expression.R

steve

2021-02-19

```r
# Names: Steven Nguyen
#
# Title: Identifying differentially expressed genes between cocaine addict deaths and, non cocaine addi

# Importing libraries
library(affy)
```

```
## Warning: package 'affy' was built under R version 4.0.3

## Loading required package: BiocGenerics

## Warning: package 'BiocGenerics' was built under R version 4.0.3

## Loading required package: parallel

##
## Attaching package: 'BiocGenerics'

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

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

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

## Loading required package: Biobase

## Warning: package 'Biobase' was built under R version 4.0.3

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

```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

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

## The following objects are masked from 'package:BiocGenerics':
##
##     combine, intersect, setdiff, union

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(ggplot2)
library(GEOquery)
```

```
## Warning: package 'GEOquery' was built under R version 4.0.3

## Setting options('download.file.method.GEOquery'='auto')

## Setting options('GEOquery.inmemory.gpl'=FALSE)
```

```r
library(limma)
```

```
## Warning: package 'limma' was built under R version 4.0.3

##
## Attaching package: 'limma'

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

```r
# 1 & 2: Importing the data from directory
GSE54839 <- getGEO("GSE54839")
```

```
## Found 1 file(s)

## GSE54839_series_matrix.txt.gz

##
## -- Column specification --------------------------------------------------------
## cols(
##   .default = col_double(),
##   ID_REF = col_character()
## )
## i Use 'spec()' for the full column specifications.

## File stored at:

## C:\Users\steve\AppData\Local\Temp\RtmpuWsqnn/GPL6947.soft
```

```r
GSE54839.expr <- exprs(GSE54839[[1]])
GSE54839.p <- pData(GSE54839[[1]])
```

```
# 3: Generating a boxplot to see if processed data is already in log2 scale
GSE54839.exprTrunc <- head(GSE54839.expr, 10)
boxplot(GSE54839.exprTrunc, main = "processed data", xlab = "Sample Name", ylab = "Expression of Samples
```
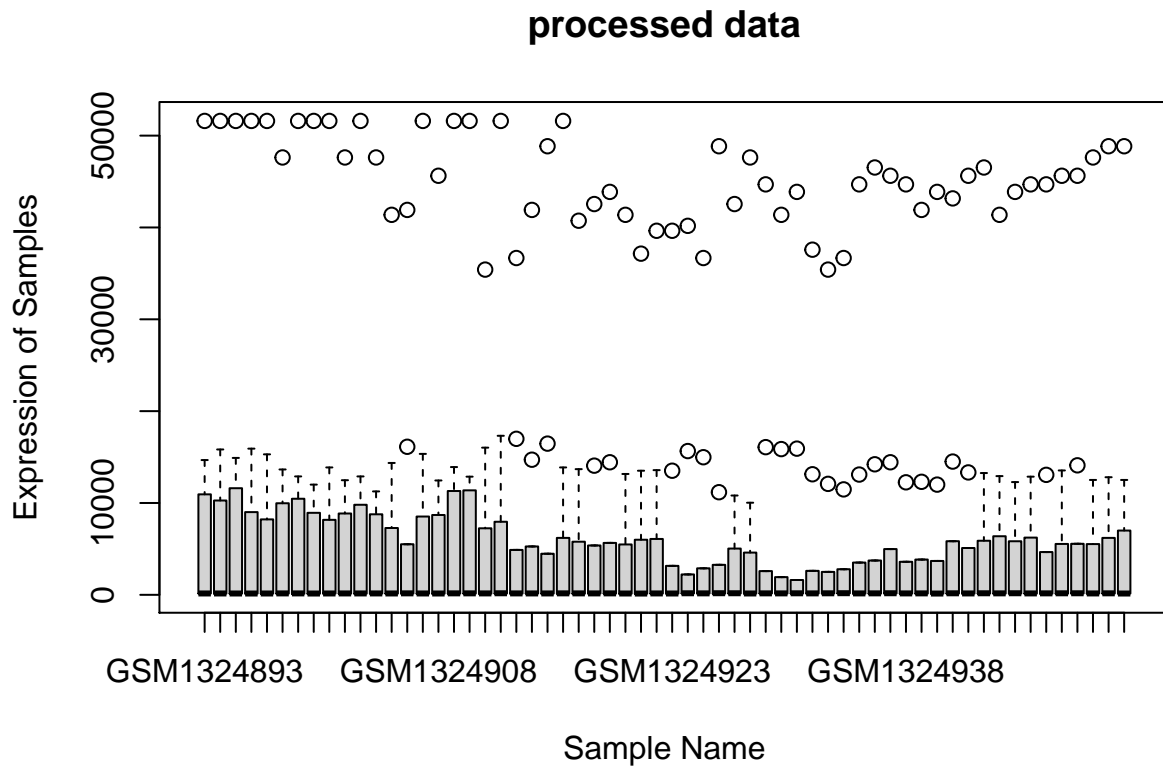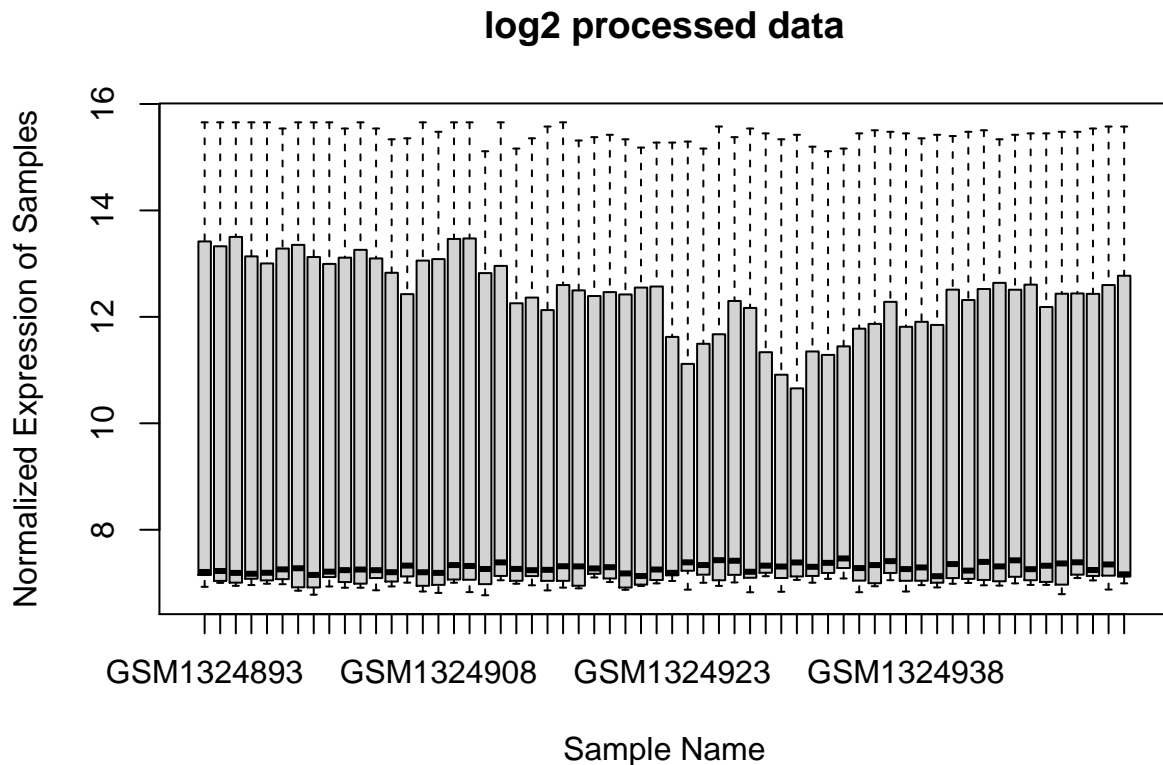
processed data


```
   # Since not in log2 scale, normalizing data with log2 scale
GSE54839.exprLog <- log2(GSE54839.expr)
GSE54839.exprLogTrunc <- head(GSE54839.exprLog, 10)
boxplot(GSE54839.exprLogTrunc, main = "log2 processed data", xlab = "Sample Name", ylab = "Normalized E
```

