DEG analysis using MAST

```r
library(Seurat)
library(ggplot2)
library(plyr)

dir.create('../res_various/res_0.3/')
```

```
## Warning in dir.create("../res_various/res_0.3/"): '../res_various/res_0.3'
## already exists
```

```r
dir.create('../res_various/res_0.3/degs_mast/')
```

```
## Warning in dir.create("../res_various/res_0.3/degs_mast/"): '../res_various/
## res_0.3/degs_mast' already exists
```

```r
subdataset <- readRDS('../tmp/crc_smc.malignantcells.Rds')
head(subdataset@meta.data, n=3); nrow(subdataset@meta.data) # 17334
```

```
##                                      nCount_RNA nFeature_RNA         Library
## AAACCTGCATACGCCG-1-PM-PS-0001-T-A1        35998         4823 PM-PS-0001-T-A1
## AAACCTGGTCGCATAT-1-PM-PS-0001-T-A1        31383         5252 PM-PS-0001-T-A1
## AAACCTGTCCCTTGCA-1-PM-PS-0001-T-A1         7302         1713 PM-PS-0001-T-A1
##                                      Patient  Sample Cell_subtype RNA_snn_res.0.1
## AAACCTGCATACGCCG-1-PM-PS-0001-T-A1     SMC01 SMC01-T         CMS2               1
## AAACCTGGTCGCATAT-1-PM-PS-0001-T-A1     SMC01 SMC01-T         CMS2               0
## AAACCTGTCCCTTGCA-1-PM-PS-0001-T-A1     SMC01 SMC01-T         CMS2               1
##                                      seurat_clusters RNA_snn_res.0.2
## AAACCTGCATACGCCG-1-PM-PS-0001-T-A1                 3               1
## AAACCTGGTCGCATAT-1-PM-PS-0001-T-A1                 5               0
## AAACCTGTCCCTTGCA-1-PM-PS-0001-T-A1                 6               1
##                                      RNA_snn_res.0.3 RNA_snn_res.0.4
## AAACCTGCATACGCCG-1-PM-PS-0001-T-A1                 2               2
## AAACCTGGTCGCATAT-1-PM-PS-0001-T-A1                 0               0
## AAACCTGTCCCTTGCA-1-PM-PS-0001-T-A1                 2               3
##                                      RNA_snn_res.0.5 RNA_snn_res.0.6
## AAACCTGCATACGCCG-1-PM-PS-0001-T-A1                 2               2
## AAACCTGGTCGCATAT-1-PM-PS-0001-T-A1                 0               0
## AAACCTGTCCCTTGCA-1-PM-PS-0001-T-A1                 5               4
##                                      RNA_snn_res.0.7 RNA_snn_res.0.8
## AAACCTGCATACGCCG-1-PM-PS-0001-T-A1                 2               2
## AAACCTGGTCGCATAT-1-PM-PS-0001-T-A1                 1               1
## AAACCTGTCCCTTGCA-1-PM-PS-0001-T-A1                 5               3
##                                      RNA_snn_res.0.9 RNA_snn_res.1
## AAACCTGCATACGCCG-1-PM-PS-0001-T-A1                 2             3
## AAACCTGGTCGCATAT-1-PM-PS-0001-T-A1                 1             1
```

```
## AAACCTGTCCCTTGCA-1-PM-PS-0001-T-A1                      4                4
##                                          RNA_snn_res.1.1 RNA_snn_res.1.2
## AAACCTGCATACGCCG-1-PM-PS-0001-T-A1                      4                5
## AAACCTGGTCGCATAT-1-PM-PS-0001-T-A1                      1                1
## AAACCTGTCCCTTGCA-1-PM-PS-0001-T-A1                      5                6
##                                          RNA_snn_res.1.3 RNA_snn_res.1.4
## AAACCTGCATACGCCG-1-PM-PS-0001-T-A1                      5                5
## AAACCTGGTCGCATAT-1-PM-PS-0001-T-A1                      2                1
## AAACCTGTCCCTTGCA-1-PM-PS-0001-T-A1                      6                6
##                                          RNA_snn_res.1.5 RNA_snn_res.1.6
## AAACCTGCATACGCCG-1-PM-PS-0001-T-A1                      5                4
## AAACCTGGTCGCATAT-1-PM-PS-0001-T-A1                      1                2
## AAACCTGTCCCTTGCA-1-PM-PS-0001-T-A1                      6                6
##                                          RNA_snn_res.1.7 RNA_snn_res.1.8
## AAACCTGCATACGCCG-1-PM-PS-0001-T-A1                      3                5
## AAACCTGGTCGCATAT-1-PM-PS-0001-T-A1                      2                2
## AAACCTGTCCCTTGCA-1-PM-PS-0001-T-A1                      6                6
##                                          RNA_snn_res.1.9 RNA_snn_res.2
## AAACCTGCATACGCCG-1-PM-PS-0001-T-A1                      3              3
## AAACCTGGTCGCATAT-1-PM-PS-0001-T-A1                     10              5
## AAACCTGTCCCTTGCA-1-PM-PS-0001-T-A1                      5              6
```

```
## [1] 17334
```

```r
subdataset@meta.data <- subdataset@meta.data[, c(1:6,10)]
head(subdataset@meta.data, n=3)
```

```
##                                      nCount_RNA nFeature_RNA         Library
## AAACCTGCATACGCCG-1-PM-PS-0001-T-A1        35998         4823 PM-PS-0001-T-A1
## AAACCTGGTCGCATAT-1-PM-PS-0001-T-A1        31383         5252 PM-PS-0001-T-A1
## AAACCTGTCCCTTGCA-1-PM-PS-0001-T-A1         7302         1713 PM-PS-0001-T-A1
##                                      Patient  Sample Cell_subtype RNA_snn_res.0.3
## AAACCTGCATACGCCG-1-PM-PS-0001-T-A1    SMC01 SMC01-T         CMS2               2
## AAACCTGGTCGCATAT-1-PM-PS-0001-T-A1    SMC01 SMC01-T         CMS2               0
## AAACCTGTCCCTTGCA-1-PM-PS-0001-T-A1    SMC01 SMC01-T         CMS2               2
```

```r
summary(subdataset$RNA_snn_res.0.3)
```

```
##    0    1    2    3    4    5    6    7
## 6609 4550 3378 1372  807  325  235   58
```

```r
Idents(subdataset) <- 'RNA_snn_res.0.3'
markers_mast <- FindAllMarkers(object = subdataset, min.pct = 0.25, logfc.threshold = 0.25, only.pos = 
```

```
## Calculating cluster 0
```

