Technology for Single Cell Protein Analysis in Immunology and Cancer Prognostics

Thesis by
Alexander Muir Sutherland

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2016

Defended June 11, 2015

Acknowledgements

I would like to thank my advisor, Professor James Heath, for giving me the opportunity to work on such amazing research projects. We were always at the cutting edge and that made every day exciting to come into work. I want to thank him for helping me grow as a scientist during my time at Caltech and guiding me down a path towards exciting discoveries.

I thank Professor David Tirrell for being both an insightful committee member, as well as a supportive figure in my time at Caltech. I thank Professor Sarkis Mazmanian, for being a great committee member and inspiring my study of immunology. I cannot remember a teacher who spoke with such ease, and deep knowledge, about a subject. Finally, I thank Professor Mark Davis for being a perceptive and knowledgeable member of my committee. I enjoyed our conversations about diagnostics (and Formula 1 aerodynamics).

A hearty thanks to those who kept the Heath Lab running: Elyse Garlock, Sherry Mathis, Kevin Kan, and Katrine Museth. A special thanks to Diane Robinson, who was so supportive and fun both inside and outside of the lab. Thank you to everyone at UCLA: Dr. Antoni Ribas, Dr. Timothy Cloughesy, Dr. Begonya Comin-Andoix, Iris Williams, and Jeff Calimim.

Dr. Habib Ahmad – you were not only a great mentor in lab but also outside of lab. Your kindness has molded my behavior to become a best person. Thank you to Dr. Jing Zhou, Dr. Jing Yu, Dr. Udi Vermesh, Dr. Ophir Vermesh, Professor Jun Wang, and Dr. Chao Ma for being great partners at the lab. A special thanks to Dr. Young Shik Shin for being incredibly helpful and knowledgeable about microfluidics.

Dr. Joseph Varghese, Ryan Henning, Ashwin Ram, Marilena Dimotsantou, JingXin Liang, Dr. Kaycie Deyle, and Dr. Jessica Pfeilsticker – you were all amazing friends. Thanks do not begin to

acknowledge how you all affected me. Julian Claudino, Chris Kelly, and Evan Fisher – you guys are great friends. Enough said. I want to thank Kailyn FitzGerald for being a big part of my time here at Caltech and for being supportive. Everyone else – Dr. John Nagarah, Dr. Ann Cheung, David Bunck, Mike Grodick, Adam Boynton, James Blakemore, Dr. Min Xue, Roshan Agrawal, So-Young Kim, Julia Sun, Dr. Songming Peng, Yapeng Su, Dr. Wei Wei, Blake Farrow, and anyone else I forgot – thanks for keeping life fun at Caltech.

Finally, I would like to thank my family. I cannot begin to thank you for your love and support. Dad, when I talk with you, I am reminded of why I fell in love with science. Your curiosity continues to inspire me to be the best scientist I can. Mom, your undying support has always kept my chin up during my time at Caltech. You were my personal cheerleader and helped me reach this point in more ways than I can describe. Kat Wolpin, you stuck by me and reminded me what family was all about. I love you all.

Abstract

The first chapter of this thesis deals with automating data gathering for single cell microfluidic tests. The programs developed saved significant amounts of time with no loss in accuracy. The technology from this chapter was applied to experiments in both Chapters 4 and 5.

The second chapter describes the use of statistical learning to prognose if an anti-angiogenic drug (Bevacizumab) would successfully treat a glioblastoma multiforme tumor. This was conducted by first measuring protein levels from 92 blood samples using the DNA-encoded antibody library platform. This allowed the measure of 35 different proteins per sample, with comparable sensitivity to ELISA. Two statistical learning models were developed in order to predict whether the treatment would succeed. The first, logistic regression, predicted with 85% accuracy and an AUC of 0.901 using a five protein panel. These five proteins were statistically significant predictors and gave insight into the mechanism behind anti-angiogenic success/failure. The second model, an ensemble model of logistic regression, kNN, and random forest, predicted with a slightly higher accuracy of 87%.

The third chapter details the development of a photocleavable conjugate that multiplexed cell surface detection in microfluidic devices. The method successfully detected streptavidin on coated beads with 92% positive predictive rate. Furthermore, chambers with 0, 1, 2, and 3+ beads were statistically distinguishable. The method was then used to detect CD3 on Jurkat T cells, yielding a positive predictive rate of 49% and false positive rate of 0%.

The fourth chapter talks about the use of measuring T cell polyfunctionality in order to predict whether a patient will succeed an adoptive T cells transfer therapy. In 15 patients, we measured 10 proteins from individual T cells (~300 cells per patient). The polyfunctional strength index was calculated, which was then correlated with the patient's progress free survival (PFS) time. 52 other

parameters measured in the single cell test were correlated with the PFS. No statistical correlator has been determined, however, and more data is necessary to reach a conclusion.

Finally, the fifth chapter talks about the interactions between T cells and how that affects their protein secretion. It was observed that T cells in direct contact selectively enhance their protein secretion, in some cases by over 5 fold. This occurred for Granzyme B, Perforin, CCL4, TNFa, and IFNg. IL-10 was shown to decrease slightly upon contact. This phenomenon held true for T cells from all patients tested (n=8). Using single cell data, the theoretical protein secretion frequency was calculated for two cells and then compared to the observed rate of secretion for both two cells not in contact, and two cells in contact. In over 90% of cases, the theoretical protein secretion rate matched that of two cells not in contact.

Table of Contents

Chapter 1: Automated Data Gathering in Single Cell Microfluidic Protein Assays

1.1 - Introduction

Microfluidic devices show great promise in advancing the proteomics field [1, 2, 3]. The miniaturization of proteomeic tests allows for the generation of large number of statistics. One caveat of increasing the number of statistics is the increased amount of man-hours necessary to gather and analyze the data. For example, analyzing four of the Heath Lab's single cell (SC) microfluidic proteomic tests per week requires approximately 14.5 hours. Furthermore, as the number of microfluidic chambers increases, the probability of human error increases as well. These random errors are difficult to detect in downstream analysis. Therefore, a method to automate SC data gathering and analysis would both save time and reduce random errors.

Computer vision has opened the doors to new technologies, with a diverse set of applications, from self-driving cars to cancer diagnostics [4]. Huang et al. used computer vision to image cells and related the cell morphology to the patient's survival time [5]. Other groups have utilized computer vision for more simple purposes; for instance, Loukas et al. have generated algorithms to determine with high accuracy if an element in a microscope image is a cell or not [6] This technology is desirable in microfluidics, as it would allow for the automation of cell counting.

Using computer vision to automate the microfluidic devices is the subject of this chapter. There are two tasks to automating the data gathering and analysis of single cell microfluidic tests: (1) automating fluorescence signal gathering from the ELISA-like test and (2) automating cell counting.

The protein fluorescence signal from the Heath Lab's SC tests can easily be applied to computer vision algorithms because the devices are organized in a regularly spaced matrix with multiple reference

markers. In contrast, automating our cell counting is more difficult because video files are used [7]. Analyzing cells from video files requires specialized software, since features must be tracked between frames [8]. Since this method of cell counting with video files was not readily adapted towards automation, a method was developed in-house to create a full-view image of the single cell microfluidic device, using image-stitching. These files were the equivalent of 500 microscope images stitched together to give a full-view of the microfluidic device's chambers. This image can then be used in an automated cell counting program.