log2 processed data


```r
# 4: Finding the number of probes and samples
ncol(GSE54839.expr)
```

```
## [1] 60
```

```r
nrow(GSE54839.expr)
```

```
## [1] 48761
```

```r
  # There are 60 samples and 48761 probes in the dataset

# 5: Extracting the column that contains the categories that we would like to compare
addiction.status <- as.character(GSE54839.p$`disease state:ch1`)
table(addiction.status)
```

```
## addiction.status
## cocaine addiction          control
##                30               30
```

```r
  # There are 30 individuals that were part of the group with cocaine addiction, and there are 30 indiv

# 6: Finding the differentially expressed probes across the 2 groups

design <- model.matrix(~-1+addiction.status)
colnames(design) <- c("control", "cocaine.addiction")

fit <- lmFit(GSE54839.exprLog, design)
contrast.matrix <- makeContrasts(control - cocaine.addiction,levels = design)
fit2 <- contrasts.fit(fit, contrast.matrix)
```

```r
fit2 = eBayes(fit2)

tt20 <- topTable(fit2,sort.by = "p", p.value = 0.20, number = nrow(GSE54839.exprLog))

# 7: Constructing a boxplot for the probe with the lowest adjusted p-value

m <- match(rownames(tt20)[1], rownames(GSE54839.exprLog))
probe <- GSE54839.exprLog[m,]
df <- data.frame(expr = probe, addiction.status = addiction.status)

means <- df %>% group_by(addiction.status) %>% summarize(mean = mean(expr))
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
means
```

```
## # A tibble: 2 x 2
##   addiction.status    mean
##   <chr>              <dbl>
## 1 cocaine addiction   9.11
## 2 control             8.61
```

```r
diff(means$mean)
```

```
## [1] -0.5029022
```

```r
tt20$logFC[1]
```

```
## [1] 0.5029022
```

```r
# convert to FC #
logFC <- tt20$logFC[1]
2**logFC
```

```
## [1] 1.417061
```

```r
## visualize ##
FC <- paste0("FC = ", round(2**logFC, 2))
main <- paste0("Expression of ", rownames(tt20)[1], ", ", FC)

ggplot(df, aes(x = addiction.status, y = expr, fill = addiction.status)) + geom_boxplot() +
  ylab("log2 expression") + xlab("Addiction status") + ggtitle(main) +
  scale_fill_manual(values = c("red", "yellow")) +
  theme_classic() + theme(legend.position = "none")
```

