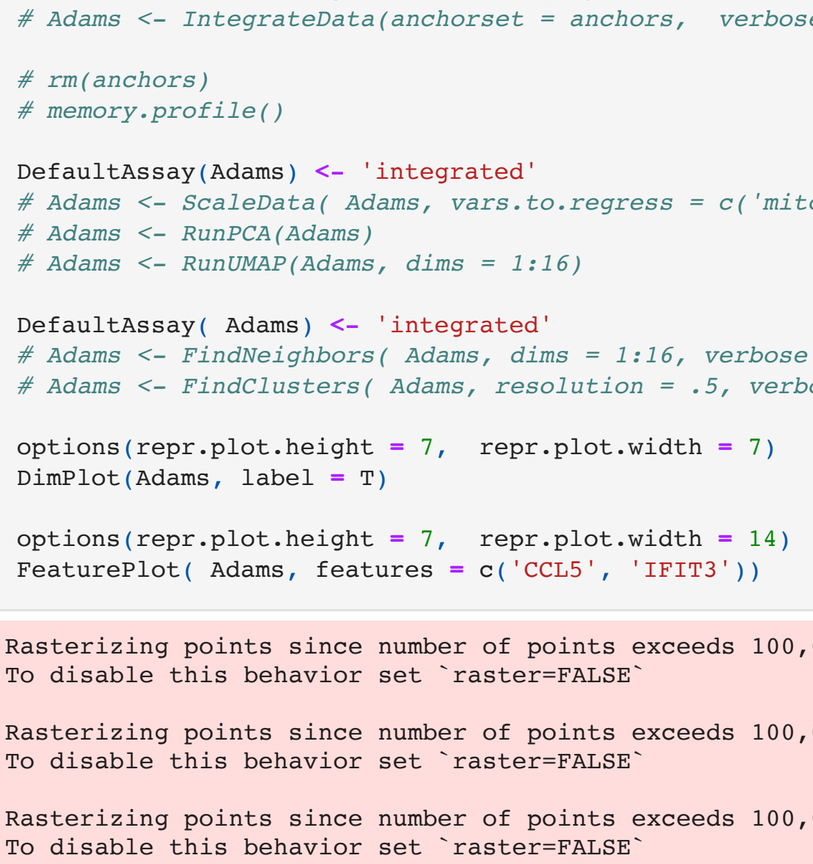
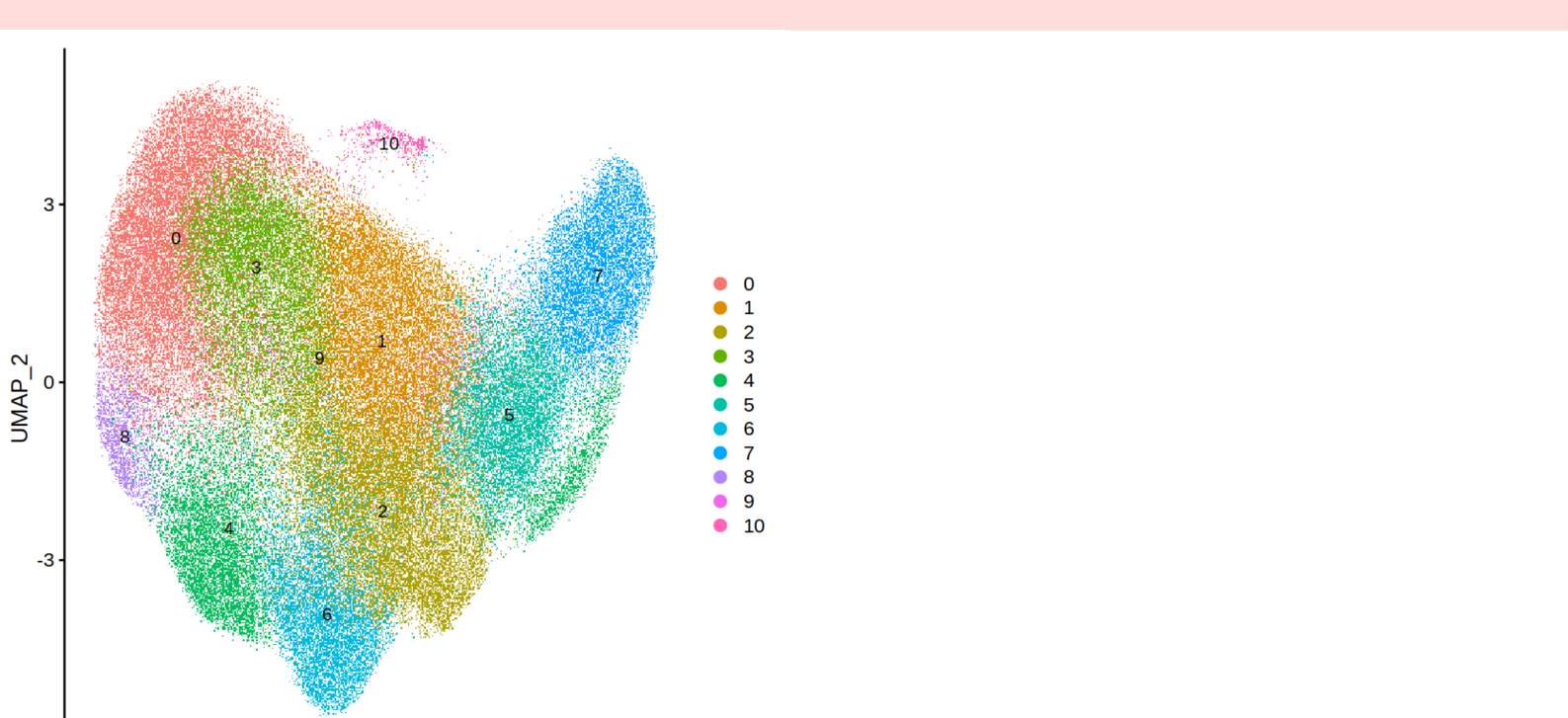
In [2]:
```
library('Seurat')
library('data.table')
library('Matrix')
library('tidyverse')
library('ggplot2')
library('ggpubr')
library('biomaRt')
library(org.Hs.eg.db)
library(stringr)
`%notin%` <- Negate(`%in%`)
library(cluster)
library(ggpubr)
library(pheatmap)
```

