Sahand_Adibnia_NucleoTIDE_Hackathon_Project_Documentation

The formation of neuromelanin in the brain is a phenomenon that is known to be unique to humans [1]. Animals do not accumulate the pigment. It is observed to form only in certain catecholaminergic neurons, namely the dopaminergic neurons of the substantia nigra pars compacta (SNpc) [2]. Interestingly, neuromelanin is closely linked to Parkinson's disease, as it is these dopaminergic neuromelanin-containing neurons that are exclusively lost in Parkinson's [2].

Currently, it is assumed that neuromelanin formed in dopaminergic neurons is a product of dopamine oxidation, a process not known to be catalyzed by any particular enzymes [3-5]. However, studies in animal models have failed to produce neuromelanin by solely increasing dopamine levels and dopamine oxidation [1]. Additionally, neuromelanin does not form in all catecholaminergic neurons. In fact, the dopaminergic neurons of the ventral tegmental area (VTA) generally lack neuromelanin [2], suggesting that some other molecular mechanism may be involved in neuromelanin formation.

The goal of this project is to use bioinformatics and gene expression data to give some insight into the formation of neuromelanin. Learning more about these molecular pathways may even provide some insight into the mechanisms that cause the selective death of neuromelanin-containing neurons in Parkinson's disease. The primary method by which this is achieved is by analyzing differences in gene expression between the SNpc, a brain region containing a large population of melanized dopaminergic neurons, and the VTA, a brain region containing a large population of non-melanized dopaminergic neurons.

The Allen Human Brain Atlas (https://human.brain-map.org/microarray/search) provides microarray expression data of numerous different brain regions in 6 neurologically and neuropathologically normal individuals, including the SNpc and VTA. It also provides a function that does differential expression analysis of different brain regions, outputting p-values and fold change values for each probe. However, it only outputs the fold change for probes that are upregulated in the target structure with respect to the contrast structure. Therefore, this search had to be done twice: once with the SNpc as the target structure and the VTA as the contrast structure, and another time with the VTA as the target structure and the SNpc as the contrast structure. This data could be directly downloaded from the site. However, it could only be downloaded 2000 probes at a time. This was rather tedious, and knowing that many probes are not significantly differentially expressed anyway, only the first 10000 probes with the highest fold change were downloaded, giving a total of 20000 probes.

READING IN P-VALUES AND FOLD CHANGE OF TOP 10000 DIFERENTIALLY EXPRESSED PROBES

```
library(readr)
library(dplyr)
```

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

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

1

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

# SNpc target, VTA contrast
files_1 <- list.files(path = "/Users/rezaadibnia/Desktop/PD NM Project/mDBR - VTA vs SNpc/SNpc target, V
SNt_VTAc_probe_list <- lapply(files_1, read_csv)


##
## -- Column specification ------------------------------------------------------
## cols(
##   id = col_double(),
##   name = col_character(),
##   `gene-id` = col_double(),
##   `gene-symbol` = col_character(),
##   `gene-name` = col_character(),
##   `entrez-id` = col_double(),
##   chromosome = col_character(),
##   `start-position` = col_character(),
##   `end-position` = col_character(),
##   p = col_double(),
##   `fold-change` = col_double()
## )


##
## -- Column specification ------------------------------------------------------
## cols(
##   id = col_double(),
##   name = col_character(),
##   `gene-id` = col_double(),
##   `gene-symbol` = col_character(),
##   `gene-name` = col_character(),
##   `entrez-id` = col_double(),
##   chromosome = col_character(),
##   `start-position` = col_character(),
##   `end-position` = col_character(),
##   p = col_double(),
##   `fold-change` = col_double()
## )
##
##
## -- Column specification ------------------------------------------------------
## cols(
##   id = col_double(),
##   name = col_character(),
##   `gene-id` = col_double(),
##   `gene-symbol` = col_character(),
##   `gene-name` = col_character(),
##   `entrez-id` = col_double(),
##   chromosome = col_character(),
##   `start-position` = col_character(),
##   `end-position` = col_character(),
##   p = col_double(),
```

```
##    'fold-change' = col_double()
## )
##
##
## -- Column specification -------------------------------------------------
## cols(
##   id = col_double(),
##   name = col_character(),
##   'gene-id' = col_double(),
##   'gene-symbol' = col_character(),
##   'gene-name' = col_character(),
##   'entrez-id' = col_double(),
##   chromosome = col_character(),
##   'start-position' = col_character(),
##   'end-position' = col_character(),
##   p = col_double(),
##   'fold-change' = col_double()
## )
##
##
## -- Column specification -------------------------------------------------
## cols(
##   id = col_double(),
##   name = col_character(),
##   'gene-id' = col_double(),
##   'gene-symbol' = col_character(),
##   'gene-name' = col_character(),
##   'entrez-id' = col_double(),
##   chromosome = col_character(),
##   'start-position' = col_character(),
##   'end-position' = col_character(),
##   p = col_double(),
##   'fold-change' = col_double()
## )
```

```r
SNt_VTAc_probes <- data.frame(bind_rows(SNt_VTAc_probe_list))

# VTA target, SNpc contrast
files_2 <- list.files(path = "/Users/rezaadibnia/Desktop/PD NM Project/mDBR - VTA vs SNpc/VTA target, SN
VTAt_SNc_probe_list <- lapply(files_2, read_csv)
```

```
##
## -- Column specification -------------------------------------------------
## cols(
##   id = col_double(),
##   name = col_character(),
##   'gene-id' = col_double(),
##   'gene-symbol' = col_character(),
##   'gene-name' = col_character(),
##   'entrez-id' = col_double(),
##   chromosome = col_character(),
##   'start-position' = col_character(),
##   'end-position' = col_character(),
##   p = col_double(),
```

```
##   `fold-change` = col_double()
## )
##
##
## -- Column specification ----------------------------------------------------
## cols(
##   id = col_double(),
##   name = col_character(),
##   `gene-id` = col_double(),
##   `gene-symbol` = col_character(),
##   `gene-name` = col_character(),
##   `entrez-id` = col_double(),
##   chromosome = col_character(),
##   `start-position` = col_character(),
##   `end-position` = col_character(),
##   p = col_double(),
##   `fold-change` = col_double()
## )
##
##
## -- Column specification ----------------------------------------------------
## cols(
##   id = col_double(),
##   name = col_character(),
##   `gene-id` = col_double(),
##   `gene-symbol` = col_character(),
##   `gene-name` = col_character(),
##   `entrez-id` = col_double(),
##   chromosome = col_character(),
##   `start-position` = col_character(),
##   `end-position` = col_character(),
##   p = col_double(),
##   `fold-change` = col_double()
## )
##
##
## -- Column specification ----------------------------------------------------
## cols(
##   id = col_double(),
##   name = col_character(),
##   `gene-id` = col_double(),
##   `gene-symbol` = col_character(),
##   `gene-name` = col_character(),
##   `entrez-id` = col_double(),
##   chromosome = col_character(),
##   `start-position` = col_character(),
##   `end-position` = col_character(),
##   p = col_double(),
##   `fold-change` = col_double()
## )
##
##
## -- Column specification ----------------------------------------------------
## cols(
```

```
##   id = col_double(),
##   name = col_character(),
##   ‘gene-id‘ = col_double(),
##   ‘gene-symbol‘ = col_character(),
##   ‘gene-name‘ = col_character(),
##   ‘entrez-id‘ = col_double(),
##   chromosome = col_character(),
##   ‘start-position‘ = col_character(),
##   ‘end-position‘ = col_character(),
##   p = col_double(),
##   ‘fold-change‘ = col_double()
## )
```

```
VTAt_SNc_probes <- data.frame(bind_rows(VTAt_SNc_probe_list))
```

CORRECTING P-VALUES AND TRANSFORMING FOLD CHANGES The definition of the fold change
of B with respect to A is B/A. Therefore, the fold change of A with respect to B is A/B. In other words,
the fold change of gene expression in the SNpc with respect to the VTA is equal to 1/(the fold change of
gene expression in the VTA with respect to the SNpc). By taking the reciprocal of the fold change values
with VTA target and SNpc contrast, we get the equivalent fold change for SNpc target and VTA contrast,
allowing us to combine the differential expression analysis of all 20000 probes into one data frame.

