1. Background

1.1 Understanding Major Depressive Disorder, Extracellular Vesicles, and miRNA

According to the World Health Organization, Major Depressive Disorder (MDD), or colloquially known as depression, is a mental disorder which results in greater than 700,000 related deaths annually (WHO, 2023). Generally, depression is characterized by a persistent depressive mood, and a loss of interest (Mayo Clinic, 2022). Given that there is no cure, treatment, or known origin for MDD, there is motivation to learn more about the factors that drive the progression of this disorder.

Extracellular vesicles (EVs) are capsules of cellular content that are periodically released by cells for various cellular mechanisms, such as communication (David & Wang, 2019). The composition of EVs details the state of the cell. miRNAs are one of the most studied contents within EVs, and they typically up/downregulate genetic expression (Ibrahim et al., 2024). There are theories that inappropriate miRNA concentrations cause dysregulation of proteins, and thus, lead to depression (Ding et al., 2023). Provided that EVs provide an avenue to study miRNAs, the authors wanted to use this connection to better expand knowledge of the interactions between these genetic regulators and MDD.

1.2 Original Paper Findings

The original authors found 8 differentially expressed miRNAs in females: miR-132-5p, let-7e-5p, miR-126-5p, miR-320b, and miR-421 to be upregulated, and miR-92a-3p, miR-885-5p, and miR-708-5p to be downregulated. Between these, they noted that miR-92a-3p had the greatest relevance to MDD. I will be attempting to reproduce their study and recover the miR-92a-3p as well. This project will be a Type B project.

2. Data

2.1 Data Type

The data is provided in Fastq format on the SRA database (study id=SRP489148). There are 40 total samples, 20 control, and 20 with MDD. All reads are paired-end reads, and each run has 2 replicates. On average, there are 6.55 million reads. Each Fastq file ranged from 247.86 megabytes to 1.84 gigabytes, and 675.75 megabases to 5.33 gigabases. In total, the dataset is 18.75 gigabytes, and 53.87 gigabases.

3. Analysis Steps

3. 1 Downloading the Files

To download the files from the SRA database, I wrote the list of accession and called the command "*cat accessions.txt | xargs -n 1 -P 4 fastq-dump --split-files.*" I directly downloaded the associated metadata for the samples and manually uploaded them to the MBIlabs-01 server.

3.2 exceRpt SmallRNA pipeline

The original authors utilized the exceRpt SmallRNA pipeline version 4.6.3 by Rob Kitchen (Rozowsky et al., 2020) for quality control. This pipeline contains a series of scripts that trim low quality reads, remove contaminant sequences, align, and calculate counts. Typically,

this resource is hosted on genboree workbench, an online resource that hosts bioinformatics tools with a graphic interface. Unfortunately, this resource was discontinued last year. The alternatives to installing these resources are through docker, a package manager that needs sudo permissions, or through an installer R script. None of these alternatives worked, as I do not have sudo permissions on the server, and the R script repeatedly echoed errors that I could not fix. Regardless, in an attempt to run a similar quality control process to the authors, I manually read and imitated the commands their pipeline runs through their GitHub scripts.

3.3 Parallel Processing of Quality Control of Fastq Files With Bash Scripting

Most of the computation steps in this course was done through R, where some commands could not run samples in parallel. If I were to compute each step for each file in series, I would not be able to finish this project in time. As a novel part of this project, I decided to learn basic bash scripting to automatically run files through various quality control steps. On the terminal, I wrote 4 bash scripts of similar structure trim_adapt.sh (remove adapters and low quality reads), align.sh (align to human reference genome), rem_rRNA.sh (remove rRNA sequences), and rem_uni.sh (remove UniVec sequences). These scripts consist of a list of accession numbers and a for loop that runs a quality control command on each file. Additionally, with the added nohup command, the scripts can run all sequences in parallel. Figures 8-11 show the bash scripts.

3.4 Quality Control Steps

To trim low quality sequences, I utilized a novel program, trimmomatic version 0.39 (Bolger et al., 2014). Copying the parameters provided by exceRpt, I cut nucleotides from ends with phred33 scores less than 3. I removed sequence lengths less than 18 nucleotides, and cut additional nucleotides with a sliding window of 4 nucleotides with a minimum phred33 quality of 20. Furthermore, I removed any illumina adapters.

To remove UniVec and rRNA sequences, I utilized a novel program, Bowtie2 version 2.4.2 (Langmead et al., 2009) with base arguments. UniVec sequences typically come from cloning genomic DNA, and are fragments of adapters and primers that should be removed (National Center for Biotechnology Information, 2017). rRNA sequences can dominate small RNA signals and are typically removed. I built the UniVec index with the provided UniVec sequences in the exceRpt GitHub repository, and the rRNA index with the Analysis of Ribosomal RNA Fragments version 2.1 (Alkin et al., 2022) R package's reference sequence. I then aligned the sequences to both indexes and took the unaligned sequences.