```
## Assuming data assay in position 1, with name et is log-transformed.
```

```
##
## Done!
```

```
## Combining coefficients and standard errors

## Calculating log-fold changes

## Calculating likelihood ratio tests

## Refitting on reduced model...

##
## Done!

## Calculating cluster 1

## Assuming data assay in position 1, with name et is log-transformed.

##
## Done!

## Combining coefficients and standard errors

## Calculating log-fold changes

## Calculating likelihood ratio tests

## Refitting on reduced model...

##
## Done!

## Calculating cluster 2

## Assuming data assay in position 1, with name et is log-transformed.

##
## Done!

## Combining coefficients and standard errors

## Calculating log-fold changes

## Calculating likelihood ratio tests

## Refitting on reduced model...

##
## Done!

## Calculating cluster 3
```

```
## Assuming data assay in position 1, with name et is log-transformed.

##
## Done!

## Combining coefficients and standard errors

## Calculating log-fold changes

## Calculating likelihood ratio tests

## Refitting on reduced model...

##
## Done!

## Calculating cluster 4

## Assuming data assay in position 1, with name et is log-transformed.

##
## Done!

## Combining coefficients and standard errors

## Calculating log-fold changes

## Calculating likelihood ratio tests

## Refitting on reduced model...

##
## Done!

## Calculating cluster 5

## Assuming data assay in position 1, with name et is log-transformed.

##
## Done!

## Combining coefficients and standard errors

## Calculating log-fold changes

## Calculating likelihood ratio tests

## Refitting on reduced model...
```

```
##
## Done!
```

```
## Calculating cluster 6
```

```
## Assuming data assay in position 1, with name et is log-transformed.
```

```
##
## Done!
```

```
## Combining coefficients and standard errors
```

```
## Calculating log-fold changes
```

```
## Calculating likelihood ratio tests
```

```
## Refitting on reduced model...
```

```
##
## Done!
```

```
## Calculating cluster 7
```

```
## Assuming data assay in position 1, with name et is log-transformed.
```

```
##
## Done!
```

```
## Combining coefficients and standard errors
```

```
## Calculating log-fold changes
```

```
## Calculating likelihood ratio tests
```

```
## Refitting on reduced model...
```

```
##
## Done!
```

```
## [1] 1314
```

```r
markers_mast <- subset(markers_mast, p_val_adj <= 0.05); nrow(markers_mast) # 1286
```

```
## [1] 1286
```

```r
markers_mast$Symbol <- unlist(lapply(markers_mast$gene, function(x) unlist(strsplit(x, split = '-ENSG'))
head(markers_mast, n=3)
```

```
##                                      p_val avg_logFC pct.1 pct.2    p_val_adj
## GAS5-ENSG00000234741.8-9     3.861405e-183 0.2870560 0.972 0.925 1.387017e-178
## SLC12A2-ENSG00000064651.14-5 2.159370e-171 0.2609346 0.855 0.690 7.756457e-167
## MYC-ENSG00000136997.20-9     8.603922e-161 0.3270763 0.770 0.605 3.090529e-156
##                              cluster                         gene  Symbol
## GAS5-ENSG00000234741.8-9           0     GAS5-ENSG00000234741.8-9    GAS5
## SLC12A2-ENSG00000064651.14-5       0 SLC12A2-ENSG00000064651.14-5 SLC12A2
## MYC-ENSG00000136997.20-9           0     MYC-ENSG00000136997.20-9     MYC
```

```r
write.table(markers_mast, '../res_various/res_0.3/degs_mast/markers.MAST.txt', sep = '\t', quote = F, co
#saveRDS(subdataset, '../tmp/crc_smc.malignantcells.Rds')
```