```
# BH-correcting p-values (Benjamini-Hochberg Procedure)
SNt_VTAc_probes$p <- p.adjust(SNt_VTAc_probes$p, method="BH")
VTAt_SNc_probes$p <- p.adjust(VTAt_SNc_probes$p, method="BH")

# Converting VTA target, SNpc contrast fold change to SNpc target, VTA contrast fold change
VTAt_SNc_probes$fold.change <- 1/(VTAt_SNc_probes$fold.change)

# Combining the data frames containing p values and fold change
p_and_FC <- data.frame(bind_rows(SNt_VTAc_probes, VTAt_SNc_probes))
p_and_FC <- p_and_FC[,c(4,10:11)]
```

VISUALIZING DIFFERENTIALLY EXPRESSED PROBES - VOLCANO PLOT

```
library(ggplot2)

# Vertical lines for log2(fold change) thresholds, horizontal line for the p-value threshold
# Top 10% of upregulated probes and bottom 10% of downregulated probes are chosen
horz_min = quantile(log2(VTAt_SNc_probes$fold.change), 0.10)
horz_max = quantile(log2(SNt_VTAc_probes$fold.change), 0.90)

# Add a column to p_and_FC data frame indicating whether probe is upregulated, downregulated, or not si
p_and_FC$differentially_expressed <- "Not Significant"
# If log2Foldchange > horz_max and pvalue < 0.05, set as "UP"
p_and_FC$differentially_expressed[log2(p_and_FC$fold.change) > horz_max & p_and_FC$p < 0.05] <- "Upregul
# If log2Foldchange < horz_min and pvalue < 0.05, set as "DOWN"
p_and_FC$differentially_expressed[log2(p_and_FC$fold.change) < horz_min & p_and_FC$p < 0.05] <- "Downreg

# Construct volcano plot
volcano_plot <- ggplot(data=p_and_FC, aes(x=log2(fold.change), y=-log10(p), col=differentially_expresse
  geom_point() +
  theme_minimal() +
```

```
  geom_vline(xintercept=c(horz_min, horz_max), col="black") +
  geom_hline(yintercept=-log10(0.05), col="black")

# The significantly differentially expressed probes are  found in the upper-left and upper-right corner
volcano_plot
```
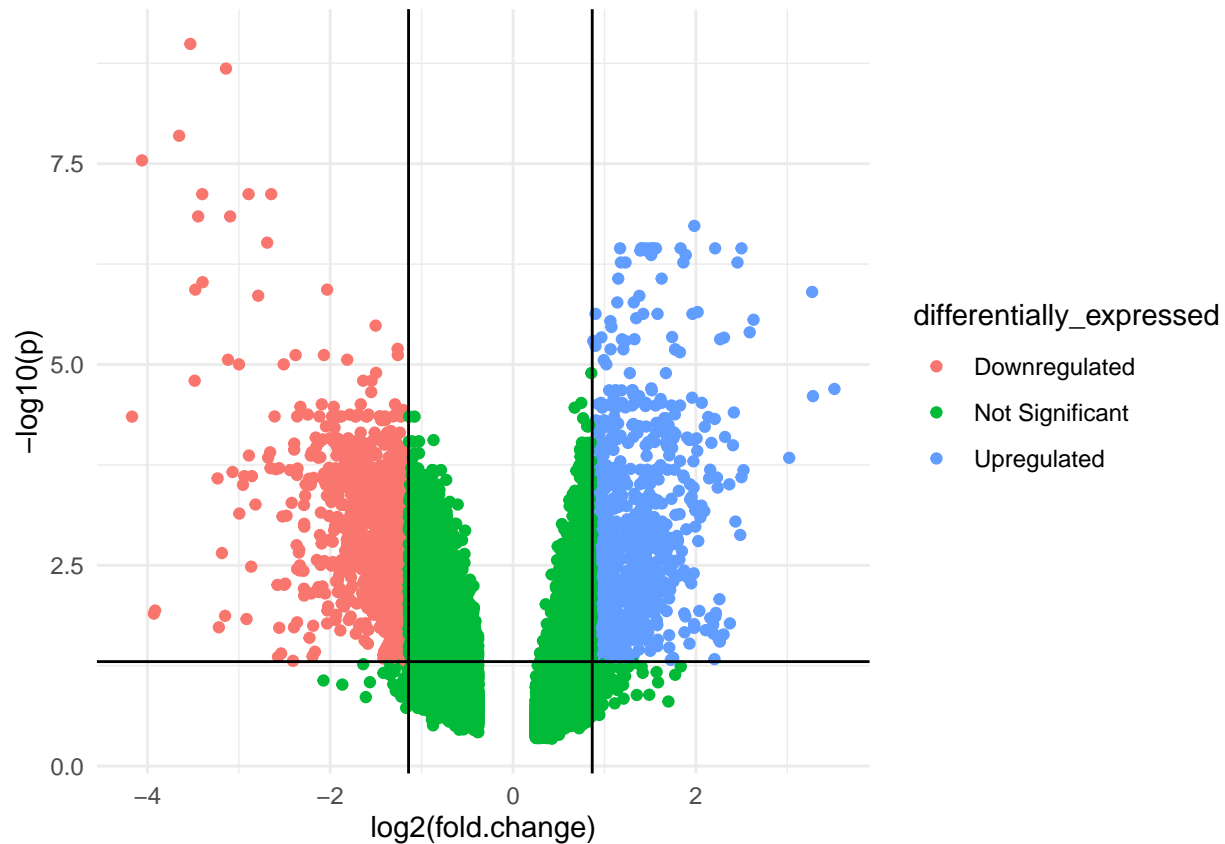

FILTERING FOR P < 0.05 and FC > 90th PERCENTILE (TOP 10%) p-values were previously corrected when constructing the volcano plot

```
library(dplyr)

# Filters for top 10% of genes with highest FC compared to all probes, not just significantly different

# SNpc target, VTA contrast
SNt_VTAc_probes <- SNt_VTAc_probes %>% filter(SNt_VTAc_probes$p < 0.05 & SNt_VTAc_probes$fold.change > 0

# VTA target, SNpc contrast
VTAt_SNc_probes <- VTAt_SNc_probes %>% filter(VTAt_SNc_probes$p < 0.05 & VTAt_SNc_probes$fold.change < 0
```

READING IN EXPRESSION DATA FOR TOP 10000 DIFFERENTIALLY EXPRESSED PROBES This data is actually not necessary to overall analysis, but is good to have for reference.

```
library(readr)
library(dplyr)
```

```r
read_csv_labeled_header <- function(file) {
  labels = c("Probe_IDs", "SN_1", "SN_2", "SN_3", "SN_4", "SN_5", "SN_6", "VTA_1", "VTA_2", "VTA_3", "V
  read_csv(file, col_names = labels)
}

# SNpc target, VTA contrast
files_expr_1 <- list.files(path = "/Users/rezaadibnia/Desktop/PD NM Project/mDBR - VTA vs SNpc/SNpc targ
SNt_VTAc_expr_list <- lapply(files_expr_1, read_csv_labeled_header)
```

```
##
## -- Column specification -----------------------------------------------
## cols(
##   Probe_IDs = col_double(),
##   SN_1 = col_double(),
##   SN_2 = col_double(),
##   SN_3 = col_double(),
##   SN_4 = col_double(),
##   SN_5 = col_double(),
##   SN_6 = col_double(),
##   VTA_1 = col_double(),
##   VTA_2 = col_double(),
##   VTA_3 = col_double(),
##   VTA_4 = col_double(),
##   VTA_5 = col_double(),
##   VTA_6 = col_double()
## )
##
##
## -- Column specification -----------------------------------------------
## cols(
##   Probe_IDs = col_double(),
##   SN_1 = col_double(),
##   SN_2 = col_double(),
##   SN_3 = col_double(),
##   SN_4 = col_double(),
##   SN_5 = col_double(),
##   SN_6 = col_double(),
##   VTA_1 = col_double(),
##   VTA_2 = col_double(),
##   VTA_3 = col_double(),
##   VTA_4 = col_double(),
##   VTA_5 = col_double(),
##   VTA_6 = col_double()
## )
##
##
## -- Column specification -----------------------------------------------
## cols(
##   Probe_IDs = col_double(),
##   SN_1 = col_double(),
##   SN_2 = col_double(),
##   SN_3 = col_double(),
##   SN_4 = col_double(),
```

```
##    SN_5 = col_double(),
##    SN_6 = col_double(),
##    VTA_1 = col_double(),
##    VTA_2 = col_double(),
##    VTA_3 = col_double(),
##    VTA_4 = col_double(),
##    VTA_5 = col_double(),
##    VTA_6 = col_double()
## )
##
##
## -- Column specification --------------------------------------------------
## cols(
##    Probe_IDs = col_double(),
##    SN_1 = col_double(),
##    SN_2 = col_double(),
##    SN_3 = col_double(),
##    SN_4 = col_double(),
##    SN_5 = col_double(),
##    SN_6 = col_double(),
##    VTA_1 = col_double(),
##    VTA_2 = col_double(),
##    VTA_3 = col_double(),
##    VTA_4 = col_double(),
##    VTA_5 = col_double(),
##    VTA_6 = col_double()
## )
##
##
## -- Column specification --------------------------------------------------
## cols(
##    Probe_IDs = col_double(),
##    SN_1 = col_double(),
##    SN_2 = col_double(),
##    SN_3 = col_double(),
##    SN_4 = col_double(),
##    SN_5 = col_double(),
##    SN_6 = col_double(),
##    VTA_1 = col_double(),
##    VTA_2 = col_double(),
##    VTA_3 = col_double(),
##    VTA_4 = col_double(),
##    VTA_5 = col_double(),
##    VTA_6 = col_double()
## )
```