With a full-view of the cells and fluorescence signal in the microfluidic device, the two automating tasks share a common goal: find and index each single cell chamber. From there, the automated cell counting algorithm must use computer vision to identify cells to obtain a cell count in the chamber, which has been demonstrated previously [9, 10]. Whereas the automated fluorescence signal analysis algorithm must locate the different fluorescent protein signals and quantify them. After these specific tasks, each program must then organize the gathered data into a readable format, suitable for downstream statistical analysis.

These two algorithms have not only saved time but also allowed us to perform tests that would previously be considered too cumbersome to perform and analyze. The complex experiment performed in Chapter 5.2.6 (T cell communication project) was made possible using this technology.

1.2 – Materials and Methods

1.2.1 – Automated Cell Counter

1.2.1.1 – Microfluidic Device

The polydimethylsiloxane (PDMS) microfluidic device was created using a special aluminum casing, in order to make the PDMS' top surface flat. The schematic for the device is shown in Figure 1. The wafer (colored blue) with the microfluidic pattern is sandwiched between two aluminum blocks.

These blocks are fastened together with screws. A 10:1 ratio of A:B Sylgard 184 PDMS (Dow Corning) is mixed and poured into the top aluminum block's window. The aluminum block casing is then placed in a vacuum chamber for 1 hour to remove bubbles from the PDMS. Next, flat plastic pieces are placed overtop the window of the top aluminum block, and the entire casing is placed in an 80 °C oven for 20-25 minutes. A second, thin layer of PDMS is made concurrently, as described in Chapter 4.2.1. The PDMS is removed from the oven and allowed to cool for 15 minutes. The cured PDMS is then removed from the aluminum casing, its control-valve holes are punched, and then it is adhered to the second, thin layer of PDMS. This is baked for 1.5 hours at 80 °C. Holes are then punched for flow-channels. The device is then adhered to a DEAL barcode substrate and baked for 2 hours at 80 °C. The device is then stored in a dessicator (< 20% humidity) for a maximum of one week.
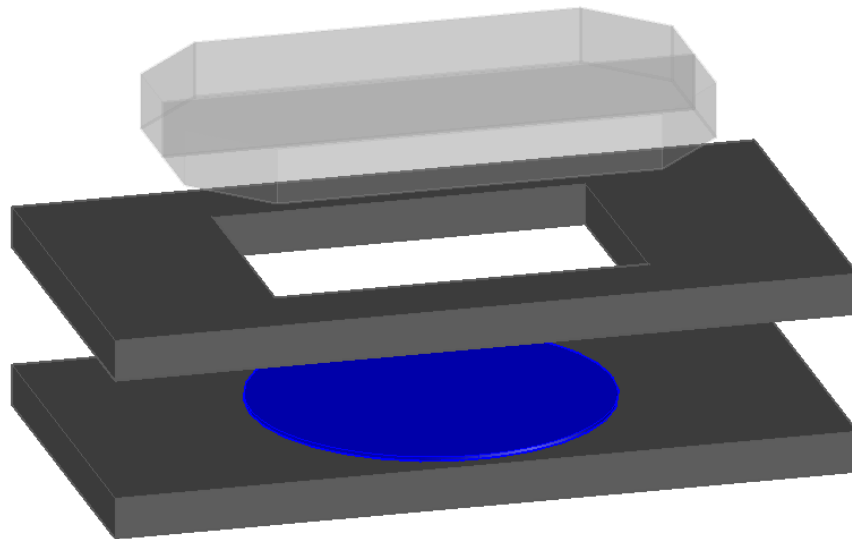

Figure 1: 3D model of aluminum casing to make flat-top PDMS. The wafer is colored blue and is placed in between the two aluminum blocks (colored dark-gray). Then plastic pieces are placed overtop the aluminum block window (colored translucent gray).

To perform automated computer vision cell counting, a full-view image of the microfluidic device must be generated. This was accomplished by using the "Stage Overview" function of the Olympus IX81 CellSens software.

The microfluidic device was secured on the stage of the Olympus IX81 and multiple tasks were performed to ensure the image quality would be sufficient for subsequent computer vision algorithms: First, if the device chambers' were not run parallel with the stage's X-axis, the device was rotated to make them parallel. Second, the device was made completely flat on the stage, which was checked by focusing on the cells on one side of the device and then making sure the cells were still in focus on the other side. The stage was adjusted to ensure the device was completely flat. Finally, the lighting (brightness/contrast) from one side the device to the other was checked for any drastic differences.

The "Stage Overview" function was then selected, and bounds were set as the top-left and bottom-right of the microfluidic chambers. A stitched image was then acquired at 10x magnification and saved as a .jpeg file. The files acquired were ~500 images stitched together to generate an image with a pixel resolution of 35k by 15k.

1.2.1.3 – Computer Vision to Isolate Microfluidic Chambers

Once the full-view image of the microfluidic device was acquired, the image was loaded into a custom written Matlab program (see Appendix A for full code). The program receives user input for the following: x and y bounds of the microfluidic chambers, rotation of the slide, start of the first single cell chamber, and the optimum binary threshold to visualize the cells. This binary threshold allows the program to convert a grayscale image into a binary image. This binary image is used to help the program identify both the chambers, as well as the cells within those chambers. Figure 2 shows the general flow of the program: stitched images of microfluidic devices were acquired as shown in Figure 2a (zoomed in

