Seurat Analysis

Performed using Galaxy

2019-10-09

```
## [1] "Minimum cells: 3"
```

```
## [1] "Minimum features: 200"
```

```
## [1] "Umi low threshold: 1"
```

```
## [1] "Umi high threshold: 20000000"
```
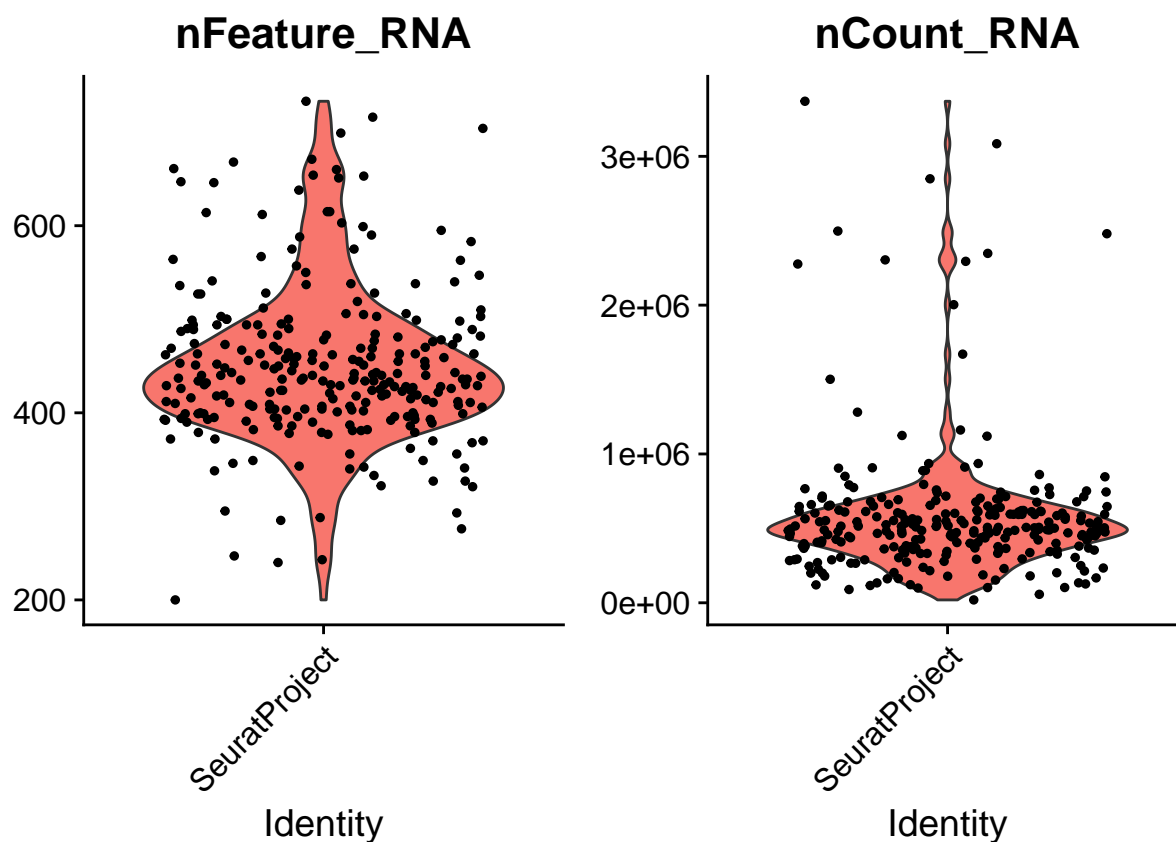
```
## [1] "Read in data, generate inital Seurat object"
```
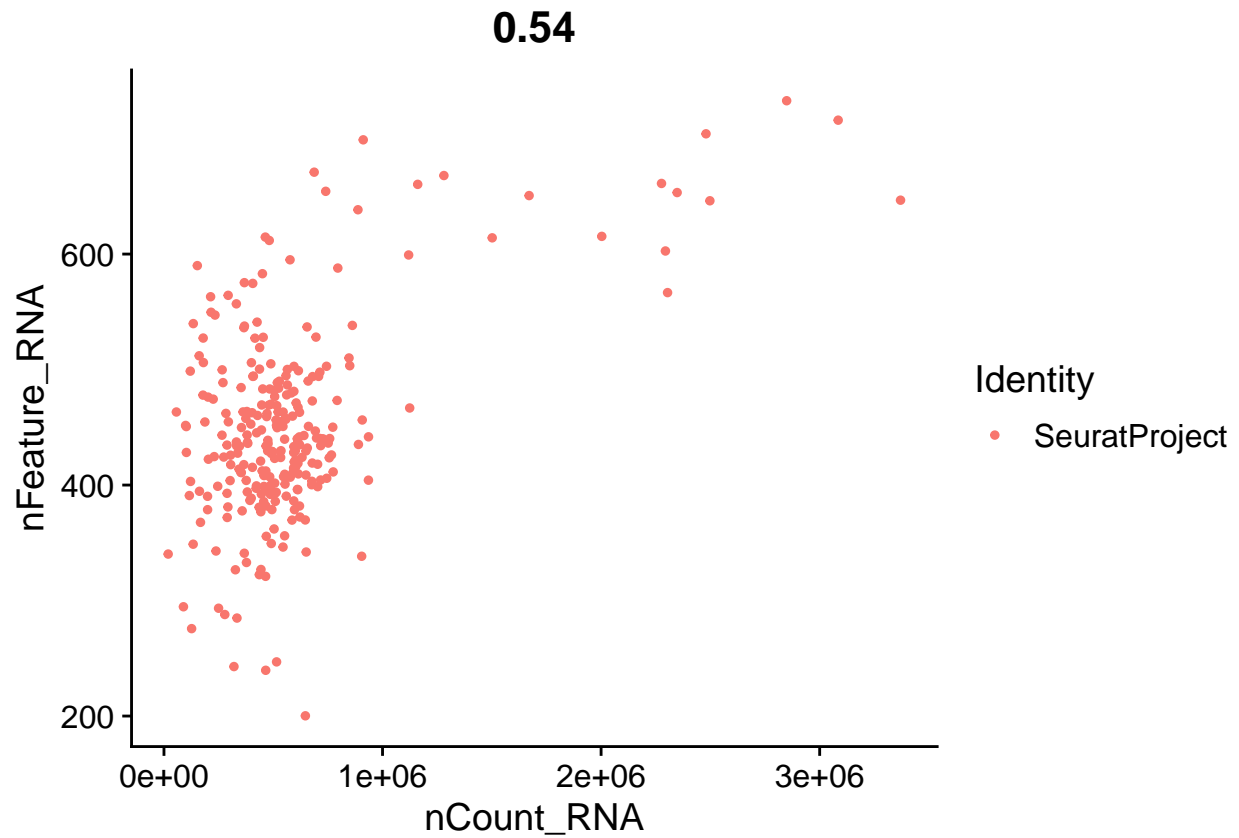
```
counts <- read.delim(params$counts, row.names=1)
seuset <- Seurat::CreateSeuratObject(counts = counts, min.cells = min_cells, min.features = min_genes)
```

```
## [1] "Raw data vizualization"
```

```
Seurat::VlnPlot(object = seuset, features = c("nFeature_RNA", "nCount_RNA"), axis="v")
```

```
Seurat::FeatureScatter(object = seuset, feature1 = "nCount_RNA", feature2 = "nFeature_RNA")
```

