

Supplementary material

Note that the rest of analyses and results generated in this project can be found online on GitHub.

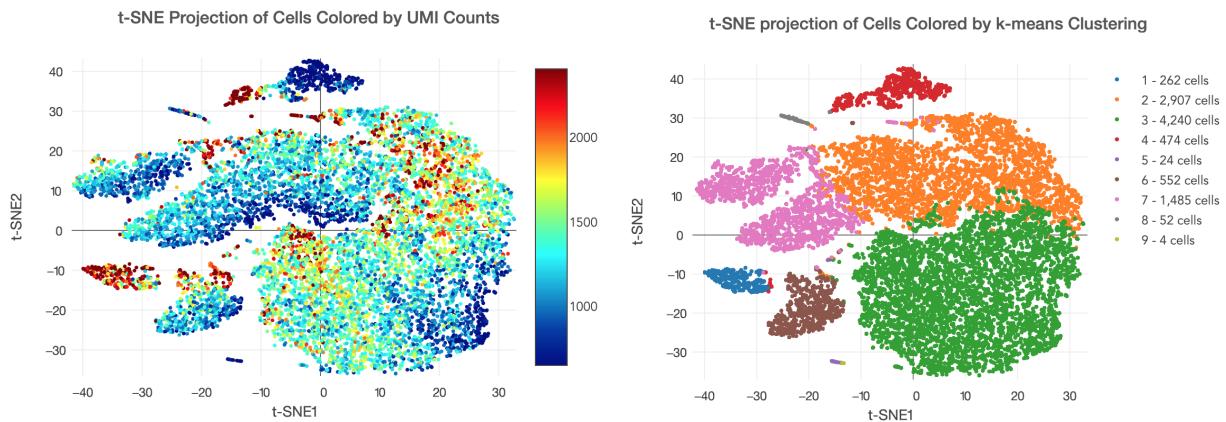


Fig. S1: t-SNE projections of the 10x-68k-PMBC data set, in the left colored by UMI counts and in the right, it shows the 68k cells clustered by K-means clustering. k numbers of clusters as 10. The plots are extracted from the summary released in 10X Genomics website.

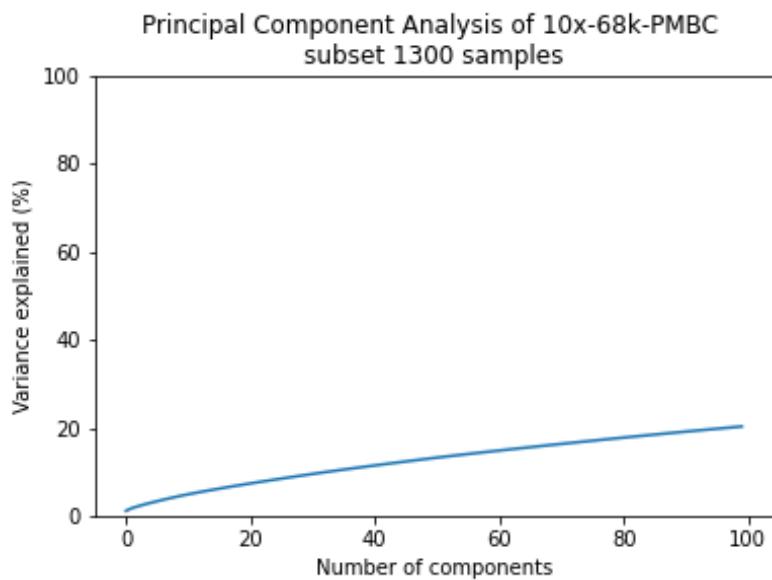


Fig. S2: Variance explained over k components resulting from applying PCA to the 2% subset from 10x-68k-PMBC data.

