Results UMAP

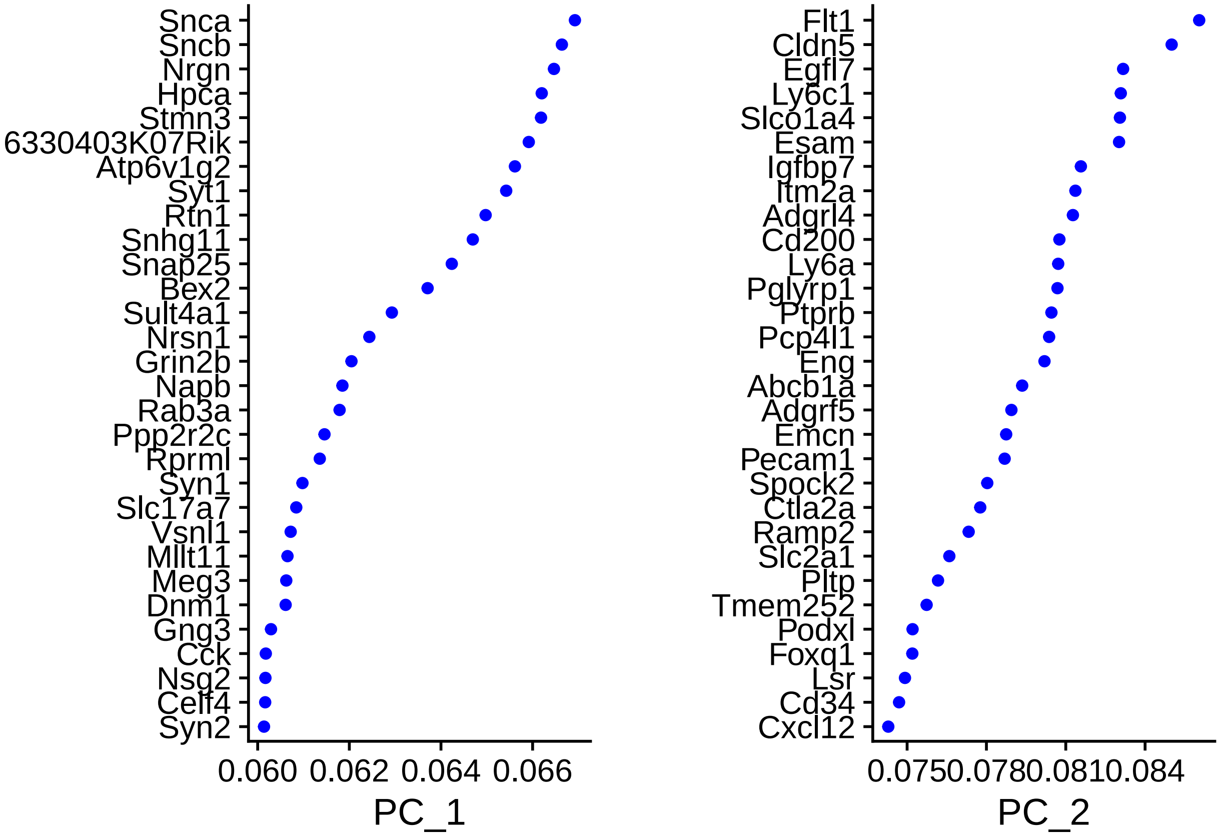
programandoconro

7/19/2019

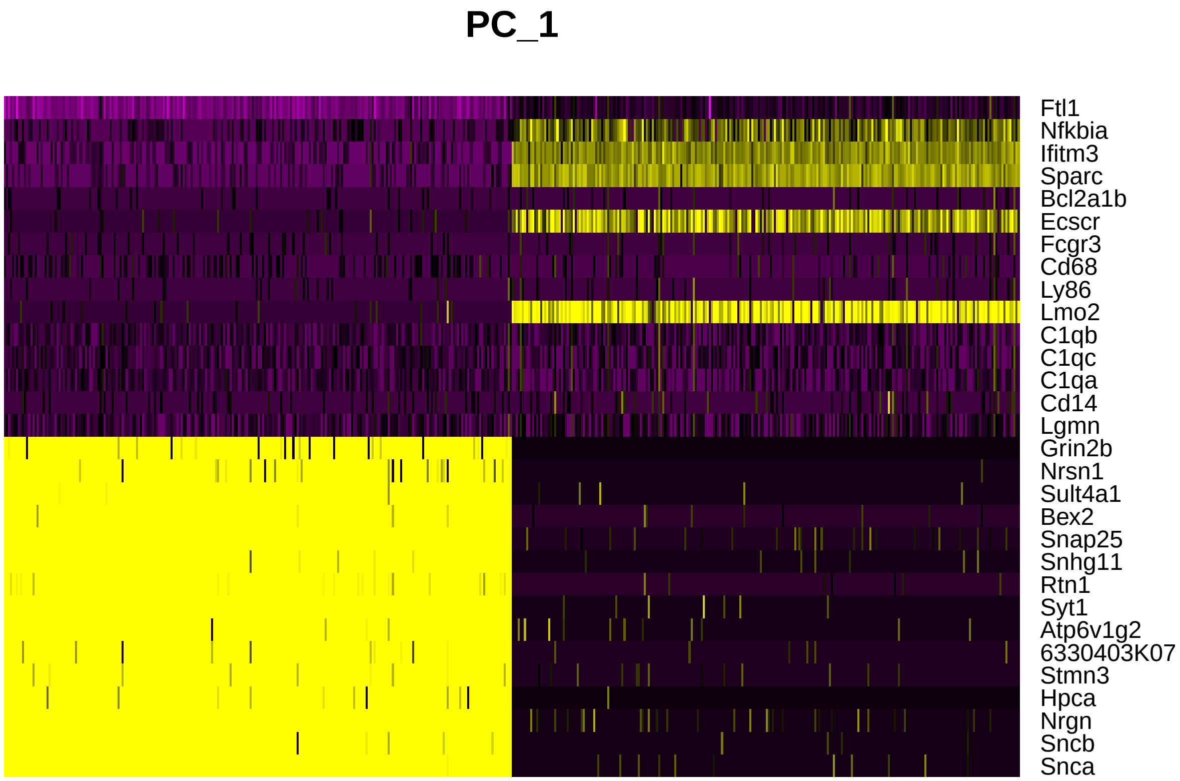
```
  ################# Requiered packages ##########################################

  library(dplyr)
  library(Seurat)
  library(cowplot)

  ############################# Exp_1 #################################################

    R048_1 <- Read10X(data.dir = "~/Dropbox/DataScience/Fiver/R048_1/")%>%
            CreateSeuratObject(min.cells = 3, min.features = 200,project = "R048 Exp.1",names.field

 # FeatureScatter(object = R048_1, feature1 = "nCount_RNA", feature2 = "nFeature_RNA")

  DMSO_1 <- Read10X(data.dir = "~/Dropbox/DataScience/Fiver/DMSO_1/")%>%
  CreateSeuratObject(min.cells = 3, min.features = 200,project = "DMSO Exp.1",names.field = 3)

 # FeatureScatter(object = DMSO_1, feature1 = "nCount_RNA", feature2 = "nFeature_RNA")

  mix_1 <- merge (x= DMSO_1,y=R048_1)%>%
    NormalizeData(verbose = FALSE)%>%
    FindVariableFeatures(selection.method = "vst", nfeatures = 2000)

 #FeatureScatter(object = mix_1, feature1 = "nCount_RNA", feature2 = "nFeature_RNA",pt.size = 1,smo