```r
SNt_VTAc_expr <- data.frame(bind_rows(SNt_VTAc_expr_list))

# VTA target, SNpc contrast
files_expr_2 <- list.files(path = "/Users/rezaadibnia/Desktop/PD NM Project/mDBR - VTA vs SNpc/VTA targ
VTAt_SNc_expr_list <- lapply(files_expr_2, read_csv_labeled_header)
```

```
##
## -- Column specification --------------------------------------------------
```

```
## cols(
##   Probe_IDs = col_double(),
##   SN_1 = col_double(),
##   SN_2 = col_double(),
##   SN_3 = col_double(),
##   SN_4 = col_double(),
##   SN_5 = col_double(),
##   SN_6 = col_double(),
##   VTA_1 = col_double(),
##   VTA_2 = col_double(),
##   VTA_3 = col_double(),
##   VTA_4 = col_double(),
##   VTA_5 = col_double(),
##   VTA_6 = col_double()
## )
##
##
## -- Column specification -----------------------------------------------------
## cols(
##   Probe_IDs = col_double(),
##   SN_1 = col_double(),
##   SN_2 = col_double(),
##   SN_3 = col_double(),
##   SN_4 = col_double(),
##   SN_5 = col_double(),
##   SN_6 = col_double(),
##   VTA_1 = col_double(),
##   VTA_2 = col_double(),
##   VTA_3 = col_double(),
##   VTA_4 = col_double(),
##   VTA_5 = col_double(),
##   VTA_6 = col_double()
## )
##
##
## -- Column specification -----------------------------------------------------
## cols(
##   Probe_IDs = col_double(),
##   SN_1 = col_double(),
##   SN_2 = col_double(),
##   SN_3 = col_double(),
##   SN_4 = col_double(),
##   SN_5 = col_double(),
##   SN_6 = col_double(),
##   VTA_1 = col_double(),
##   VTA_2 = col_double(),
##   VTA_3 = col_double(),
##   VTA_4 = col_double(),
##   VTA_5 = col_double(),
##   VTA_6 = col_double()
## )
##
##
## -- Column specification -----------------------------------------------------
```

```
## cols(
##   Probe_IDs = col_double(),
##   SN_1 = col_double(),
##   SN_2 = col_double(),
##   SN_3 = col_double(),
##   SN_4 = col_double(),
##   SN_5 = col_double(),
##   SN_6 = col_double(),
##   VTA_1 = col_double(),
##   VTA_2 = col_double(),
##   VTA_3 = col_double(),
##   VTA_4 = col_double(),
##   VTA_5 = col_double(),
##   VTA_6 = col_double()
## )
##
##
## -- Column specification -----------------------------------------------------
## cols(
##   Probe_IDs = col_double(),
##   SN_1 = col_double(),
##   SN_2 = col_double(),
##   SN_3 = col_double(),
##   SN_4 = col_double(),
##   SN_5 = col_double(),
##   SN_6 = col_double(),
##   VTA_1 = col_double(),
##   VTA_2 = col_double(),
##   VTA_3 = col_double(),
##   VTA_4 = col_double(),
##   VTA_5 = col_double(),
##   VTA_6 = col_double()
## )
```

```r
VTAt_SNc_expr <- data.frame(bind_rows(VTAt_SNc_expr_list))
```

JOINING EXPRESSION AND PROBE DATA FRAMES INTO ONE Note that probe data frames contain p-values and fold changes.

```r
library(dplyr)

SNt_VTAc_expr <- SNt_VTAc_expr %>% inner_join(SNt_VTAc_probes, by = c("Probe_IDs" = "id"))
VTAt_SNc_expr <- VTAt_SNc_expr %>% inner_join(VTAt_SNc_probes, by = c("Probe_IDs" = "id"))

expr <- data.frame(bind_rows(SNt_VTAc_expr, VTAt_SNc_expr))
```

After filtering for probes that had p-values < 0.05 and significant fold changes, the probes needed to be aligned to individual genes. To do this, the expression data for all probes in all brain regions in all 6 donors had to be downloaded from Allen Human Brain Atlas (AHBA). The probes lacking Entrez IDs were first filtered out. Conveniently, AHBA provides a presence/absence flag that determines whether a certain probe's expression is well above the background levels or not. This flag was used to filter out any probes that are not expressed above the background in less than 1% of all samples. There may be multiple probes that align to the same gene, but only one probe per gene can be selected. Therefore, a systematic method was used

to align probes to genes using the collapseRows function from the WGCNA package, similar to methods previously used in the literature [6]:

- If there is only one probe for a gene, that one is chosen
- If there are two probes for one gene, the one with the max variance across all samples is chosen (method = maxRowVariance)
- If there more than two probes for one gene, the probe with the highest connectivity (summed adjacency) is chosen (connectivityBasedCollapsing = TRUE)

ALIGNING PROBES TO GENES

```
# Map probes to genes
# First download 'Complete normalized microarray datasets' for all six donors from http://human.brain-m
# Store the downloaded files in a AHBA download directory
library(WGCNA)
```

```
## Loading required package: dynamicTreeCut
```

```
## Loading required package: fastcluster
```

```
##
## Attaching package: 'fastcluster'
```

```
## The following object is masked from 'package:stats':
##
##     hclust
```

```
##
```

```
##
## Attaching package: 'WGCNA'
```

```
## The following object is masked from 'package:stats':
##
##     cor
```

```
########## Load data ##########

# Set donor names
donorNames <- c("donor9861", "donor10021", "donor12876", "donor14380", "donor15496", "donor15697")
names(donorNames) <- donorNames

# Location of AHBA directories
ahba_download <- "/Users/rezaadibnia/Desktop/AHBA Full Data/Microarray" # Where downloaded files are st

# Read in probe info (files are the same for each donor)
probeInfo <- read.csv(paste0(ahba_download, "/normalized_microarray_donor9861/Probes.csv")) # Same for

# Read expression data for each donor (probes x samples)
brainExpr <- lapply(donorNames, function(d){
  file1 <- paste0(ahba_download, "/normalized_microarray_", d, "/MicroarrayExpression.csv")
```

```
  e <- read.csv(file1, header = FALSE)
  rownames(e) <- e[,1]
  e <- e[,-1]
  file2 <- paste0(ahba_download, "/normalized_microarray_", d, "/SampleAnnot.csv")
  sample_annotation <- read.csv(file2)
  colnames(e) <- sample_annotation$structure_id
  e
})


########## Filter probes based on concatenated expression data from all donors ##########

# Probes with missing Entrez IDs
probes_missing_entrezID <- is.na(probeInfo$entrez_id) # same in all donors
print(paste(sum(probes_missing_entrezID), "probes with missing entrez IDs"))


## [1] "10521 probes with missing entrez IDs"

# Probes with expression well above background in at least 1% of samples in all donors
pa_call <- lapply(donorNames, function(d){
  file <- paste0(ahba_download, "/normalized_microarray_", d, "/PACall.csv")
  pa <- read.csv(file, header = FALSE)
  rownames(pa) <- pa[,1]
  pa[,-1]
})
pa_concat <- Reduce(cbind, pa_call)
sum <- rowSums(pa_concat)
low_presence <- sum < 0.01*ncol(pa_concat)
print(paste(sum(low_presence), "probes present in <1% of samples"))


## [1] "6068 probes present in <1% of samples"

# Concatenate data across all donors and filter probes by ID
expr_concat <- Reduce(cbind, brainExpr)
expr_concat <- expr_concat[!(probes_missing_entrezID | low_presence),]
probeInfo <- probeInfo[!(probes_missing_entrezID | low_presence), ]

probe2gene <- collapseRows(expr_concat,
                           rowGroup = probeInfo$entrez_id,
                           rowID = probeInfo$probe_id,
                           method = "maxRowVariance",
                           connectivityBasedCollapsing = TRUE
                           )
selected_probes <- probe2gene$selectedRow
entrez_id <- probeInfo$entrez_id[selected_probes]
entrez_order <- order(entrez_id)

probeInfo <- probeInfo[selected_probes, ]
probeInfo <- probeInfo[entrez_order,]
```

ALIGNING SIGNIFICANTLY DIFFERENTIALLY EXPRESSED PROBES TO GENES

```r
library(dplyr)

# Combine data frames by probe ID to filter out all probes not corresponding to a gene
final_expr <- expr %>% inner_join(probeInfo, by = c("Probe_IDs" = "probe_id"))