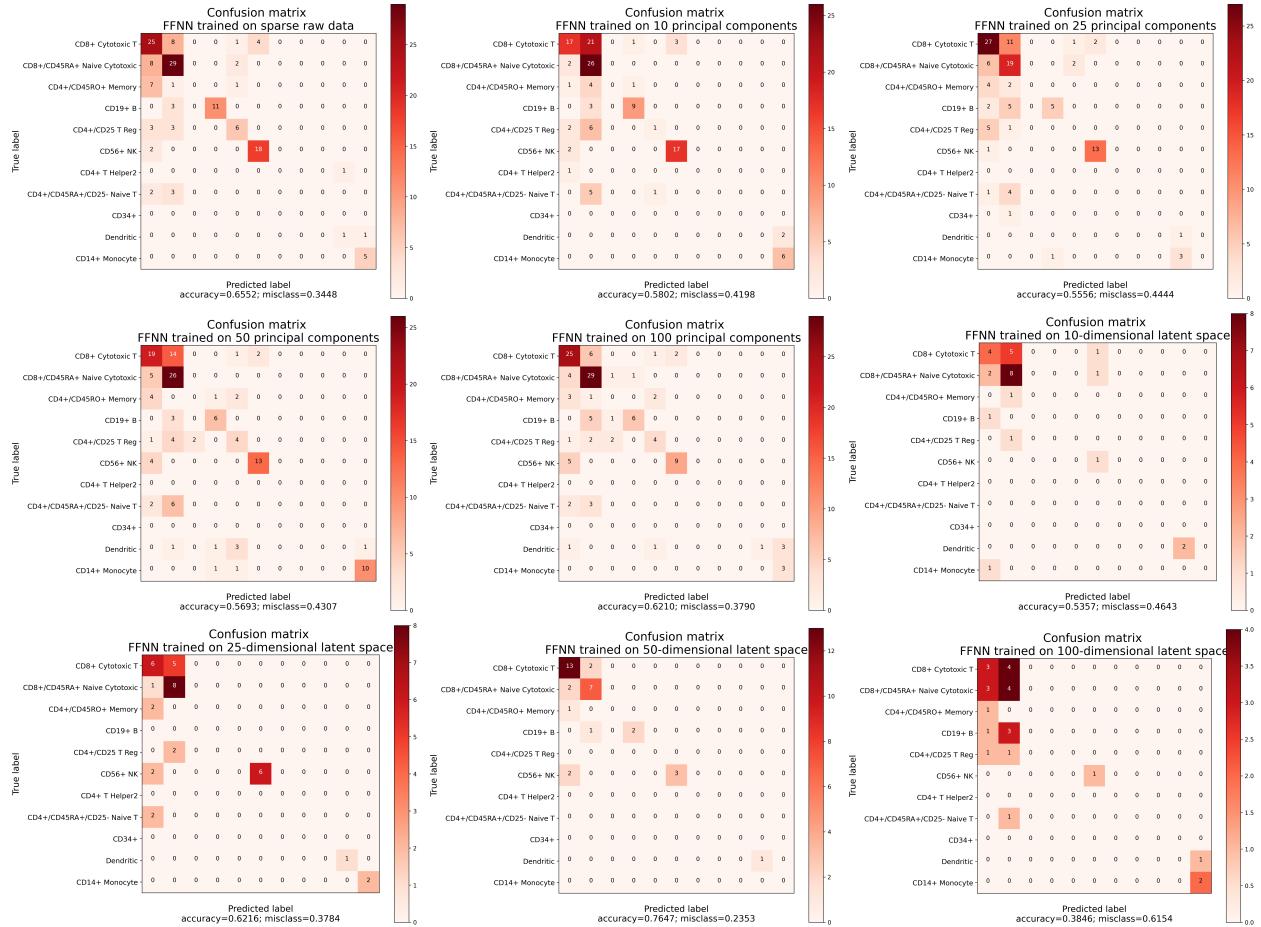


Fig. S3: Confusion matrices from each FFNN model test set for the subset with 1,300 cells.