# FeatureScatter(object = mix_1, feature1 = "nCount_RNA", feature2 = "nFeature_RNA",pt.size = 1,smo
  rm(R048_1,DMSO_1)#remove objects to reduce RAM load

  ######################## PCA exp_1 #################################################

  all.genes <- rownames(x = mix_1)

  mix_1 <- ScaleData(object = mix_1, features = all.genes)
  mix_1 <- RunPCA(object = mix_1, features = VariableFeatures(object = mix_1))
```

```
## PC_ 1
## Positive:  Snca, Sncb, Nrgn, Hpca, Stmn3, 6330403K07Rik, Atp6v1g2, Syt1, Rtn1, Snhg11
##     Snap25, Bex2, Sult4a1, Nrsn1, Grin2b, Napb, Rab3a, Ppp2r2c, Rprml, Syn1
##     Slc17a7, Vsnl1, Mllt11, Meg3, Dnm1, Gng3, Cck, Nsg2, Celf4, Syn2
## Negative:  Ftl1, Nfkbia, Ifitm3, Sparc, Bcl2a1b, Ecscr, Fcgr3, Cd68, Ly86, Lmo2
##     C1qb, C1qc, C1qa, Cd14, Lgmn, Hexb, Ifi30, Ctsd, Plin2, BC028528
##     Selplg, Fcrls, Anxa2, Ms4a6d, Lilrb4a, Lyz2, Dab2, C5ar1, Pim1, Atf3
## PC_ 2
## Positive:  Flt1, Cldn5, Egfl7, Ly6c1, Slco1a4, Esam, Igfbp7, Itm2a, Adgrl4, Cd200
##     Ly6a, Pglyrp1, Ptprb, Pcp4l1, Eng, Abcb1a, Adgrf5, Emcn, Pecam1, Spock2
##     Ctla2a, Ramp2, Slc2a1, Pltp, Tmem252, Podxl, Foxq1, Lsr, Cd34, Cxcl12
