1.     Raw sequencing data (fastq format): Sorry, but the raw sequencing data is as BAM files. Is this okay? The BAM files were uploaded in the shared Dropbox folder.

a.     WT\_TFH

b.     PRDM1KO\_TFH

c.      WT\_TH1

d.     BCL6KO\_TH1

e.     DKO\_TH1

f.      DKO\_TFH

g.     WT\_Naive

h.     Optional: different KO naïve: I don’t have KO naïve samples.

i.       Optional: names of papers in which there are data sets you would like to use: Is this question for TF motif data set? If so, I would like to use the same data set from Dapeng’s 2018 immunity paper (maybe Homer?) for Anlaysis 1-(1). For Analysis 1-(2), I would like to combine as many as TF data sets including Homer, curated CisBP and JASPAR + others if you have any further idea. I attached curated CisBP and JASPAR motif data sets as .meme file. The references for these are Bingfei’s 2017 NI paper and Justin’s 2017 Nature paper.

j.       Optional: Anything other sample you think would be helpful for the analysis

2.     Sample description

a.     Library preparation method: Nuclei were isolated from 5x104 of SM CD4 T cells from each KO mouse. Tagmented DNA were purified by MinElute PCR purification kit and subjected to PCR using indexing primers (dual indexing, Ad1.1-1.21 and Ad2.1-2.21) by KAPA Real-Time Library amplification kit (cycles 12). PCR amplicons were purified using AmPureXP beads, and fragment size-distribution were quantified by TapeStation. High throughput Sequencing were performed by HiSeq4000 (Paired end; 100 cycles).

b.     Sample names and biological replicates: I generated this in a separate xls file. The ATAC-Seq were performed by 3 bilogical replicates.

c.      Reference genome (I suppose Mus Musclus - mm10): Yes, mm10.

3.     Gene names of different groups you would like to use for subsetting: As for Analysis 1-(3) and Analysis 3, I attached Tfh-associated genes, Th1-associated genes, Group I, II, III and Group IV gene list as xls file. The analysis priority will be Tfh-associated genes > Group IV genes > Group II genes > Th1-associated genes, Group I and Group III genes.

4.     Optional: corresponding RNA-seq data: I have multiple data types. I can give you normalized count data from DEseq output and also TPM data. These data (.tsv) are uploaded in the shared Dropbox folder. Do you need pre-filtered data? Because some genes has only zero count or less values than rarely expressed genes in CD4 T cells, e.g. CD8 and CD19. I also have pre-filtered normalized count data. I will upload it if you need it.

(Fastq files and sample description. If you have analysis results such as transcript count, TPM and DEseq output, we can use those too.)