for clarity). This is converted to binary using a user-defined threshold (Figures 2b). The computer plots the average intensity of the binary image by row and by column (Figure 2b), and uses this information to crop out a chamber (Figure 2c).
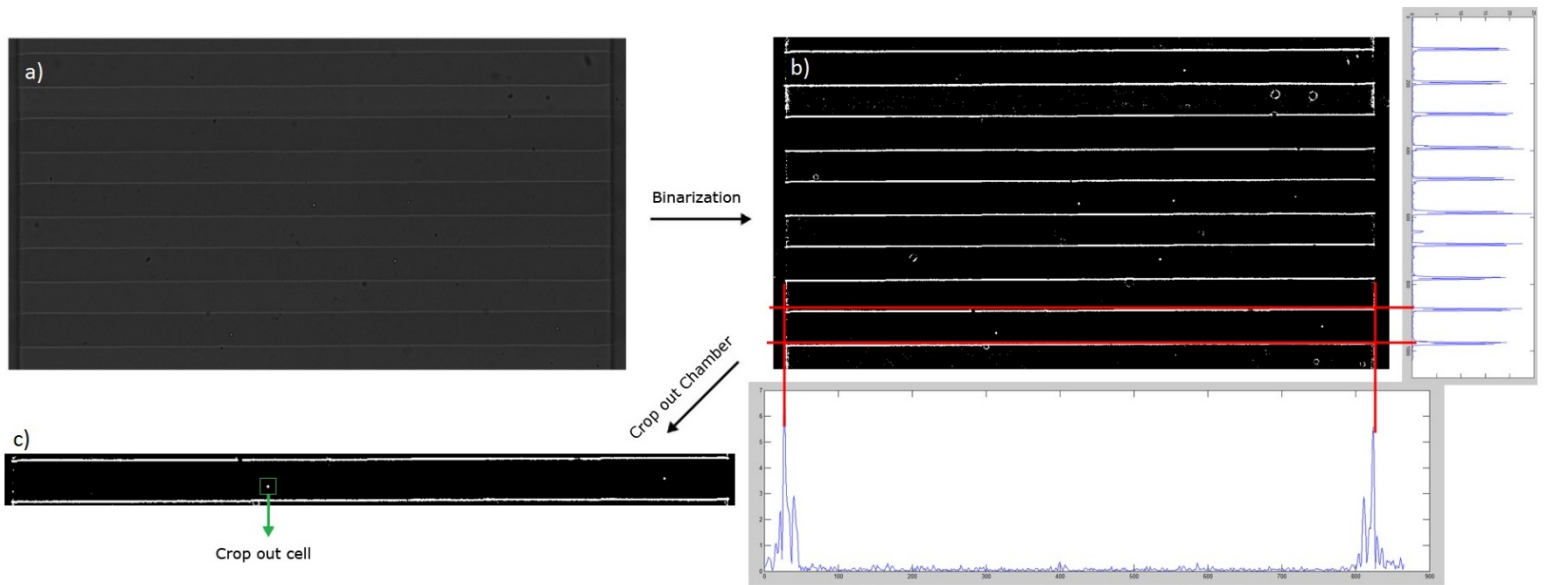
Figure 2: Automated Single Cell Detection flow of chamber cropping. a) The stitched full-view image of five chambers. b) Binary imaged of the same chambers, with plots of average intensity along both axes. c) Cropped out chamber with cell boxed in green.

1.2.1.4 – Detection of Cells in Single Cell Chambers

Once a single cell chambers is located by the program, the number of cells is determined using two cell-identification methods. First, a feature finding method identifies circular regions-of-interest (ROIs) using the binary image of the single cell chamber. If the cells size is known, one can gate these ROIs by the cell size. The ROIs found using this method are put into a list of "possible cells." Next, the program crops out these "possible cells" regions from the original grayscale image. One characteristic feature of T cells imaged on light microscopes is a distinct white nucleus/cytoplasm, with a dark membrane (Figure 3). The program takes advantage of these characteristic by taking a cross-section of

intensity from the cell (Figure 3a). This intensity graph will show one large peak, bounded by two minor troughs (Figure 3b).
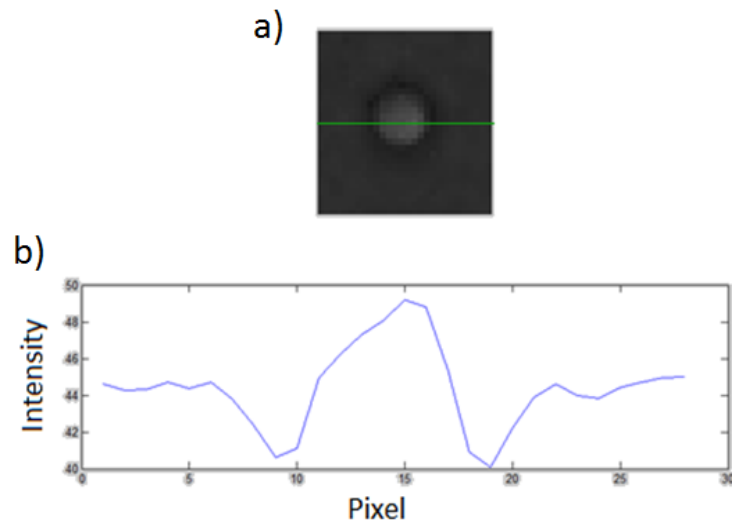
Figure 3: Cell detection method. a) Grayscale image of cropped single cell. Green line shows where the cross-section trace was taken. b) Cross-section trace of the cell in part a. The trace shows the characteristic peak in the cell's center, bordered by two troughs.

1.2.2 – Automated Genepix Analysis

1.2.2.1 – Protein Signal Gathering

The fluorescent protein signal from the single cell tests produce a multilayered .tif file. This .tif file is fed into a custom written Matlab program (full code in Appendix B).

The goal of this Matlab program is to isolate each single cell chamber, extract the protein signal data from that chamber, and organize the data for output. The general flow of gathering the protein signal from these images is shown in Figure 4. First, the user identifies the four corners of the device by clicking on the reference signals. The program uses this information to cut up the "blocks" of the
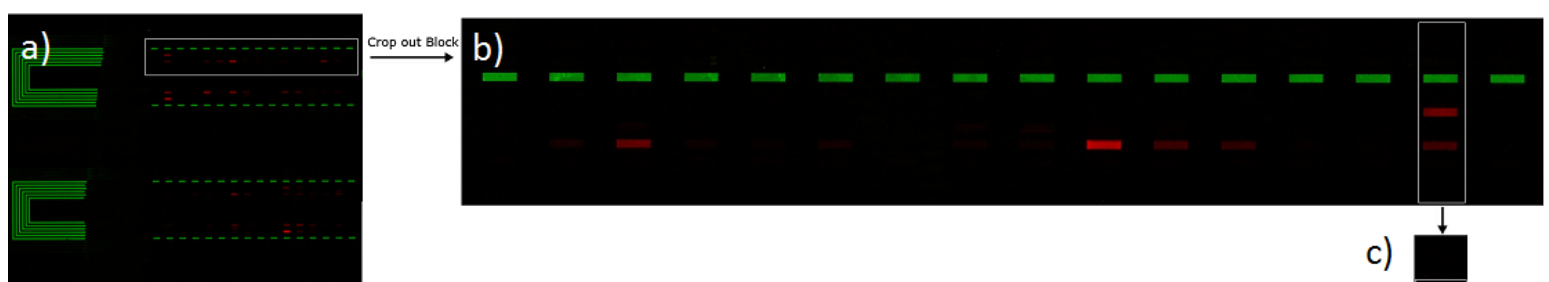
Figure 4: Flow of Automated Genepix program flow. a) The full single cell device is cropped into blocks. b) A single block with 16 features is then further chopped up into chambers. c) A single chamber is analyzed by locating the reference and then all proteins relative to that reference. The protein levels are quantified and stored in a dataset.

chamber (Figure 4b). These blocks each contain 16 single cell chambers, each with a reference signal. The program locates these reference signals using a rectangular feature finding method [regionprops(' BoundingBox')], and ensures 16 of these features exist. If 16 were not found, the program shifts the block's frame from side-to-side in order to find any feature that may have been cut off. If this fails, the program ignores the current block.

When 16 reference features are successfully found, their locations are used to isolate each single cell chamber (Figure 4c). These chambers are cropped out and stored as individual chambers objects. The user inputs an excel file with the order of proteins, as well as the location of the reference signal. This allows that program to identify the protein locations and identities in the chambers. Because the protein signals are uniformly patterned using barcodes, offset distances from the reference signal can be used to identify each protein. The protein signals are cropped out, averaged, and saved for each chamber. The program performs an outlier detection by removing any outlier pixel values using the "removeoutliers" function [11].