# Remove unnecessary and repeated columns
final_expr <- final_expr[,-c(14, 18:21, 24:29)]
final_expr$fold.change <- sort(final_expr$fold.change, decreasing = TRUE)

final_expr[,c(15, 17:18)]
```

```
##      gene.symbol          p fold.change
## 1          AGTR1 2.017857e-05 11.44100000
## 2          FEZF1 1.250000e-06  9.65900000
## 3          LMOD3 1.452381e-04  8.11800000
## 4           NPNT 2.771875e-06  6.19600000
## 5         CELSR1 3.971429e-06  6.01800000
## 6          RSPO2 2.067010e-04  5.74200000
## 7           PYGL 3.581818e-07  5.65300000
## 8         ATP2A3 9.025070e-04  5.40100000
## 9           OC90 3.955056e-05  5.33600000
## 10         SRPX2 1.012987e-04  5.30000000
## 11        SLC6A3 8.355197e-03  4.78400000
## 12         F2RL2 2.553991e-04  4.69400000
## 13       SLC18A2 1.238258e-02  4.65600000
## 14         CBLN1 3.581818e-07  4.62900000
## 15         ENPP1 4.760417e-05  4.61700000
## 16          CCL3 4.660821e-02  4.60700000
## 17          FOXA1 2.041451e-04  4.45900000
## 18         NTSR1 2.030387e-02  4.30700000
## 19       SLC35D3 6.686930e-04  4.26400000
## 20          TPBG 3.000000e-05  4.18400000
## 21       ALDH1A1 5.714286e-04  4.17400000
## 22       FAM19A3 8.000000e-04  4.14100000
## 23         HTR1F 5.796178e-04  4.06500000
## 24         SSTR1 2.223077e-06  4.05500000
## 25         RDH12 1.191358e-04  4.04300000
## 26          SDC1 6.574924e-04  3.98500000
## 27       SLC10A4 1.899718e-02  3.94700000
## 28          SMPX 3.409639e-04  3.93300000
## 29         GFRA1 1.603448e-04  3.92700000
## 30         MATN3 2.343333e-06  3.89900000
## 31          VAV3 4.333333e-07  3.70500000
## 32         RGAG1 2.157395e-02  3.67100000
## 33          ZAR1 1.256367e-02  3.66400000
## 34     LOC441052 2.567442e-04  3.65600000
## 35        CYP2S1 3.234568e-05  3.54900000
## 36          EBF3 1.588235e-04  3.54800000
## 37         KCNJ6 7.291667e-04  3.53700000
## 38         PROK2 2.684887e-03  3.53100000
## 39         RGS16 7.448071e-04  3.42900000
## 40         APOC1 7.365079e-05  3.42700000
## 41          RGS4 6.468085e-06  3.42000000
```

```
## 42          COL12A1 6.520833e-06  3.41900000
## 43          LRRC55 4.188889e-05  3.41000000
## 44          FAM70A 8.308824e-05  3.37400000
## 45          IL1RAPL2 2.784810e-03  3.36400000
## 46          B3GALT1 4.555556e-06  3.34100000
## 47             DRD2 1.095541e-04  3.34100000
## 48            PTPN3 1.725838e-03  3.33200000
## 49            KLHL1 4.703704e-04  3.32400000
## 50              FST 3.763889e-03  3.31500000
## 51             SV2C 1.351515e-04  3.26900000
## 52            APOC2 3.085106e-03  3.26500000
## 53           CADPS2 5.669811e-05  3.25100000
## 54              LPO 1.177489e-02  3.20600000
## 55             HPGD 9.809783e-04  3.20100000
## 56           AKAP14 2.144112e-03  3.20000000
## 57          GALNTL6 1.575314e-03  3.19900000
## 58           FAM71F1 4.663158e-05  3.15000000
## 59           RESP18 5.129151e-03  3.13600000
## 60            IZUMO2 1.714851e-03  3.12100000
## 61              NEB 1.797732e-03  3.09800000
## 62           GPR161 9.615385e-04  3.09100000
## 63            ITPR1 8.523810e-07  3.08500000
## 64     LOC100130344 1.161369e-03  3.06200000
## 65           CYP27C1 3.804945e-03  3.05600000
## 66             GBE1 6.939227e-03  3.02300000
## 67         TMEM200C 2.041451e-04  3.01600000
## 68              EN1 2.668731e-02  2.99800000
## 69            CNIH3 2.343333e-06  2.99100000
## 70              PGR 3.664773e-03  2.98800000
## 71          FLJ30901 1.161369e-03  2.97500000
## 72            GSG1L 5.223368e-04  2.97100000
## 73           OSBPL3 1.533755e-03  2.96200000
## 74             RERG 3.581818e-07  2.95600000
## 75           ST8SIA6 1.128713e-03  2.93400000
## 76              RET 1.192740e-02  2.93100000
## 77             KLF5 1.124378e-03  2.92500000
## 78            BEND4 5.801887e-03  2.90900000
## 79              TUB 4.710145e-04  2.89000000
## 80             LMX1B 1.359551e-03  2.86800000
## 81           KLHL13 2.096774e-05  2.86400000
## 82            SCUBE1 5.653207e-03  2.86100000
## 83             SIAH3 1.981818e-05  2.85300000
## 84            CAB39L 4.333333e-07  2.85100000
## 85             GPR26 5.894118e-03  2.84700000
## 86           C16orf93 3.581818e-07  2.83600000
## 87             FNDC1 3.674856e-02  2.83000000
## 88            GPRC5A 5.392039e-03  2.80800000
## 89           SLC2A13 9.225352e-05  2.78700000
## 90           PRKAR2B 3.815385e-07  2.75000000
## 91            CHRNB3 1.739766e-02  2.74500000
## 92             CLDN1 7.314410e-03  2.74200000
## 93            EFCAB5 3.017467e-04  2.73800000
## 94           SLITRK4 3.000000e-05  2.73700000
## 95            ST3GAL1 3.000000e-05  2.69600000
```

```
## 96            ZNF503 2.343333e-06  2.68400000
## 97             VGLL3 1.474026e-03  2.67800000
## 98              SNCA 1.731141e-03  2.67300000
## 99            TRIM58 6.387316e-03  2.65200000
## 100           GABRA4 2.784810e-03  2.63400000
## 101             LMO3 3.815385e-07  2.62600000
## 102            NCAPG 1.020725e-03  2.62400000
## 103             ZW10 2.612500e-02  2.61400000
## 104             OXTR 1.880819e-03  2.61200000
## 105            PCDH8 8.926554e-04  2.57400000
## 106            GPR98 1.927273e-03  2.55000000
## 107            PITX1 2.647607e-02  2.54400000
## 108             OPN3 2.814465e-03  2.52500000
## 109             EYA4 3.803596e-03  2.52000000
## 110             ANK1 3.012987e-05  2.51900000
## 111             DPYS 4.443022e-03  2.51700000
## 112    LOC100131532 5.012376e-03  2.51300000
## 113         C17orf82 4.846154e-06  2.51200000
## 114            SLC2A4 4.032476e-03  2.51200000
## 115            CACNB3 4.076923e-04  2.50900000
## 116    LOC100144602 9.419890e-04  2.50000000
## 117          C9orf135 4.180929e-02  2.49700000
## 118             RGNEF 1.899718e-02  2.49200000
## 119             SEZ6L 8.696884e-04  2.48500000
## 120           OSBPL10 1.772031e-03  2.48300000
## 121             INSM2 2.671614e-02  2.47800000
## 122             SLIT1 2.382182e-02  2.47500000
## 123            IGFBP4 2.209688e-02  2.47400000
## 124             ZWINT 2.026279e-02  2.47300000
## 125               SNN 5.669811e-05  2.47200000
## 126             KCND3 3.240964e-05  2.47000000
## 127            CHRNA6 1.359551e-03  2.46900000
## 128             SYT17 2.096774e-05  2.46500000
## 129            EPS8L1 2.164879e-02  2.46400000
## 130             PRMT8 2.261670e-02  2.45400000
## 131               LXN 2.074176e-02  2.44200000
## 132             TRPC6 1.420500e-02  2.43500000
## 133           PLEKHA5 1.279630e-05  2.42500000
## 134             RIMS1 3.767313e-03  2.41500000
## 135            FKBP1B 4.262136e-02  2.41400000
## 136              BNC2 2.085048e-02  2.41200000
## 137             KANK4 1.579167e-03  2.40400000
## 138          ARHGAP11A 6.939227e-03  2.40300000
## 139           CLPTM1L 2.477932e-02  2.39500000
## 140               PID1 9.463087e-05  2.39300000
## 141               DRP2 1.684211e-02  2.39100000
## 142             OLFM3 3.098291e-04  2.37600000
## 143              OR7G1 2.870813e-02  2.36900000
## 144               PKIB 5.796178e-04  2.36600000
## 145                GLA 4.393723e-02  2.36100000
## 146               AMPH 2.163624e-02  2.36100000
## 147               NTN1 1.511777e-03  2.34800000
## 148             CORO2A 5.368421e-07  2.34700000
## 149               VAT1 3.423529e-05  2.34300000
```