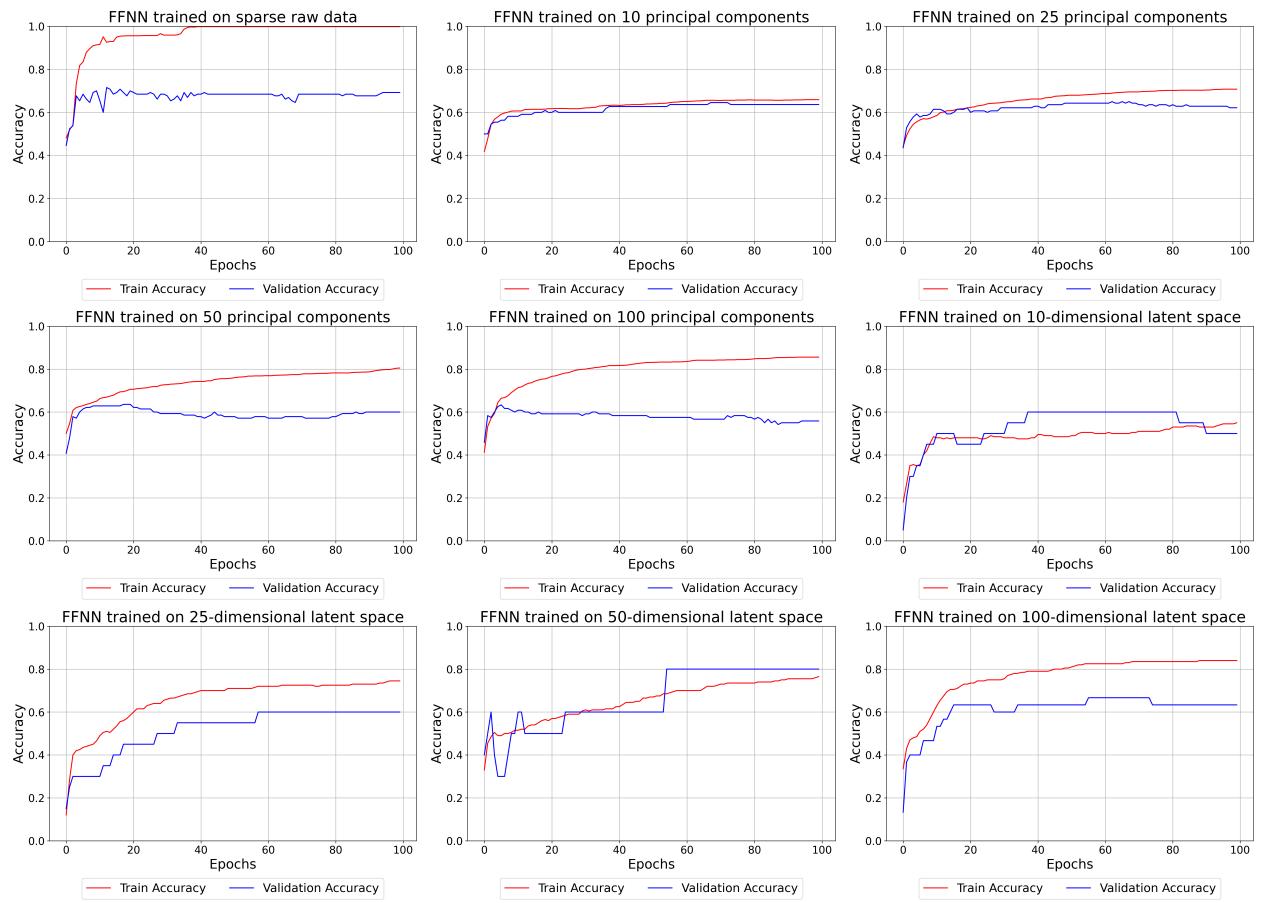


Fig. S4: Learning curves from each FFNN model for the subset with 1,300 cells.

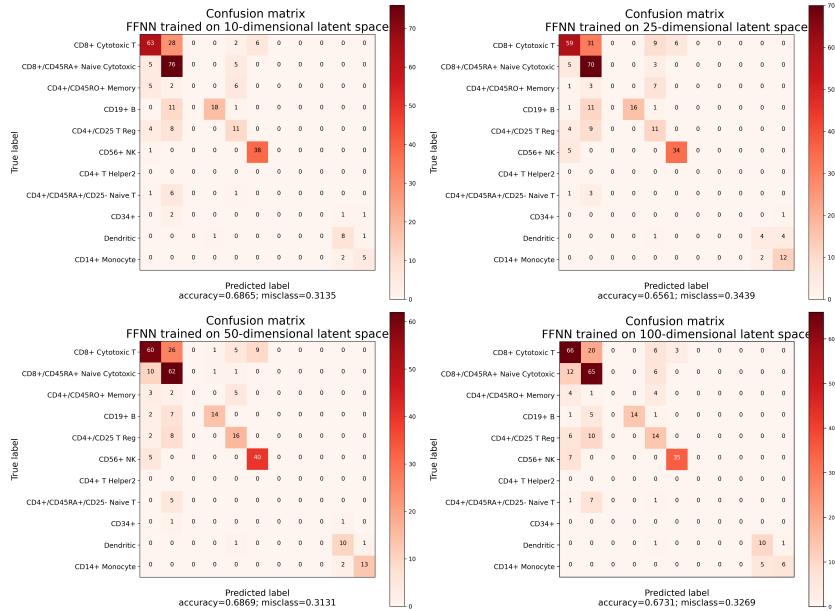


Fig. S5: Confusion matrices from each FFNN model test set for the subset with 15,000 cells.

