

## cell\_type\_assignment

Changhua Yu

8/15/2019

### Demonstration for assigning cell type

- The cell marker used is derived from Jay Shendure Groups Cell Atlas Project

## Reading in required package and R scripts

```
source("../R/celltype_assign.R")
```

```
atlas_markers = Build.Ref.Markers(path = "../data/jay_cell_markers.txt", sep = '\t', splitter = c("[[:p
head(atlas_markers)
```

```
##      gene      celltype
## 1  foxj1      Ciliated
## 2 scgb1a1 Clara (Club cell)
## 3  krt15 Clara (Club cell)
## 4 scbg3a2 Clara (Club cell)
## 5  cyp2f2 Clara (Club cell)
## 6   chad Clara (Club cell)
```

## Reading Seurat Output Used also as ShinyApp interactive session input

```
seurat.out.neg = readRDS('../ShinyDiff_multi/input/WTIKCD45NEG_out.rds')
```

### Build the experimental marker from conserved marker list of seurat output

```
# filter only the positive fold change conserve markers with a max_pval threshold of 0.05
exp.markers.neg = seurat.out.neg$conserved.markers %>% filter(wni_avg_logFC > 0 & max_pval <= 0.05)
# rename the first row as gene for the pipeline to run
colnames(exp.markers.neg)[1] = "gene"

# Insert a column "avg_logFC" as the average of fold change of a marker across 4 different conditions,
exp.markers.neg = exp.markers.neg %>% mutate(avg_logFC = (wni_avg_logFC+wti_avg_logFC+kii_avg_logFC+kin_avg_logFC)/4)
head(exp.markers.neg)
```

```
## # A tibble: 6 x 25
##   gene   wni_p_val wni_avg_logFC wni_pct.1 wni_pct.2 wni_p_val_adj kii_p_val
##   <chr>     <dbl>      <dbl>    <dbl>    <dbl>      <dbl>      <dbl>
## 1 Ptprb 1.33e-149      1.57    0.746    0.209      2.77e-145  0.
## 2 Cd93  3.02e-148      1.31    0.811    0.237      6.30e-144  0.
## 3 Sema~ 6.90e-122      1.51    0.835    0.389      1.44e-117 6.39e-304
```

```
## 4 Sema~ 1.84e-103      1.17      0.392      0.073      3.83e- 99 0.
## 5 Plvap 9.50e- 95      1.20      0.63      0.197      1.98e- 90 0.
## 6 Hpgd  2.63e- 76      1.32      0.662      0.286      5.47e- 72 0.
## # ... with 18 more variables: kii_avg_logFC <dbl>, kii_pct.1 <dbl>,
## #   kii_pct.2 <dbl>, kii_p_val_adj <dbl>, wti_p_val <dbl>,
## #   wti_avg_logFC <dbl>, wti_pct.1 <dbl>, wti_pct.2 <dbl>,
## #   wti_p_val_adj <dbl>, kini_p_val <dbl>, kini_avg_logFC <dbl>,
## #   kini_pct.1 <dbl>, kini_pct.2 <dbl>, kini_p_val_adj <dbl>,
## #   max_pval <dbl>, minimump_p_val <dbl>, cluster <dbl>, avg_logFC <dbl>
```

```
# creat a experiemental marker set
exp.markers.neg = Build.Exp.Markers(exp.markers.neg)
head(exp.markers.neg)
```

```
##      gene      wni_p_val wni_avg_logFC wni_pct.1 wni_pct.2 wni_p_val_adj
## 1  ptprb 1.328580e-149      1.568486      0.746      0.209 2.768362e-145
## 2   cd93 3.022382e-148      1.306909      0.811      0.237 6.297737e-144
## 3  sema3c 6.903665e-122      1.506889      0.835      0.389 1.438517e-117
## 4  sema3g 1.838285e-103      1.165673      0.392      0.073 3.830435e-99
## 5   plvap 9.496584e-95      1.201729      0.630      0.197 1.978803e-90
## 6   hpgd 2.626964e-76      1.316952      0.662      0.286 5.473804e-72
##      kii_p_val kii_avg_logFC kii_pct.1 kii_pct.2 kii_p_val_adj wti_p_val
## 1  0.000000e+00      1.362327      0.956      0.316 0.000000e+00      0
## 2  0.000000e+00      1.432443      0.973      0.303 0.000000e+00      0
## 3  6.39005e-304      1.119915      0.913      0.442 1.331495e-299      0
## 4  0.000000e+00      1.131309      0.628      0.143 0.000000e+00      0
## 5  0.000000e+00      1.469009      0.861      0.270 0.000000e+00      0
## 6  0.000000e+00      1.220732      0.941      0.538 0.000000e+00      0
##      wti_avg_logFC wti_pct.1 wti_pct.2 wti_p_val_adj      kini_p_val
## 1      1.417246      0.974      0.314      0 0.000000e+00
## 2      1.449010      0.986      0.300      0 0.000000e+00
## 3      1.110820      0.900      0.427      0 1.023374e-237
## 4      1.101606      0.604      0.129      0 8.477935e-166
## 5      1.495852      0.888      0.263      0 1.379157e-272
## 6      1.151542      0.930      0.509      0 5.114002e-116
##      kini_avg_logFC kini_pct.1 kini_pct.2 kini_p_val_adj      max_pval
## 1      1.4267667      0.878      0.261      0.000000e+00 1.328580e-149
## 2      1.4366569      0.918      0.264      0.000000e+00 3.022382e-148
## 3      1.3219349      0.872      0.436      2.132404e-233 6.903665e-122
## 4      0.9204982      0.401      0.096      1.766547e-161 1.838285e-103
## 5      1.3459533      0.763      0.245      2.873750e-268 9.496584e-95
## 6      1.0359701      0.732      0.407      1.065605e-111 2.626964e-76
##      minimump_p_val cluster avg_logFC sample
## 1      0      0 1.443706      1
## 2      0      0 1.406255      1
## 3      0      0 1.264890      1
## 4      0      0 1.079772      1
## 5      0      0 1.378136      1
## 6      0      0 1.181299      1
```

