AML2

jtrincado

2022-02-08 10:52:01

Contents

1. Put together both 34 and 38 libraries. Apply QC and dimensionality reduction. 1
2. Get the LSC6 score 5
3. Predict the class of the cells using the markers and the expression of the BM cells form Van_Galen paper 7
4. Project the predictions from Velten onto our UMAP 11
Clusters 0 and 4 seems the most enriched among the LSC. They bare a similar proportion of HSC-like predicted cells. We will consider both as LSC enriched clusters 12

1. Put together both 34 and 38 libraries. Apply QC and dimensionality reduction.

```
## Modularity Optimizer version 1.3.0 by Ludo Waltman and Nees Jan van Eck
##
## Number of nodes: 10151
## Number of edges: 342059
##
## Running Louvain algorithm...
## Maximum modularity in 10 random starts: 0.8229
## Number of communities: 17
## Elapsed time: 1 seconds
```

```
## CD34_AAACCCAAGTGGACGT-1 CD34_AAACCCAAGTGGATAT-1 CD34_AAACCCACAGAGTAAT-1
##                       4                       4                       4
## CD34_AAACCCACATAGGTTC-1 CD34_AAACCCATCTATACTC-1
##                       8                       4
## Levels: 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
```

```
##
##           0    1    2    3    4    5    6    7    8    9   10   11   12   13
##   CD34 1327    4    3   10  961    1    2    5  473  465  389    4    3  250
##   CD38    0 1027 1010  991    0  898  835  486    2    7    5  324  282    2
##
##          14   15   16
##   CD34  180   71   77
##   CD38    0   57    0
```

```
## Warning: The default method for RunUMAP has changed from calling Python UMAP via reticulate to the R-
## To use Python UMAP via reticulate, set umap.method to 'umap-learn' and metric to 'correlation'
## This message will be shown once per session
```

UMAP_2
4
0
−4
−10
−5
0
5
UMAP_1
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16

Phase
UMAP_2
4
0
−4
−10
−5
0
5
UMAP_1
G1
G2M
S

```
## Check for expression of malignant marker for t(8;21) RUNX1T1
```

2. Get the LSC6 score

```
## [1] "CD34"   "SPINK2" "SOCS2"  "FAM30A" "ADGRG1" "DNMT3B"
```

Elsayed_LSC_score
UMAP_2
UMAP_1
4
0
−4
−10
−5
0
5
3
2
1
0

3. Predict the class of the cells using the markers and the expression of the BM cells form Van_Galen paper

```
## Performing PCA on the provided reference using 1821 features as input.
## Projecting cell embeddings
## Finding neighborhoods
## Finding anchors
##  Found 2846 anchors
## Filtering anchors
##  Retained 1351 anchors
## Finding integration vectors
## Finding integration vector weights
## Predicting cell labels
##
##         HSC Prog  GMP ProMono Mono  cDC  pDC earlyEry lateEry ProB    B Plasma
##   CD34  539 2977  233      40   15  138   23        5       0   51  203      0
##   CD38    0   12  603    1591 3564  151    1        0       2    0    0      0
```

```
##
##            T  CTL   NK
##    CD34    1    0    0
##    CD38    1    0    1
```

prediction
UMAP_2
UMAP_1
4
0
−4
−10
−5
0
5
GMP
ProMono
ProB
B
lateEry
pDC
cDC
HSC
Prog
T
NK
Mono
earlyEry
HSC
Prog
GMP
ProMono
Mono
cDC
pDC
earlyEry
lateEry
ProB
B
T
NK

count
1000
500
0
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Clusters
Predicted_cell_type
HSC
Prog
GMP
ProMono
Mono
cDC
pDC
earlyEry
lateEry
ProB
B
T
NK

4. Project the predictions from Velten onto our UMAP

```
##
##                    B cells                      blasts
##                         85                           2
##             dendritic cells Eosinophil-basophil-mast prog
##                        152                         241
##                   erythroid                   HSCs & MPPs
##                        420                         132
##                      LMPPs                 Megakaryocyte
##                        485                          94
##                  Monocytes                  Myeloblasts
##                       3590                        2310
##                 Myelocytes                     NK cells
##                       2607                           1
##                    T cells
##                          3
```

Clusters 0 and 4 seems the most enriched among the LSC. They bare a similar proportion of HSC-like predicted cells. We will consider both as LSC enriched clusters