

Supplement 3A: Principal component analysis



PCA R Code for Supplementary Material

```
# Load necessary libraries
library(readr)
library(ggplot2)
# Step 1: Load normalized count data
normalized_counts <- read_csv("~/Documents/normalized_counts.csv")
# Step 2: Transpose data for PCA
pca_input <- as.data.frame(normalized_counts[, -1])
rownames(pca_input) <- normalized_counts$Genes
pca_input_t <- t(pca_input)
# Step 3: Run PCA
pca_result <- prcomp(pca_input_t, scale. = TRUE)
# Step 4: Create sample metadata
sample_metadata <- data.frame(
  Sample = colnames(normalized_counts)[-1],
  BrainRegion = sapply(strsplit(colnames(normalized_counts)[-1], "_"), function(x) x[1]),
  Treatment = sapply(strsplit(colnames(normalized_counts)[-1], "_"), function(x) x[2]),
  Sex = sapply(strsplit(colnames(normalized_counts)[-1], "_"), function(x) x[4])
)
# Step 5: Create PCA dataframe
pca_scores <- as.data.frame(pca_result$x)
pca_scores$Sample <- rownames(pca_scores)
pca_scores <- merge(pca_scores, sample_metadata, by = "Sample")
# Step 6: Plot by Sex
ggplot(pca_scores, aes(x = PC1, y = PC2, color = Sex)) +
  geom_point(size = 3) +
  theme_minimal() +
  ggtitle("PCA: Colored by Sex")
# Step 7: Plot by Brain Region
ggplot(pca_scores, aes(x = PC1, y = PC2, color = BrainRegion)) +
  geom_point(size = 3) +
  theme_minimal() +
  ggtitle("PCA: Colored by Brain Region")
# Step 8: Plot by Treatment
ggplot(pca_scores, aes(x = PC1, y = PC2, color = Treatment)) +
  geom_point(size = 3) +
  theme_minimal() +
  ggtitle("PCA: Colored by Treatment")
# Step 9: Plot by Brain Region + Treatment
ggplot(pca_scores, aes(x = PC1, y = PC2, color = interaction(BrainRegion, Treatment))) +
  geom_point(size = 3) +
  theme_minimal() +
  ggtitle("PCA: Colored by Brain Region + Treatment")
```