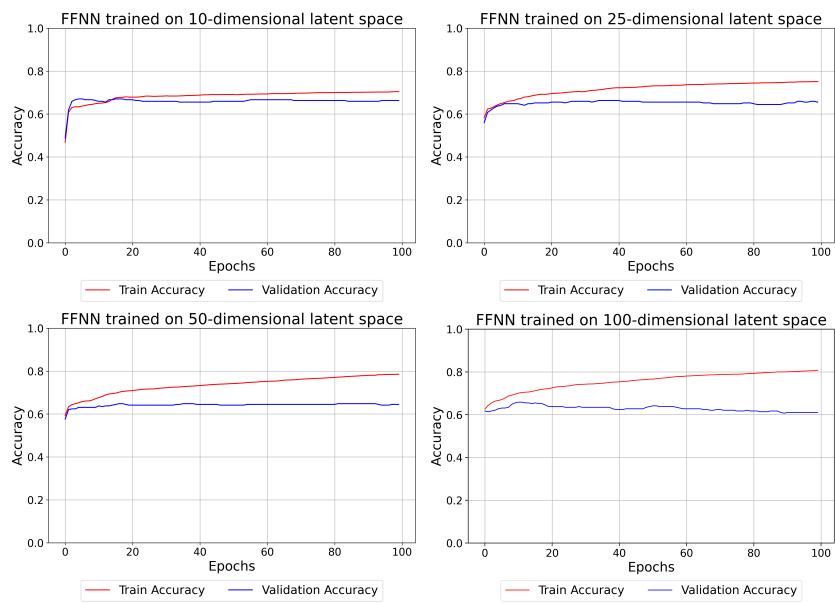


Fig. S6: Learning curves from each FFNN model for the subset with 15,000 cells.

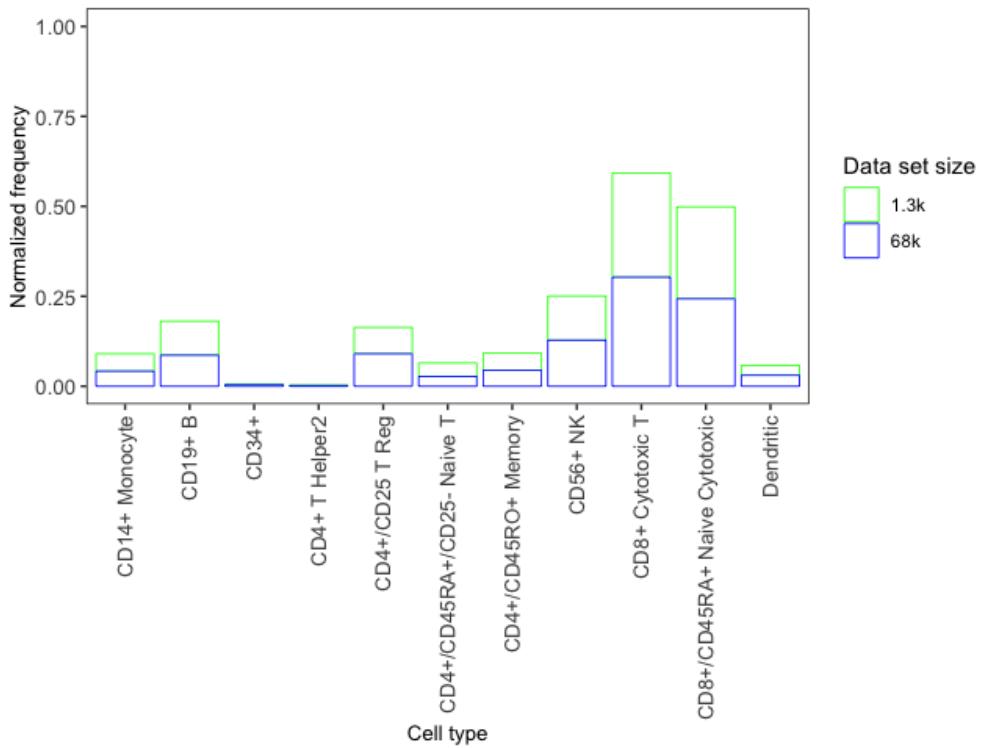


Fig. S7: Normalized number of cells per cell type in both the entire 10X-PBMC-68k data set and the 2% subset that was used in the study.