To check the quality of the sequences, I ran them under FastQC version 0.12.1 (Babraham Bioinformatics, 2023). Finally, I removed all miRNAs that have counts less than 10 in 30% or greater number of samples – as specified by the original authors.

3.5 Aligning and Getting Count Data

Using the hg38 human genome, I created an alignment index with STAR. I then used the Genocode basic gene annotation GTF file (Gencode, 2024) to align the fastq files to the index To calculate counts of miRNA I used ShortReads version 2.0.1 (Morgan et al., 2009).

3.6 Principal Component Analysis (PCA)

Another novel part of my analysis is conducting PCA. PCA is a dimension reduction technique that attempts to capture the variance of the data through vectors called principal components (PCs). Given that there were no variables in the metadata, the principal components can be used as covariates in downstream analysis steps and as variables to remove outliers. Firstly, I plotted the variance that each PC explains and take the top 5 PCs that explain the most variance. Then, I removed samples that have PCs values 3 standard deviations greater from the mean. Additionally, I plotted each PC against MDD/Control to see if they are autocorrelated, as subsequent analysis requires them to be independent.

3.7 Differential Expression Analysis and Functional Enrichment

I used DESeq2 (Love et al., 2014) to determine if there are any differentially expressed miRNAs. As stated by the authors, I used the median of ratios normalization method, and I also used the negative binomial general linear model to calculate differential expression. miRNAs with Benjamin Hochfield adjusted P-values < 0.05 were considered significantly expressed.

For functional enrichment, I used the novel program, miRinGO (no released version) (Sayed & Park, 2023). This is a web-based R-shiny application which finds the top GO terms based on a pool of significant miRNAs. The universe this application uses for the search is based on the location the sample is taken from, I used the genes affected by miRNA in the brain. For KEGG analysis, I used the novel program, miRNApath version 4.0 (Chiromatzo et al., 2007). Like miRinGO, this is a web-based application which finds the most significant KEGG results.

4. Results

4.1 Quality Control Results

The initial trimming with trimmomatic resulted in an average of 98.14% reads from both ends surviving per fastq file (min=97.76%, max=98.44%). An average of 0.15% of the reads aligned to UniVec sequences (min=0.04%, max=0.51%) per fastq file, and were removed. An additional 0.033% of sequences (min=0%, max=0.09%)per fastq file mapped to rRNA sequences and were removed. At this point, FastQC reported no issues with adapter contamination and a satisfactory read quality. In all samples, STAR failed to align an average of 51.08% of reads per fastq file (min=39.42%, max=57.56%) due to short sequence lengths. Overall, an average 27.72% unique reads (min=25.42%, max=36.51%) mapped to human miRNAs per fastq file. Using the hsa.gff3 file, an average of 2,224 reads aligned per fastq file (min=262, max=11303), which corresponds to an average of 0.27% alignment per fastq file (min=0.1%, max=0.3%). Finally, after removing miRNAs with less than 10 reads in more than 30% of samples, we discarded 2593 miRNAs. Our final sample contained 32 total miRNAs among the 40 subjects.

4.2 PCA results

The top 5 PCs explain 15.33%, 12.23%, 8.38%, 8.00%, and 7.81% of the variance. PCs beyond the first five explain five or less percent of the variance and were not included in subsequent analysis. Figure 1 shows a plot between PC1 and PC2, coloured by if they were MDD or control. There are no obvious signs of clustering between samples between MDD/Control. Figure 2 shows the autocorrelation matrix between each PC and the sample group

(MDD/control). The sample group did not correlate linearly with any of the top 5 PCs, and so, I used all five PCs for differential expression analysis as covariates.

Figure 3 shows the density plots of all samples under PC1-5. In PC1, sample SRR27929326 was an outlier. In PC2, SRR27929316 was an outlier. In PC3, samples SRR27929314 and SRR27929311 were outliers. In PC4 SRR27929318 was an outlier. There were no outliers in PC5. I removed all 5 outlier sequences and continued with 35 samples.

4.3 Differential Expression Analysis and Functional Enrichment Results

Running the DESeq command resulted in no significantly differentiated miRNA, as all samples had an adjusted P-value greater than 0.98. Figure 4 is a volcano plot of the miRNAs that passed quality control, with non-significant samples coloured black. To move on with functional enrichment analysis, I choose the most significant miRNA, hsa-miR-191-5p ($p < 0.1$).