0.54



```
## [1] "Filter and normalize for UMI counts"
```

```
seuset <- subset(seuset, subset = `nCount_RNA` > low_thresholds & `nCount_RNA` < high_thresholds)
seuset <- Seurat::NormalizeData(seuset, normalizeation.method = "LogNormalize", scale.factor = 10000)
```

```
## [1] "Variable Genes"
```

```
seuset <- Seurat::FindVariableFeatures(object = seuset, selection.method = "mvp")
Seurat::VariableFeaturePlot(seuset, cols = c("black", "red"), selection.method = "disp")
```

Non−variable count: 822
Variable count: 105
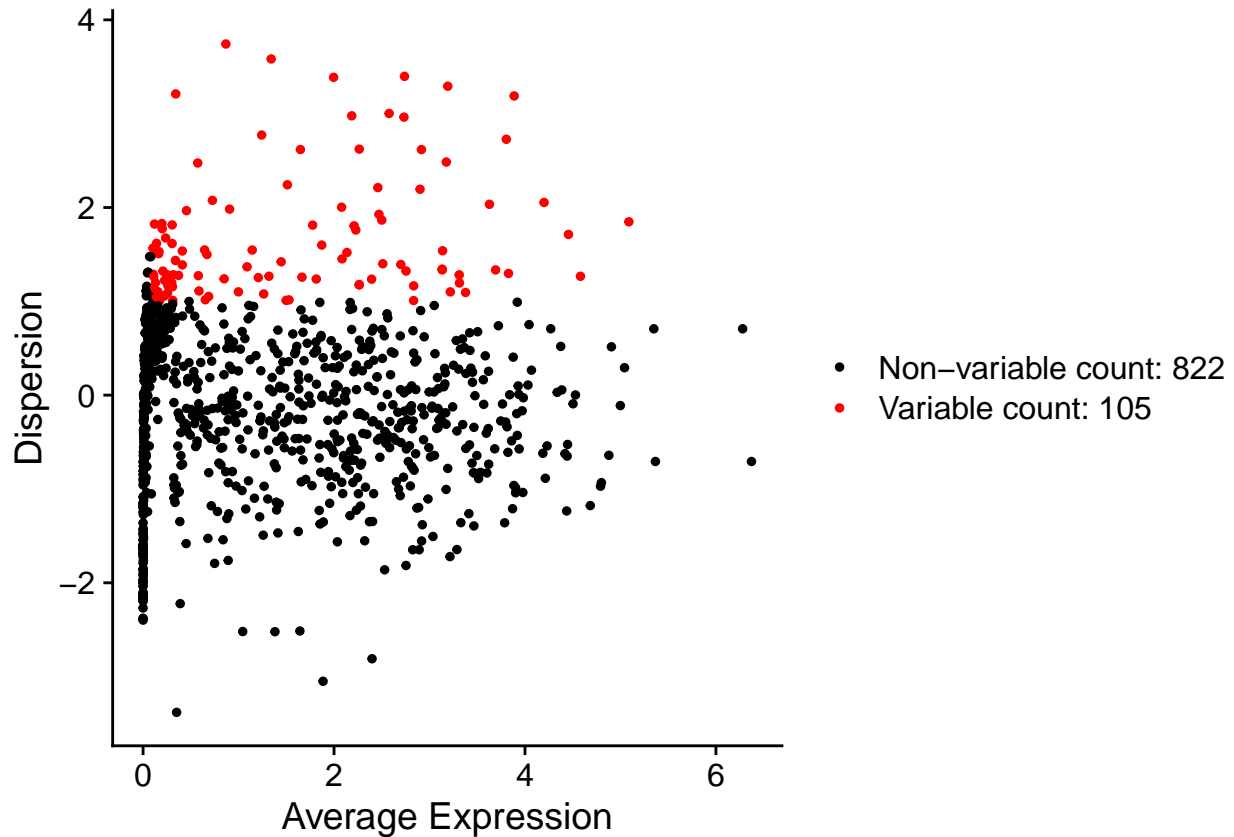
```r
seuset <- Seurat::ScaleData(object = seuset, vars.to.regress = "nCount_RNA")
```

```
## Regressing out nCount_RNA
```

```
## Centering and scaling data matrix
```

```
## [1] "PCA Visualization"
```

```r
seuset <- Seurat::RunPCA(seuset, npcs=10)
```

```