Expression of ILMN_1774077, FC = 1.42
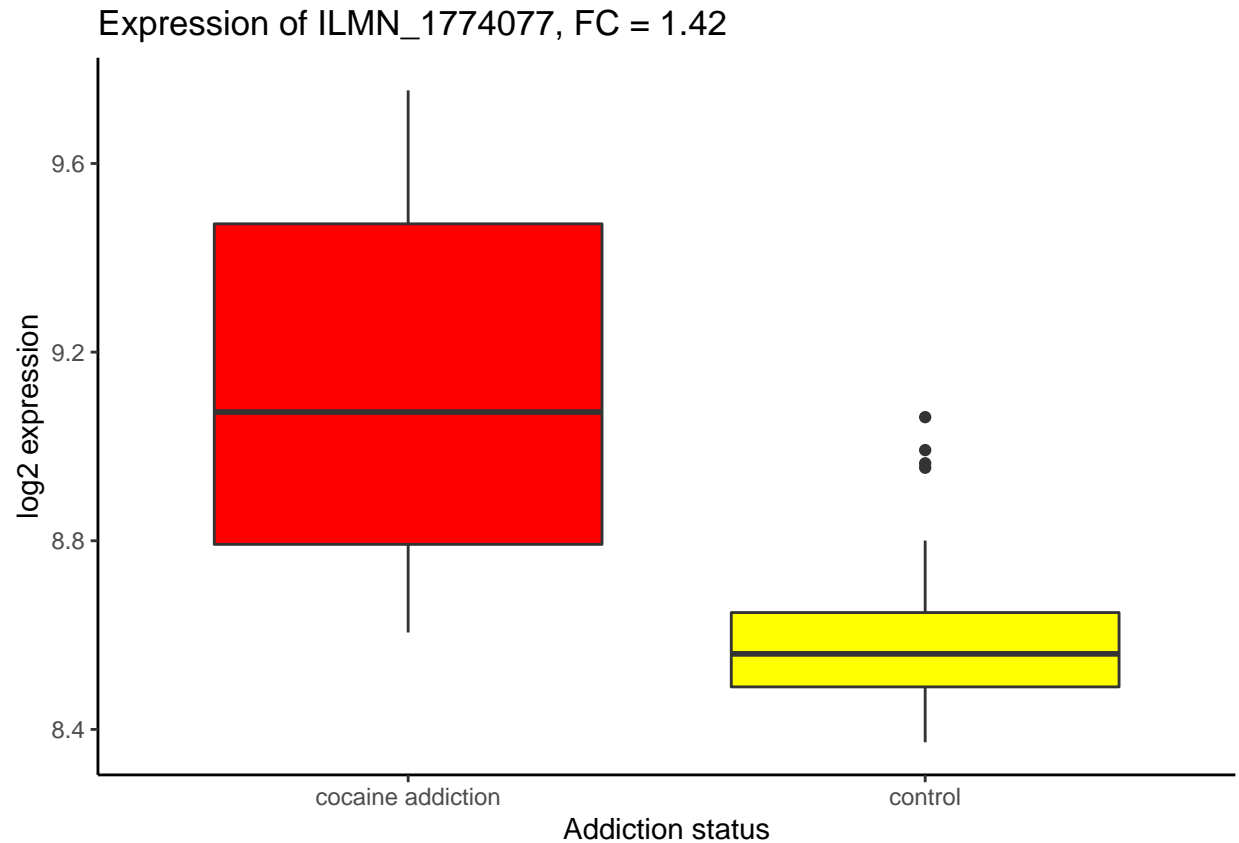

```
# 8: Using R to output the annotation of the data
platform2 <- annotation(GSE54839[[1]])

# 9: Downloading the platform (GPL) for this data
pl2 <- getGEO(platform2)
```

```
## Using locally cached version of GPL6947 found here:
## C:\Users\steve\AppData\Local\Temp\RtmpuWsqnn/GPL6947.soft
```

```
pl2 <- Table(pl2)

probe <- rownames(tt20)[1]
m <- match(probe, pl2$ID)
pl2$`ILMN_Gene`[m]
```

```
## [1] "GBP2"
```

```
# 10: Finding the gene names corresponding to all probes, and creating a table

match <- match(rownames(tt20), rownames(GSE54839.exprLog))
ILMNGene <- pl2$`ILMN_Gene`[match]

df20 <- data.frame(tt20)
df220 <- df20 %>% select(logFC,adj.P.Val)
df220["ILMN_Gene"] <- ILMNGene

head(df220, 5)
```

```
##                   logFC   adj.P.Val ILMN_Gene
## ILMN_1774077  0.5029022 0.000346728      RGL4
## ILMN_2103919 -0.3447804 0.002191241  ITGB1BP1
## ILMN_1695706  0.4978619 0.004252562    COX4I2
## ILMN_1667825  0.4422986 0.004252562 HS.537387
## ILMN_1886493  0.3352531 0.004252562     BOLA2
```

```r
# 11: Using DAVID to identify Gene Ontology and KEGG pathways that are associated with the differential

probe <- rownames(tt20)
m <- match(probe, pl2$ID)
genes<-pl2$ILMN_Gene[m]
keep <- genes!=""
genes <- genes[keep]
genes <- strsplit(genes, " /// ")
genes <- unlist(genes)
genes <- unique(genes)
write.table(genes, row.names = FALSE, quote = FALSE)
```

```
## x
## GBP2
## LRFN3
## H3F3B
## MLKL
## HS.535028
## MCL1
## SHMT1
## SLC39A8
## NXT1
## KCNB1
## EMP1
## CCDC109B
## HIST2H2AA3
## C19ORF4
## VCAM1
## CHSY1
## TNFRSF10B
## TMEM191B
## GSTO1
## C12ORF52
## ADAM17
## CLIC1
## HS.562318
## CCL2
## RGS16
## GAPDH
## LRRFIP2
## CDKN1A
## C14ORF152
## ENTPD4
## ATP2A3
## C7ORF43
## CAPS
## SEZ6L
## PATL1
```

```
## CERK
## LOC730012
## DNAJC15
## LOC649999
## COMTD1
## AURKA
## RDH12
## FER1L3
## COPS7A
## SCARA3
## HS.544112
## RELL2
## BTG3
## AXUD1
## BTG1
## HIST2H2AC
## PIB5PA
## EVI5L
## C12ORF10
## CCNL1
## NFKB1
## C16ORF59
## AP3D1
## GFPT2
## SDC1
## LRRC32
## HMGN1
## FAM69B
## SLITRK1
## PCDHB9
## SAT1
## CXCL10
## TUBG1
## RAB9B
## RIPK2
## C1ORF142
## MAST1
## RAB3IP
## SAP18
## FOSB
## TM4SF1
## JUN
## ECSIT
## MED8
## ATP5G1
## LOC440359
## LOC729378
## OCIAD2
## PORCN
## F3
## CSDA
## FKBP2
## LOC731835
## C9ORF125
```

```
## SPG3A
## SERPINA3
## COMMD5
## APOL2
## EVL
## SLC1A1
## UNC5A
## RPL37
## FLJ36070
## MGST3
## PSME3
## SOCS2
## LOC283514
## SNX5
## AP3S2
## CAMK1G
## HS.538770
## CD44
## LOC650128
## C8ORFK32
## CTHRC1
## PTPN3
## LOC653841
## OVGP1
## LOC728492
## BTBD6
## PTBP1
## LOC643882
## TAP1
## DPYS
## RNF165
## ANGPT1
## IRAK3
## LOC649095
## MESP1
## FLJ44968
## MDH2
## CD93
## PRMT8
## SDHA
## RRAD
## NFKBIZ
## NDUFV1
## UBE3C
## FIG4
## SULT2B1
## ATP6V0B
## UQCR
## ACTR1B
## IRF7
## STAG3L4
## MXD1
## PVALB
## CSF2RA
```

```