Using this miRNA as my significant gene in miRinGO, and my universe as the genes affected by miRNAs in the brain, I generated the top GO terms plotted in figure 5. The two most significant GO terms found were nervous system development (id=0007268, $p < 0.001$), and chemical synaptic transmission (id=0007399, $p < 0.001$). I continued using hsa-miR-191-5p to find the most significant KEGG pathway with miRPATH. The top pathway that is related to MDD was thyroid hormone signalling pathway (8 related genes, false discovery rate = 0.0387).

5. Discussion

5.1 Comparison of Results

The authors had 8 significantly expressed miRNAs in females, my results found none. Figure 7 shows the overlap between the miRNAs that passed the quality control filters and the genes the authors found significant. As the venn diagram shows, the significant genes that the authors found for females passed our quality control steps. Additionally, the authors ran GO analysis for their results, and their top terms were “neurotransmitter transport” and “regulation of transmitter levels.” One of my GO terms, “chemical synaptic transmission” is related to neurotransmitters, as they are the basis of neuronal communication (Sheffler et al., 2023). My most significant KEGG pathways also did not relate to neurotransmitters.

5.2 Reason for Differing Results

The reason I had different results than the authors of the original paper is likely due to the in-accessibility of the ExceRpt SmallRNA pipeline. I tried to imitate as much of the original pipeline as I could, but they had more internal steps and quality control steps that I could not imitate. Additionally, the paper mentioned that they had more variables to act as covariates, whereas the metadata they provided contained no useful covariates. The authors utilized a combination of PCs and other variables in the DESeq differential analysis step, whereas I only had PCs. Finally, for gene ontology, the authors never specified what they used for the universe, which could explain the different significant results on top of the differing significant miRNAs.

Statement of Contribution

I attended presentations from 11:30 AM to 2:45 PM. I believed I asked around 3-5 questions.

6. Appendix

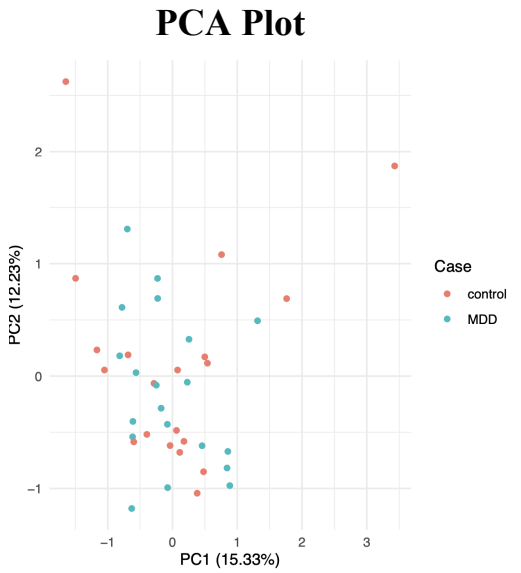

Figure 1. Scatterplot of the first two PCs. PC1 explains 15.33% of the variance of the dataset, and PC2 explains 12.33%. The samples are coloured by MDD/control. There is no obvious evidence of clustering.

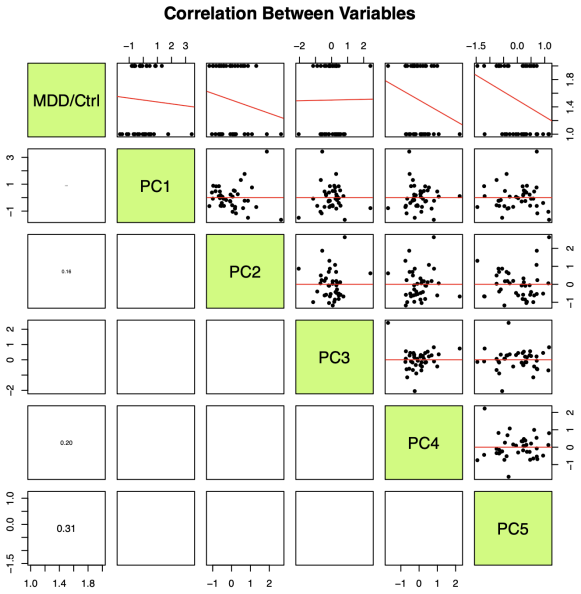

Figure 2. Autocorrelation plot between Principal components and MDD/Control. Box at row n, column m, depicts information between box at n, n and box m, m. Boxes below the diagonal of the matrix are the R^2 correlation values between the variables, where larger sized text indicates a

stronger correlation. Boxes above the diagonal plot are a scatterplot of the samples with the relevant two variables as the axis.