## PC_ 1
## Positive:  AA619741, St6gal1, Nlrp9b, Pkd2l2, Jazf1, Nlrp2, Siah2, Zfp735, Nanos2, Bnc2
##     Trp63, Flrt3, Dnajb4, Ctnna3, Rab3d, Dnahc8, Tesc, Kbtbd8, Smarca2, Casp8
##     4930503L19Rik, Angel2, Gng12, Tnfaip8, Epc1, Gstt1, Wdr20a, Laptm4b, Cldn15, Mthfd2
## Negative:  Tmsb10, Enpep, Capg, Gcdh, Tcn2, Cwf19l1, Tor1b, Dapk1, Ccm2, Faah
##     Ahsg, Cd109, Irak4, 6030422M02Rik, Slc18a1, Slc20a1, Hspa5, Cyp4f39, Slamf9, Serpina3m
##     Gabra1, 1700019A02Rik, Gm4956, Itih5, Marveld1, BC049762, 9430060I03Rik, Car14, Spin4, 4930563E2
## PC_ 2
## Positive:  Enpep, Capg, Tjp2, Tmsb10, Slc20a1, Dapk1, Rbl1, Faah, Manba, Arfgef1
##     Tcn2, Ccdc50, Laptm4b, Gng12, Cyp4f39, Tnfaip8, Itih5, Slamf9, Casp8, Slc18a1
##     Dpys, Dnajb4, Mfsd9, Serpina3m, Cd109, Esrp1, Flrt3, Tesc, Tmem51, Ccm2
## Negative:  Obox6, Gm13718, Larp4, Tor1b, Ypel5, Tekt2, Fem1c, Gm4850, Tnfrsf19, Kbtbd8
##     Rab3d, Arid5a, Il22ra2, Nags, Duoxa2, Gatad2a, Trpc6, Marveld1, Gm16515, 4930563E22Rik
##     Wasl, Spata2L, Bnc2, Smurf2, Siah2, Jazf1, Trim40, Tbcel, Myh14, Dnahc8
```