## TNFSF14
## FRMPD1
## RNF141
## PNKD
## CXXC4
## EIF2B2
## CACNA2D2
## ZADH1
## SPRYD5
## CHCHD7
## TFRC
## TMEM25
## C19ORF46
## ANXA6
## BDNF
## ARAF
## ITGB1BP1
## ATF3
## IER3
## LOC728229
## HEY1
## GSTM1
## PTTG2
## FLJ41046
## TIMM22
## BMP7
## STX11
## LOC400451
## COL4A1
## EDG3
## LOC145853
## MAP2K3
## FBXL15
## TMEM141
## ABCG4
## ABCB9
## C6ORF117
## PPP2R5B
## B4GALT4
## POPDC3
## DKFZP434N035
## COL8A1
## C8ORF70
## DLK2
## THAP1
## CORO2A
## GRK6
## LOC653857
## RBM23
## KLF11
## TMBIM4
## PDE1A
## RRAGA
## PRKCG
```

```
## LARGE
## AP3B2
## RBM17
## SNCB
## MAOB
## MRPS15
## BCCIP
## LOC442117
## SMPD1
## SHD
## ABCF3
## SYT7
## KIAA1754
## RNASE1
## TRIB1
## C1ORF102
## UHMK1
## SETD4
## GSTM2
## CHSY3
## EIF4ENIF1
## FSD1
## CH25H
## EXTL2
## NPNT
## NBPF11
## EHD4
## EIF2S3
## PSMB7
## PELI1
## TUBB6
## LOC653199
## UPF3A
## SLC6A3
## MSN
## MYH9
## GALNT9
## ZNF385D
## NUDT14
## ANK1
## LOC85389
## TMEM103
## CYC1
## DNASE1L3
## ATP5B
## CDAN1
## PLA1A
## RABGGTA
## MRRF
## ICMT
## C3ORF39
## BAG3
## APOC1
## PI4KAP2
```

```
## USP27X
## LOC642755
## RAB11FIP5
## PABPC1
## ST3GAL5
## ARFGAP3
## GPR4
## DCAKD
## ADM
## CLEC4G
## GADD45B
## HLA-B
## GBP1
## CBLN2
## HS.562947
## SYNGR1
## PLAUR
## RET
## PITPNM1
## TUSC2
## ARHGDIG
## SCN3B
## MAMDC2
## ETS1
## TMEM31
## IL4R
## HS.572444
## GLTSCR2
## UPP1
## ANP32B
## GARNL4
## MARCH3
## PRRT3
## HS.561493
## TMEM16F
## SOD2
## PKIB
## PCDH8
## ASPH
## SH2B3
## SMAD3
## SMAD7
## HERC5
## TNRC4
## TGIF1
## C21ORF56
## FOS
## C10ORF10
## ANKRD34A
## PNOC
## PRPF8
## ROPN1
## CACNG2
## GALT
```

```
## LOC730744
## ZNF229
## KCNJ6
## SV2A
## LECT1
## C6ORF159
## GEM
## ZFP36
## HS.544050
## RASGRP2
## KIAA0427
## HS.554410
## CEND1
## RPL12
## SMARCC1
## CCDC136
## SLC4A5
## PDLIM3
## RAXL1
## PCDH21
## DOLK
## SLC24A6
## HS.452398
## SLITRK3
## C6ORF168
## RGS7BP
## LOC653086
## RPL4
## CDK5
## SEMA3E
## PPAPDC3
## DAPP1
## VSTM2L
## RNF215
## ACCN1
## ATP1A1
## MAP4K4
## CAMK2B
## SRGN
## ZC3HAV1
## HS.434307
## GSTTP2
## PTDSS1
## ASCC2
## GPX4
## IL17D
## HIST1H2BD
## LYSMD4
## LOC220686
## LPHN1
## ITGA5
## FOXQ1
## CHRM5
## CLEC2L
```

```
## GNAI3
## DTWD2
## LOC650832
## CTNNAL1
## PGAM5
## POLR1E
## SLIT2
## LOC730051
## MAGI1
## FOXA2
## C12ORF24
## C21ORF70
## SERPING1
## AGRP
## HS.383564
## COX5B
## MADD
## PLCB4
## SPTBN4
## RHOC
## SCAMP5
## C20ORF82
## ENTPD6
## SERF1B
## LOC284701
## FOSL1
## C3ORF18
## NUDT18
## GPR12
## HS.566032
## ROBO2
## UQCC
## ATRN
## CYP2S1
## ME3
## UQCRC1
## DSE
## ATCAY
## DRD2
## WDR41
## HS.562122
## TRA2A
## DEPDC6
## ZFHX2
## FAHD1
## TNMD
## MCOLN1
## OR13C9
## ICA1
## CBLN1
## HS.566010
## HSPA6
## ALG3
## LOC644852
```

```