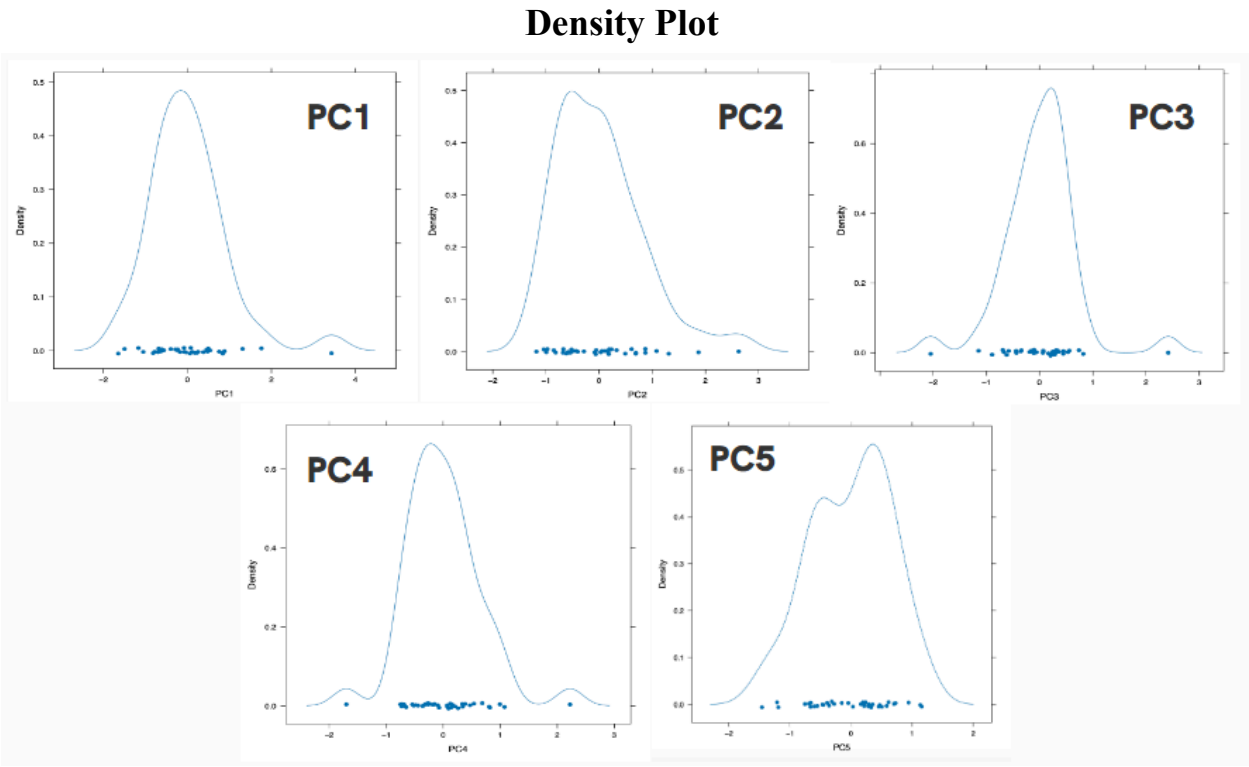

Figure 3. Density plot of all samples for Principal components (PC) 1 to 5. Samples are plotted as the rug. In PC1, sample SRR27929326 was an outlier. In PC2, SRR27929316 was an outlier. In PC3, samples SRR27929314 and SRR27929311 were outliers. In PC4 SRR27929318 was an outlier. There were no outliers in PC5.

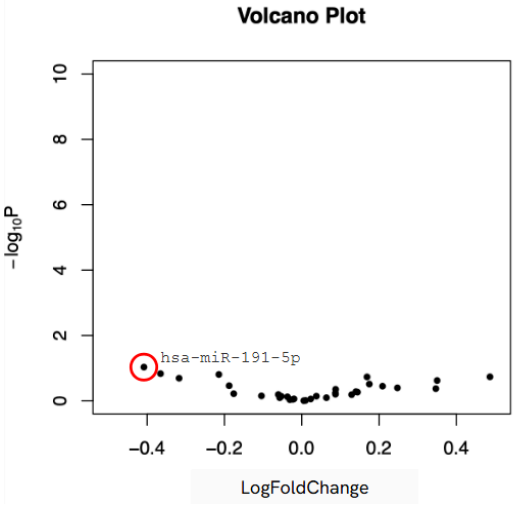

Figure 4. Volcano plot of the potentially 32 differentially expressed miRNA after running DESeq command. Lowest P-value (p=0.09) miRNA is highlighted with a red circle, hsa-miR-191-5p.