The automated program's output was compared to a reference dataset. This reference dataset was analyzed using the Genepix Pro "Block" arrays, using three circular features per barcode feature. This method is called the "reference" because it is the standard for data analysis in the Heath Lab. Using the program's output, as well as the reference, multiple quality metrics were calculated:

1) $\text{Average Error} = \sum_{i=0}^{n} |x_i^{Automated} - x_i^{Reference}|$

where x_i is the fluorescence signal of the output for the ith feature and n is the total number of features analyzed.

2) Percent of automated signals within range of reference (example: percent of cells within 20% of the reference signal) =

$$\sum_{i=0}^{n} \frac{\begin{cases} 1 \; if \; \dfrac{|x_i^{Automated} - x_i^{Reference}|}{x_i^{Reference}} < 0.2 \\[4mm] 0 \; if \; \dfrac{|x_i^{Automated} - x_i^{Reference}|}{x_i^{Reference}} > 0.2 \end{cases}}{n}$$

After comparing the hand-analyzed data to the automated program's data, the data that did not match between the two were investigated further. This involved using the genepix program to get the most accurate value possible by using the rectangle tool for each feature (Accurate Value). This is considered more accurate than the reference method because the accurate method uses 45% more data. Using this "accurate value" for each point, we were able to determine if the hand-analysis or automated program was closer to the true value.

1.2.2.2 – Local Background Subtraction

Because the devices had non-uniform background signals, a local background subtraction method was created to generate background values for each block down the slide. This was done by averaging the signal for proteins from zero cell chambers in its row, as well as the two neighboring rows. This was performed for all rows of the device, generating background levels for every protein per row.

1.3 – Results

1.3.1 - Automated Cell Counter

Using this program, the accuracy of cell detection was 99.7% for over 3900 data points when user input was performed in the case of questionable cells. The cells that were missed were obscured by foreign debris or were unnaturally stretched out, which negated the program's ability to measure a

clean trace of the cell's cross section. The average time required was 7±4 minutes without user input and 14±10 minutes with user input.

1.3.2 - Automated Genepix Analysis

The automated genepix program was used to analyze a total of 11 single cell device (~100,000 protein signal data points) and the average percent error was calculated at 5.8%. These errors were broken down into the magnitude of the signal for which the error occurred on: 90% of the errors occurred for signals less than 400 fluorescence intensity units, 9% occurred between 400 and 2000, and 1% were from signals greater than 20,000 fluorescence intensity units.

Furthermore, a subset of these errors were scrutinized in depth by comparing the automated and reference values to the accurate value. Of the subset of errors that were scrutinized, the automated program was closer to the accurate value 64% of the time. Therefore, it was concluded that these errors were not problematic, as the automated program was closer to the actual value more often than the reference.

A second method of accuracy was used, where the percent of signals within a certain percentage range of the reference was calculated. 99.94% of protein signals has a percent difference of 20% or less of the reference signal's value. This was tested for multiple percent differences (Figure 5).

Fraction of matching features with a certain percentage
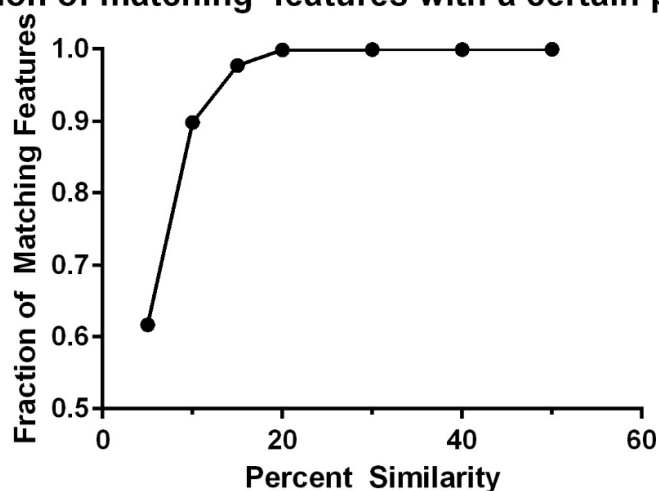
Figure 5: Graph showing the fraction of features that matched within a certain percent similarity between the Automated and Reference datasets. For instance, a fraction of 0.9 feature values matched within a 10% similar.

A representative graph of background level vs. row using the standard method and local background subtraction is shown in Figure 6. Implementation of local background subtraction improved the number of detected proteins detected as significant by an average of 29% in a full single cell test. In a photocleavable experiment to identify Jurkat cells (Chapter 3), using local background subtraction improved the fraction of cells found by 61%.
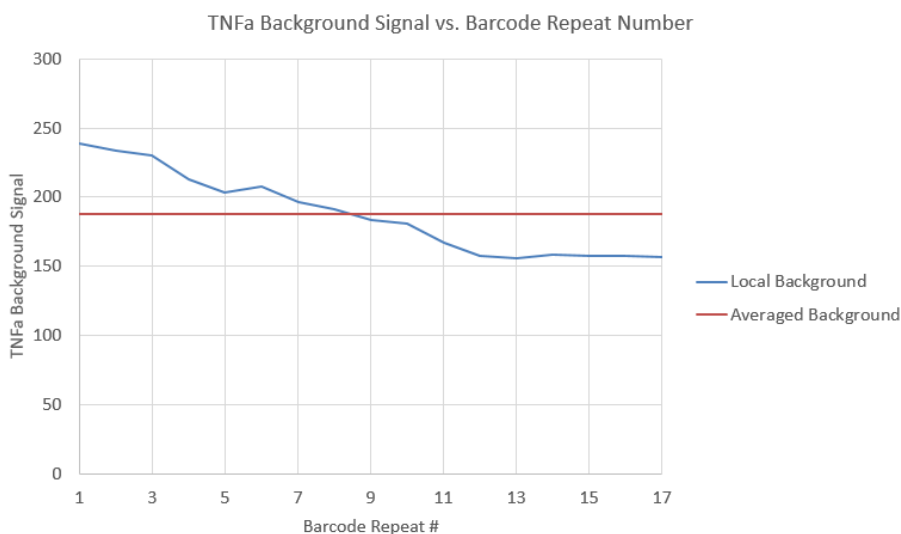
Figure 6: Background subtraction methods compared. Averaging the background for the entire device yields the red line, which stays constant for every block repeat. The blue line shows the local background value, which monotonically changes from one side of the device to the other.

1.4 – Discussion

The advantage of these programs is not only that time is saved but also that errors become systematic and it allows us to perform experiments that were previously cumbersome to analyze. For instance, performing the experiments in Chapter 5.2.6 without automated would have required approximately 4.8 hours extra processing time per device. See Table 1 for timing differences. Additionally, because a stitched image is taken instead of videos, the labor required to run the experiment is reduced, allowing the experimentalist to run multiple devices to increase statistics. Both programs have another advantage in that, in the case of the program's failure, the user can simply analyze the data by hand.

Table 1: Timing differences between reference analysis method and automated analysis.

