/mnt/rstor/genetics/Wynshaw-BorisLab/lab139/R-4.1.2/bin/R

library(Seurat)  
library(data.table)  
library(Matrix)  
library(sctransform)  
library(dplyr)  
library(RColorBrewer)  
library(ggplot2)

library(patchwork)

library(DoubletFinder)

library(ROCR)

library(fields)

library(KernSmooth)

library(parallel)

setwd("/mnt/rstor/genetics/Wynshaw-BorisLab/lab139/analysis\_w\_newref/clay\_16wk\_si\_dob20210816/")

clay1\_16wk\_si <- readRDS("/mnt/rstor/genetics/Wynshaw-BorisLab/lab139/analysis\_w\_newref/clay\_16wk\_si\_dob20210816/2022\_12\_02\_clay1\_cellbender\_afterRead10x\_h5.rds")

clay1\_16wk\_si <- CreateSeuratObject(counts = clay1\_16wk\_si, project = "clay1\_16wk\_si")

clay1\_16wk\_si[["percent.mt"]] <- [PercentageFeatureSet](https://satijalab.org/seurat/reference/PercentageFeatureSet.html)(clay1\_16wk\_si, pattern = "^MT-")

clay1\_16wk\_si <- [subset](https://rdrr.io/r/base/subset.html)(clay1\_16wk\_si, subset = nFeature\_RNA > 1250 & nFeature\_RNA < 7500 & percent.mt < 15)

clay1\_16wk\_si <- SCTransform(clay1\_16wk\_si, vst.flavor = "v2", vars.to.regress = "percent.mt")

setwd("/mnt/rstor/genetics/Wynshaw-BorisLab/lab139/analysis\_w\_newref/clay\_16wk\_si\_dob20220211/")

clay2\_16wk\_si <- readRDS("/mnt/rstor/genetics/Wynshaw-BorisLab/lab139/analysis\_w\_newref/clay\_16wk\_si\_dob20220211/clay\_16wk\_si\_dob20220211\_raw\_feature\_bc\_matrix\_cellbender\_filtered.rds")

clay2\_16wk\_si <- CreateSeuratObject(counts = clay2\_16wk\_si, project = "clay2\_16wk\_si")

clay2\_16wk\_si[["percent.mt"]] <- [PercentageFeatureSet](https://satijalab.org/seurat/reference/PercentageFeatureSet.html)(clay2\_16wk\_si, pattern = "^MT-")

clay2\_16wk\_si <- [subset](https://rdrr.io/r/base/subset.html)(clay2\_16wk\_si, subset = nFeature\_RNA > 1250 & nFeature\_RNA < 6000 & percent.mt < 15)

clay2\_16wk\_si <- SCTransform(clay2\_16wk\_si, vst.flavor = "v2", vars.to.regress = "percent.mt")

merge.list <- list(clay1\_16wk\_si, clay2\_16wk\_si)

features <- [SelectIntegrationFeatures](https://satijalab.org/seurat/reference/SelectIntegrationFeatures.html)(object.list = merge.list, nfeatures = 3004)

features <- features[-grep(pattern = "ZsGreen|TdTomato|Cre|Dre", x = features)]

merge.list <- [PrepSCTIntegration](https://satijalab.org/seurat/reference/PrepSCTIntegration.html)(object.list = merge.list, anchor.features = features)

merge.anchors <- FindIntegrationAnchors(object.list = merge.list, normalization.method = "SCT", anchor.features = features)

merge.combined <- IntegrateData(anchorset = merge.anchors, normalization.method = "SCT")

merge.combined <- [RunPCA](https://satijalab.org/seurat/reference/RunPCA.html)(merge.combined, verbose = FALSE)