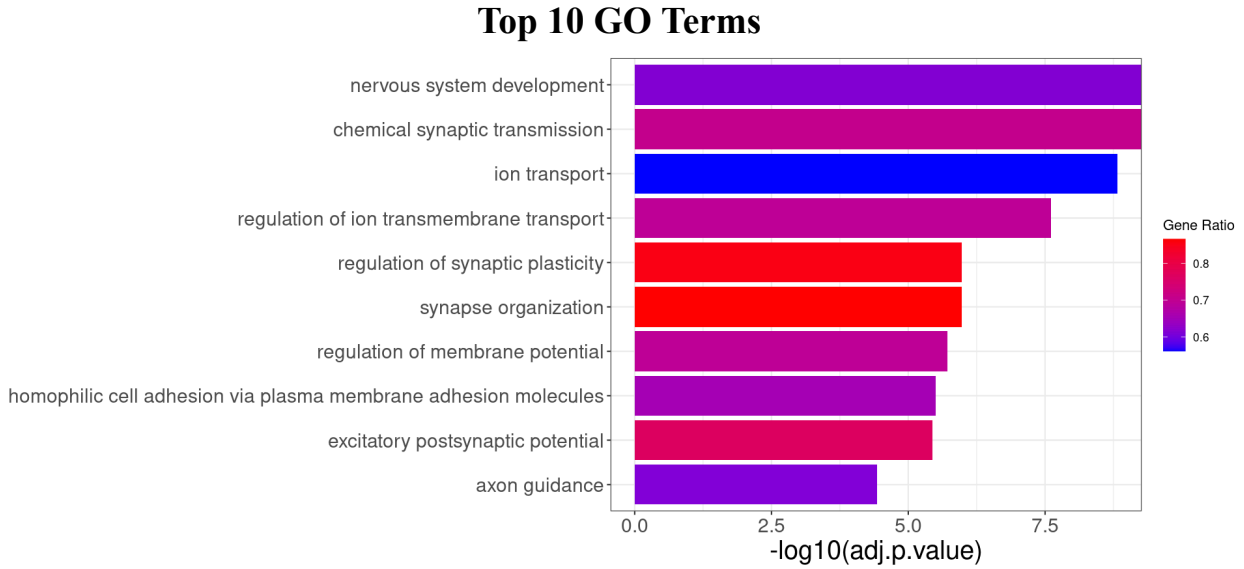

Figure 5. Bar plot of the top 10 gene ontology terms of hsa-miR-191-5p with the universe set to the genes affected by miRNA in the brain.

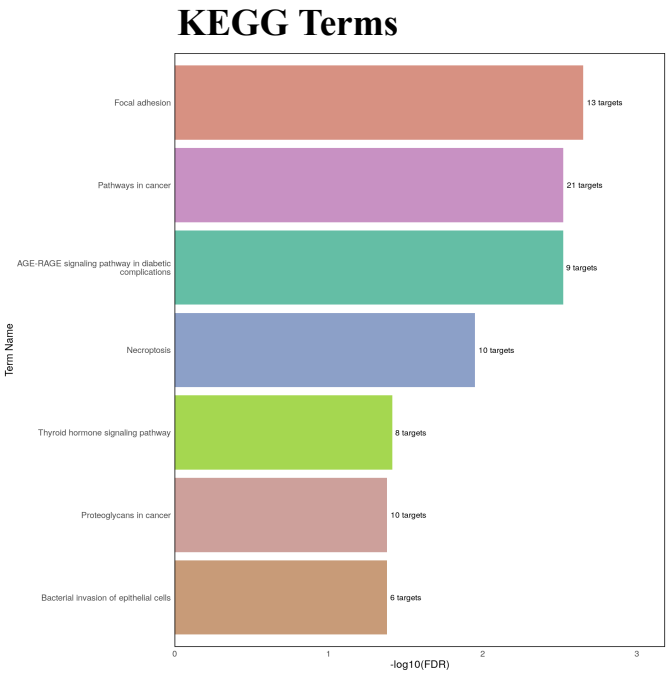

Figure 6. Barplot of the top KEGG pathways of hsa-miR-191-5p. The number of targets indicates how many genes are affected by the miRNA in that specific pathway.