Time per device (Cell Counting)	Reference Analysis (hrs)	Automated Analysis (hr)	Time Saved with Automated Analysis (hrs)
Making PDMS	1.25	1.375	-0.175
Capturing Video/Image	0.25	0.167	0.0833
Analyzing Cell Counts	2.5	0.1	2.4
Analyzing Fluor Counts	2.5	0.1	2.4
TOTAL	6.5	1.74	4.8

Furthermore, the automated cell counting program has been extended to analyzing fluorescence intensity in addition to counting cells (Chapter 5.2.6). This has allowed us to phenotype our cells on chip, which increases the information gained from each cell in a single cell experiment.

The cell counting program performed extremely well with healthy cells; however, the program does not account for dead cells. This could become an issue when cell viability is low and the user wants to account for chambers with dead or apoptotic T cells. Because dead/apoptotic T cells have lower contrast than healthy T cells on light microscopes, it is difficult to identify them in a sensitive manner. This is because their detection will require the program to accept lower contrast features as potential dead cells, which would consequently increase the amount of noise accepted. One option for detecting

dead/apoptotic cells would be to perform a fluorescent stain before the experiment, and then using the fluorescent cell detection method developed for Chapter 5.2.6. Live/dead stains could be used for dead cells and Annexin V for apoptotic cells.

1.5 – Conclusions

Two automated programs have been developed that automated cell counting and fluorescence signal gathering in single cell microfluidics tests. The cell counting algorithm reduced the time required to analyze a single device from 2.5 to 0.1 hours, without any appreciable loss in accuracy. The automated fluorescence signal gathering program was comparable in accuracy to manual analysis, differing by 5.8% on average. This has demonstrated the capability of automation in single cell data gathering.

References (Chapter 1)

[1]   J. Heath et al., "A clinical microchip for evaluation of single immune cells reveals high functional heterogeneity in phenotypically similar T cells," *Nature Medicine,* vol. 17, p. 738–743, 2011.

[2]   H. Ploegh et al., "A microengraving method for rapid selection of single cells producing antigen-specific antibodies," *Nature Biotechnology,* vol. 24, pp. 703 - 707, 2006.

[3]   R. Fan et al., "High-throughput Secretomic Analysis of Single Cells to Assess Functional Cellular Heterogeneity," *Analytical Chemistry,* vol. 85, no. 4, p. 2548–2556, 2013.

[4]   M. Birdsall, "Google and ITE The Road Ahead for Self-Driving Cars," *ITE JOURNAL-INSTITUTE OF TRANSPORTATION ENGINEERS,* vol. 84, no. 5, pp. 36-39, 2014.

[5]   K. Huang et al., "Identifying survival associated morphological features of triple negative breast cancer using multiple datasets," *J Am Med Inform Assoc,* vol. 20, p. 680–687, 2013.

[6]   C. Loukas et al., "An image analysis-based approach for automated counting of cancer cell nuclei in tissue sections," *CYTOMETRY PART A,* vol. 55A, no. 1, pp. 30-42, 2003.

[7]   J. Heath et al., "Multifunctional T-cell Analyses to Study Response and Progression in Adoptive Cell Transfer Immunotherapy," *Cancer Discovery,* vol. 3, no. 4, pp. 418-429, 2013.

[8]   R. Chatterjee et al., "Cell tracking in microscopic video using matching and linking of bipartite graphs," *COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE,* vol. 112, no. 3, pp. 422-431,

2013.

[9]   B. Roysam et al., "Improved Automatic Detection and Segmentation of Cell Nuclei in Histopathology

Images," *Biomedical Engineering, IEEE Transactions,* vol. 57, no. 4, pp. 841-852, 2010.

[10] F. Theis et al., "An automatic method for robust and fast cell detection in bright field images from

high-throughput microscopy," *BMC Bioinformatics,* vol. 14, p. 297, 2013.

[11] V. Petaccio, "Remove Outliers," Mathlab Central, 30 7 2009. [Online]. Available:

http://www.mathworks.com/matlabcentral/fileexchange/24885-remove-outliers. [Accessed 15 5

2015].

Chapter 2: Predicting Glioblastoma Patient Response to Anti-Angiogenic Therapy using Statistical Leaning Methods

Abbreviations

LR: Logistic Regression
GBM: Glioblastoma Multiforme
kNN: k-Nearest Neighbors
RF: Random Forests
DEAL: DNA-Encoded antibody libraries
CV: Coefficient of Variation
AUC: Area under the Curve

2.1 - Introduction

Using biomarkers to diagnose before bulk symptoms arise has received extensive attention in the medical community [1, 2, 3, 4]. For example, in pregnancy tests, human chorionic gonadotropin (hCG) can be detected in the user's urine after its levels double in blood/urine every ~72 hours after conception [5]. This application shows the impactful promise in diagnosing conditions in the body before bulk symptoms are displayed.

In cancer diagnostics and prognostics, multiple biomarkers are typically necessary due to the subtlety of the disease [6]. This is because the state of cancer is not as drastically transformative to the body as something like pregnancy. Cancer is constantly kept in check by nutrient limitations via blood vessels, DNA repair mechanisms, basement membrane sequestration, and the immune system [7]. Therefore, in order to diagnose a cancer before it becomes metastatic and/or displays symptoms, one must look for subtle alterations in general cell functioning.

Cancer diagnostics using biomarkers began with the discovery that free DNA was present in the bloodstream of cancer patients, which included mutated gene products from the tumor [8]. However, the presence of mutated genomic material does not necessarily imply functional changes in the gene's

product: protein levels [1]. Genes must be translated and transcribed into proteins in order to exact a function on the body [9]. Therefore, in order to determine if functional changes are present in the body, one must detect proteins as a biomarker.

Proteins have been used as biomarkers for a wide array of applications, including diagnostic for Alzheimer's disease, rheumatoid arthritis, and different forms of cancer [10, 11]. Many methods exist for protein detection, including 2-D PAGE, mass spectrometry, surface-enhanced laser desorption/ionization, and enzyme linked immunosorbant assays (ELISAs) [12, 13, 14, 15]. Each method has its own advantage, with ELISAs being very accurate and selective for a small number of protein targets [14]. A multiplexed ELISA-platform was previously developed in the Heath lab, and showed high accuracy in detecting blood proteins [13]. A multiplexed and accurate detection system is advantageous in cancer diagnostics where malignant transformations are preceded by subtle changes in specific proteins [16, 6], especially if the tumor is difficult to assess, such as in the case of Glioblastoma Multiforme (GBM) [17].

GBM is one of the most deadly forms of cancer, with a post-diagnosis median survival time of 12-18 months with standard care [18, 19]. Surgical resection, radiation, and temozolomide have improved survival; however, reoccurrence is common [20]. Bevacizumab, a VEGF blocking antibody, has shown promise in treating these cases by altering tumor vasculature [21].

Over a century ago, it was observed that tumors promote blood vessel growth, but only in 1971 did anti-angiogenic drugs become a topic in the research community [22, 23]. For three decades, antiangiogenic drugs, like Bevacizumab, have shown promise as anti-tumor drugs in both preclinical and clinical trials [24]. Though the exact mechanism is not fully established, Bevacizumab can reduce tumor bulk. This can inhibit tumor vasculature growth, and in certain cases, normalize the vasculature's structure [25, 26].