[ElbowPlot](https://satijalab.org/seurat/reference/ElbowPlot.html)(merge.combined, 50)

merge.combined <- [FindNeighbors](https://satijalab.org/seurat/reference/FindNeighbors.html)(merge.combined, dims = 1:40)

merge.combined <- [FindClusters](https://satijalab.org/seurat/reference/FindClusters.html)(merge.combined, resolution = 1.5)

merge.combined <- [RunUMAP](https://satijalab.org/seurat/reference/RunUMAP.html)(merge.combined, dims = 1:40)

p2 <- [DimPlot](https://satijalab.org/seurat/reference/DimPlot.html)(merge.combined, reduction = "umap", label = TRUE, repel = TRUE, pt.size = 3)

setwd("/mnt/rstor/genetics/Wynshaw-BorisLab/lab139/analysis\_w\_newref/paper\_rerun/")

[saveRDS](https://rdrr.io/r/base/readRDS.html)(merge.combined, file = "/mnt/rstor/genetics/Wynshaw-BorisLab/lab139/analysis\_w\_newref/paper\_rerun/clay16wkSI\_combined\_beforePrepSCT\_cellbender\_ltsremoved.rds")

merge.combined.prep <- [PrepSCTFindMarkers](https://satijalab.org/seurat/reference/PrepSCTFindMarkers.html)(merge.combined)

[saveRDS](https://rdrr.io/r/base/readRDS.html)(merge.combined.prep, file = "/mnt/rstor/genetics/Wynshaw-BorisLab/lab139/analysis\_w\_newref/paper\_rerun/clay16wkSI\_combined\_afterPrepSCT\_cellbender\_ltsremoved.rds")

merge.combined.markers <- [FindAllMarkers](https://satijalab.org/seurat/reference/FindAllMarkers.html)(merge.combined.prep, assay = "SCT", only.pos = TRUE, min.pct = 0.25, logfc.threshold = 0.25)

write.csv(merge.combined.markers, file = "clay16wkSI\_combined\_markers\_cellbender\_ltsremoved.csv")

#Supplemental Table 1 = "clay16wkSI\_combined\_markers\_cellbender\_ltsremoved.csv"

pt <- table(Idents(merge.combined), merge.combined$orig.ident)

pt <- as.data.frame(pt)

pt$Var1 <- as.character(pt$Var1)

write.csv(pt, file = "clay16wkSI\_combined\_proportions\_cellbender\_ltsremoved.csv")

merge.combined <- readRDS("/mnt/rstor/genetics/Wynshaw-BorisLab/lab139/analysis\_w\_newref/paper\_rerun/clay16wkSI\_combined\_beforePrepSCT\_cellbender\_ltsremoved.rds")

merge.combined.prep <- readRDS("/mnt/rstor/genetics/Wynshaw-BorisLab/lab139/analysis\_w\_newref/paper\_rerun/clay16wkSI\_combined\_afterPrepSCT\_cellbender\_ltsremoved.rds")

[DefaultAssay](https://mojaveazure.github.io/seurat-object/reference/DefaultAssay.html)(merge.combined.prep) <- "RNA"

Idents(object = merge.combined.prep) <- "seurat\_clusters"

ul <- [subset](https://rdrr.io/r/base/subset.html)(x = merge.combined.prep, idents = c(1,3,4,7,8,15,18,19))

tdtom\_ul <- subset(x = ul, subset = TdTomato > 0)

UL\_zsgreenpos <- [subset](https://rdrr.io/r/base/subset.html)(x = tdtom\_ul, subset = ZsGreen > 0)

UL\_zsgreenneg <- [subset](https://rdrr.io/r/base/subset.html)(x = tdtom\_ul, subset = ZsGreen == 0)

UL\_zsgreenpos [["zsgreen"]] <- "pos"

UL\_zsgreenneg [["zsgreen"]] <- "neg"

zsgreen\_posneg <- merge(x = UL\_zsgreenpos, y = UL\_zsgreenneg, project = "zsgreen\_posneg", merge.data = TRUE)

Idents(object = zsgreen\_posneg) <- "zsgreen"