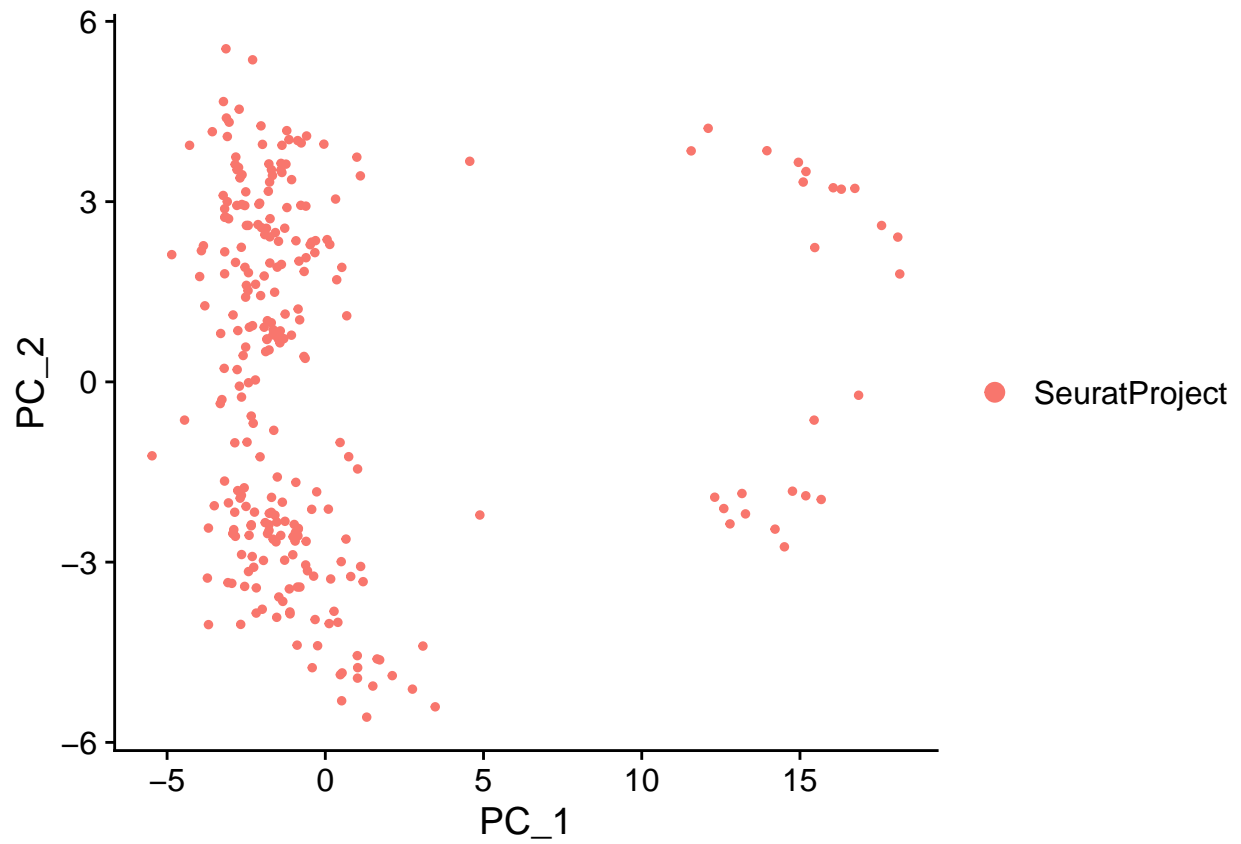
```
## PC_ 3
## Positive:  Casp8, Mfsd9, Dnajb4, Ypel5, Cldn15, Manba, Ctnna3, Trp63, Nanos2, Dnahc8
##      Trim40, Gstt1, Zfp735, St6gal1, Larp4, Mthfd2, Nlrp9b, Sap130, Myh14, Flrt3
##      9130024F11Rik, Gcdh, Zbtb40, Calb1, AA619741, Cwf19l1, F3, Gabra1, Irak4, Tnfaip8
## Negative:  Arid5a, Gm4850, Hspa5, Duoxa2, Fem1c, Arfgef1, Trpc6, Dapk1, Gm13718, Smurf2
##      Tor1b, Ccdc50, Capg, Tjp2, Tmsb10, Faah, Serpina3g, Wasl, Gng12, Enpep
##      Slc18a1, Spata2L, Gm16515, Rbp1, Siah2, Tcn2, Gatad2a, Nags, Cyp4f39, Rbl1
## PC_ 4
## Positive:  Car14, 1700019A02Rik, Serpina3m, 6030422M02Rik, Trpc6, Cd109, Gabra1, Gcdh, F3, 4930563E2
##      Ahsg, Zfp811, Hspa5, Sap130, Tsc22d1, Gm4850, Nags, Duoxa2, Mpst, Flrt3
##      Fem1c, Smarca2, Arfgef1, Tnfaip8, Jazf1, Mthfd2, Bnc2, Tnfrsf19, Manba, Spin4
## Negative:  Itih5, Ebi3, Rbp1, Slc18a1, Serpina3g, Mef2c, Faah, Cwf19l1, Larp4, Tbcel
##      