## ATXN7L3
## WDR1
## SLC9A7
## OGDHL
## CNTNAP1
## ID2
## OCRL
## SLC6A15
## CAMKV
## CEBPD
## GDF3
## SERTAD3
## SLC24A1
## PTP4A2
## TMEM41A
## CEP68
## C1ORF71
## LOC285359
## DEF8
## NIPSNAP1
## CCDC44
## FLJ20464
## EGLN2
## HIG2
## BID
## MFSD3
## RGS20
## DRG1
## GNA12
## EGR1
## NKTR
## SLITRK4
## LOC339344
## OPRL1
## DDX41
## GPD1L
## IL17RB
## HS.103173
## ASGR2
## TRIM3
## CALD1
## C2ORF40
## PKNOX1
## CECR7
## RNF26
## GDAP1L1
## ATP6V0D1
## IRF1
## C22ORF39
## TESK1
## LGI2
## VAV3
## RNF5P1
## PIP5K1B
```

```
## EHD1
## CNIH3
## HS.146766
## NTHL1
## MGC39606
## GBA
## EPB49
## PTPRR
## RIMS3
## SYNGR3
## YIPF2
## LIMS1
## HS.584478
## CD38
## CLP1
## CACNG7
## ZNF179
## CCDC24
## HS.4290
## LOC285307
## PSMC2
## NEDD9
## HSPB1
## CCL8
## HS.561963
## C6ORF136
## CPLX1
## HS.387982
## CADPS2
## IL18R1
## ALKBH6
## TH
## VCL
## FLJ20581
## ALDH2
## HSPA1B
## CAND2
## KIAA0513
## PPTC7
## DENR
## HS.568055
## SLC22A18
## PAPPA
## VARS
## SFXN4
## FAM71E1
## GAL3ST3
## HS.388347
## CACNB3
## PCDH10
## BXDC1
## NPHP3
## EPHA8
## CLTB
```

```
## KIF3C
## ATP2B2
## SLC2A11
## BLZF1
## KEL
## HS.531296
## FAM73B
## HS.545303
## TNFRSF6B
## ADAMTS9
## SERPINB3
## STMN2
## LRFN5
## DNAJB1
## LOC345630
## THY1
## LOC647310
## TTC8
## PTTG1IP
## FAM86C
## C10ORF65
## CTPS
## CDK2
## RPL5
## PPM1H
## COX10
## YES1
## CHURC1
## NR2F6
## ADAM8
## NFATC2IP
## LRFN2
## LRRN3
## LOC606724
## B4GALNT1
## BTRC
## CNTNAP5
## HSPD1
## SPAG16
## TIPARP
## REL
## PEX11B
## NMNAT2
## EVC
## HS.142910
## HS.547458
## CORO1A
## DNM1
## FAM80A
## HTRA2
## MX2
## TUBB2A
## NECAB3
## FAM43B
```

```
## ARHGAP11A
## AGTR1
## EBF3
## TUSC1
## FAM89A
## AHR
## CRY1
## CDC42SE1
## LOC643918
## HS.542163
## COPS3
## AKR1B1
## FAM108C1
## HS.551420
## ENO2
## FRAP1
## ST3GAL1
## DIRAS2
## HK1
## SERPINB1
## C6ORF1
## ATP6V1F
## SOCS3
## SERTAD1
## LOC149134
## SEC61A2
## PPP4R1
## HS.552896
## RAP1GAP
## DBN1
## DUSP5
## GSTT2
## NT5C3L
## DHDDS
## POU4F1
## PVRL1
## CTSH
## TSTA3
## LDB2
## GPR26
## NCDN
## TMEM38A
## HIST3H2A
## UBE2J1
## KIAA1853
## C1ORF54
## SPRY2
## CLCN6
## HS.444405
## FLJ37464
## RNF208
## LOC648689
## TSEN34
## PMS2L2
```

```
## HS.184721
## TIMP1
## CEP170
## PSMD4
## SERPINH1
## TMEM55B
## SOX12
## CLIP3
## LOC654103
## HARS
## PTPRCAP
## RLTPR
## PPP1R3C
## ZNF219
## SNX19
## LOC643011
## RTN3
## EHD3
## DHX30
## CDK5R2
## PRPH
## EML3
## PPP1R15A
## CLPTM1L
## FHOD1
## CKAP5
## ECE2
## ZNF313
## HS.543482
## GBF1
## DUSP26
## DRAM
## KPTN
## PDZD7
## LIMK1
## KTI12
## CEBPB
## SNRPB
## CLSTN3
## HS.550452
## MYBPH
## CNNM1
## C20RF34
## TMEM130
## GNAO1
## HSDL2
## HS.578993
## P4HA1
## GRIN1
## LOC653489
## UBE2O
## HS3ST2
## PIN1
## HAPLN4
```

```
## LOC647509
## MAP1LC3A
## C12ORF11
## LOC647834
## HS.22689
## LOC643453
## ADAMTS1
## IER5
## DPP6
## SERINC3
## SLC25A37
## C1ORF83
## CISD1
## FBXW5
## ISG20
## BRSK1
## NEFL
## CCNA1
## LOC650681
## DDX5
## GOT2
## AP1S1
## TCP11L2
## MAP4
## LOC727922
## PACRG
## MED4
## HS.540022
## USP47
## TMEM137
## INPP4A
## HS.579963
## FBL
## LOC388275
## MYOZ2
## RAB37
## CCDC110
## ARHGEF11
## B3GALT4
## NRD1
## NRIP3
## EXTL1
## KATNAL2
## TP53I13
## IMPDH1
## C6ORF154
## DDX1
## RNASE4
## GZMK
## MAFF
## LDLR
## RAB12
## CBS
## ANKRD46
```