## Negative:  Mt1, Lgals1, Ctsd, Mt3, Gfap, Apoe, Aldoc, Fabp5, Nupr1, C1qb
##     C1qc, C1qa, Cd68, Ly86, Tubb2b, Fcgr3, Bcl2a1b, Ndrg2, Fabp7, Prdx6
##     Clu, Slc1a2, S100b, Sdc4, Cd9, Lgmn, Fcrls, Ctsb, Dbi, Ifi30
## PC_ 3
## Positive:  S100a16, Ptn, Dbi, Fabp7, Tubb2b, Gpr37l1, Atp1a2, Ptprz1, Mt3, Scd2
##     S100a1, Gpm6b, S100b, Ndrg2, Aldoc, Atp1b2, Cspg5, Clu, Cnn3, Olig1
##     Tuba1a, Serpina3n, Cpe, Plpp3, Gfap, Slc1a2, Mmd2, Chst2, Htra1, Luzp2
## Negative:  Ly86, Cd68, Fcgr3, C1qb, C1qc, Bcl2a1b, C1qa, Lgmn, Hexb, Ftl1
##     Cd14, Fcrls, Ifi30, Cd83, Selplg, Mef2c, Timp2, Ms4a6d, Cx3cr1, Lilrb4a
##     C5ar1, Tlr2, Lyz2, Gpr34, Ccl3, Fcgr1, Tmem119, Ccrl2, Rgs1, Ccl4
## PC_ 4
## Positive:  Gfap, Chchd10, Slc1a2, Clu, Aldoc, Mt3, Gstm1, Fxyd1, Nupr1, Cpe
##     Igfbp2, Fam107a, Acsbg1, Aldh1l1, Mlc1, Rbp1, Slc25a18, Apoe, Tst, Sparcl1
##     Aqp4, Prdx6, Gja1, Slc4a4, Timp1, Glul, Mrps6, Fgfr3, Ndrg2, Lcn2
## Negative:  Pdgfra, Lhfpl3, C1ql1, Sox10, Qpct, Cacng4, Tnr, Pllp, Plppr5, Matn4
##     Pbk, 1700086L19Rik, Hmgn2, Gpr17, Nxph1, Lockd, Vcan, Sulf2, Olig2, Marcksl1