Gm16515, Casp8, Ccm2, Tekt2, Slamf9, Gm13718, Cyp4f39, Rbl1, Col10a1, Zbtb40
##      BC049762, Dapk1, Obox6, 9130024F11Rik, Lct, Dnajb4, Irak4, Mfsd9, Esrp1, Gm4956
## PC_ 5
## Positive:  Slamf9, Esrp1, Gm4956, Tjp2, Enpep, Wasl, Dpys, Faah, Col10a1, Gm13718
##      Tmsb10, 9130024F11Rik, Zfp735, Cyp4f39, Siah2, Gm16515, Tesc, Kbtbd8, Calb1, Dapk1
##      Rab3d, AA619741, St6gal1, BC049762, Lct, Slc20a1, Nlrp9b, Gng12, Nanos2, Ccdc50
## Negative:  Serpina3g, Ebi3, Rbp1, Mef2c, Car14, Irak4, Serpina3m, 1700019A02Rik, Tnfaip8, Spin4
##      Cd109, Tsc22d1, Trpc6, Manba, Tbcel, Gcdh, Ypel5, Itih5, Marveld1, Rbl1
##      Gabra1, 6030422M02Rik, Smarca2, Ccm2, Slc18a1, Zfp811, Casp8, Mthfd2, F3, 4930503L19Rik
```
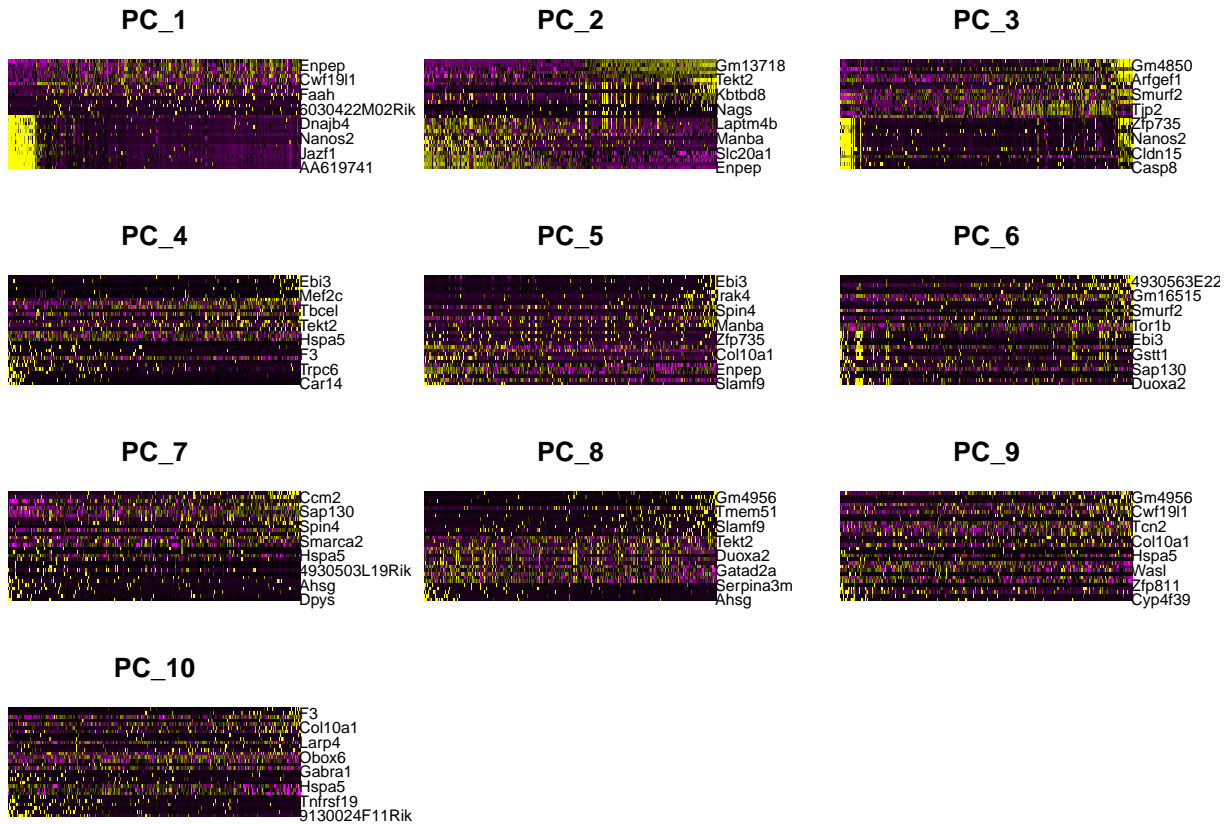
```r
Seurat::VizDimLoadings(seuset, dims = 1:2)
```
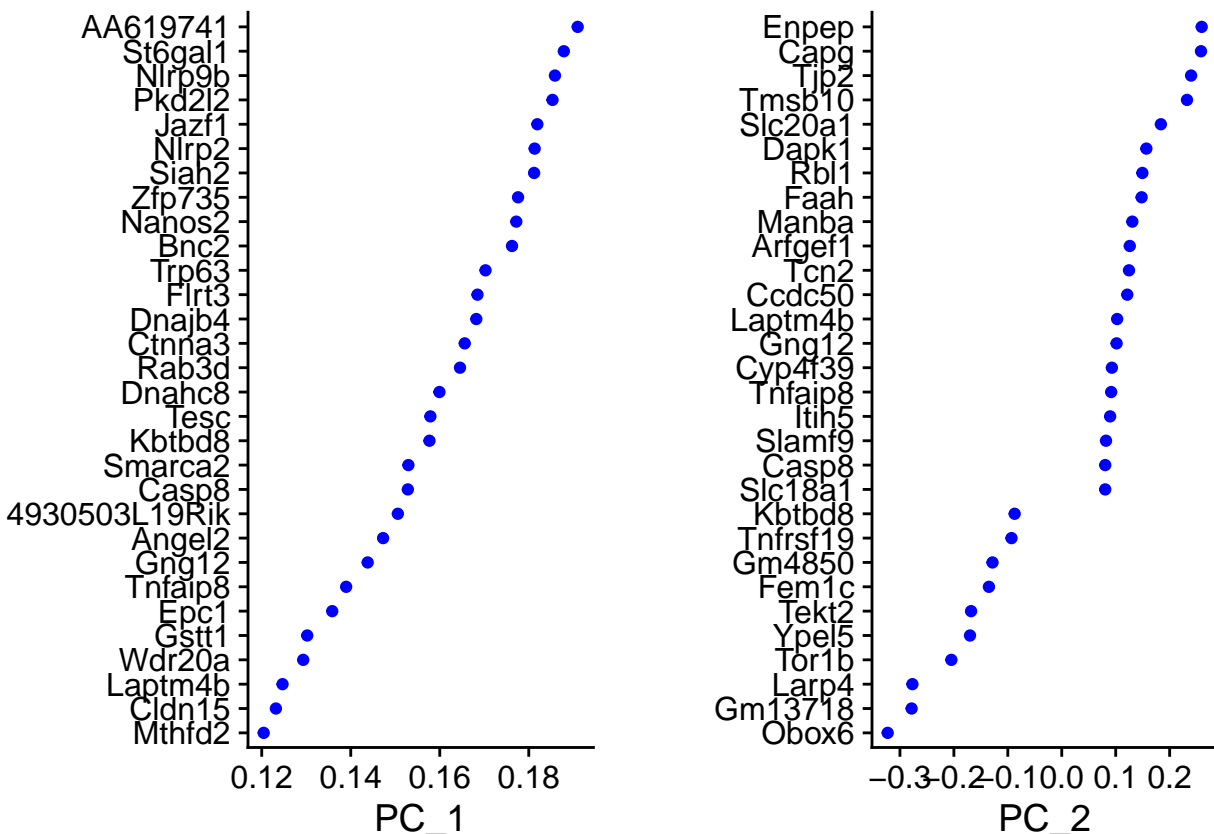
```
Seurat::DimPlot(seuset, dims = c(1,2), reduction="pca")
```



```
Seurat::DimHeatmap(seuset, dims=1:10, nfeatures=30, reduction="pca")
```