```
## LOC399947
## CCDC51
## SLC16A3
## MYH7
## NDRG4
## ATM
## FIBP
## LOC647000
## CACNG3
## DOLPP1
## Septin 6
## HSPA1L
## MAGI3
## CYFIP2
## ADAM29
## MFSD5
## BCL6
## CMTM6
## FBXO44
## OCIAD1
## PITRM1
## SORBS1
## SLC10A4
## STXBP1
## HSPA1A
## CTGLF3
## LOC643968
## SLC7A4
## STAM
## PARD6A
## AES
## COL7A1
## RPL8
## RFPL1S
## GAS6
## IGSF3
## POLR3GL
## TRAK2
## LIN7B
## SSTR1
## ATP2B3
## DYNLRB2
## YARS
## LRRTM1
## ATP6V1B2
## CCNK
## RANBP3
## TWIST1
## TMEM51
## NLRP8
## MBTPS1
## RAB6B
## LOC644278
## NPTN
```

```
## DYM
## LOC728729
## APOL5
## CCL20
## KLHL13
## LOC201140
## SAE1
## PNN
## NEFM
## RPUSD1
## MED10
## ZCCHC14
## HS.577474
## EN1
## LOC645688
## S100A10
## ACOT7
## ARRDC4
## TNFAIP1
## HPS6
## TXNDC13
## ITM2C
## RBM9
## DRD1IP
## GLTPD1
## ZNF329
## FJX1
## BBOX1
## SVOP
## CBARA1
## RN7SK
## TP53BP1
## LOC388135
## EIF3F
## RHBDF1
## FTH1
## AJAP1
## NME1
## SLC35C1
## C18ORF17
## ST5
## NFIL3
## CEP27
## CEP70
## PIGA
## HS.574124
## SYT17
## SLC25A22
## LOC644596
## PLEKHG1
## SRD5A1
## FXR2
## HLA-E
## STAT1
```

```
## C21ORF63
## HS.543076
## RAB11FIP1
## RASL10A
## KIF1A
## DET1
## LOC286208
## C1ORF216
## PRKCZ
## KIAA0196
## MKNK2
## KIAA0040
## NEK2
## MTA3
## CPLX2
## EPHA2
## ZNF224
## GPR137
## PTGS2
## ZNF678
## HOOK2
## SSX4
## DYDC1
## HS.567036
## C20ORF100
## INPP4B
## HS.563926
## LOC388610
## LOC387856
## PSEN2
## FLJ45248
## CLYBL
## OR4K15
## AKR1C3
## NFATC3
## KIAA0672
## ACHE
## CCRK
## C1ORF85
## HS.156525
## HIST2H2BF
## HS.417262
## LOC643286
## STXBP2
## FOXRED1
## HS.566642
## SERGEF
## DNAH11
## LRIG2
## NP
## ROD1
## ATP13A2
## ADFP
## TMEM18
```

```
## HS.134065
## BFSP2
## STOML1
## CALM3
## CPNE4
## HIST1H3H
## INSIG1
## HS.541932
## KIAA0406
## SRI
## KCNMB2
## C10ORF84
## GFRA1
## ODC1
## DIRAS1
## FLJ20125
## LOC646256
## SNRPE
## FLJ20160
## DSCR1L1
## VARS2
## GRM1
## ABHD14A
## NOS3
## SLC2A6
## BARD1
## GABRB3
## FAM46C
## SCYL1
## CASP4
## LOC651621
## MSI2
## MGST1
## CAMK2A
## LOC643075
## PPP1R15B
## CLN5
## CNN3
## ASPSCR1
## OSR2
## CRYGD
## PARD6G
## TROVE2
## APITD1
## FGF9
## USP49
## ERN1
## PPM1E
## SLC2A13
## PIP5K1C
## JPH4
## SMU1
## EFEMP1
## PDLIM1
```

```
## F8A1
## YTHDC1
## C10ORF104
## AGK
## LILRB3
## SEC11A
## PIGF
## GBA2
## DNAJC5
## GPR162
## BAT1
## KCND3
## TTC1
## HS.542027
## BNIP2
## FDX1L
## HS.297557
## ENO3
## RUNX1T1
## LOC645261
## LOC729985
## LOC642356
## C14ORF173
## TMEM160
## LOC652730
## ANKRD2
## PTPN5
## KLF6
## GALNT11
## HS.554507
## HS.131714
## CTNND1
## TAGLN3
## CENPF
## MRPL41
## HS.520781
## MRPS12
## PTPRN
## TBC1D24
## MRPS18C
## LRCH4
## HS.570617
## DLK1
## HS.537685
## WTAP
## AMT
## FGF2
## SLC32A1
## TMEM110
## RGS4
## PDPN
## HP1BP3
## LOC652612
## MICALL2
```

```