```
## 150      HIST1H2AM 1.000000e-03  2.33700000
## 151         CLTCL1 2.700893e-04  2.33000000
## 152       TNFRSF19 1.349315e-03  2.32700000
## 153         UNC13C 1.313084e-03  2.32700000
## 154     ST6GALNAC5 3.327769e-02  2.31500000
## 155          CRYGD 5.245098e-03  2.30900000
## 156           INSM1 4.464875e-02  2.30000000
## 157           UGGT1 1.440789e-03  2.28900000
## 158         POU3F2 4.853659e-06  2.28300000
## 159           NEFH 1.927273e-03  2.28100000
## 160           DOK6 2.096774e-05  2.27600000
## 161          ROBO2 3.056277e-04  2.27300000
## 162        B4GALT6 1.473913e-03  2.27200000
## 163         CLSTN2 6.856187e-03  2.27200000
## 164        FAM135B 2.684887e-03  2.25900000
## 165        CCDC85A 5.796178e-04  2.25100000
## 166            MCC 3.737255e-04  2.24700000
## 167          CPLX2 1.349315e-03  2.23900000
## 168          PRKCD 1.265823e-02  2.23700000
## 169       C18orf34 9.126679e-03  2.23400000
## 170            NOV 1.739766e-02  2.22900000
## 171           PBX1 5.420000e-05  2.21900000
## 172         GRAMD4 4.076923e-04  2.21500000
## 173          PADI1 1.407767e-02  2.20900000
## 174         GPRIN3 2.095563e-02  2.20200000
## 175            KMO 3.177711e-03  2.20000000
## 176          PDZD7 3.295964e-02  2.19300000
## 177          CSDC2 1.938406e-03  2.19200000
## 178           ENO3 3.012987e-05  2.19100000
## 179        ATP6V1B2 1.910946e-03  2.18800000
## 180        C22orf25 3.883197e-02  2.18700000
## 181         KIF26B 2.087719e-05  2.17500000
## 182          PLA2G7 3.327769e-02  2.17100000
## 183         ELAVL2 1.362416e-03  2.17100000
## 184          CTXN3 5.704142e-03  2.17000000
## 185       LINC00261 4.499295e-02  2.16600000
## 186           CUX2 2.348039e-04  2.16100000
## 187         COMTD1 4.663278e-03  2.14900000
## 188            MMD 1.103275e-03  2.14900000
## 189         LRRC3B 2.597015e-05  2.14700000
## 190          NRXN3 2.718978e-02  2.13800000
## 191        ANKRD29 5.859649e-05  2.13400000
## 192         C3orf80 2.825746e-03  2.13000000
## 193      LOC730091 2.628099e-03  2.12900000
## 194          KIF3C 3.151515e-03  2.12400000
## 195           SOX6 3.284024e-03  2.12200000
## 196          FABP3 1.927273e-03  2.11900000
## 197         SEMA3E 3.594080e-02  2.11700000
## 198           TAL2 8.120301e-05  2.11100000
## 199        ZNF385D 1.190821e-03  2.11100000
## 200        RNF144B 3.228700e-03  2.10900000
## 201          DAPL1 2.397172e-02  2.10800000
## 202          STYK1 3.036131e-02  2.10600000
## 203           FUT9 2.848297e-03  2.10400000
```

```
## 204         PCSK1 8.815920e-03  2.10400000
## 205        PIP5K1B 1.684211e-02  2.10100000
## 206         RELL2 2.161247e-02  2.10000000
## 207        SLC44A5 1.069182e-02  2.09900000
## 208          EPN3 3.548559e-02  2.09600000
## 209       SH3BGRL2 6.468085e-06  2.09600000
## 210          SYT1 5.796178e-04  2.09100000
## 211          GPX3 1.731141e-03  2.09000000
## 212         TRPM2 4.556962e-02  2.08700000
## 213         RNF43 7.660695e-03  2.08000000
## 214         MYOM1 2.292079e-04  2.07900000
## 215         CNTN4 3.258929e-03  2.07800000
## 216       ANKRD34A 2.223693e-02  2.07300000
## 217          VAV2 5.051903e-04  2.06600000
## 218           BSN 7.663158e-03  2.06300000
## 219         GLRA3 9.026290e-03  2.05600000
## 220         TPGS2 5.555556e-04  2.05300000
## 221         FXYD7 2.684887e-03  2.05000000
## 222      LOC158696 1.984626e-02  2.04600000
## 223         RCAN2 1.927273e-03  2.04100000
## 224         TTC39A 5.281174e-03  2.04000000
## 225         WIPF3 3.142530e-02  2.04000000
## 226        DYNC1I1 3.375551e-02  2.03900000
## 227         CHEK1 1.984626e-02  2.03600000
## 228        CREB3L1 1.857651e-02  2.03500000
## 229       TMEM159 2.700893e-04  2.03300000
## 230        DENND2A 1.020725e-03  2.03300000
## 231          PPM1J 1.851852e-02  2.03000000
## 232         CERS6 1.731141e-03  2.02800000
## 233        NEDD4L 3.409639e-04  2.02500000
## 234       TMEM241 8.189300e-03  2.02500000
## 235       FAM126A 8.240081e-03  2.02300000
## 236        SLC36A1 3.161180e-02  2.02000000
## 237       C15orf59 5.859649e-05  2.01900000
## 238          HMMR 1.376221e-02  2.01600000
## 239          CALN1 2.700893e-04  2.01500000
## 240         MAN1C1 5.669811e-05  2.01500000
## 241          RGS6 5.012376e-03  2.01100000
## 242         CRYBB1 4.359712e-02  2.01000000
## 243          EPB49 4.735632e-02  2.00800000
## 244         STK17A 1.684211e-02  2.00300000
## 245         AHCYL2 1.268908e-02  2.00100000
## 246        FAM207A 2.367056e-02  1.99800000
## 247         MCM10 8.922925e-03  1.99600000
## 248         NYAP2 8.588589e-03  1.99600000
## 249        COL11A1 4.959184e-05  1.99500000
## 250        PGM2L1 7.234973e-03  1.99500000
## 251          OR2W3 2.512594e-02  1.99300000
## 252          SOGA2 8.820000e-06  1.99100000
## 253           GBA2 1.131742e-02  1.99100000
## 254       TMEM182 1.933896e-02  1.99000000
## 255           CMAS 1.324885e-03  1.98500000
## 256       MGC50722 2.074176e-02  1.98300000
## 257      LINC00152 9.025070e-04  1.98000000
```

17

```
## 258       C1orf135 2.089041e-02  1.97900000
## 259          RSPO4 1.013021e-03  1.97900000
## 260          EPHA5 4.781421e-02  1.97600000
## 261          SPCS3 4.731183e-04  1.97100000
## 262          TEX15 1.871013e-02  1.96900000
## 263          VWA5A 3.760446e-03  1.95900000
## 264           RTN2 1.734019e-02  1.95800000
## 265         SLC7A8 2.268041e-03  1.95700000
## 266           QPRT 5.859649e-05  1.95600000
## 267         CAMK1G 8.587174e-03  1.95200000
## 268          CADPS 4.076923e-04  1.95200000
## 269          TTPAL 6.897590e-04  1.94800000
## 270        RARRES1 3.556976e-02  1.94700000
## 271          KCNQ3 3.411079e-03  1.93900000
## 272        C10orf10 3.497831e-02 1.93900000
## 273         FAM54A 2.373868e-02  1.93500000
## 274           NPM2 4.894366e-04  1.93300000
## 275       TMEM132B 4.501880e-02  1.93300000
## 276          SEPT5 9.382716e-03  1.93000000
## 277          CDK14 2.292079e-04  1.92900000
## 278        C2orf55 2.223693e-02  1.92800000
## 279          RAB3A 5.842167e-03  1.92800000
## 280          FABP7 5.542312e-03  1.92600000
## 281          RAB3C 7.756813e-03  1.92400000
## 282        JAKMIP1 5.957983e-05  1.92000000
## 283         SPATS2 1.280335e-02  1.91800000
## 284           GALT 4.006039e-02  1.91600000
## 285        RUNDC3A 4.052974e-02  1.91600000
## 286        FAM110C 3.840991e-02  1.91500000
## 287         LRRC8A 2.348039e-04  1.91100000
## 288          SNX16 1.691837e-03  1.90600000
## 289           DPCD 2.897694e-02  1.90500000
## 290       OR7E156P 3.511039e-02  1.90500000
## 291           GARS 1.684211e-02  1.89900000
## 292          PPTC7 4.729242e-04  1.89900000
## 293        SLITRK5 1.508991e-02  1.89800000
## 294        CACNA2D2 8.922925e-03 1.89300000
## 295        KIAA1377 3.702665e-03 1.89200000
## 296        CEACAM4 7.352941e-03  1.89100000
## 297          MAP3K5 1.691837e-03  1.89000000
## 298           WNT3 2.402707e-03  1.89000000
## 299         TWIST1 8.954057e-03  1.88900000
## 300         LPGAT1 3.522788e-02  1.88600000
## 301        ANKRD34C 9.126679e-03 1.88400000
## 302           PCLO 3.308715e-03  1.88300000
## 303          ALPK3 3.589608e-02  1.88100000
## 304         DLGAP3 1.476826e-02  1.87800000
## 305         CYB5R1 5.202454e-03  1.87800000
## 306            HK1 6.180160e-03  1.87600000
## 307          LIN7A 1.869159e-03  1.87600000
## 308        ALDH4A1 6.071429e-03  1.87500000
## 309 LOC100129781 8.954057e-03    1.87400000
## 310          ANXA6 3.512337e-03  1.87300000
## 311          GRIA3 2.343333e-06  1.87000000
```