PC_1

Enpep
Cwf19l1
Faah
6030422M02Rik
Dnajb4
Nanos2
Jazf1
AA619741

PC_2

Gm13718
Tekt2
Kbtbd8
Nags
Laptm4b
Manba
Slc20a1
Enpep

PC_3

Gm4850
Arfgef1
Smurf2
Tjp2
Zfp735
Nanos2
Cldn15
Casp8

PC_4

Ebi3
Mef2c
Tbcel
Tekt2
Hspa5
F3
Trpc6
Car14

PC_5

Ebi3
Irak4
Spin4
Manba
Zfp735
Col10a1
Enpep
Slamf9

PC_6

4930563E22
Gm16515
Smurf2
Tor1b
Ebi3
Gstt1
Sap130
Duoxa2

PC_7

Ccm2
Sap130
Spin4
Smarca2
Hspa5
4930503L19Rik
Ahsg
Dpys

PC_8

Gm4956
Tmem51
Slamf9
Tekt2
Duoxa2
Gatad2a
Serpina3m
Ahsg

PC_9

Gm4956
Cwf19l1
Tcn2
Col10a1
Hspa5
Wasl
Zfp811
Cyp4f39

PC_10

F3
Col10a1
Larp4
Obox6
Gabra1
Hspa5
Tnfrsf19
9130024F11Rik

```
seuset <- Seurat::JackStraw(seuset, dims=10, reduction = "pca", num.replicate = 100)
