Background

A major advancement in cancer immunotherapy is adoptive transfer of antigen-specific T cells. To get an effective response, the two factors that play a big role are the number of transferred T cells and their differentiation state. Interleukin-2, IL-2, is a big interest in cancer immunotherapy because it was discovered as a T-cell growth factor. Il-2 can promote the expansion of T cells, but it also drives terminal differentiation and causes the expression of co-inhibitory receptors like TIM-3 and PD-3. TIM-3 and PD-3 are hallmarks of T cell exhaustion in chronic viral infection and cancer (Mo et al. 1). The goal is to be able to get the desirable properties of IL-2 whiling minimizing the negative ones. The paper shows that H9T, which is an engineered interleukin-2 partial agonist which promotes $CD8^+$ T cells with out terminal differentiation. Engineering cytokine variants with distinctive properties is a promising strategy for forming new molecules with translational potential. In figure 1 data was focused on the differential effects of H9T versus H9 and Il-2 on $CD8^+$ cells. The data used to make the four subplots (figure 1 c-f) in figure 1 was source data that investigated the expansion fold and percent TIM-3 for H9T versus H9 and Il-2. For the second figure the data that was used was ATAC seq data that was from mouse $CD8^+$ T cells expanded for six days (Mo et al. 1)

Data Analysis and Structure

For figure 1, each of the subplots contained different csv files. For figure 1c, I prepared the data by grouping the columns together. The data contained information for the for each factor: IL-2, H9T, and H9. The first grouped together each of the columns that corresponded to each factor by creating a new variable and calling all the columns that attribute with IL-2 for example and then repeated the same process for H9T and H9. Next, for each of the new variables the mean was taken along axis 1 because I wanted to find the mean of the data based on the rows, so I would have the mean expansion fold for each of the times. Then I calculated the standard error for each so then I can add in the error bars in the plot. Then using mat plot libs plt.errorbar the means and standard errors for Il2, H9, and H9T were plotted. The same process was used to create figure 1e, but instead of using expansion fold data the plot was looking into TIM-3 %. Also, for figure 1b, time was not on the x-axis like how it was for figure 1c, it was concentration so the log10 of the concretion data had to be taken using numpy. For plots 1d and 1e, a slightly different process was taken. The process was the same for grouping the data together based on if it was IL-2, H9, or H9T data, taking the means and standard error. The difference in the process was that figures 1d and 1e contain line of best fits in their figures. Starting with figure 1d, the means and standard error for the points were already plotted. Figure 1d was similar with figure 1c in that the x-axis is not time but concentration so log10 had to be taken of the data to match the figures. The issue was finding the correct line of best fit to use on the data. Originally, a lot of the code that I had tried focused more on a linear line of best fit which was not correct because the data is not a linear relationship, it has a curve to it. When looking at the original figure to mean it looked like it was a log curve. I tried creating code where I wrote out the log function of a line and calculated the residuals and then tried to use scipy.optimize but I ran into trouble that it was not giving the lines(they would either be straight or not code at all) I needed or would not go in negative values. So, I decided to then try just plotting a polynomic curve based on the data. To do this I used np.poly fit on the data and set it to a degree of 2. Then to make my x I used np.linespace and then used np.polyval to evaluate the polynomial at specific values where I put my x and the results from np.poly fit in. Then I just repeated the same process for the other two lines and plots them on the data. They are not perfect

to the original, but they are close there is more of a curve to these lines because it was a polynomial, but after consulting with the professor I was told they don't have to look perfect and to try my best and it was fine to leave it as it as long as there are similarities in the shape. The same process was used for figure 1f, the only difference was that the way the H9T points plotted on the graph, it resembled more of a linear relationship so I used np.polyfit again but with a degree of 1 because it was linear and plotted a line of best fit based on the m and b values it gave me using mx + b to calculate y.

For figure 2h, the ATAC seq dataset was loaded into the notebook. I prepared the data by removing all the columns in the file but the first four columns which were: normalized. D0, normalized.D6.IL2, normalized. D6.H9, and normalized D6.H9T. The rest of the columns were removed because they were not relevant to the PCA plot, they were data for other subfigures in the paper. The next step was to transform the data so that the shape of the data goes from 55200 rows and 4 columns to 4 rows and 55200 columns. Based on how the data was and it was labelled, the data was already normalized but a log transformation was still needed to be performed. Before a log transformation is preformed, I created a nest for loop that goes through the range of all the columns and then all the rows so that I can check if anywhere in the data there are zero values. If there was a zero value in the data, I consulted with the professor, and he agreed that the value should just be changed to a 1 because the paper did not explicitly say what they did with the values. The next step is do the log transformation on the data, the paper did not explicitly say what log transformation what preformed so with the professor we performed a log10, log2, and log20 transformation and were never able to achieve the same axis's as the plot had, so the professor can me permission to just leave it at a log2 transformation. The next step was to find the pca components using sklearn and I did a pca for two components. The variance explained by each of the pca components ended up being 77% for the first principal component and 19.2% in the second pca component, which is alittle off from the original figure because it was 83.4% explained variance for the first principal component and 15.4% for the second. After consulting with the professor, he said that this was ok because the values I got were close enough to the original figure and the difference could be due to our uncertainty of what log transformation they performed with the data. The next part in my data was to plot the two principal components and color each point approximately. When the data was plotted, all the points were in the correct spots and colored correctly but as we knew before the axis were off by a factor of about 100 and it was a mirror image of the pca plot that was in the figure. After consulting with the professor again, we decided that since the placement of the points and the shape of the pca were similar with the original that it was appropriate and met expectations for this project to leave it like that. The only change that was made was that we flipped the x-axis from going from -200 to 300 to go from 300 to -200 to show that the shape and point placements are the same as the original. The only issue is the axis's but the paper did not explain on what log transformation they performed which is why the axis's are off, but it was checked with the professor for us to come to that conclusion. The last thing that was added to the code was adding annotation marks to each of the points so they can easily be identified in the plot.

