Get Interpatient Clonogroups

Lincoln Harris

9.12.19

Which clonogroups have cells from multiple patients

read in Ashley's cDNA plate metadata csv

```
cDNA_plate_metadata <- read.csv("../rawdata/cDNA_plate_metadata.csv", header = T)
```

read in TCR_summary.txt from TRACER output note that TCR_summary.txt has had all header lines removed

```
tcr_summary_df <- read.csv("../TCR_analysis/filtered_TCRAB_summary/TCR_summary1.txt", header = F, na.st:
```

define new df to store interpatient matches

```
matches_df <- data.frame(matrix(ncol = 3, nrow = 1))
colnames(matches_df) <- c('Clonal_group', 'Num_patients', 'Group_size')
```

Show me the clonogroups that have cells from multiple patients
for loop does the trick here

```
for(i in 1:nrow(tcr_summary_df)){
  curr_clonogroup <- tcr_summary_df[i,]
  group_size <- length(which(!(is.na(tcr_summary_df[i,]))))
  patientsList <- c()
  for(j in 1:length(curr_clonogroup)){
    curr_cell <- tcr_summary_df[i,j]
    curr_cell <- as.character(curr_cell)
    curr_plate <- strsplit(curr_cell, "_")[[1]][2]
    curr_index <- which(cDNA_plate_metadata$plate == curr_plate)
    curr_patient <- cDNA_plate_metadata$patient_id[curr_index]
    patientsList <- c(patientsList, curr_patient)
  }
  if(length(unique(patientsList)) > 1){
    currLen <- length(unique(patientsList))
    clonalGroupNum <- i
    rowToAdd <- c(clonalGroupNum, currLen, group_size)
    matches_df <- rbind(matches_df, rowToAdd)
  }
}
```

```
matches_df
```

```
  Clonal_group Num_patients Group_size
1           NA           NA         NA
2            2            2         32
3            4            2          2
4            8            2          2
5           22            2          2
6           59            3          3
7           61            2          2
8           85            2          5
9           89            2          4
```

```
10          90          2          2
11          96          2          2
12         110          2          2
13         119          2          2
14         126          2          2
15         137          2          2
16         140          2          2
17         146          2          3
18         164          2          3
19         168          3          3
```

```r
matches_df1 <- matches_df[-1,]
print(matches_df1, row.names = F)
```

```
 Clonal_group Num_patients Group_size
            2            2         32
            4            2          2
            8            2          2
           22            2          2
           59            3          3
           61            2          2
           85            2          5
           89            2          4
           90            2          2
           96            2          2
          110            2          2
          119            2          2
          126            2          2
          137            2          2
          140            2          2
          146            2          3
          164            2          3
          168            3          3
```