CellFindR Vignette

Code

Welcome to CellFindR! An add-on to Seurat to help with cell discovery analysis in R. We will go through some standard methods to augment Seurat workflow and create a directory of files to help with your analysis.

load required packages:

note, there are a few ways to load the object, you can download the CellFindR.tar.gz file and use devtools (install_packages('file directory')) to ensure that all the other subpackages will be also installed.

The foolproof method would be to go into the R file and just run all the functions included in that file along with Seurat.

Hide

Hide

```
library(Seurat)
library(dplyr)
library(Matrix)
library(CellFindR)
```

We begin by setting our working directory and importing our data:

Hide

Hide

```
working_directory <- '/Users/kyu/E14.5/'
project_name <- 'E14.5'
```

We load our data using our created function load_tenx that follows the initial steps for Seurat object initiation. Note our matrix count files from Seurat are in the mm10 folder so we will point to there.

Notable design choices that can be altered is the nGene cut off for cells at 7000 cells and we do not filter out the mitochondria genes These parameters can be changed depending on the dataset. Other parameters such as minimum 3 cells must express the gene and 100 genes must be expressed by a cell are standard protocol for Seurat that are followed.

Hide

Hide

```
setwd(working_directory)
load_tenx('mm10', res = 0.7, proj_name = project_name, cutoff= 7000, mito = FALSE)
```

The above step should take a few minutes to run! Now you should have a .Robj in the same directory as your matrix.mtx file! This will be the R object that holds all the detailed information that you need! Some basic batch control plots are also generated. A histogram of the number of genes expressed in each cell is made and a violin plot + histogram of the percent mitochondrial genes for each gene is also generated. When you load it the variable name will be tenx.

The following creates a directory with unique timestamp for CellFindR algorithm generation:

File failed to load: file:///Users/kyu/Desktop/CellFindR%20Vignette_files/extensions/MathZoom.js

Hide

Hide

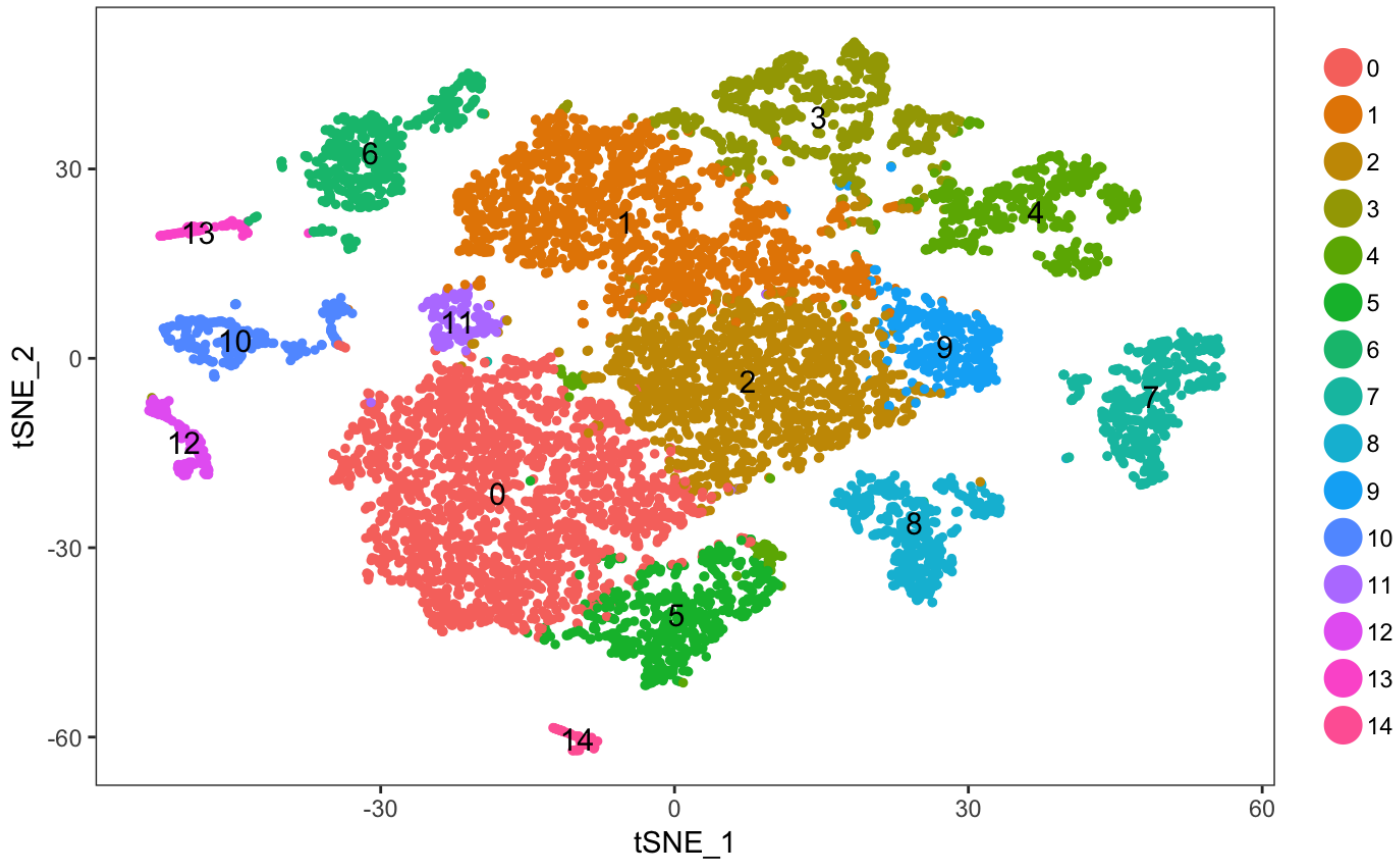
```
dir_add <- paste(working_directory, 'output', Sys.time(),'/', sep ='')
dir.create(dir_add)
file_dest <- paste(dir_add, project_name,'_', sep = '')
```