```
## 312           C1S 5.374150e-04  1.86600000
## 313          TAB3 5.043263e-03  1.86500000
## 314         APOL2 2.411802e-02  1.86500000
## 315       SMARCA1 3.388248e-02  1.85900000
## 316          SHOC2 8.029046e-03  1.85400000
## 317       SH3BGRL3 5.796178e-04  1.85200000
## 318        CCDC69 2.897694e-02  1.85100000
## 319         PDE6H 4.761905e-02  1.84900000
## 320         MSL3P1 1.268908e-02  1.84900000
## 321       ATP6V1G1 1.160494e-03  1.84700000
## 322          PSD3 1.511777e-03  1.84700000
## 323         MPZL1 4.106464e-04  1.84700000
## 324       RAP1GDS1 3.522788e-02  1.84500000
## 325        AHNAK2 3.874936e-02  1.84400000
## 326         FSTL4 8.827038e-03  1.84300000
## 327          CLTB 8.173375e-03  1.84300000
## 328          MSI2 3.804945e-03  1.84100000
## 329       ANKRD50 8.954057e-03  1.83900000
## 330          CDC6 4.057683e-02  1.83900000
## 331          GFER 1.230769e-02  1.83600000
## 332     LOC284801 1.047359e-02  1.83500000
## 333        GPR152 2.711761e-02  1.83500000
## 334 LOC100292489 3.701195e-04  1.83300000
## 335         SNX25 2.700893e-04  1.83200000
## 336        DPH3P1 2.691131e-02  1.83100000
## 337        DLGAP4 1.603448e-04  1.83100000
## 338       FLJ43860 1.539056e-02  1.82800000
## 339        PCNXL2 9.672131e-04  1.82600000
## 340       C11orf74 3.702665e-03  1.82500000
## 341        DIRAS3 1.420000e-08  0.45269353
## 342          PENK 1.592593e-05  0.45228403
## 343           CRH 1.169231e-06  0.45228403
## 344          SCGN 1.433333e-07  0.45105999
## 345       NEUROD2 2.226277e-03  0.44984256
## 346         CALB2 2.070000e-09  0.44943820
## 347        BAIAP3 1.433333e-07  0.44903458
## 348         CHODL 2.182266e-04  0.44822949
## 349           SST 3.147651e-04  0.44822949
## 350         PTPRT 7.571429e-08  0.44762757
## 351           IHH 3.290000e-03  0.44682752
## 352           KIT 1.448276e-04  0.44662796
## 353          NXPH2 1.952941e-04  0.44642857
## 354          MYH7 7.571429e-08  0.44622936
## 355         NPTX1 2.010526e-04  0.44603033
## 356           VIP 5.547945e-03  0.44503783
## 357        CARTPT 4.370925e-02  0.44503783
## 358        DMRTA2 1.898254e-02  0.44424700
## 359        SORCS3 9.956522e-06  0.44365572
## 360        ADRA1D 5.527523e-03  0.44365572
## 361       ARHGAP36 5.279188e-04  0.44326241
## 362         SLC6A4 1.870200e-02  0.44326241
## 363         HTR1E 9.642857e-05  0.44306602
## 364         EGFL6 7.631579e-06  0.44189129
## 365         F13A1 1.620295e-02  0.44169611
```

```
## 366        RPRM 2.030848e-03  0.44169611
## 367         NGB 3.363636e-05  0.44169611
## 368        CALB1 7.530864e-03  0.44150110
## 369        UNC5D 4.297521e-04  0.44130627
## 370     SLITRK6 3.147651e-04  0.44130627
## 371        KCNJ3 4.205128e-05  0.44111160
## 372         TPH2 4.292683e-02  0.44091711
## 373     RASGRP1 6.803170e-03  0.44091711
## 374     HIST1H1A 1.429577e-04  0.44033465
## 375        NDST3 6.978261e-04  0.44033465
## 376     KIAA1239 1.692525e-03  0.44014085
## 377        SYTL5 3.129032e-05  0.43975374
## 378       PLXNC1 8.561644e-05  0.43936731
## 379        SOX14 3.147651e-04  0.43917435
## 380      ADCYAP1 1.677445e-02  0.43917435
## 381         DRD5 1.169231e-06  0.43859649
## 382      ONECUT1 7.693920e-04  0.43840421
## 383        HTR2C 5.153061e-04  0.43840421
## 384         WNT4 4.431373e-05  0.43802015
## 385        KCNQ5 2.882788e-03  0.43725404
## 386       SORCS1 4.854881e-04  0.43687200
## 387        GULP1 2.362791e-04  0.43630017
## 388         NOS1 4.006116e-04  0.43572985
## 389        LPPR5 6.096774e-05  0.43572985
## 390        CALCR 5.004010e-03  0.43572985
## 391        NPTX2 4.085366e-04  0.43516101
## 392          FEV 6.803170e-03  0.43459365
## 393       NECAB2 3.211382e-03  0.43421624
## 394        PDGFD 5.582524e-04  0.43421624
## 395         COCH 5.593220e-04  0.43383948
## 396        C1QL3 1.027689e-02  0.43365134
## 397        GATA3 2.751606e-03  0.43365134
## 398       CHRNA3 2.081737e-03  0.43365134
## 399        CBLN4 3.479532e-03  0.43346337
## 400      TMEM215 2.038462e-03  0.43346337
## 401        DMBT1 8.778626e-04  0.43308792
## 402      COL21A1 8.291667e-05  0.43308792
## 403        DACT1 9.074468e-05  0.43308792
## 404       SPINK2 1.517592e-02  0.43252595
## 405       PCP4L1 8.084677e-04  0.43233895
## 406        S1PR5 4.205128e-05  0.43233895
## 407       KCNAB1 1.237705e-04  0.43177893
## 408     C10orf82 8.618677e-04  0.43159258
## 409      SLC30A3 2.255730e-03  0.43159258
## 410        LAMB1 8.212851e-04  0.43084877
## 411        BEAN1 1.358209e-04  0.42955326
## 412        GALR1 9.648799e-04  0.42936883
## 413        USH1C 1.653595e-04  0.42900043
## 414         GRM8 1.836735e-03  0.42863266
## 415       SPHKAP 1.271318e-04  0.42863266
## 416         LGR5 1.271318e-04  0.42844901
## 417       TSPAN8 2.720430e-03  0.42808219
## 418        CDH22 4.456140e-05  0.42771600
## 419      SLC17A6 9.463199e-03  0.42698548
```

```
## 420      SEMA3D 1.945455e-04  0.42680324
## 421        CNR1 3.987730e-04  0.42662116
## 422     PCDH11Y 5.536585e-04  0.42625746
## 423       MYO1E 4.693333e-04  0.42589438
## 424        OXR1 1.653595e-04  0.42589438
## 425        LRP2 9.074468e-05  0.42571307
## 426      NIPAL4 3.503185e-04  0.42553191
## 427     TCERG1L 4.881890e-04  0.42517007
## 428        DGKK 2.084158e-04  0.42480884
## 429        HN1L 1.392857e-04  0.42462845
## 430     MADCAM1 4.447439e-04  0.42426814
## 431       ZMAT4 2.447796e-03  0.42390844
## 432      SEMA5A 9.969697e-05  0.42372881
## 433      IGFBP3 2.223587e-03  0.42354934
## 434        RBP4 5.225653e-03  0.42337003
## 435         EHF 3.115523e-04  0.42301184
## 436        TAC1 8.353081e-03  0.42069836
## 437      MBOAT1 4.830718e-03  0.42052145
## 438       FNDC9 1.825613e-03  0.42034468
## 439     PPP1R17 5.901521e-03  0.42016807
## 440       RPRML 1.665313e-02  0.41999160
## 441     PCDH11X 8.747592e-04  0.41893590
## 442   LINC00473 2.819149e-03  0.41893590
## 443       DACH1 1.000000e-04  0.41858518
## 444       SCN1A 4.431373e-05  0.41841004
## 445        UGT8 3.207237e-04  0.41718815
## 446        AQP3 2.497162e-03  0.41701418
## 447        EGR1 2.999703e-02  0.41649313
## 448   LOC149134 6.518692e-04  0.41631973
## 449       CCNE2 1.952941e-04  0.41459370
## 450       ENPP2 3.110294e-04  0.41442188
## 451       MYO5B 7.286432e-03  0.41356493
## 452       NTNG1 5.976190e-04  0.41288192
## 453       PRKG2 1.377676e-03  0.41220115
## 454        FA2H 4.854881e-04  0.41186161
## 455       FRMD5 2.060914e-04  0.41135335
## 456        PMCH 1.126472e-02  0.41034058
## 457       CHRM2 4.600505e-03  0.40983607
## 458       PCSK6 8.747592e-04  0.40899796
## 459     ZSCAN12 1.355906e-03  0.40866367
## 460      POU4F1 1.303738e-02  0.40816327
## 461    TMEM200A 1.963351e-03  0.40783034
## 462        GNG4 9.909256e-04  0.40766408
## 463        ANLN 1.955556e-04  0.40716612
## 464        SYT6 2.089172e-03  0.40600893
## 465         SCD 7.045455e-05  0.40502228
## 466       FGF14 2.381166e-04  0.40502228
## 467        RYR2 1.059649e-03  0.40469446
## 468        NGEF 1.972678e-04  0.40322581
## 469         ALK 7.173448e-04  0.40322581
## 470      RAD54L 2.214022e-03  0.40290089
## 471        LDB3 1.271318e-04  0.40225261
## 472        GAD2 2.494279e-03  0.40160643
## 473        SGK2 2.339623e-04  0.40064103
```