In [11]:
```
# counts <- readMM('/home/kevin/Data/Fibrosis/GSE136831/GSE136831_RawCounts_Sparse.mtx.gz')
lung_gene <- readRDS('lung_uni.RDS')
```

In [88]:
```
# meta.data <- data.frame(read_tsv(file = '/home/kevin/Data/Fibrosis/GSE136831/GSE136831_AllCells.Samples.CellType.MetadataTable.txt.gz'))
# barcodes <- read_tsv(file = '/home/kevin/Data/Fibrosis/GSE136831/GSE136831_AllCells.cellBarcodes.txt.gz', col_names = FALSE)
# features <- as.data.frame(read_tsv(file = '/home/kevin/Data/Fibrosis/GSE136831/GSE136831_AllCells.GeneIDs.txt.gz'))
# features <- features$HGNC_EnsemblAlt_GeneID

# colnames(counts) <- rownames(meta.data)
# rownames(counts) <- features
# counts <- counts[, meta.data$Disease_Identity != 'COPD']
# meta.data <- meta.data[meta.data$Disease_Identity != 'COPD', ]
# counts <- counts[ rowSums( counts > 0) > 0, ]
# counts <- counts[ lung_gene, ]

# counts <- counts[, meta.data$Manuscript_Identity %in% c('Macrophage_Alveolar', 'Macrophage', 'cMonocyte') ]
# meta.data <- meta.data[meta.data$Manuscript_Identity %in% c('Macrophage_Alveolar', 'Macrophage', 'cMonocyte'), ]

# counts <- counts[, meta.data$Subclass_Cell_Identity %in% c('Macrophage', 'Macrophage_Alveolar', 'Monocyte')]
# meta.data <- meta.data[ meta.data$Subclass_Cell_Identity %in% c('Macrophage', 'Macrophage_Alveolar', 'Monocyte'),]

# Adams <- CreateSeuratObject( counts = counts, min.cells = 10, min.features = 300)
# Adams$mito.ratio <- PercentageFeatureSet( Adams, pattern = 'MT-') / 100
# Adams$heam.ratio <- (PercentageFeatureSet( Adams, pattern = 'HBB') + PercentageFeatureSet( Adams, pattern = 'HBA')) / 100
# Adams$ribo.ratio <- (PercentageFeatureSet( Adams, pattern = 'RPL') + PercentageFeatureSet( Adams, pattern = 'RPS')) / 100
# Adams <- Adams[, Adams$heam.ratio < .001 & Adams$ribo.ratio < .3 & Adams$mito.ratio < .15]
# Adams <- Adams[, Adams$nFeature_RNA < 5000 & Adams$nFeature_RNA > 300]
# Adams@meta.data$orig.ident <- meta.data[ colnames(Adams), 'Subject_Identity']
# Adams@meta.data <- cbind( Adams@meta.data, meta.data[colnames(Adams),  c('Disease_Identity', 'Subject_Identity', 'Manuscript_Identity')])
# Adams <- Adams[, Adams$orig.ident != '244C']

# Adams <- SplitObject( Adams, split.by = 'Disease_Identity')

# for(i in 1: length(Adams)) {
#     Adams[[i]] <- NormalizeData(Adams[[i]])
#     Adams[[i]] <- FindVariableFeatures(Adams[[i]] )
# }

# features <- SelectIntegrationFeatures(object.list = Adams, nfeatures = 2000)
# anchors <- FindIntegrationAnchors(object.list = Adams, anchor.features = features, verbose = T)
# Adams <- IntegrateData(anchorset = anchors,  verbose = T)

# rm(anchors)
# memory.profile()

DefaultAssay(Adams) <- 'integrated'
# Adams <- ScaleData( Adams, vars.to.regress = c('mito.ratio', 'nFeature_RNA', 'orig.ident'))
# Adams <- RunPCA(Adams)
# Adams <- RunUMAP(Adams, dims = 1:16)

DefaultAssay( Adams) <- 'integrated'
# Adams <- FindNeighbors( Adams, dims = 1:16, verbose = F)
# Adams <- FindClusters( Adams, resolution = .5, verbose = F)

options(repr.plot.height = 7,  repr.plot.width = 7)
DimPlot(Adams, label = T)

options(repr.plot.height = 7,  repr.plot.width = 14)
FeaturePlot( Adams, features = c('CCL5', 'IFIT3'))
```

```
Rasterizing points since number of points exceeds 100,000.
To disable this behavior set `raster=FALSE`

Rasterizing points since number of points exceeds 100,000.
To disable this behavior set `raster=FALSE`

Rasterizing points since number of points exceeds 100,000.
To disable this behavior set `raster=FALSE`
```





In [91]:
```
# Adams <- Adams[, Idents(Adams) %notin% c(9, 10)]
saveRDS(Adams, '/mnt/storage/projects/fibrosis_kevin/Adams_macrophage_Mar_4.RDS')
```