seuset <- Seurat::ScoreJackStraw(seuset, dims = 1:10)
Seurat::JackStrawPlot(seuset, dims = 1:10)
```
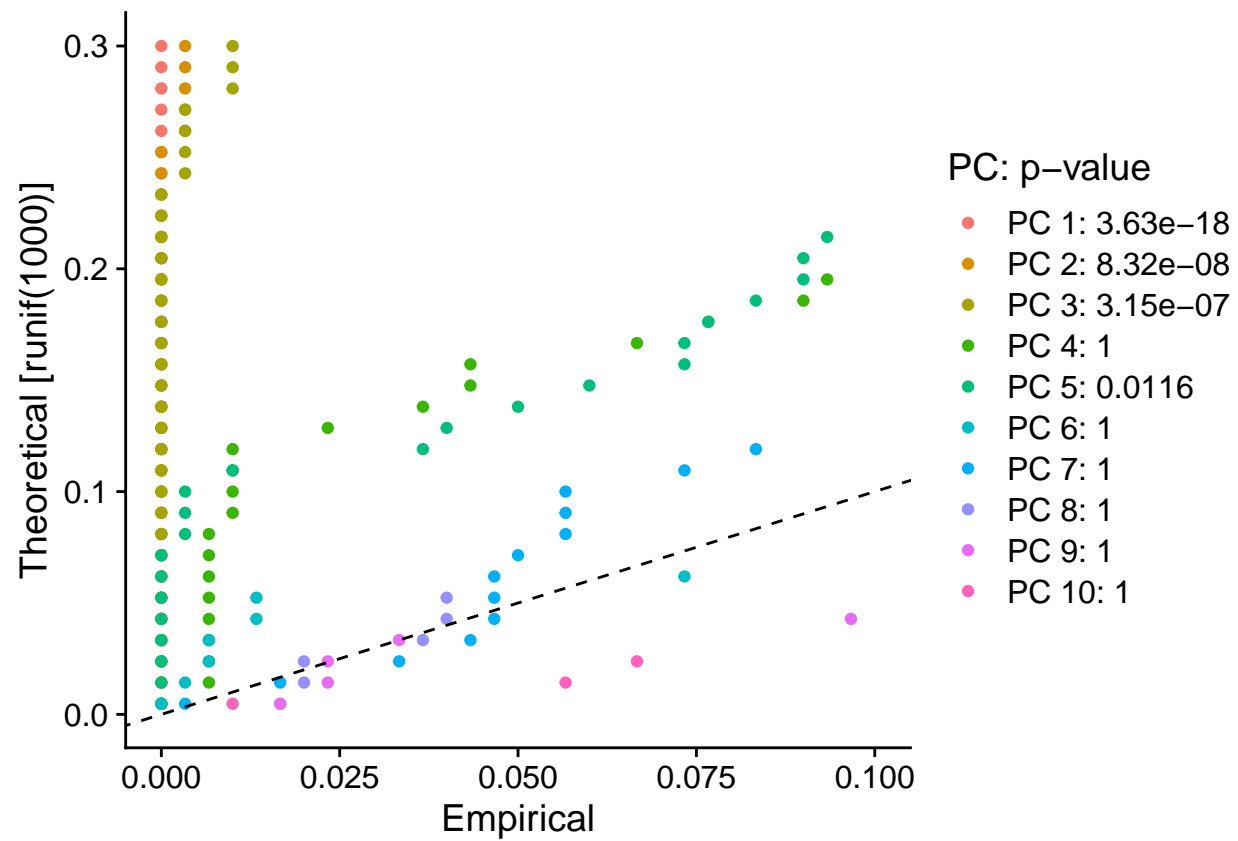
```
Seurat::ElbowPlot(seuset, ndims = 20, reduction = "pca")
```
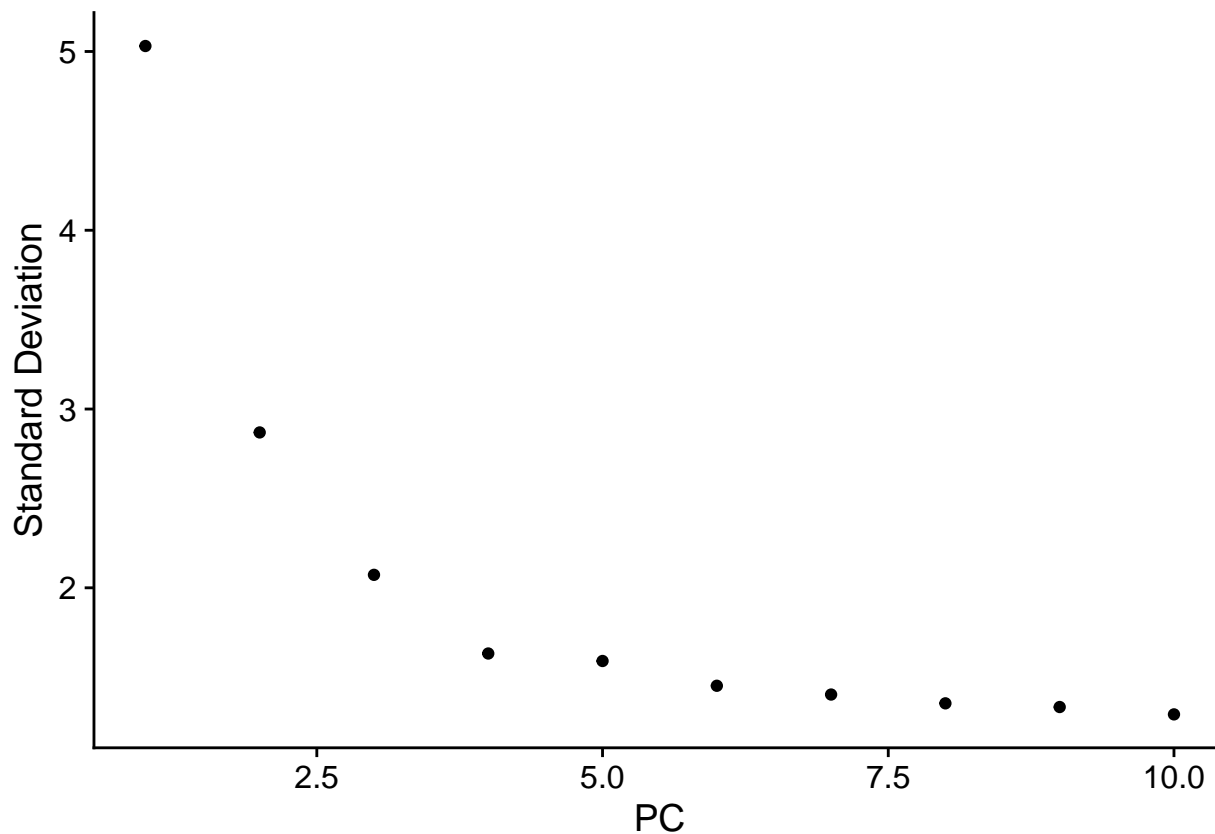
```
## [1] "tSNE"
```

```
seuset <- Seurat::FindNeighbors(object = seuset)
```

```
## Computing nearest neighbor graph
```

```
## Computing SNN
```

```
seuset <- Seurat::FindClusters(object = seuset)
```

```
## Modularity Optimizer version 1.3.0 by Ludo Waltman and Nees Jan van Eck
##
## Number of nodes: 268
## Number of edges: 9101
##
## Running Louvain algorithm...
## Maximum modularity in 10 random starts: 0.6565
## Number of communities: 5
## Elapsed time: 0 seconds
```