```
## 474       C21orf91 1.358209e-04  0.40032026
## 475        SOX2-OT 2.421053e-03  0.40032026
## 476         ADRA1B 1.825613e-03  0.40016006
## 477           MOBP 8.291667e-05  0.39968026
## 478          C1QL1 1.160535e-03  0.39968026
## 479          TMTC2 3.147651e-04  0.39936102
## 480           KLK6 2.570833e-04  0.39904230
## 481        CNTNAP4 2.732283e-04  0.39777247
## 482  LOC100292909 1.057338e-02  0.39698293
## 483           SYN2 8.618677e-04  0.39666799
## 484        ADAMTS9 1.436170e-03  0.39635355
## 485          SAMD3 6.786427e-03  0.39478879
## 486      LOC344595 2.010526e-04  0.39447732
## 487        TMEM139 5.949458e-03  0.39447732
## 488        TMEM233 9.973046e-03  0.39339103
## 489           ZEB2 4.431373e-05  0.39292731
## 490       A3GALT2P 3.252525e-03  0.39277298
## 491        ADAMTS4 4.911003e-03  0.39215686
## 492             TF 1.366352e-03  0.39154268
## 493        ADAMTS1 8.519389e-03  0.39138943
## 494          CNDP1 4.776490e-03  0.39032006
## 495        EMILIN2 1.838509e-04  0.39016777
## 496        TMEM144 2.214022e-03  0.38910506
## 497         PRSS3P2 1.281810e-02  0.38744673
## 498          SMAD9 4.158590e-03  0.38729667
## 499        PCOLCE2 7.885010e-04  0.38714673
## 500         CAPN13 8.618677e-04  0.38699690
## 501          KLHL4 5.288221e-04  0.38684720
## 502           PLP1 1.377676e-03  0.38624952
## 503          SSTR2 4.693333e-04  0.38595137
## 504           SIM1 4.137760e-02  0.38565368
## 505          KCNA5 1.522388e-03  0.38550501
## 506          RAB3B 2.977444e-04  0.38550501
## 507          KCNJ2 2.726190e-04  0.38491147
## 508           NNAT 2.378698e-03  0.38476337
## 509         FIGNL2 4.164306e-04  0.38476337
## 510            FOS 3.722414e-02  0.38476337
## 511        SLC35F4 2.552301e-04  0.38417211
## 512        NEUROD6 1.234983e-02  0.38299502
## 513          TRPM6 1.953757e-04  0.38284839
## 514           STC1 2.214022e-03  0.38167939
## 515           OCA2 1.115681e-02  0.38124285
## 516          EVI2A 4.933333e-05  0.38095238
## 517        CREB3L3 1.318224e-02  0.38066235
## 518           RBP7 1.916329e-03  0.37965072
## 519          SYNJ2 3.147651e-04  0.37950664
## 520            GDA 6.635199e-03  0.37936267
## 521          GPR37 6.454483e-03  0.37921881
## 522           NMBR 2.497175e-03  0.37878788
## 523         IGFBP5 3.283283e-03  0.37764350
## 524          KCNG1 3.919753e-04  0.37764350
## 525          MAP6D1 3.147651e-04  0.37707391
## 526        SLC45A3 4.417344e-04  0.37707391
## 527          KLK10 1.903861e-02  0.37622272
```

```
## 528        GPR62 1.377676e-03  0.37593985
## 529       ADRA1A 4.164306e-04  0.37593985
## 530         ERMN 3.147651e-04  0.37551633
## 531      LDLRAP1 6.836283e-04  0.37523452
## 532      C21orf62 5.332817e-03  0.37453184
## 533          TOX 8.939828e-03  0.37397158
## 534        ABCA8 8.353081e-03  0.37174721
## 535         GAD1 1.906203e-02  0.37133309
## 536       YEATS2 4.342561e-03  0.37091988
## 537          MOG 7.949029e-03  0.37064492
## 538        ENPP6 2.231834e-02  0.37064492
## 539    LOC653051 2.453390e-04  0.37050760
## 540        OPRM1 1.746916e-02  0.37037037
## 541         CDH4 4.771784e-03  0.36954915
## 542  LOC100129122 2.677596e-03  0.36616624
## 543        ITGA2 7.959596e-04  0.36603221
## 544       BTBD16 1.184928e-02  0.36603221
## 545      STEAP1B 2.728249e-03  0.36483035
## 546      SLC31A2 5.821768e-03  0.36469730
## 547      HIST1H3A 3.495146e-04  0.36416606
## 548          ARC 2.689633e-02  0.36350418
## 549       DMRTC2 1.764331e-04  0.36337209
## 550      RASGRF1 2.284360e-04  0.36271309
## 551    LOC643475 1.461405e-02  0.36218761
## 552      CCDC152 5.657100e-03  0.36192544
## 553        SUSD2 6.746032e-05  0.36192544
## 554       B3GAT2 3.522618e-03  0.36101083
## 555      SHROOM4 1.597315e-04  0.35958288
## 556      SLC17A8 1.754665e-02  0.35906643
## 557         PLD1 2.384956e-04  0.35829452
## 558      B3GALTL 1.025093e-02  0.35829452
## 559      SEC14L5 1.318224e-02  0.35778175
## 560        RAMP3 2.666455e-02  0.35739814
## 561       SPOCK1 1.972826e-04  0.35727045
## 562         RTP1 4.776490e-03  0.35676061
## 563       CARNS1 1.441048e-02  0.35561878
## 564        COL5A2 5.332817e-03  0.35448423
## 565         MEIS2 1.119863e-03  0.35435861
## 566        BACE2 1.237705e-04  0.35373187
## 567        BCAS1 6.965066e-04  0.35260931
## 568     C10orf128 1.785714e-03  0.35260931
## 569       C4orf32 1.736695e-03  0.35236082
## 570        PHLDB1 6.361502e-04  0.35223670
## 571         ST18 3.089598e-03  0.35198874
## 572       HAPLN4 1.322049e-02  0.35026270
## 573         FOXA2 9.074468e-05  0.35014006
## 574        HTR1A 1.853715e-02  0.34952814
## 575        GCHFR 4.150442e-03  0.34904014
## 576        HOXD3 1.076280e-02  0.34806822
## 577        TAS2R5 3.089598e-03  0.34770515
## 578         CLCA4 2.255730e-03  0.34650035
## 579        SEMA4D 3.160336e-02  0.34638033
## 580         NELL1 1.126472e-02  0.34447124
## 581         GRIK1 2.284360e-04  0.34176350
```