## PRRT2
## LNX1
## ENTPD3
## ADCY5
## HS.561819
## BSCL2
## DAB2
## NR1I2
## HS.162932
## HS.122864
## ETF1
## IDH3B
## KIAA0514
## XKR4
## RGAG1
## AP2M1
## FAM134A
## KIFAP3
## TMEM2
## S100B
## LOC642351
## BCKDHB
## AIRE
## LOC198437
## RUFY3
## DKK1
## C12ORF45
## ZNF486
## MLYCD
## SCCPDH
## C12ORF39
## ZNF223
## ELF1
## RPS4Y1
## HS.584143
## CPEB3
## IER5L
## NEFH
## HMOX1
## LOC650668
## DMRT2
## POLDIP2
## ZNF621
## LOC407835
## LOC401010
## LONP2
## LRWD1
## SLIT1
## ERF
## ELAVL4
## ADAD2
## VDAC3
## ZNF831
## AACS
```

```
## S100A9
## APOO
## WDR74
## GDPD2
## EPHA10
## PADI4
## CSDC2
## HPGD
## C16ORF57
## PPM1J
## CECR6
## HTR2A
## RKHD1
## C5AR1
## LRRC49
## SEZ6L2
## COPS5
## CYP2B6
## KIAA1618
## ATP6V0E1
## NDUFA11
## WWOX
## PJA1
## CD8A
## LOC645317
## ANGPT2
## ZC3H12A
## GATA2
## PHGDH
## WDR37
## ADNP2
## COG7
## IL8
## C16ORF80
## KIAA0363
## PLD3
## GBGT1
## KCTD2
## TBC1D19
## CD320
## LOC653584
## PLOD2
## OKL38
## PPA2
## ARHGEF6
## HS.502672
## MPZL2
## LARS
## MGC13379
## GDPD5
## NAV2
## KCNQ2
## CRMP1
## ZNF14
```

```
## PAQR6
## SRGAP3
## CREB5
## TYW1
## TIMM10
## PHIP
## KIAA1543
## MIA3
## FAM70B
## ALDOA
## HS.579350
## PRKD1
## LOC389833
## HS.574226
## ZNF721
## SYP
## PTPLAD2
## RIMS1
## NLF2
## TUBA1B
## CYB561
## ABCA1
## SYT3
## HS.136279
## ASNS
## KRT126P
## CHRFAM7A
## LDOC1
## PPP1R14C
## CLEC4GP1
## SCG2
## CYBRD1
## AKAP13
## HS.518888
## LSMD1
## C20ORF75
## FBXO16
## DKFZP781N1041
## ABI3BP
## HS.541888
## COPG
## HS.569948
## STX2
## CXCL16
## DCLRE1C
## CCDC32
## IL1A
## FBXL2
## ELFN2
## KLHDC8A
## PLEKHJ1
## PDXP
## ARF1
## MORG1
```

```
## RHOT1
## PEMT
## NPHP4
## TRAPPC5
## SYT9
## TMEM47
## HS.162371
## SLC39A3
## ADORA3
## GSTM3
## NME2
## TRPM6
## CCDC85A
## KIAA1324
## SLC12A5
## HS.550282
## APLN
## DNAJC12
## LOH11CR2A
## KCNMB1
## PRDX5
## SCD5
## LMX1A
## SPTLC3
## VGLL3
## ANKRD50
## LOC283871
## N4BP2
## RAB40C
## LOC644195
## STK3
## AP1M1
## MAGEC2
## LOC440704
## ATP6V1E2
## HS.44677
## TBC1D3G
## B4GALT5
## LATS1
## PKP4
## CADM1
## TMEM37
## PKD1L2
## MAPT
## HS.282397
## FILIP1L
## XYLT2
## STXBP6
## FGF1
## DIDO1
## GPR161
## DCTN1
## PDXK
## WNK1
```

```
## LOC644150
## CHD1
## TERF2
## PLCD3
## ATP6V1C2
## CMTM1
## FGF4
## NDUFA8
## CCDC50
## C3ORF60
## DBNDD1
## CDC42
## STON1
## CGI-96
## JAM2
## IKBKAP
## SELE
## GJD3
## ERAP1
## DDC
## HS.541235
## IFI16
## ABHD11
## SYN1
## TFPI
## HS.7023
## LOC646664
## POLR2H
## RNMTL1
## SLC22A17
## RAD51L3
## EML4
## HS.486010
## CMTM4
## MAP1A
## IQCK
## BSN
## RHBDD2
## PIWIL1
## ZNHIT4
## BASP1
## MRP63
## YARS2
## HDHD3
## PLA2G2D
## SPINT2
## SYTL4
## HS.536770
## CDKN2AIP
## SERTAD2
## MIDN
## HSPB8
## CCNB1IP1
## LOC341457
```

```
