

# Figure\_1.rmd

Kaitlyn Lagattuta

2024-11-18

```
suppressPackageStartupMessages(source("utils.R"))
```

```
## Warning: package 'tidyr' was built under R version 4.3.2
```

```
## Warning: package 'ggplot2' was built under R version 4.3.2
```

## Figure 1A

```
sref = readRDS("data/sref_combat_full_authTplusNKT_20hPCs_tcrfilt0607_nvargenes200_sampTH1.instTH0.5.p  
md = sref$meta_data  
md$UMAP1 = sref$umap$embedding[,1]  
md$UMAP2 = sref$umap$embedding[,2]  
clust = readRDS("data/metadata_moreclust2_combat_full_authTplusNKT_20hPCs_tcrfilt0607_nvargenes200_samp  
md = left_join(md, clust[,c("X", "cl0.5")])
```

```
## Joining with 'by = join_by(X)'
```

```
md = label_invariants(md)  
md$clust = paste("A", as.numeric(as.character(md$cl0.5))+1, sep="")  
F1A = my_cluster_umap(md, "clust")  
F1A
```

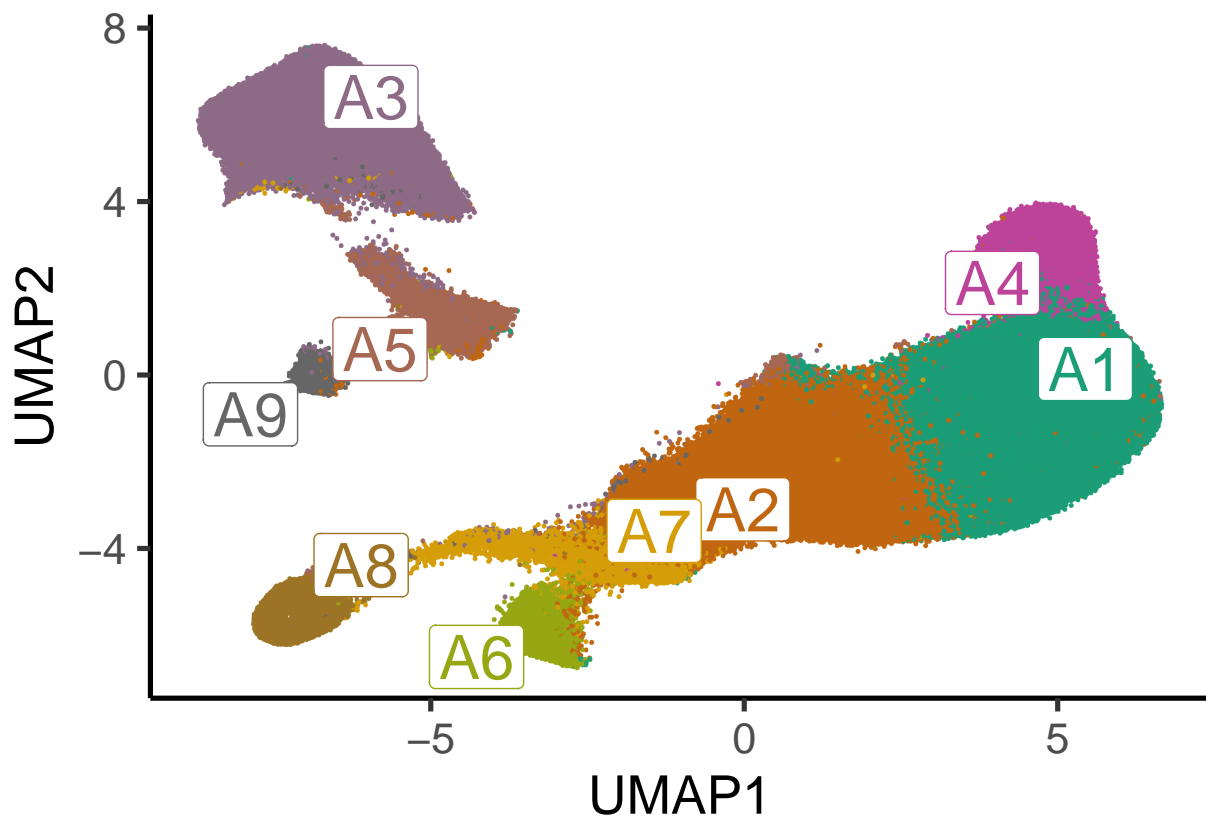


Figure 1B

```
g = ggplot(md[order(md$my_abMAIT),], aes(UMAP1, UMAP2, color=my_abMAIT))
g = g + scale_color_manual(values=c("gray90", "darkgoldenrod2"))
F1B = my_umap_theme(g, aes=12) + labs(color="MAIT TCR")
F1B
```

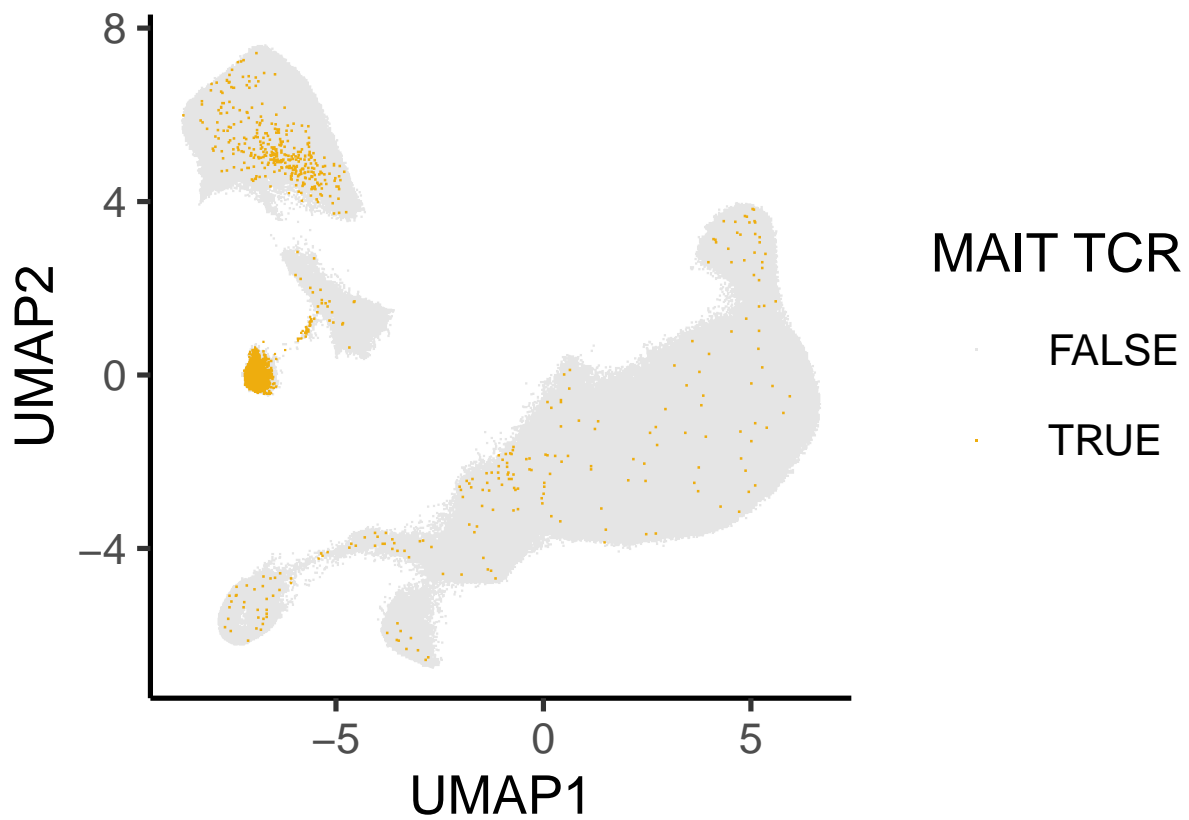


Figure 1C

```
g = ggplot(md[order(md$my_abNKT),], aes(UMAP1, UMAP2, color=my_abNKT))
g = g + scale_color_manual(values=c("gray90", "red"))
F1C = my_umap_theme(g, aes=12) + labs(color="NKT TCR")
F1C
```