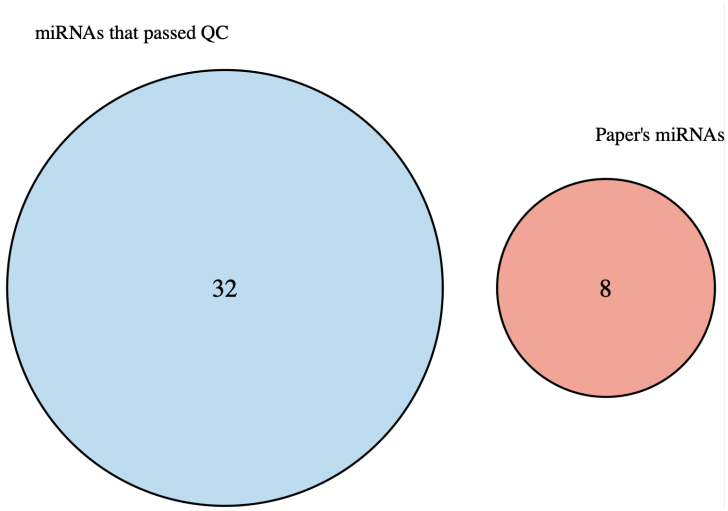

Figure 7. Venn diagram of miRNAs that the original authors found significant (n=8), with the number of miRNAs that passed quality control (n=32).

```bash
#!/bin/bash

samples=(
SRR27929302
SRR27929303
SRR27929304
SRR27929305
SRR27929306
        )

# Run Trimmomatic for each sample with nohup
for sample in "${samples[@]}"; do

  nohup trimmomatic PE -threads 16 -phred33 \
    "${sample}_1.fastq" "${sample}_2.fastq" \
    "${sample}_1_paired.fastq" "${sample}_1_unpaired.fastq" \
    "${sample}_2_paired.fastq" "${sample}_2_unpaired.fastq" \
    ILLUMINACLIP:adapters.fa:2:30:10:8:true \
    SLIDINGWINDOW:4:20 \
    LEADING:3 \
    TRAILING:3 \
    MINLEN:18 > "${sample}_trimmomatic.log" 2>&1 &

done
```

Figure 8. Bash script to run trimmomatic in parallel and in the background. Other samples were in other scripts so as to not overload the server's RAM.

```bash
#!/bin/bash

samples=(
        SRR27929302
SRR27929303
SRR27929304
SRR27929305
SRR27929306
SRR27929307
SRR27929308
SRR27929309
SRR27929310
SRR27929311
        )

for sample in "${samples[@]}"; do

  nohup bowtie2 -x ../uni_ind/UniVecInd \
    -1 "${sample}_1_paired.fastq" -2 "${sample}_2_paired.fastq" \
    --un-conc-gz "${sample}_noUni_%.fastq.gz" \
    --al-conc-gz "${sample}_UniVec_%.fastq.gz" \
    --threads 2 \
    -S "${sample}_univec_aligned.sam" > "${sample}_univec_filtering.log" 2>&1 &

done
```

```bash
#!/bin/bash

samples=(SRR27929312
SRR27929313
SRR27929314
SRR27929315
SRR27929316
SRR27929317
SRR27929318
SRR27929319
SRR27929320
SRR27929321
        )

for sample in "${samples[@]}"; do

  nohup bowtie2 -x ../rna_ind/rRNA_index \
    -1 "${sample}_noUni_1.fastq.gz" -2 "${sample}_noUni_2.fastq.gz" \
    --un-conc-gz "${sample}_clean_%.fastq.gz" \
    --al-conc-gz "${sample}_rRNA_%.fastq.gz" \
    --threads 12 \
    -S "${sample}_rRNA.sam" > "${sample}_rRNA_filtering.log" 2>&1 &

done
```

Figure 9. Left bash script to remove UniVec sequences. Right bash script to remove rRNA sequences. Both scripts utilize Bowtie2 to align to a Bowtie2 index.

```bash
#!/bin/bash

samples=(
        SRR27929302
        SRR27929303
        SRR27929304
        SRR27929305
        SRR27929306
        SRR27929307
        SRR27929308
        SRR27929309
        SRR27929310
        SRR27929311
        )

# Define paths
GENOME_DIR="../genome/gen"
THREADS=16

for sample in "${samples[@]}"; do

  nohup STAR --runThreadN ${THREADS} \
    --genomeDir ${GENOME_DIR} \
    --readFilesIn "${sample}_clean_1.fastq.gz" "${sample}_clean_2.fastq.gz" \
    --readFilesCommand gunzip -c \
    --outFileNamePrefix "${sample}_" \
    --outSAMtype BAM SortedByCoordinate \
    --outBAMcompression 10 \
    --limitBAMsortRAM 20000000000 > "${sample}_STAR.log" 2>&1 &

done
```