Conclusion

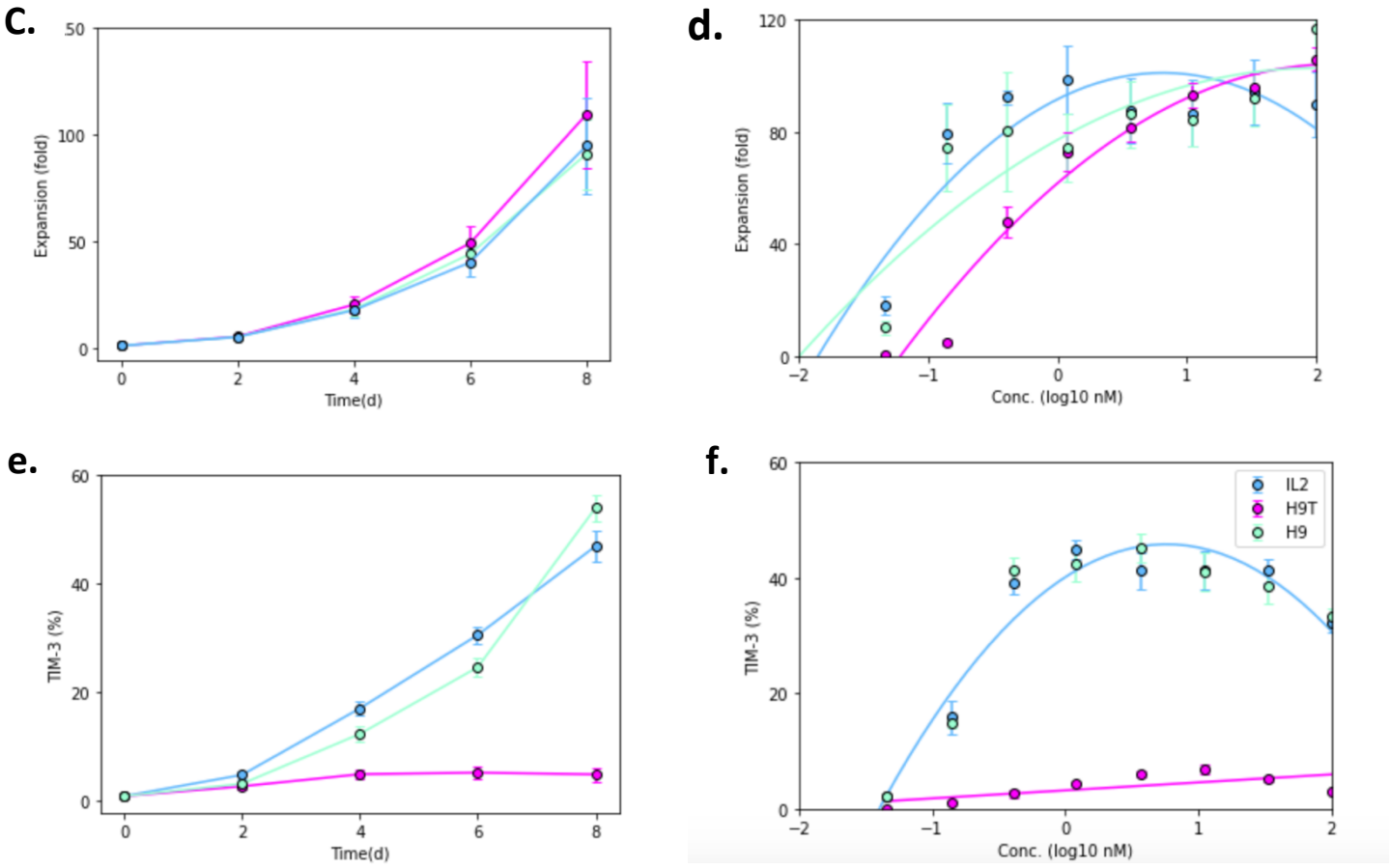
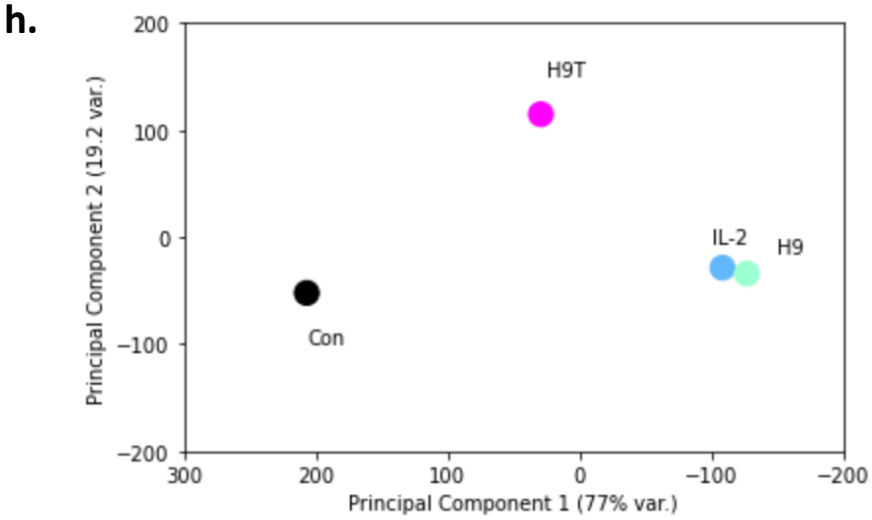