##     2810417H13Rik, Bcas1, Cdca3, Lims2, Top2a, Ugdh, Dpysl3, H2afz, Lrrn1, Sema3d
## PC_ 5
## Positive:  Top2a, 2810417H13Rik, Birc5, Pbk, Ccna2, Cdca3, Mki67, Cks1b, Tubb5, Prc1
##     Nusap1, Cdca8, Ube2c, Spc25, Cdk1, Racgap1, Smc2, Tpx2, Smc4, Cenpf
##     F3, Slc1a3, Tuba1b, Rrm2, Ccnb1, Spc24, Hmmr, Knstrn, Cenpe, Mt2
## Negative:  Mog, Opalin, Nkx6-2, Ermn, Mag, Cldn11, Plp1, Klk6, Ptgds, Mobp
##     Gjb1, Fa2h, Tmem88b, Tspan2, Ppp1r14a, Tmem125, Aspa, Gjc2, Efnb3, Tubb4a
##     Ugt8a, Hapln2, Mbp, Mal, Apod, Tnni1, Lgi3, Grb14, Tmeff2, Stmn4
```

```
VizDimLoadings(object = mix_1, dims = 1:2, reduction = 'pca')
```



```
# DimPlot(object = mix_1, reduction = 'pca')

 DimHeatmap(object =mix_1, dims = 1, cells = 500, balanced = TRUE)
```

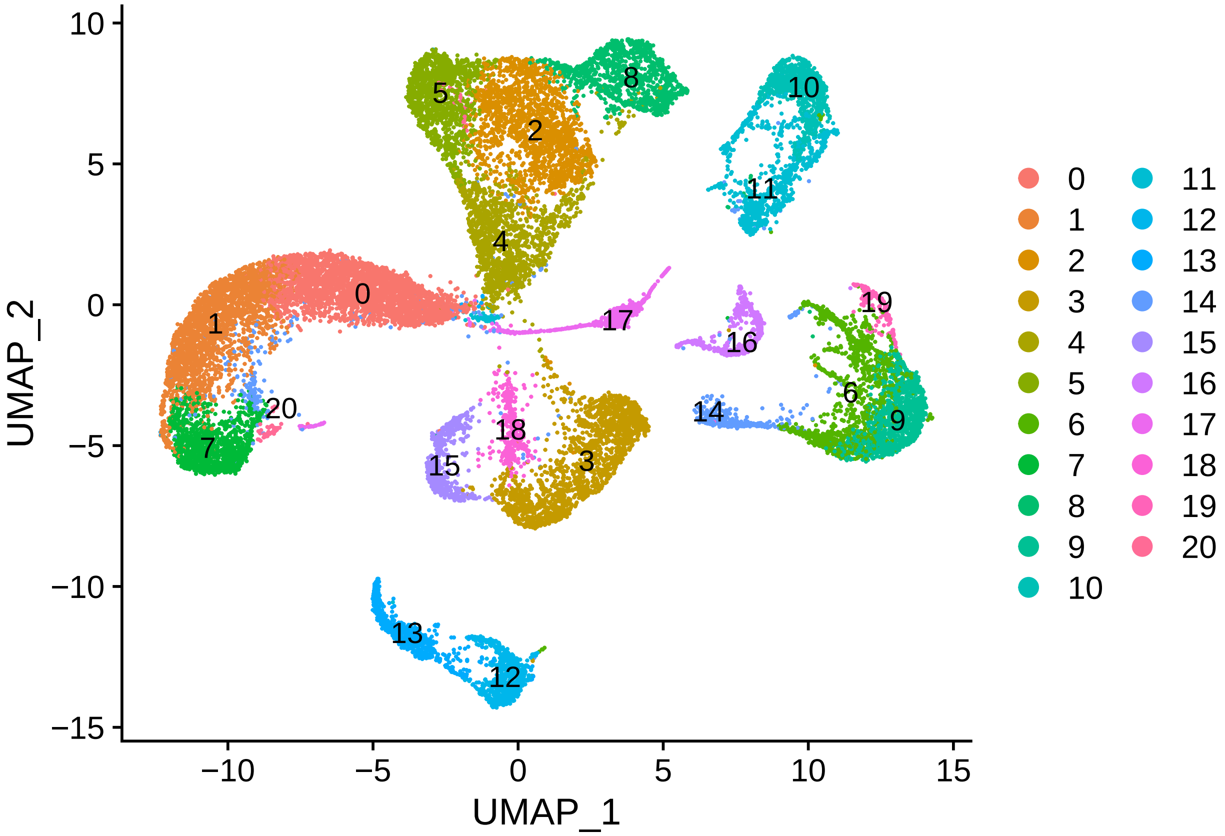


```
    # DimHeatmap(object =mix_1, dims = 1:15, cells = 500, balanced = TRUE)

     # ElbowPlot(object = mix_1)

      ################### Clusters exp_1 ################################################

      mix_1 <- FindNeighbors(object = mix_1, dims = 1:10)
      mix_1 <- FindClusters(object = mix_1, resolution = 0.5)
```