Using Bevacizumab to treat GBM requires time to determine whether the drug is working, typically on the order of months to see if it is effective [27]. Furthermore, GBM is a difficult tumor to assess: two reasons for this are false-positive assessments and differences in scan interpretation between doctors [21]. Therefore, it is difficult for doctors to determine if Bevacizumab is working. During this time, the drug may have no effect, costing money to the patient for an ineffective drug (Bevacizumab is ~$100,000 per year), and wasting time that could be spent trying other drugs [28]. Furthermore, anti-angiogeneic drugs come with side effects, including thromboembolism, gastrointestinal perforation, and hypertension [29, 30]. Therefore, creating a prognostic method to determine if the drug is working at an early stage is desirable.

In this project, a multiplexed ELISA diagnostic platform is applied to prognose if GBM patients responding to a Bevacizumab treatment. Two mathematical models were developed which were used to predict whether a patient had a growing tumor at that time point.

2.2 – Materials and Methods

2.2.1 - DNA-Encoded Antibody Libraries

DNA-Encoded antibody libraries (DEAL) were used as a capture antibody for the ELISA [31]. These conjugates are monoclonal antibodies linked to a single stranded DNA (ssDNA). The ssDNA is unique for each antibody, which will hybridize to its unique complement ssDNA. The complementary strands, which will be discussed in more detail in Chapter 2.2.2, were immobilized to a surface. DEAL antibody conjugates are then immobilized to this surface by hybridization, much like in a standard 96 well-plate ELISA (See Appendix A, Figure 1).

The synthesis of a DEAL antibody conjugate has been described in detail elsewhere [31]. Briefly, 100 µL of antibody at 0.5-1 mg/mL and 50 µL of 33mer ssDNA (IDT DNA, see appendix X for ssDNA sequences used) at 200 µM were desalting using a 0.5mL Zeba column (Life Technologies). Next, 1uL of

200mM succinimidyl 4-formylbenzoate (S-4FB, Solulink) in N,N-dimethylformamide (DMF) was mixed with the DNA, and 2.2 μL of 40mM succinimidyl 4-hydrazinonicotinate acetone hydrazone (SANH, Solulink) in DMF with the antibody. Furthermore, 10 μL of DMF was added to the DNA solution to increase S-4FB solubility. The solutions were incubated at room temperature for 2 hrs.

After the incubation, both antibody and DNA solutions were desalted again with a Zeba column and buffer exchanged into pH = 6.0 citrate buffer. The solutions were then mixed and allowed to incubate at room temperature for 16-18 hours. The mixture was then purified in a fast protein liquid chromatography (ÄKTAFPLC, GE Healthcare).

2.2.2 - Spotted DNA Microarrays

Patterning ssDNA for the DEAL antibody conjugate to bind to was accomplished using a non-contact microarray printer (Sprint Inkjet Microarrayer, Arrayjet). ssDNA was mixed in either ArrayJet Nucleic Acid Printing Buffer or 50% Dimethylsulfoxide (DMSO) in water.

For the ArrayJet Nucleic Acid Printing Buffer, 10 μL of 400 uM DNA in water was mixed 15 μL of deionized water and 25 mL of 2x Arrayjet Nucleic Acid Printing Buffer. The solution was mixed thoroughly and loaded into the ArrayJet Microarray printer. The ssDNA was printed on poly-L-Lysine slides (Thermo Scientific) and allowed to sit overnight to incubate. Next, the slides were exposed to 60 mJ/cm^2 using a VWR UV-Crosslinker [32]. The slides were then stored in a dessicator.

Slides patterned with 50% DMSO in water were patterned in a similar fashion to the ArrayJet nucleic acid printing buffer; however, ssDNA was printed on GAPS II slides (Corning) and were then baked for 2 hrs at 70 °C after overnight incubation, rather than exposing to UV. They were also stored in a dessicator until use.

Prior to their use in an ELISA, the slides were washed with either PBS buffer or 0.02% sodium dodecyl sulfate (SDS) in water. Slides were dunked 5x in the PBS or SDS solution and then three times in deionized water.

2.2.3 - PDMS microwells

Polydimethylsiloxane (PDMS) was used to create microwells for the ELISAs. PDMS was made using Sylgard 184 (Dow Corning) at a ratio of 1 (Part B) to 10 (Part A). The PDMS was mixed thoroughly and poured onto a chlorotrimethylsilane treated silicon wafer placed in a petri dish. The PDMS was then cured for 2 hrs at 80 °C and cut out with a scalpel. Microwells were created using a 3mm biopsy punch or cut out with a scalpel.

Prior to testing, the PDMS was washed with 70% ethanol in water and dried. The bottom surface was cleaned with scotch tape before adhering to the spotted DNA microarray slide.

2.2.4 - Running an ELISA test

A blocking solution was made by mixing ~450 mg of Bovine Serum Albumin (Sigma) in 15 mL of PBS (Irvine Scientific). 30-40 µL of this 3% BSA solution was added to each microwell and allowed to incubate for 1 hr at 37 °C. The BSA was removed by pipetting and 30 µL of the DEAL antibody conjugates was added after diluting DEAL conjugates 1:50 in 3% BSA (final concentration of 100 ug/mL). This was incubated for 1.5 hours at 37 °C. The wells were then washed three times with BSA, and 30uL of the protein solution or patient serum was added to the well. This was incubated for 1 hr. The wells were then washed three times with BSA and incubated with BSA for 15 min. Biotinylated secondary antibodies were diluted 1:150 and added to the wells after washing. This was then incubated for 1 hr on an orbital shaker at 37 °C. The wells were washed again with BSA three times. A 0.5 mg/mL solution of Streptavidin-Cy5 (eBioscience) and 10 µM solution of Cy3 conjugated ssDNA was added at a 1:100 dilution. The ssDNA was complementary to a specific ssDNA barcode, which would then become the

reference feature. Finally, the wells were washed three times with BSA and the PDMS was removed from the slide. The slides were then dunked in PBS/0.05% Tween-20 three times and deionized water once. Slides were then dried on a slide spinner and scanned on a Genepix 4400A scanner.

Data was gathered using the Genepix Pro 7 program using the block finding method. The blocks consisted of circle regions of interest (ROI) that were aligned with the spotted DNA features. The average of these blocks was calculated automatically and exported to an excel file for data analysis.

2.2.5 - Optimizing Blood Protein Measurements

The selectivity of the ELISAs was optimized by performing cross-talk checks for all DEAL antibody conjugates. This was accomplished by running an ELISA test, as described in Chapter 2.2.4, using a mixture of recombinant proteins. The recombinant proteins tested were MIF, VEGFR2, HGF, mouse IL-2 (mIL-2), IL-2, IL-13, VEGF, IL-6, mIL-3, CXCL13, and TGFa. The protein concentrations were all 10 ng/mL, except HGF, which was at 50 ng/mL. This ELISA test determined cross-reactivity of the secondary detection antibodies, which are polyclonal and tend to be promiscuous. In the test, DEAL antibody conjugates for all proteins were added to 11 different wells (one per protein), along with the mixture of all recombinant proteins. During the secondary antibody step, only one secondary antibody was added per well. The slides were scanned and the signal of each protein was measured in each well. Any signal found from a protein that exceeded an acceptable background level (determined from the average of non-contaminated spots plus two standard deviations) was deemed "cross-talk."