Figure 1 sought to understand the differential effects of H9T versus IL-2 and H9 on $CD8^+$ T cells. Figure 1c was looking into the expansion of mouse pre-activated $CD8^+$ cells with 10nM of the three cytokines: IL-2, H9, and H9T using eight mice. In figure 1d they were looking into the expansion of mouse pre-activated $CD8^+$T cells cultured with a range of concentrations of the three cytokines for eight days. The experimenters used three mice for each cytokine to perform the experiment. The conclusion from these first two graphs is that the H9T could expand the activated mouse $CD8^+$T cells like IL-2 and H9. When looking at both plots the lines are very close together especially in figure 1c, there is not a significant difference present. In figure 1d the three lines show a similar trend, no drastic difference they only difference is that H9T required a higher concentration to reach maximal proliferation. When looking at the figures IL2 and H9 reach their high point and begin to flatten out when H9T hadn't quite reached the leveling off period yet. For figures 1e and 1f the goal was to look of the expression of TIM-3%. In figure 1e the researchers were looking into the expression of TIM-3% from mouse pre-activated $CD8^+$ cells with 10 nM of each of the three cytokines using an n = 4 mice. In figure 1f the researchers were looking at the expression of TIM-3% from the mouse pre-activated $CD8^+$ cells with a range of concentrations of the three cytokines for eight days using an n = 3 mice. The conclusion from these two graphs is that H9T made significantly lower expression of TIM-3 than IL-2 and H9. It was previously known that IL-2 induces the expression of co-inhibitory receptors like TIM-3 so it made sense that IL-2 created more TIM-3% as the concentration of IL-2 increased. H9 is a IL-2 superkine, so it makes sense that they act in similar ways and both produce TIM-3. It is an

important result that H9T does not create a lot of TIM-3 because TIM-3 is a hallmark of T cell exhaustion, and it could provide hope for H9T becoming a promising cytokine for cancer immunotherapies.



For figure 2, the goal was to better understand the transcriptomic differences, so an assay was performed for transposase-accessible chromatin using ATAC-seq. Figure 2h, preforms a PCA analysis on the ATAC-seq peaks of mouse $CD8^+$ cells expanded for six days. The black dot that is labelled "Con" is the control baseline from day 0, the other three dots was the normalized day 6 results. When looking at the PCA plot it makes sense that Il-2 and H9 were grouped closer together because they had similar results to expansion fold and TIM%. Additionally, it makes sense that H9T is in a group of its own because it produces different results in the experiments for example like TIM-3 expression, so it makes sense that is in a different group then Il-2 and H9. The conclusion the researchers reached for ATAC-seq data is that there was distinct chromatin accessibility at the *Tcf7* and *Havcr2* loci in the cells that were cultured in the H9T in comparison to Il-2 and H9, which also makes sense based on the distinct groupings in the PCA plot.

A potential interesting direction for the study is to understand more how H9T can be used in cancer immunotherapies, bring a clinical aspect into the research. It would also be interesting to investigate other new cytokine variants that are engineered with distinctive properties and translational potential and see how it will affect cancer therapies and how many doors it could open.

Overall, the results demonstrate that by engineering and changing the signals of IL-2 researchers can create H9T that can support T cell proliferation while limiting glycolysis and exhaustion.

References

1. Mo, F., Yu, Z., Li, P. *et al.* An engineered IL-2 partial agonist promotes $CD8^+$ T cell stemness. *Nature* **597,** 544–548 (2021). https://doi.org/10.1038/s41586-021-03861-0 https://www.nature.com/articles/s41586-021-03861-0#Sec32