```
## Modularity Optimizer version 1.3.0 by Ludo Waltman and Nees Jan van Eck
##
## Number of nodes: 22563
## Number of edges: 736253
##
## Running Louvain algorithm...
## Maximum modularity in 10 random starts: 0.9422
## Number of communities: 21
## Elapsed time: 3 seconds
```

```
      # find all markers of cluster 1

       #  cluster1.markers <- FindMarkers(object = mix_1, ident.1 = 1, min.pct = 0.25)

      # find all markers distinguishing cluster 5 from clusters 0 and 3

       # cluster5.markers <- FindMarkers(object = mix_1, ident.1 = 5, ident.2 = c(0, 3), min.pct = 0.25

      # find markers for every cluster compared to all remaining cells, report only the positive ones
     # pbmc.markers <- FindAllMarkers(object = mix_1, only.pos = TRUE, min.pct = 0.25, logfc.threshold
     # pbmc.markers %>% group_by(cluster) %>% top_n(n = 2, wt = avg_logFC)
```

```
####################### UMAP exp_1 ######################################################
 #Need to install pip on Mac and run in console: sudo apt-get install pip
 #Install the package with: pip install umap-learn

 mix_1 <- RunUMAP(object = mix_1, dims = 1:10)

  DimPlot(object = mix_1, reduction = 'umap',label = T)
```

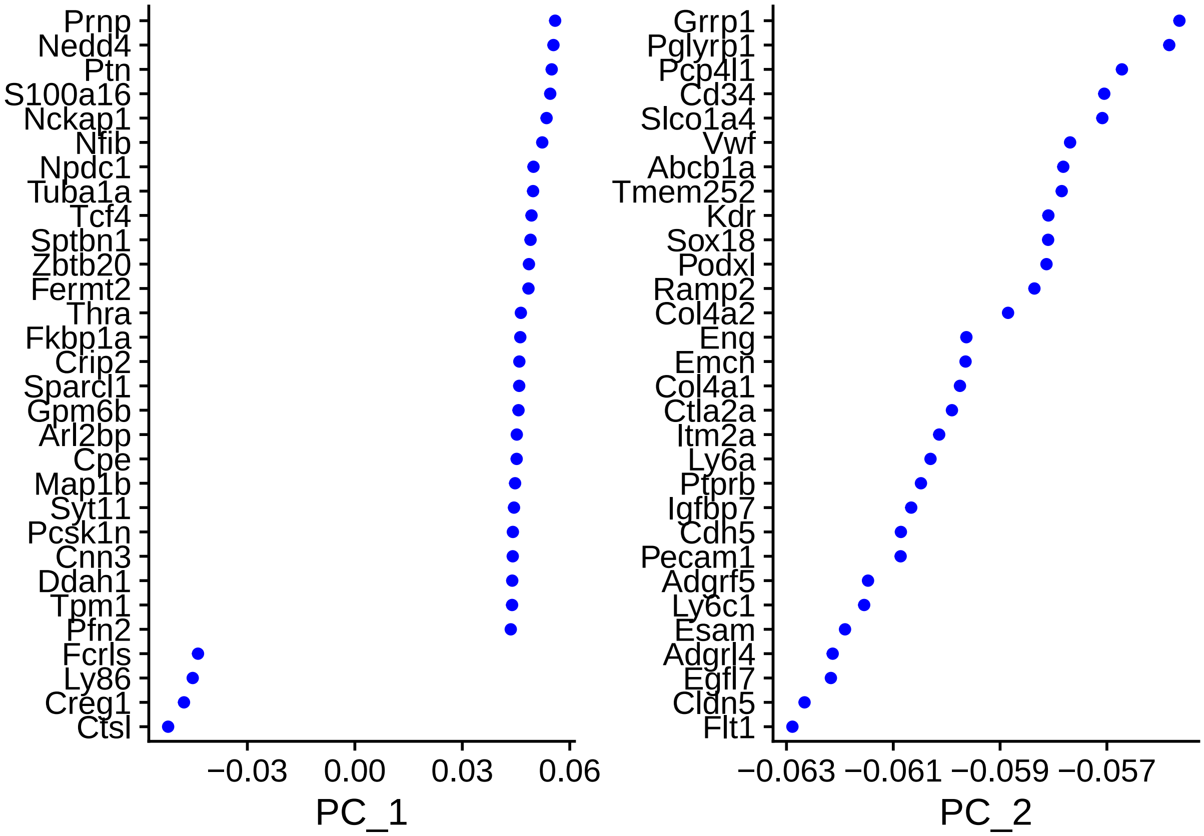


```
 ##########################################################################################
```

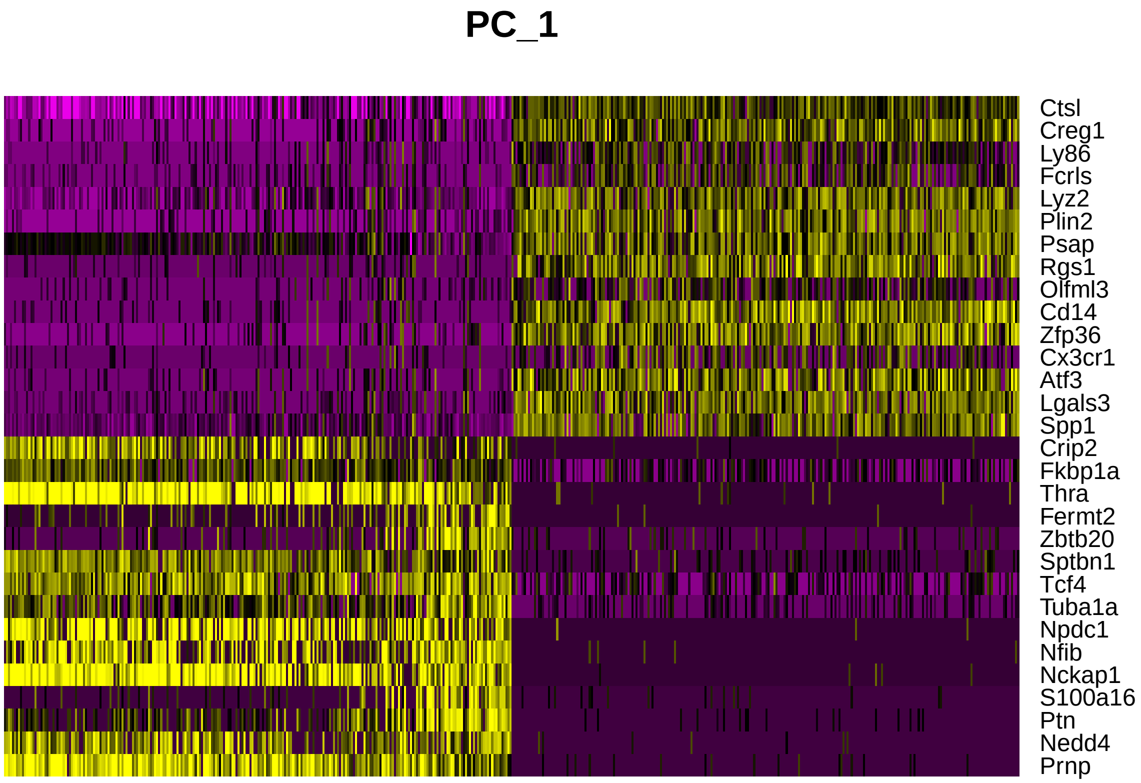
```
    rm(mix_1,all.genes)
   #############################  Exp_2 ################################################

     R048_2 <- Read10X(data.dir = "~/Dropbox/DataScience/Fiver/R048_2/")%>%
             CreateSeuratObject(min.cells = 30, min.features = 2000,project = "R048 Exp.2")