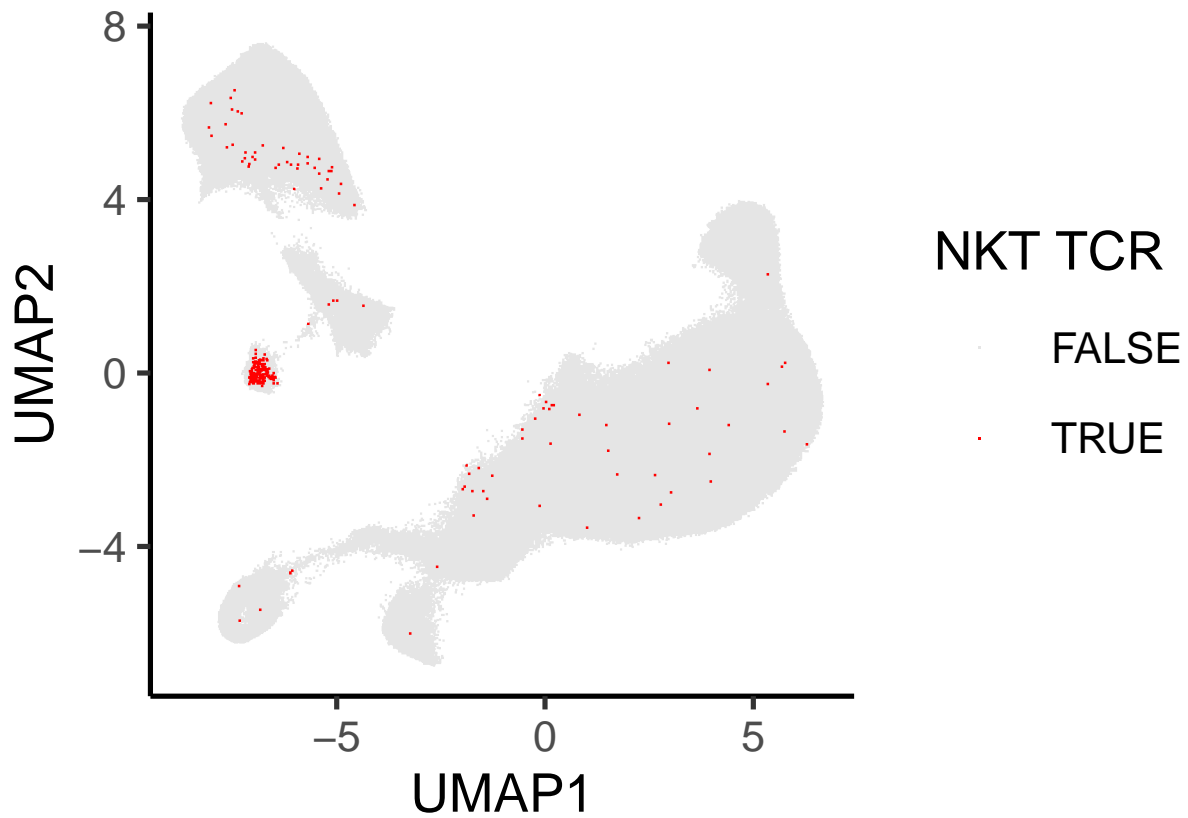


Figure 1D

```
tcr_info = process_tcr_info()
md = left_join(md, tcr_info[,c("barcode_id", "clonotype")], by=c("X"="barcode_id"))
F1D = highlight_twins(md, tcr_info)
```

```
## 'summarise()' has grouped output by 'COMBAT_ID'. You can override using the
## '.groups' argument.
```

