Add metadata features from tracer

Lincoln Harris

7.13.18

This is the workflow for reading in TRACER output to an existing Seurat object

load libraries

```
library(Seurat)
```

```
Loading required package: ggplot2
```

```
Loading required package: cowplot
```

```

Attaching package: 'cowplot'
```

```
The following object is masked from 'package:ggplot2':

    ggsave
```

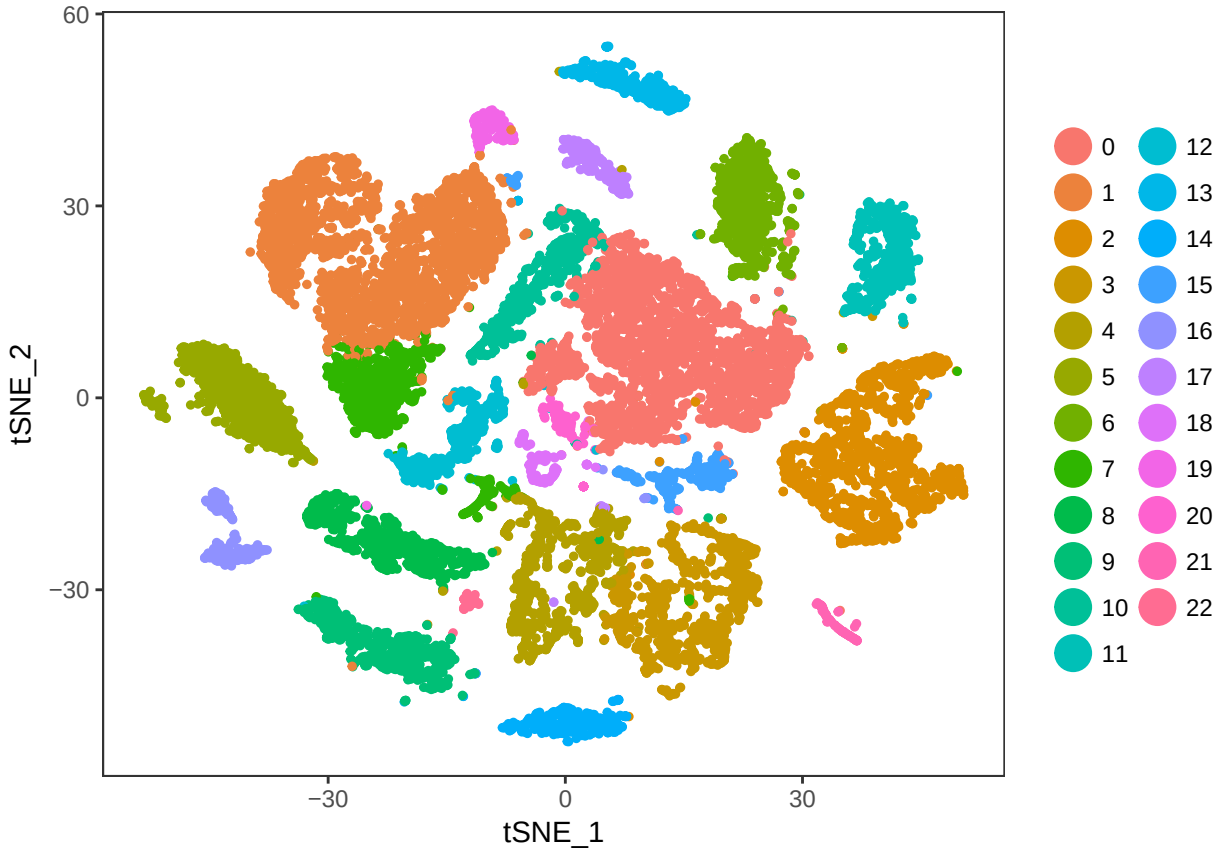
```
Loading required package: Matrix
```

Load Seurat obj

```
#load("/Users/lincoln.harris/Desktop/04_tiss_subset_181016.RData")
load("../rawdata/04_tiss_subset_181016_ALL.RData")
```

```
dim(tiss_subset@meta.data)
```

```
[1] 17516    46
```




```
colnames(tiss_subset@meta.data)
```

```
 [1] "nGene"                         "nReads"
 [3] "orig.ident"                    "well"
 [5] "plate"                         "sample_name"
 [7] "sample_type"                   "patient_id"
 [9] "DOB"                           "gender"
[11] "race"                          "smokingHx"
[13] "histolgy"                      "driver_gene"
[15] "driver_mutation"               "secondary_mutation"
[17] "Notes"                         "stage.at.dx"
[19] "pathlogy_review"               "biopsy_date"
[21] "biopsy_type"                   "biopsy_site"
[23] "biopsy_timing"                 "treatment_status"
[25] "treatment_navie"               "treatment_type"
[27] "treatment"                     "infections"
[29] "pfs"                           "date_of_death"
[31] "sort_date"                     "percent.ercc"
[33] "free_annotation"               "cell_ontology_class"
[35] "percent.ribo"                  "main_seurat_id_cluster"
[37] "S.Score"                       "G2M.Score"
[39] "Phase"                         "immune_annotation"
[41] "general_annotation"            "immune_subtype_annotation"
[43] "T_cell_subtype_annotation"     "MF_cell_subtype_annotation"
[45] "Final_immune_annotation"       "epithelial_subannotation"
```

Read in Tracer data

```
# A/B summary
tracer_summary <- read.csv("../TCR_analysis/filtered_TCRAB_summary/cell_data.csv", header = T)
```

Lets define a new metadata df and add some new cols

```
meta_edit <- tiss_subset@meta.data

meta_edit$A_productive <- NA
meta_edit$A_productive <- as.vector(meta_edit$A_productive)

meta_edit$B_productive <- NA
meta_edit$B_productive <- as.vector(meta_edit$B_productive)

meta_edit$clonal_group_AB <- NA
meta_edit$group_size_AB <- NA
```

Make sure cell IDs look the same

```
tracer_summary$cell_name <- gsub("[.]", "_", tracer_summary$cell_name)
```

find cell name matches btwn meta_edit and tracer_summary this match() function is so much more efficient that looping!!

```
match_vec <- match(row.names(meta_edit), tracer_summary$cell_name) # meta_edit first
match_vec1 <- match(tracer_summary$cell_name, row.names(meta_edit)) # tracer_summary first

length(match_vec)
```

```
[1] 17516
```

```
length(match_vec1)
```

```
[1] 2952
```

```
head(match_vec)
```

```
[1] NA NA NA NA NA NA
```

```
length(unique(match_vec))
```

```
[1] 2953
```

add tracer_summary info to meta_edit, based on cell name matches

```
for (i in 1:length(match_vec)){
  currIndex <- match_vec[i]
  meta_edit$A_productive[i] <- as.vector(tracer_summary$A_productive[currIndex])
  meta_edit$B_productive[i] <- as.vector(tracer_summary$B_productive[currIndex])
}
```

now export to csv

```
write.csv(meta_edit, "metadata_with_assembled_TCRs.csv")
```