  # FeatureScatter(object = R048_2, feature1 = "nCount_RNA", feature2 = "nFeature_RNA")

   DMSO_2 <- Read10X(data.dir = "~/Dropbox/DataScience/Fiver/DMSO_2/")%>%
   CreateSeuratObject(min.cells = 30, min.features = 2000,project = "DMSO Exp.2")

  # FeatureScatter(object = DMSO_2, feature1 = "nCount_RNA", feature2 = "nFeature_RNA")

   mix_2 <- merge (x= DMSO_2,y=R048_2)%>%
     NormalizeData(verbose = FALSE)%>%
     FindVariableFeatures(selection.method = "vst", nfeatures = 2000)

  # FeatureScatter(object = mix_2, feature1 = "nCount_RNA", feature2 = "nFeature_RNA",pt.size = 1,sm
# FeatureScatter(object = mix_2, feature1 = "nCount_RNA", feature2 = "nFeature_RNA",pt.size = 1,sm
 rm(RO48_2,DMSO_2)#remove objects to reduce RAM load

 ######################### PCA exp_2 #################################################

 all.genes <- rownames(x = mix_2)

 mix_2 <- ScaleData(object = mix_2, features = all.genes)
 rm(all.genes)

 mix_2 <- RunPCA(object = mix_2, features = VariableFeatures(object = mix_2))
```

```
## PC_ 1
## Positive:  Prnp, Nedd4, Ptn, S100a16, Nckap1, Nfib, Npdc1, Tuba1a, Tcf4, Sptbn1
##     Zbtb20, Fermt2, Thra, Fkbp1a, Crip2, Sparcl1, Gpm6b, Arl2bp, Cpe, Map1b
##     Syt11, Pcsk1n, Cnn3, Ddah1, Tpm1, Pfn2, Tsc22d1, Scd2, Selenom, Uchl1
## Negative:  Ctsl, Creg1, Ly86, Fcrls, Lyz2, Plin2, Psap, Rgs1, Olfml3, Cd14
##     Zfp36, Cx3cr1, Atf3, Lgals3, Spp1, Wfdc17, Msr1, Id2, Apoe, Ccrl2
##     Tgfbi, Tmem119, Ccl3, Rgs2, Ifi27l2a, Ccl4, Cd83, Npl, Ms4a7, Gpr34
## PC_ 2
## Positive:  Syt11, Ckb, Gpm6b, Pcsk1n, Anks1b, Mapt, Kif1a, Ank2, Gnao1, Ank3
##     Olig1, Aplp1, Fez1, Pcdh9, Ncam1, Shisa4, Clip3, Elavl3, Fam171b, Tmod2
##     Sv2a, Serpina3n, Trim2, Pfn2, Slc24a2, Cnp, Tubb4a, Ncald, Phyhipl, Stmn4