DefaultAssay(zsgreen\_posneg) <- "SCT"

zsgreen\_ULneuron.markers <- FindMarkers(object = zsgreen\_posneg, logfc.threshold = 0, assay = "SCT", ident.1 = "pos", ident.2 = "neg", recorrect\_umi = FALSE)

write.csv(zsgreen\_ULneuron.markers, file = "clay16wkSI\_red\_v\_yellow\_markers\_rna\_cellbender\_noLTS\_thresh0.csv")

#Supplemental Table 2 = "clay16wkSI\_red\_v\_yellow\_markers\_rna\_cellbender\_noLTS\_thresh0.csv"

[DefaultAssay](https://mojaveazure.github.io/seurat-object/reference/DefaultAssay.html)(merge.combined.prep) <- "RNA"

Idents(object = merge.combined.prep) <- "seurat\_clusters"

ul <- [subset](https://rdrr.io/r/base/subset.html)(x = merge.combined.prep, idents = c(1))

tdtom\_ul <- subset(x = ul, subset = TdTomato > 0)

UL\_zsgreenpos <- [subset](https://rdrr.io/r/base/subset.html)(x = tdtom\_ul, subset = ZsGreen > 0)

UL\_zsgreenneg <- [subset](https://rdrr.io/r/base/subset.html)(x = tdtom\_ul, subset = ZsGreen == 0)

UL\_zsgreenpos [["zsgreen"]] <- "pos"

UL\_zsgreenneg [["zsgreen"]] <- "neg"

zsgreen\_posneg <- merge(x = UL\_zsgreenpos, y = UL\_zsgreenneg, project = "zsgreen\_posneg", merge.data = TRUE)

Idents(object = zsgreen\_posneg) <- "zsgreen"

DefaultAssay(zsgreen\_posneg) <- "SCT"

zsgreen\_ULneuron.markers <- FindMarkers(object = zsgreen\_posneg, logfc.threshold = 0, assay = "SCT", ident.1 = "pos", ident.2 = "neg", recorrect\_umi = FALSE)

write.csv(zsgreen\_ULneuron.markers, file = "clay16wkSI\_red\_v\_yellow\_cluster1\_markers\_rna\_cellbender\_noLTS\_thresh0.csv")

#Supplemental Table 3 = "clay16wkSI\_red\_v\_yellow\_cluster1\_markers\_rna\_cellbender\_noLTS\_thresh0.csv"

[DefaultAssay](https://mojaveazure.github.io/seurat-object/reference/DefaultAssay.html)(merge.combined.prep) <- "RNA"

Idents(object = merge.combined.prep) <- "seurat\_clusters"

ul <- [subset](https://rdrr.io/r/base/subset.html)(x = merge.combined.prep, idents = c(1,3,4,7,8,15,18,19))

tdtom\_ul <- subset(x = ul, subset = TdTomato > 0)

set.seed(5)

UL\_zsgreenpos <- tdtom\_ul[, sample(colnames(tdtom\_ul), size = 1118, replace=F)]

toRemove <- WhichCells(UL\_zsgreenpos)

UL\_zsgreenneg <- tdtom\_ul[,!colnames(tdtom\_ul) %in% toRemove]

UL\_zsgreenpos [["zsgreen"]] <- "pos"

UL\_zsgreenneg [["zsgreen"]] <- "neg"

zsgreen\_posneg <- merge(x = UL\_zsgreenpos, y = UL\_zsgreenneg, project = "zsgreen\_posneg", merge.data = TRUE)

Idents(object = zsgreen\_posneg) <- "zsgreen"

DefaultAssay(zsgreen\_posneg) <- "SCT"

zsgreen\_ULneuron.markers <- FindMarkers(object = zsgreen\_posneg, logfc.threshold = 0, assay = "SCT", ident.1 = "pos", ident.2 = "neg", recorrect\_umi = FALSE)

write.csv(zsgreen\_ULneuron.markers, file = "clay16wkSI\_random\_ULneuron\_markers\_rna\_cellbender\_noLTS\_thresh0.csv")

#data for Supplemental Fig. 5D = "clay16wkSI\_random\_ULneuron\_markers\_rna\_cellbender\_noLTS\_thresh0.csv"