The reproducibility of the ELISAs was tested by running tests with the same sample under different conditions. These different conditions included: different microwells in the PDMS, different DNA microarray slides on the same day, different slides on different days, and different test operators. The standard recombinant sample with the 11 proteins from Chapter 2.2.5 was used for the tests. The

protein intensities were measured under each of these conditions and the coefficient of variation (CV%) was calculated to determine how much variation each condition incurred.

2.2.6 - Running GBM patient samples

Blood samples were collected at UCLA using standard phlebotomy techniques. 10mL collection tubes coated with ACD-A anticoagulant were used (BD Vacutainer). The tubes were centrifuged for 15 minutes at 1500 xg and plasma was collected and aliquot at 200 µL into cryovials. The cryovials were frozen at -80 °C until the day of testing.

The ELISAs were performed as described in Chapter 2.2.4, with the proteins listed in Appendix B, Table 1. Patient tumors were classified by Dr. Tim Cloughesy as "Growth" and "No Growth" using brain scans.

2.2.7 - Statistical Learning

The data consisted of n=92 samples from 20 different patients and p=35 proteins (see full list of proteins in Appendix X). There were two different types of data; "Raw" and "Background Subtracted." If not specified explicitly, the "Background Subtracted" dataset was used for analysis.

2.2.7.1 - Distribution comparisons

To determine if a protein was expressed at significantly different levels between patients who were responding to the drug (No Growth) versus patients not responding (Growth), a non-paired Mann-Whitney U-Test was performed using the ranksum() function in Matlab. The p-values were calculated and scaled using the Bonferroni correction [33].

2.2.7.2 - Logistic Regression

To generate a statistical learning model, logistic regression was used in R (R Project for Statistical Computing). Logistic regression fits binary data to a logit curve. It is an extension of ordinary linear

regression, with the addition of a link function and no necessity for normally distributed data [34]. The

logistic curve was fit to the patient's response to the drug, where "Growth" = 1 and "No Growth" = 0. An

algorithm was developed around the logistic curve fitting as shown schematically in Figure 1. After

measuring the protein levels for all patients, the data was pre-processed. This included background

subtraction, checking for missing data, and removing patients that lacked drug-response data. At this

point, the algorithm randomly picks a subset with 30% of the total protein predictors. The patients were

then split into two different sets (without replacement): the training and test set. The training set was

80% of the samples and the test was the remaining 20%. The training set was plugged into the glm()

function in R. Predictions for the test set were then calculated using the predict() function, which were

then compared to their actual values to obtain prediction quality values (e.g., accuracy, sensitivity,

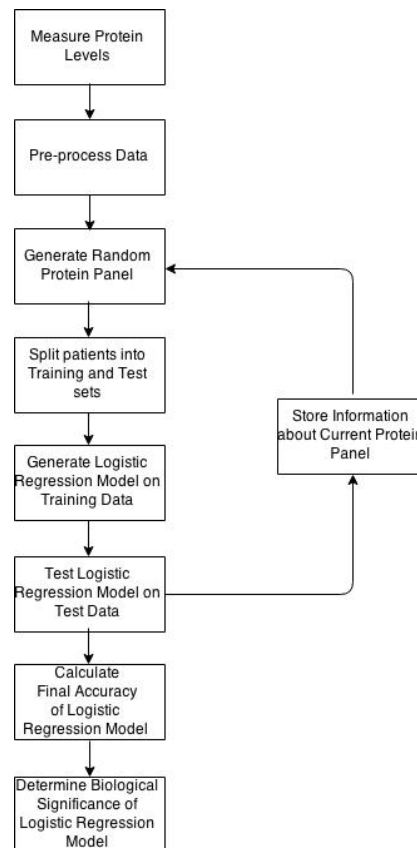specificity, positive predictive value, area under the curve, etc.).

Figure 1: Flow chart of Logistic regression program. This method was used to find the optimum panel of proteins to predict whether patients had growing or shrinking tumors in response to Bevacizumab.

The program then stores information about these quality parameters about the current model, including accuracy, sensitivity, specificity, positive and negative predictive value, and AUC was calculated by integrating underneath a Receiver Operating Characteristic (ROC) curve, which is a plot of 1-specificity vs. sensitivity. This accuracy was then stored in a dataset along with which proteins were used in that LR model. The algorithm then creates a new randomized set of training and test samples, as well as a new random subset of predictors. This process was then repeated 10,000 times to build a large dataset.

These results were then analyzed to determine which proteins were present in protein panels that predicted with an accuracy higher than that of the average accuracy of protein panels. This led to a subset panel of the best predicting 3-5 proteins, which were then tested using the logistic regression model on a test set using a bagging method to obtain a final prediction for each patient [34].

Throughout this algorithm, the studentized residual (rstudent) of each patient was logged and patients with rstudent values larger than 2.5-3 in more than 50% of models were removed as outliers [34].

2.2.7.3 - Ensemble method

An ensemble model was created to improve the accuracy of prediction with the cost of interpretability [35]. The ensemble model used three different statistical learning methods to predict whether a patient would respond to the drug: k-Nearest Neighbors (kNN), Logistic Regression (LR), and Random Forrest (RF). The models were trained in a similar fashion to the LR model in Chapter 2.2.6.2; however, the experimental flow was slightly different (Figure 2). After obtaining predictions from each model, they were correlated to determine how similar the model predictions were. If correlations exceed 0.7-0.8 for the models, then their average, or ensemble prediction, has little difference from the

individual predictions. The predictions were then averaged and weighted to obtain a final prediction.

This was used to calculate quality of prediction values.
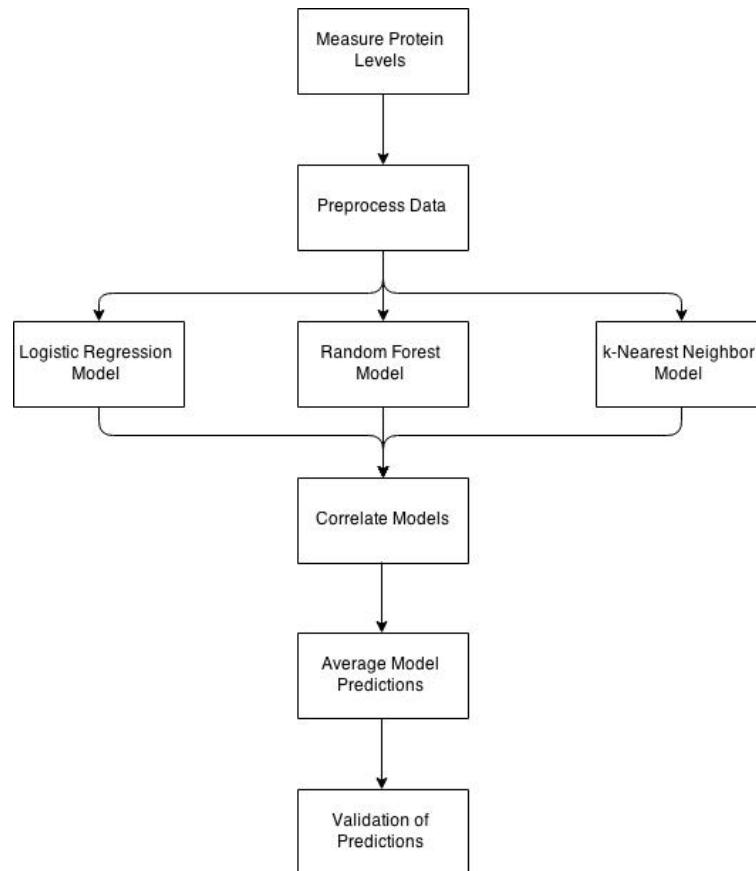
Figure 2: Flow chart for ensemble learning method. This method was used to determine how to predict whether a patient had a growing/ non-growing tumor in response to Bevacizumab treatment.