```
## 582       C18orf56 9.054670e-03  0.34153005
## 583           BTG3 1.271318e-04  0.34048349
## 584          PTPRK 1.564327e-03  0.33875339
## 585          PTPRH 5.589354e-03  0.33772374
## 586         SLC5A11 3.781049e-03  0.33658701
## 587          PROX1 2.467062e-02  0.33658701
## 588          CMTM5 2.959184e-02  0.33568312
## 589          HOXD1 2.373567e-02  0.33400134
## 590         TRIM59 6.454483e-03  0.33277870
## 591        TP53TG5 6.375000e-06  0.33123551
## 592            GEM 6.836283e-04  0.33101622
## 593          STAC2 2.173760e-02  0.33057851
## 594        C12orf76 1.955381e-03  0.32959789
## 595         CDKN1C 1.128205e-03  0.32754668
## 596         RASSF2 2.779923e-04  0.32690422
## 597           PLK5 3.359684e-03  0.32658393
## 598 LOC100289058 1.266603e-02  0.32626427
## 599           CD44 1.245981e-03  0.32499188
## 600          KCNC2 2.151899e-03  0.32414911
## 601           GRM7 2.388235e-03  0.32310178
## 602        C22orf42 9.085747e-03  0.32216495
## 603        EPB41L4B 1.162791e-03  0.32113038
## 604         ATP11A 3.627760e-04  0.32071841
## 605           NDE1 7.505330e-04  0.32010243
## 606           MEI1 4.498106e-02  0.31816736
## 607           GAS7 5.332817e-03  0.31766201
## 608          TEX26 6.129944e-03  0.31705770
## 609         HAPLN2 1.163836e-02  0.31675641
## 610          CDH13 1.687943e-03  0.31645570
## 611         EGFLAM 3.362286e-02  0.31230481
## 612        TMEM125 4.386266e-03  0.31046259
## 613          KCTD8 1.952941e-04  0.31017370
## 614          SNTG2 2.794896e-02  0.30921459
## 615         TSPAN18 2.121982e-03  0.30778701
## 616          SHOX2 3.601408e-02  0.30590395
## 617           GLDN 1.160580e-02  0.30581040
## 618         DSCAML1 3.843750e-04  0.30571691
## 619         KCNIP1 5.755771e-03  0.30284676
## 620           GLTP 4.164306e-04  0.30165913
## 621        TMEM235 1.772152e-02  0.30147724
## 622          KCNC1 1.242206e-02  0.29994001
## 623        GALNT14 1.377676e-03  0.29904306
## 624          PTPRD 2.284360e-04  0.29832936
## 625         SLC26A9 9.539474e-03  0.29824038
## 626         STXBP6 1.678521e-03  0.29806259
## 627         CHRNA5 2.835545e-02  0.29717682
## 628           HMSD 1.596708e-02  0.29568303
## 629       KCNJ2-AS1 7.286432e-03  0.29498525
## 630          GRIP2 8.012048e-03  0.29359953
## 631         NKX6-1 5.332817e-03  0.28977108
## 632 LOC100270746 3.359684e-03  0.28910090
## 633           VSTM1 3.792833e-02  0.28776978
## 634         TCP11L2 1.071804e-03  0.28719127
## 635         ELOVL1 1.167665e-02  0.28661508
```

```
## 636          GJB1 1.473368e-02  0.28344671
## 637         CHST3 1.984334e-03  0.28272547
## 638     C17orf110 7.847946e-03  0.28248588
## 639       CCDC147 4.956849e-02  0.27995521
## 640         KCNE4 4.911003e-03  0.27831895
## 641         CRHBP 2.045124e-02  0.27777778
## 642        ZNF608 1.359712e-04  0.27762354
## 643          ENC1 1.377676e-03  0.27196084
## 644      POM121L1P 4.046553e-03  0.26917900
## 645         PDE1C 5.083665e-03  0.26490066
## 646          TJP2 1.267739e-02  0.26184865
## 647         SPSB2 6.831461e-04  0.26143791
## 648         SLED1 7.906504e-04  0.25839793
## 649          CHN2 2.214022e-03  0.25813113
## 650        NKX6-2 2.089172e-03  0.25693731
## 651 LOC100131170 6.129944e-03  0.25588536
## 652         APLNR 1.468804e-02  0.25529742
## 653       GPIHBP1 4.910714e-03  0.25342119
## 654         EDIL3 4.398907e-04  0.25233409
## 655         CDK18 1.237705e-04  0.24764735
## 656        ADRA2A 5.755771e-03  0.24420024
## 657        SH3TC2 1.332725e-02  0.24414062
## 658         DMRT2 1.963351e-03  0.24207214
## 659         PRR5L 6.803170e-03  0.23640662
## 660        DEPDC7 1.678521e-03  0.23457659
## 661        ZNF488 7.471980e-03  0.23424690
## 662      HIST1H4A 6.454483e-03  0.23293734
## 663        N4BP2L1 1.306836e-03  0.23201856
## 664        COL25A1 3.637218e-03  0.22904260
## 665         LLGL1 1.131498e-02  0.21877051
## 666         VAT1L 9.780622e-04  0.21105952
## 667      PPP1R14A 2.563818e-03  0.20942408
## 668         OPCML 3.670412e-03  0.20673971
## 669         ACTG2 4.430077e-02  0.20508614
## 670         ERBB3 7.022107e-03  0.19948135
## 671           BOK 3.414873e-03  0.19762846
## 672        SLC6A17 1.903861e-02  0.19447686
## 673        ATP10B 1.051797e-02  0.19230769
## 674         ACSL1 2.214022e-03  0.19029496
## 675        RBFOX1 1.677445e-02  0.18993352
## 676         SEMA3C 6.216028e-03  0.18670650
## 677        ANKRD65 2.779923e-04  0.17636684
## 678        ARHGEF2 2.213343e-02  0.17571604
## 679         GLRA2 1.816399e-02  0.16980812
## 680         WIPF1 2.453390e-04  0.16829350
## 681         PLIN3 1.838509e-04  0.16758840
## 682         GOLT1A 2.121539e-02  0.16553551
## 683         PRRG1 1.952941e-04  0.15987210
## 684         RTN4R 2.030848e-03  0.15900779
## 685         MYH11 2.978787e-02  0.15615240
## 686        NKX2-2 2.453390e-04  0.13732491
## 687     LOC728401 1.724690e-02  0.13468013
## 688        ZNF735 2.629228e-02  0.12899897
## 689         GREM1 3.584035e-02  0.11924636
```

```
## 690        BLID 2.329749e-03  0.11708231
## 691   KIAA1024L 1.140974e-02  0.11341726
## 692    KIAA1683 2.552301e-04  0.10998680
## 693        ATF3 1.471225e-02  0.09184423
## 694       P2RX7 1.175299e-02  0.08987956
## 695   MATN1-AS1 2.563818e-03  0.08946144
## 696        CLMN 2.226277e-03  0.07949758
```

Among the 696 genes found to be significantly upregulated or downregulated in the SNpc relative to the VTA were genes linked to dopamine, including dopamine receptor 2 (DRD2, BH-corrected p = 1.095541e-04, FC = 3.341), dopamine receptor 5 (DRD5, BH-corrected p = 1.169231e-06, FC = 0.2442002), dopamine active transporter (SLC6A3, BH-corrected p = 8.355197e-03, FC = 4.784) and vesicular monoamine transporter (SLC18A2, BH-corrected p = 0.01238258, FC = 4.656). The SLC18A2 gene codes for the VMAT2 enzyme, which transports dopamine into monoaminergic vesicles to prevent it from oxidizing [1]. VMAT2 activity has been shown to be inversely proportional to neuromelanin levels, so one may assume that higher SLC18A2 expression would contribute to less neuromelanin. However, the opposite case is evident here. With fold change = 4.656, SLC18A2 is actually upregulated in the SNpc, which contains neuromelanin, with respect to the VTA, which does not. This suggests that there may be more at play in neuromelanin formation than just the action of VMAT2.

Ideally, further analysis should be done on these 696 genes to determine which molecular pathways they are involved in. This includes gene set enrichment analysis and assigning gene ontology terms to groups of these genes. Due to time limitations, this could not be accomplished within the hackathon.

REFERENCES

[1] M. Vila, "Neuromelanin, aging, and neuronal vulnerability in Parkinson's disease," Mov. Disord., vol. 34, no. 10, pp. 1440–1451, 2019, doi: 10.1002/mds.27776.

[2] A. Kouli, K. M. Torsney, and W. Kuan, "Parkinson's Disease: Etiology, Neuropathology, and Pathogenesis," pp. 3–26, 2018, doi: https://doi.org/10.15586/codonpublications.parkinsonsdisease.2018.ch1.

[3] J. Segura-Aguilar and I. Paris, "Handbook of neurotoxicity," Handb. Neurotox., vol. 1–3, pp. 1–2371, 2014, doi: 10.1007/978-1-4614-5836-4.

[4] P. Muñoz, S. Huenchuguala, I. Paris, and J. Segura-Aguilar, "Dopamine oxidation and autophagy," Parkinsons. Dis., vol. 2012, 2012, doi: 10.1155/2012/920953.

[5] A. Herrera, P. Muñoz, H. W. M. Steinbusch, and J. Segura-Aguilar, "Are Dopamine Oxidation Metabolites Involved in the Loss of Dopaminergic Neurons in the Nigrostriatal System in Parkinson's Disease?," ACS Chem. Neurosci., vol. 8, no. 4, pp. 702–711, 2017, doi: 10.1021/acschemneuro.7b00034.

[6] A. Keo et al., "Transcriptomic signatures of brain regional vulnerability to Parkinson's disease," Commun. Biol., vol. 3, no. 1, pp. 1–12, 2020, doi: 10.1038/s42003-020-0804-9.