F1D

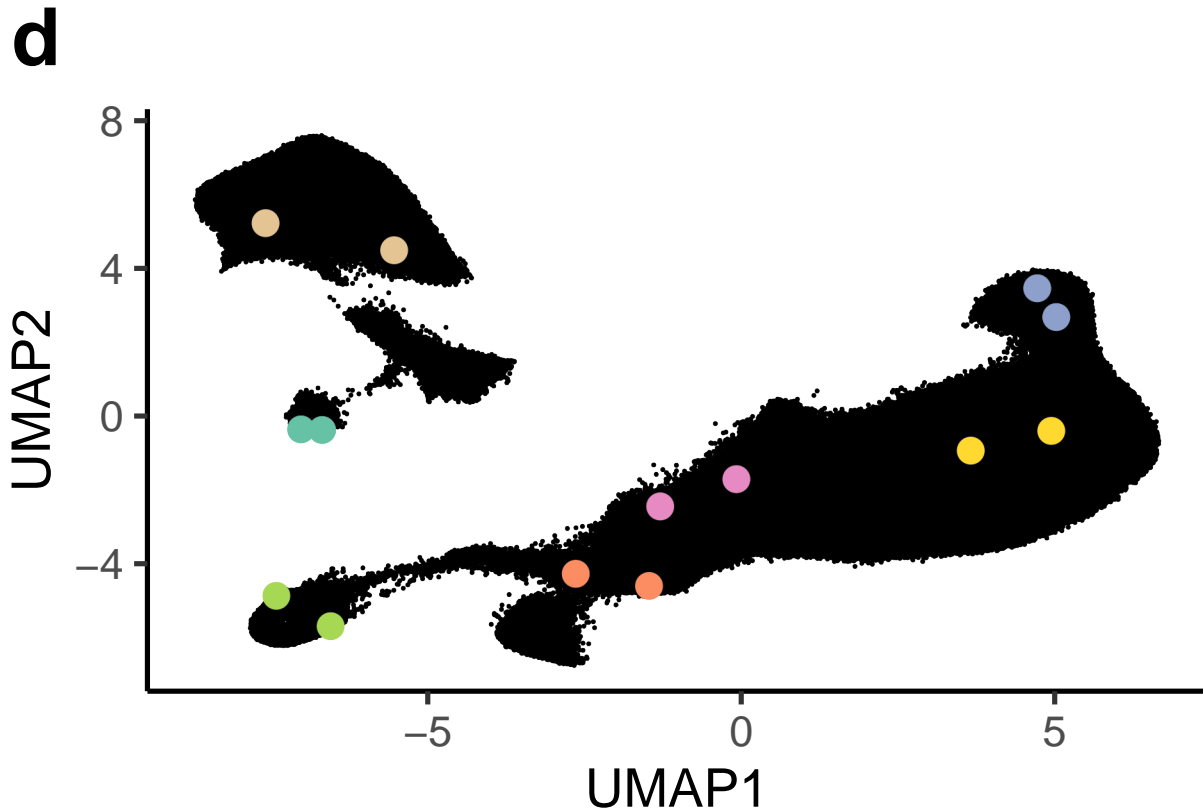


Figure 1E

```
twin_res1 = twin_analysis(tcr_info)
```

```
## 'summarise()' has grouped output by 'COMBAT_ID'. You can override using the
## '.groups' argument.
```

```
##
## Exact binomial test
##
## data: sum(clustdf$obs) and nrow(tt)/2
## number of successes = 80, number of trials = 115, p-value < 2.2e-16
## alternative hypothesis: true probability of success is greater than 0.2165559
## 95 percent confidence interval:
## 0.6174432 1.0000000
## sample estimates:
## probability of success
## 0.6956522
##
## [1] 6.137422e-28
```

```
F1E = twin_res1$lineplot  
F1E
```