Lets plot a tSNE, we can load the object now from our directory instead of running it again! Tenx TSNE of the initial groups!

Hide

Hide

```
load(paste(working_directory, 'mm10/', project_name, '.Robj', sep = ''))
#run clustering for big groups, for time and computation sake, if we are satisfied with
 the first order groups generated initially by Seurat we can keep the res and do not nee
d to run.
#res <- find_res(tenx)
#print(res)
#tenx <- FindClusters(tenx,pc.use = 1:10,resolution = res ,print.output = 0,save.SNN =
 T)
TSNEPlot(tenx, do.label = TRUE)
```



Runs CellFindR clustering into the directory file-dest: This process will take a quite a few minutes as it iterates through all the clusters appropriately. The output files are as listed: assign_test.csv = Cell ID with cluster identification. res_test.csv = How each subgroup resolution was chosen.

tSNE of each subgroupings labeled with cluster identification matrix of each subgrouping with top differential genes and expression. directory folders for each subgroup that created further divisions.

label cells with newly generated cluster IDs.

Hide

Hide

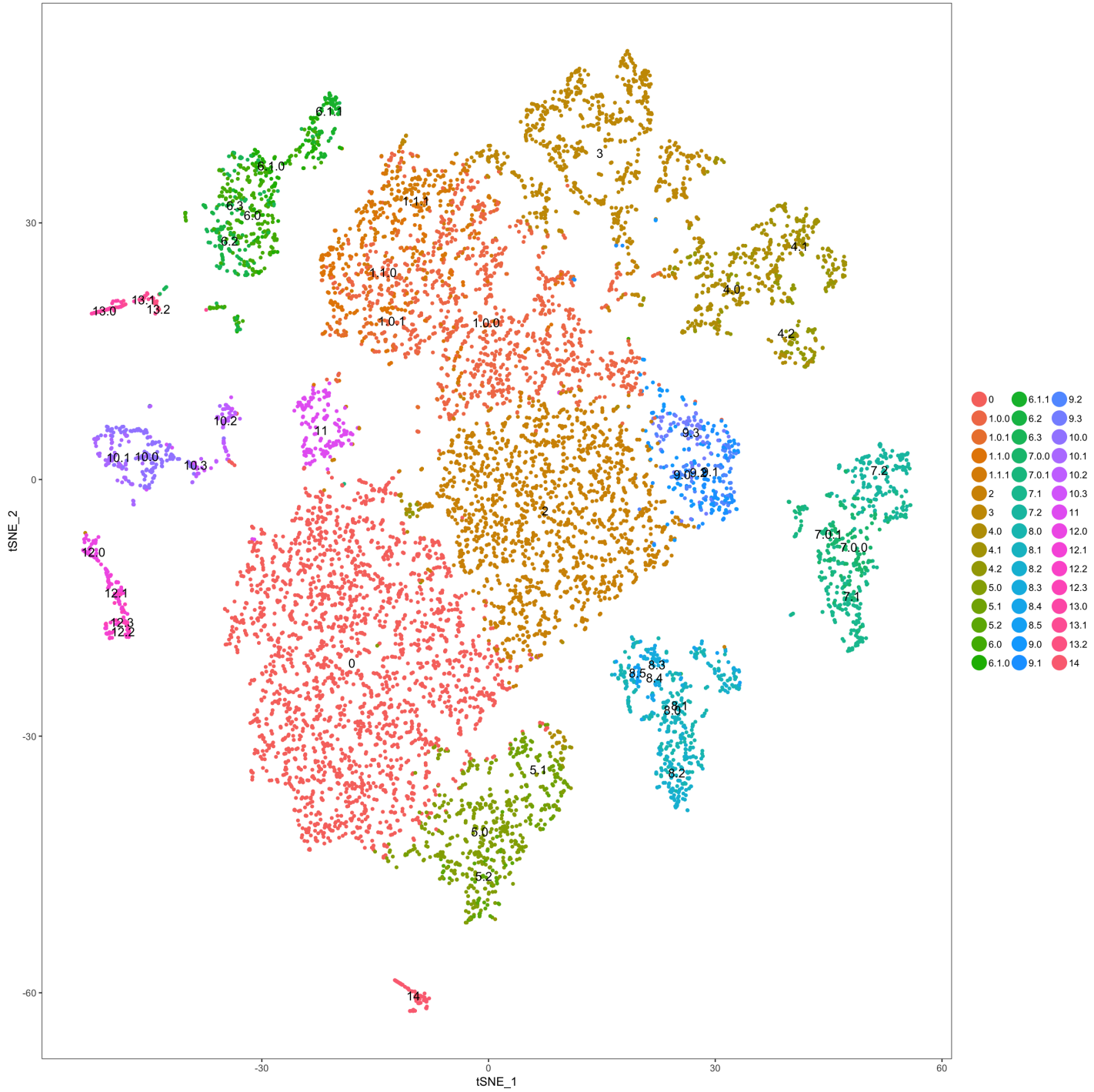
```
tenx@data.info <- tenx@data.info[order(row.names(tenx@data.info)),]
label_cells <- read.csv(paste(dir_add, "/assign_test.csv", sep = ''), header = TRUE)
label_cells <- label_cells[order(label_cells$"cellnames"),]
tenx@data.info$finalcluster <- label_cells$clusterid
tenx<-SetAllIdent(tenx, id = 'finalcluster')
TSNEPlot(tenx, do.label = TRUE)
pdf('./TSNE_all_groups.pdf', width = 15, height = 15)
```

Hide

Hide

```
TSNEPlot(tenx, do.label = TRUE)
dev.off()
```

```
quartz_off_screen
                2
```



Creates the stat matrix and differential gene matrix for all the CellFindR Groups

Hide

Hide

```
# get all the matrix
getmatrix(tenx, ex_loc, 'all_alg_group_matrix')
get_stats(tenx, ex_loc,project_name)
```

File failed to load: file:///Users/kyu/Desktop/CellFindR%20Vignette_files/extensions/MathZoom.js