Figure 10. Bash script to align cleaned fastq files to hg38 with STAR.

```bash
#!/bin/bash

samples=(
        SRR27929302
        SRR27929303
        SRR27929304
        SRR27929305
        SRR27929306
        SRR27929307
        SRR27929308
        SRR27929309
        SRR27929310
        )

GTF="hsa.gff3"  # Path to miRNA annotation (GFF/GTF)
THREADS=16

for sample in "${samples[@]}"; do

  nohup featureCounts \
    -T ${THREADS} \
    -a ${GTF} \
    -F gff \
    -o "${sample}_miRNA_counts.txt" \
    -t miRNA \
    -g "Name" \
    -M \
    --fraction \
    --primary \
    -O \
    -s 0 \
    "${sample}_Aligned.sortedByCoord.out.bam" > "${sample}_featureCounts.log" 2>&1 &

done

do awk BEGIN {OFS=\"\t\"} NR==1 || FNR>1 *_miRNA_counts.txt > combined_miRNA_counts.txt
```

Figure 11. Bash script to get miRNA counts from aligned samples with SubReads.

```r
# load files
count_files <- list.files(pattern = "*_miRNA_counts.txt$", full.names = TRUE)

# Initialize an empty list to store data
count_list <- list()

# Read each file and store it in the list
for (file in count_files) {
  sample_name <- gsub("*_miRNA_counts.txt", "", basename(file))
  df <- read.delim(file, header=TRUE, row.names=1, comment.char = "#")
  count_list[[sample_name]] <- setNames(df[, ncol(df)], rownames(df))
}

# Combine into a single data frame
final_counts <- as.data.frame(do.call(cbind, count_list))
colnames(final_counts) <- names(count_list)  # Set column names to sample names
```

Figure 12. R Code to load in miRNA counts files and turn them all into a singular matrix

```r
# Combine into a single data frame
final_counts <- as.data.frame(do.call(cbind, count_list))
colnames(final_counts) <- names(count_list)  # Set column names to sample names

# remove low counts
min_count <- 10
min_samples <- ceiling(0.3 * ncol(final_counts))  # 30% of total samples

# keep only miRNAs with ≥10 counts in at least 30% of samples (exceRpt steps)
filtered_counts <- final_counts[rowSums(final_counts >= min_count) >= min_samples, ]

# read meta and change rownames to Runs
metadata <- read.csv("SraRunTable.csv", header=TRUE, row.names=1)
metadata$cell_type <- as.factor(metadata$cell_type)

# load into dds (already noranmlized internall)
dds <- DESeqDataSetFromMatrix(countData = filtered_counts,
                              colData = metadata, # row names look good
                              design = ~ cell_type)  # mdd/not is specified in this
```

Figure 13. R Code to turn counts matrix into a DESeq2 object.

```r
# PCA
vsd <- varianceStabilizingTransformation(dds, blind=F)
pca <- prcomp(t(assay(vsd)))

pdf('scree.pdf')
plot(summary(pca)$importance[2,1:10]*100, ylab="% Variance Explained", xlab="PC", type="b")
dev.off()

# Extract PC1-PC5 which had vairance > ~5%
pca_data <- pca$x[, 1:5]

# Add PCs to the colData of the DESeq2 object
colData(dds) <- cbind(colData(dds), pca_data)
```

Figure 14. R code to compute principal component analysis for the dataset and add principal component analysis data back into the DESeq2 object to use as covariates.

```r
# remove any with greater 2sd from norm (95% interval)
pdf('density.pdf')
densityplot(PC1, pch=19)
densityplot(PC2, pch=19)
densityplot(PC3, pch=19)
densityplot(PC4, pch=19)
densityplot(PC5, pch=19)
dev.off()

# plot sample respect to PC
PC1 <- as.numeric(pca$x[,1])
PC2 <- as.numeric(pca$x[,2])
PC3 <- as.numeric(pca$x[,3])
PC4 <- as.numeric(pca$x[,4])
PC5 <- as.numeric(pca$x[,5])

sample_outliers=c()
alloutliers=c()
for(i in 1:5){
  a<-subset(rownames(pca$x), pca$x[,i] > (mean(pca$x[,i])+3*sd(pca$x[,i])))
  b<-subset(rownames(pca$x), pca$x[,i] < (mean(pca$x[,i])-3*sd(pca$x[,i])))
  out<-c(a,b)
  sample_outliers <- c(sample_outliers,out)
  print(paste("outliers in PCA",i,":",sep=""))
  print(sample_outliers)
  alloutliers=c(alloutliers,sample_outliers)
  sample_outliers=c()
}

# drop these then re-run
dds_filtered <- dds[, !colnames(dds) %in% alloutliers]
```

Figure 15. R code to find outliers within the PCAs and remove them from the DESeq2 object if they are three standard deviations from the mean.

```r
# run with covariates
design(dds_filtered) <- formula(~ cell_type + PC1 + PC2 + PC3 + PC4 + PC5)

ddsRes <- DESeq(dds_filtered)
res <- results(ddsRes, contrast = c("cell_type", "MDD", "control"))
res_sig <- res[res$pvalue == min(res$pvalue),] # Adjusted Pvalues were not significant
res_df <- as.data.frame(res)
```