## Run Assign.Cluster.Type for cell type assignment

```
celltype.atlas.neg = Assign.Cell.Type(ref.markers = atlas_markers,exp.markers = exp.markers.neg)
head(celltype.atlas.neg)
```

```
##      sample cluster      type1.name
## sample      1      0 Hematopoetic progenitor
## sample1     1      1 Hematopoetic progenitor
## sample2     1      2      endothelial cell
## sample3     1      3      Type II pneumocyte
## sample4     1      4      Smooth muscle
## sample5     1      5      unknown
##
## sample                                cd93-1.41 kit-0
## sample1                             cd93-1.61 kit-0
## sample2                             kdr-1.66 pecam1-
## sample3 sftpc-3.72 sftpa1-3.57 sftpb-3.31 scd1-2.95 slc34a2-2.68 s100g-1.9 ppp1r14c-1.36 egfl6-1.21
## sample4                                cald1-0.95 smtn-0
## sample5
##      type1.fc      type2.name      type2.gene type2.fc
## sample  2.720455 endothelial cell  eng-0.69 pecam1-0.63 1.312117
## sample1 2.993784 endothelial cell  pecam1-0.54 eng-0.51 1.048688
## sample2 3.501031      Mono/Mac/DC  ly6c1-1.48 mertk-0.6 2.080064
## sample3 21.110164 epithelial cell   epcam-1 cdh1-0.91 1.904849
## sample4 2.139155      B cell        ebf1-1.79 1.789907
## sample5 0.000000      unknown              0.000000
##
##      type3.name      type3.gene type3.fc
## sample      podocyte  pde3a-0.86 0.8602076
## sample1      B cell   cd200-0.88 0.8828012
## sample2 Hematopoetic progenitor cd34-1.23 cd38-0.62 1.8477205
## sample3      B cell   egfl6-1.21 1.2050201
## sample4      fibroblast des-1.28 1.2810840
## sample5      unknown              0.0000000
```

```
# save assignment list to a desire directory
write.csv(celltype.atlas.neg,file = "../MAP3K3/output/190815/CD45NEG_atlas.csv")
```