## C17ORF28
## CLIC6
## ANTXR1
## FBXO2
## ABCA6
## PER2
## DGKQ
## FLJ10213
## HS.510096
## C8ORF76
## CNGA3
## C2CD2L
## RUNDC3A
## HS.439057
## GNG5
## STK39
## LOC728006
## LMO3
## HS.581946
## OSBPL10
## OR2T34
## MPPE1
## NDUFAF2
## SNAPC2
## LOC649970
## FAM71F1
## SERPINE1
## TMEM60
## SEC31B
## HS.204433
## SLC8A3
## CEACAM6
## SNPH
## HS.352677
## SNAP23
## ABCF2
## LOC654042
## CCL5
## ALDH5A1
## FLJ43692
## ZNF266
## TNFRSF10A
## HCFC1R1
## NDUFA3
## OR7E156P
## FYN
## FSTL3
## HS.580825
## KLC1
## OPA3
## CAPNS1
## AASS
## HNRPDL
## TRIM48
```

```
## RIMS2
## RNF157
## CHI3L2
## MOCS1
## ATP6V0C
## LOC645458
## KBTBD8
## MRPS6
## TPR
## RGS6
## SEMA4B
## EXOSC1
## IMP4
## AMOTL1
## ST3GAL2
## ALAS2
## LINGO1
## HAGH
## PPP2R5D
## HS.452445
## PCDH20
## HSPA12A
## LRRCC1
## SV2C
## ADAM23
## JUNB
## HRSP12
## GEMIN6
## RAB3C
## NSMCE2
## N6AMT2
## SMO
## ALS2CR14
## OTUD7A
## ERP29
## NUP93
## TRAF6
## TBC1D22B
## C10ORF58
## PDE1B
## TMEM203
## CCDC84
## HS.547824
## HS.538480
## BRUNOL6
## RCAN2
## LYNX1
## HS.120287
## SOX2
## LGALS3BP
## TNNT2
## DOCK3
## TOM1
## LINGO2
```

```
## PTPLA
## C14ORF156
## FLJ22184
## CDKN2D
## APEH
## HS.345154
## KRTCAP3
## NDEL1
## ALDH1A1
## KCNE1L
## RBBP6
## PKM2
## EVI2B
## HS.149038
## CCDC99
## HIST1H2BK
## DDHD2
## ANKRD29
## ABCA5
## RAB3D
## KRT222P
## SLC35F2
## SPON1
## RNF126
## NINJ1
## HAAO
## TMEM32
## IGFBP4
## INHA
## RGS9
## GMEB1
## LOC649817
## BMP2KL
## SNX7
## MAP1S
## KIAA1688
## TDRD5
## HS.130919
## UPP2
## BIRC3
## CBFB
## USP5
## FOXA1
## PANX1
## NPM2
## KIF5C
## HS.541072
## AQP11
## RPS26L
## LOC650298
## MRCL3
## NLGN4X
## PDLIM5
## C1ORF152
```

```
## TARBP2
## HS.90221
## SRA1
## HS.530207
## CCDC56
## YIF1B
## IL24
## SRPX
## GCN5L2
## HS.558212
## LOC651933
## HS.562901
## HS.124848
## GPR177
## LYPLA3
## SLC25A41
## EPDR1
## HS.565092
## LOC728944
## TNFRSF4
## TIMM23
## HMP19
## FERMT2
## ZNF280B
## RPL36AL
## STK36
## HS.566242
## SLC35D3
## TUBB
## CKMT1A
## OPCML
## LOC440345
## CMAS
## DLG2
## THBD
## SLC9A2
## OSBPL3
## RHPN1
## SOCS7
## CAPRIN2
## CA5B
## EGR2
## ST7OT4
## DYRK2
## RNF123
## ONECUT2
## LPO
## NKD2
## GINS2
## DUOX1
## ABHD8
## COX19
## KAL1
## FBLN7
```

```
## TCEA3
## HS.551391
## CNOT6
## C10ORF132
## MS4A6A
## HS.540296
## SEMA4F
## OBFC1
## NEK11
## IFNAR2
## LPPR4
## GAD2
## VMO1
## APEX2
## DPM2
## HS.575038
## SP110
## TUBB3
## SLC7A8
## HS.369978
## FLJ46309
## WNT10B
## HS.538692
## RBMS1
## TAF15
## LOC158160
## RCN1
## D4S234E
## LOC653555
## HS.151444
## NFKBIA
## PARP6
## NAMPT
## LOC643995
## TTC5
## ERCC3
## AGPAT7
## LOC642527
## CD79A
## CHN1
## C13ORF23
## SETX
## TRAPPC1
## HEY2
## C9ORF109
## HS.187621
## ACTA1
## GPR89B
## LRRC15
## LOC653610
## FBXO32
## KIAA1622
## PRMT7
## SPTB
```

```
## MLL5
## FLJ11286
## FLJ25404
## TRMT11
## HS.562356
## CREG2
## RALYL
```

12:

We analzyed the GSE54839 dataset, where this dataset compares 2 groups, one group being the control (
have a cocaine addiciton. In this dataset, there are 60 samples and 48761 probes included. Each group
we found there to be 1673 differentially expressed probes in GSE54839, and the names of the top 3 gen