20240627\_Matt\_Dufort\_notes\_on\_hashtags

1. It would not surprise me if the population with lots of cells is actually the CD3/CD28 stimulation, rather than the islet-specific cells as currently labeled.  I would very much expect to get more cells from that stimulation than from islet peptides or CEFX.
   1. A: The islet and CD3/CD28 hashtags were swapped.
2. I use three different transformations of the hashtag counts (after splitting them out from the antibody tags - I usually put the hashtags as a separate assay in Seurat). First, just log (count + 1).  Second, log (cpm + 1).  Third, a variant of centered log-ratio (CLR), calculated across hashtags within each subject.  That third one is a little more complicated and doesn’t often outperform the others, so I think we should skip it for now.  
   Also, CLR can be calculated two different ways - by marker across cells, or by cell across markers.  I think the way Seurat does it is the first (meaning they isolate each marker, and calculate the geometric mean of that marker in each cell), which I don’t think makes as much sense at least for the hashtags.
3. I plot the histograms of those values for each hashtag in each sample/pool, to see which measure gives the cleanest bimodal distribution.  Most often it’s logCPM.  I also use those plots to determine the thresholds for positive/negative for each.
4. I set thresholds for positive/negative for each hashtag.  It’s much easier to see the peaks and valleys if you limit the y-axis.  For most human samples, I have had to set separate thresholds for the different hashtags.  In some of the early experiments, I actually set different thresholds for each hashtag in each sample due to variation in staining and sequencing.  From looking at your plots with the current data, I would increase the threshold for CD3/CD28, and decrease the thresholds slightly for CEFX- and islet-specific.
5. If the hashtags prove to be correct, I would search common TCRs from the islet-specific for hits elsewhere.  It strikes me as unlikely to get THAT many islet-specific cells.