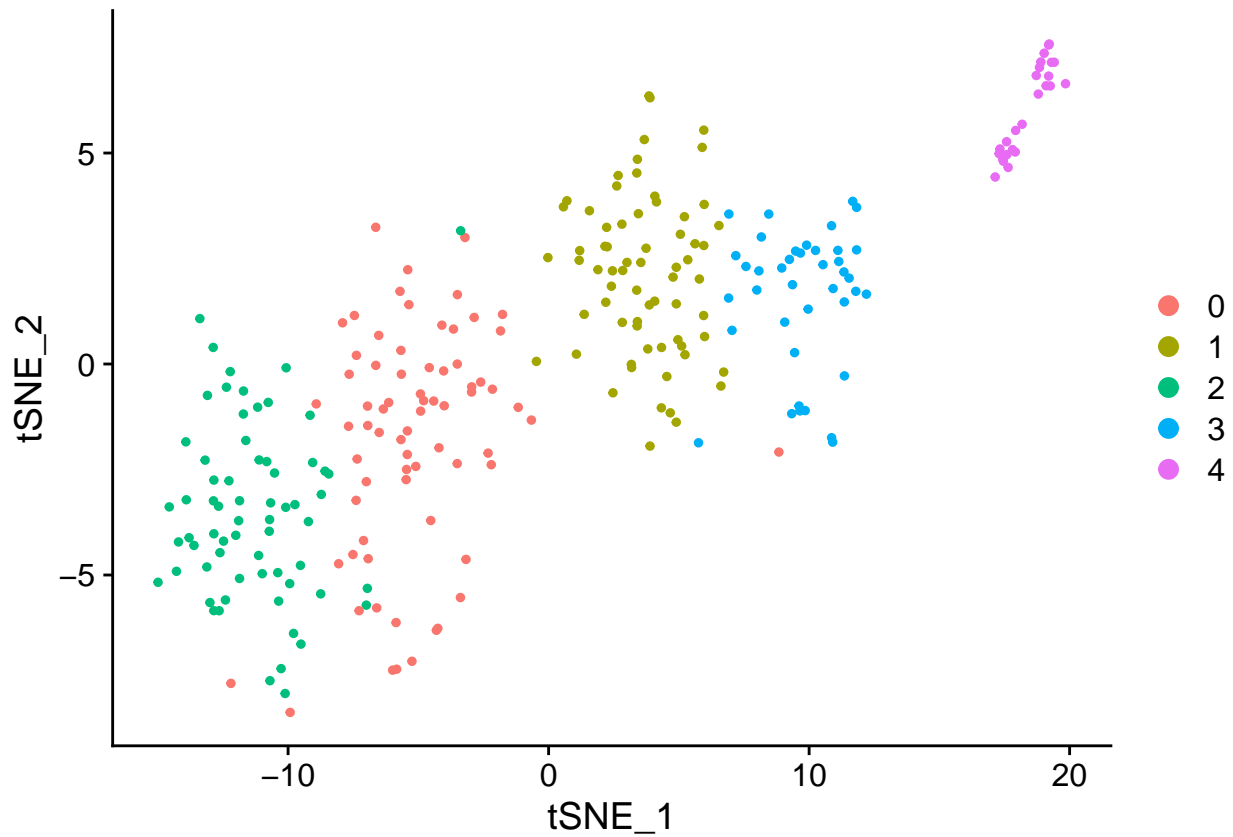
```
seuset <- Seurat::RunTSNE(seuset, dims = 1:10, resolution =0.6)
Seurat::DimPlot(seuset, reduction="tsne")
```

[1] "Marker Genes"

```
markers <- Seurat::FindAllMarkers(seuset, only.pos = TRUE, min.pct = min_pct, logfc.threshold = logfc_t
```

```
## Calculating cluster 0
```

```
## Calculating cluster 1
```

```
## Calculating cluster 2
```

```
## Calculating cluster 3
```

```
## Calculating cluster 4
```

```
top10 <- dplyr::group_by(markers, cluster)
top10 <- dplyr::top_n(top10, 10, avg_logFC)
Seurat::DoHeatmap(seuset, features = top10$gene)
```