Figure 16. R code to change the formula for DESeq2 and calculate differentially expressed miRNAs

7. References

Alkan, F., Wilkins, O. G., Hernández-Pérez, S., Ramalho, S., Silva, J., Ule, J., & Faller, W. J. (2022). Identifying ribosome heterogeneity using ribosome profiling. *Nucleic Acids Research* (Vol. 50, Issue 16, pp. e95–e95). Oxford University Press (OUP). https://doi.org/10.1093/nar/gkac484

Babraham Bioinformatics. (2023, January 3). *FastQC.* https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* (Vol. 30, Issue 15, pp. 2114–2120). Oxford University Press (OUP). https://doi.org/10.1093/bioinformatics/btu170

Chiromatzo, A. O., Oliveira, T. Y., Pereira, G., Costa, A. Y., Montesco, C. A., Gras, D. E., Yosetake, F., Vilar, J. B., Cervato, M., Prado, P. R., Cardenas, R. G., Cerri, R., Borges, R. L., Lemos, R. N., Alvarenga, S. M., Perallis, V. R., Pinheiro, D. G., Silva, I. T., Brandão, R. M., Cunha, M. A., … Silva, W. A., Jr (2007). miRNApath: a database of miRNAs, target genes and metabolic pathways. *Genetics and molecular research* (Vol. 6, Issue 4, p. 859–865).

Ding, R., Su, D., Zhao, Q., Wang, Y., Wang, J.-Y., Lv, S., & Ji, X. (2023). The role of microRNAs in depression. *In Frontiers in Pharmacology* (Vol. 14). Frontiers Media SA. https://doi.org/10.3389/fphar.2023.1129186

Doyle, L., & Wang, M. (2019). Overview of Extracellular Vesicles, Their Origin, Composition, Purpose, and Methods for Exosome Isolation and Analysis. *Cells* (Vol. 8, Issue 7, p. 727). MDPI AG. https://doi.org/10.3390/cells8070727

Gencode. (2024, October 17). *Release 47 (GRCh38.p14).* Human. https://www.gencodegenes.org/human/

Ibrahim, P., Denniston, R., Mitsuhashi, H., Yang, J., Fiori, L. M., Żurawek, D., Mechawar, N., Nagy, C., & Turecki, G. (2024). Profiling Small RNA From Brain Extracellular Vesicles in Individuals With Depression. *International Journal of Neuropsychopharmacology* (Vol. 27, Issue 3). Oxford University Press (OUP). https://doi.org/10.1093/ijnp/pyae013

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* (Vol. 10, Issue 3). Springer Science and Business Media LLC. https://doi.org/10.1186/gb-2009-10-3-r25

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. In Genome Biology (Vol. 15, Issue 12). Springer Science and Business Media LLC. https://doi.org/10.1186/s13059-014-0550-8

Mayo Clinic. (2022, October 14). *Depression (major depressive disorder).* Diseases and Conditions. https://www.mayoclinic.org/diseases-conditions/depression/symptoms-causes/syc-20356007

Morgan, M., Anders, S., Lawrence, M., Aboyoun, P., Pagès, H., & Gentleman, R. (2009). ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics* (Vol. 25, Issue 19, pp. 2607–2608). Oxford University Press (OUP). https://doi.org/10.1093/bioinformatics/btp450

National Center for Biotechnology Information (2017, November 28). *The UniVec Database*. https://www.ncbi.nlm.nih.gov/tools/vecscreen/univec/#Overview

Rozowsky, J., Kitchen, R. R., Park, J. J., Galeev, T. R., Diao, J., Warrell, J., Thistlethwaite, W., Subramanian, S. L., Milosavljevic, A., & Gerstein, M. (2019). exceRpt: A Comprehensive Analytic Platform for Extracellular RNA Profiling. *Cell Systems* (Vol. 8, Issue 4, pp. 352-357.e3). Elsevier BV. https://doi.org/10.1016/j.cels.2019.03.004

Sayed, M., & Park, J. W. (2023). miRinGO: Prediction of Biological Processes Indirectly Targeted by Human microRNAs. *Non-Coding RNA* (Vol. 9, Issue 1, p. 11). MDPI AG. https://doi.org/10.3390/ncrna9010011

World Health Organization. (2023, March 13). *Depressive disorder (depression)*. https://www.who.int/news-room/fact-sheets/detail/depression

Sheffler, Z. M., Reddy, V., & Pillarisetty, L. S. (2025). Physiology, neurotransmitters. *StatPearls*. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/30969716