## Negative:  Flt1, Cldn5, Egfl7, Adgrl4, Esam, Ly6c1, Adgrf5, Pecam1, Cdh5, Igfbp7
##     Ptprb, Ly6a, Itm2a, Ctla2a, Col4a1, Emcn, Eng, Col4a2, Ramp2, Podxl
##     Sox18, Kdr, Tmem252, Abcb1a, Vwf, Slco1a4, Cd34, Pcp4l1, Pglyrp1, Grrp1
## PC_ 3
## Positive:  Syt1, Slc17a7, Snap25, Snhg11, Dnm1, Myt1l, Grin2b, Camk2b, Syn2, Adcy1
##     Ndrg4, Scn2a, Grin1, Slc6a17, Sv2b, Celf4, Atp6v1g2, Plppr4, Rasgrp1, Thy1
##     Hpcal4, Nrxn3, Slc4a10, Nptx1, Sult4a1, Syp, Pde1a, Gabra1, Pcsk2, Vsnl1
## Negative:  Mog, Mag, Cldn11, Nkx6-2, Ermn, Mobp, Ppp1r14a, Plp1, Tmem88b, Tspan2
##     Car2, Opalin, Mal, Ugt8a, Sept4, Efnb3, Gpr37, Cnp, Mbp, Apod
##     Trf, Enpp2, Aspa, Gjb1, Gjc2, Fa2h, Lgi3, Grb14, Tubb4a, Klk6
## PC_ 4
## Positive:  Slc24a2, Syt1, Slc17a7, Spock1, Dnm1, Grin2b, Snap25, Tmem151a, Sv2b, Srcin1
##     Snhg11, Hpcal4, Adcy1, Dnajc6, Grin1, Syp, Ppp2r2c, Sv2a, Eno2, Nrgn
##     Rbfox3, Myt1l, Camk2a, Gabra1, Napb, Camk2b, Snca, Sult4a1, Pacsin1, Nptx1
## Negative:  Ptprz1, Fabp7, Gpr37l1, Tubb2b, Atp1b2, Cspg5, Atp1a2, F3, Gm3764, Slc1a3
##     Tril, Luzp2, Mt3, Ndrg2, Timp4, Ednrb, Aqp4, Mlc1, Tmem47, Scrg1
##     Slc4a4, Acsbg1, Fgfr3, Mmd2, Mdk, Igfbp2, Sox9, Clu, Ntrk2, Mtss1l
## PC_ 5
## Positive:  Pla2g7, Prdx6, Mgst1, Chchd10, Gstm1, Fabp5, Slc4a4, Mlc1, Apoe, Psap
##     Aqp4, Clu, Gpnmb, Thbs1, Slc6a11, Aldoc, Fxyd1, Ecm1, Vegfa, Fgfr3
##     Clec4d, Lgals3, Lyz2, Gjb6, Aldh1l1, Cldn10, Acsl6, Fam107a, Acsbg1, Gfap
## Negative:  Pclaf, Birc5, Top2a, Mki67, Tubb5, Stmn1, Cdk1, Pbk, Ube2c, Ccna2
##     Cdca3, H2afz, Cdca8, Prc1, Tuba1b, Smc2, Tpx2, Cenpf, Racgap1, Cks1b
##     Cks2, Selenoh, Ccnb1, Tmpo, Spc24, Nusap1, Rrm2, Hmgb2, H2afx, Tk1
```

```
VizDimLoadings(object = mix_2, dims = 1:2, reduction = 'pca')
```



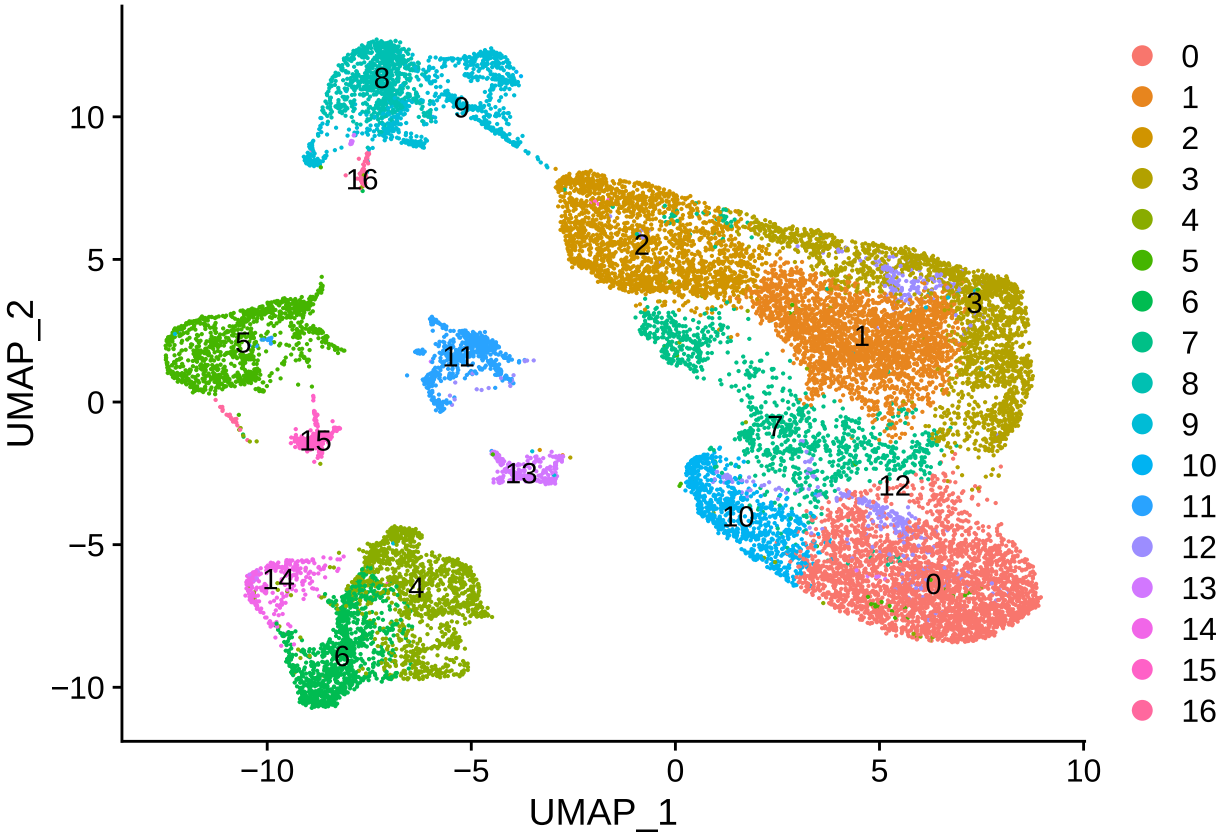
```
# DimPlot(object = mix_2, reduction = 'pca')

DimHeatmap(object =mix_2, dims = 1, cells = 500, balanced = TRUE)
```



```
#  DimHeatmap(object =mix_2, dims = 1:15, cells = 500, balanced = TRUE)

#  ElbowPlot(object = mix_2)

  ################### Clusters exp_2 #################################################

  mix_2 <- FindNeighbors(object = mix_2, dims = 1:10)
  mix_2 <- FindClusters(object = mix_2, resolution = 0.5)
```

```
## Modularity Optimizer version 1.3.0 by Ludo Waltman and Nees Jan van Eck
##
## Number of nodes: 19093
## Number of edges: 603438
##
## Running Louvain algorithm...
## Maximum modularity in 10 random starts: 0.9207
## Number of communities: 17
## Elapsed time: 2 seconds
```

```
  ####################### UMAP exp_2 ############################################################
  #pip install umap-learn

  mix_2 <- RunUMAP(object = mix_2, dims = 1:10)
  DimPlot(object = mix_2, reduction = 'umap',label = T)
```



```
  ######################### cluster biomarkers exp_2 #######################################

  # find all markers of cluster 1

   #cluster1.markers <- FindMarkers(object = mix_2, ident.1 = 1, min.pct = 0.25)

  # find all markers distinguishing cluster 5 from clusters 0 and 3

   #cluster5.markers <- FindMarkers(object = mix_2, ident.1 = 5, ident.2 = c(0, 3), min.pct = 0.25)

  # find markers for every cluster compared to all remaining cells, report only the positive ones

  #  pbmc.markers <- FindAllMarkers(object = mix_2, only.pos = TRUE, min.pct = 0.25, logfc.threshol
#  pbmc.markers %>% group_by(cluster) %>% top_n(n = 2, wt = avg_logFC)
```