2.3 – Results

2.3.1 - Spotted DNA Microarrays

Using the ArrayJet Nucleic Acid printing buffer showed good DNA print quality. The average

signal levels ranged from 20,000 – 45,000 fluorescence units (scanned at 450 gain and 15% power), with

CV% across the slide ranging from 10-25%. However, the signal changed (or "aged") significantly over

the course of one month, changing by ~32% per week on average. Therefore, emphasis was placed on

using 50% DMSO as a printing buffer.

The slide "aging" issue with 50% DMSO was much lower than with ArrayJet. Over the course of one week, the signal of 50% DMSO changed by only 2.7%. Furthermore, the average signal ranged from 30,000 – 40,0000 fluorescence units, which was sufficient for the blood tests. The CV% of DMSO was low as well, ranging from 3-5%.
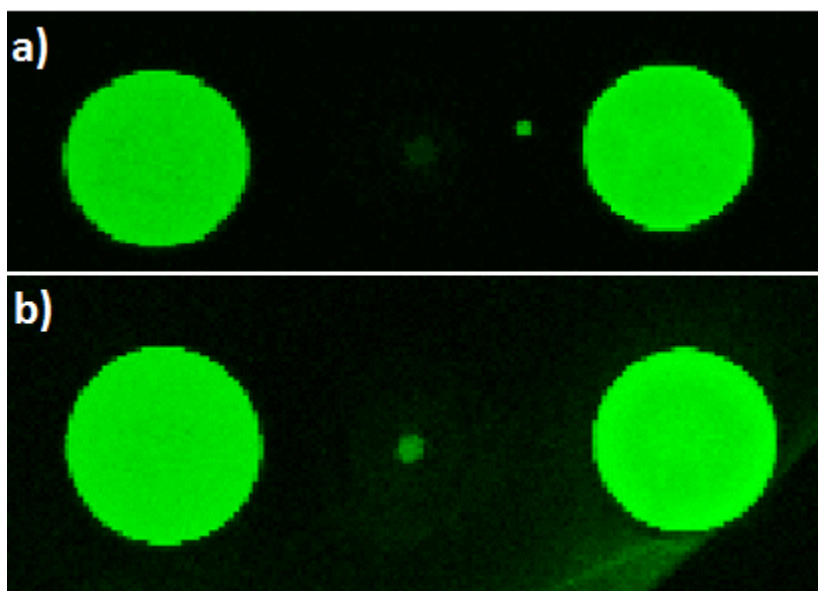
Figure 3: Images of printed DNA microarrays under different washing conditions. a) Washed with 0.02% SDS. b) Washed once with PBS.

With the 50% DMSO printed DNA, washing with 0.02% SDS reduced streaking of the DNA compared to PBS (Figure 3). Therefore, 0.02% SDS was used as the primary wash solution.

2.3.2 - Optimizing Blood Protein Measurements

The selectivity of the tests was optimized by running a cross-talk check with all proteins. Figure 4 shows a truncated list of proteins from the cross-talk check. VEGF and VEGFR2 have nonspecific background, which can be subtracted out before data analysis. Otherwise, it can be visually observed that only mIL-2 and IL-2 crosstalk. The crosstalk of IL-2 and mIL-2 was then tested under different temperature conditions (Figure 5). It was observed that at room temperature, the cross-talk was 7.8%, whereas at 37 °C, the cross-talk was only 2.3%. Therefore, tests were run at 37 °C from then on.

Crosstalk of proteins in truncated protein panel
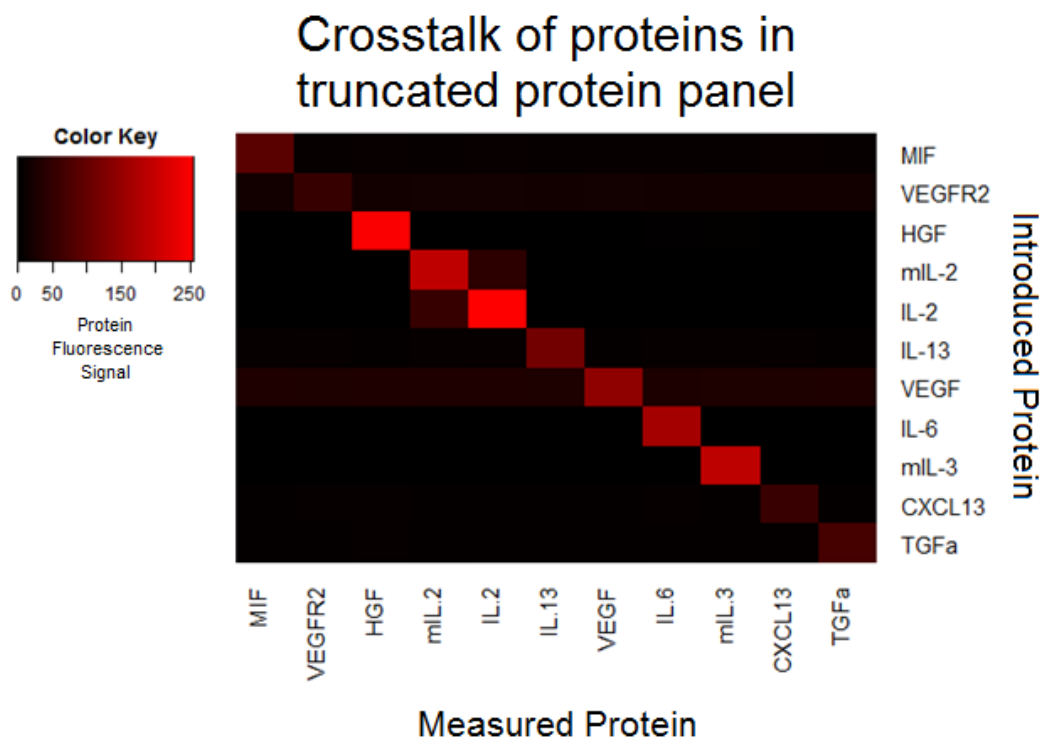
Figure 4: Truncated panel of protein crosstalk. The diagonal shows the signal of each protein measured at 10 μg/mL (or 50 μg.mL for HGF). The off diagonal component show the protein cross-talk. The recombinant protein that was introduced in each well is shown in the rows. The protein that was measured for is shown in the columns. Ex: When VEGF was introduced, all proteins shown slight signal when measured.
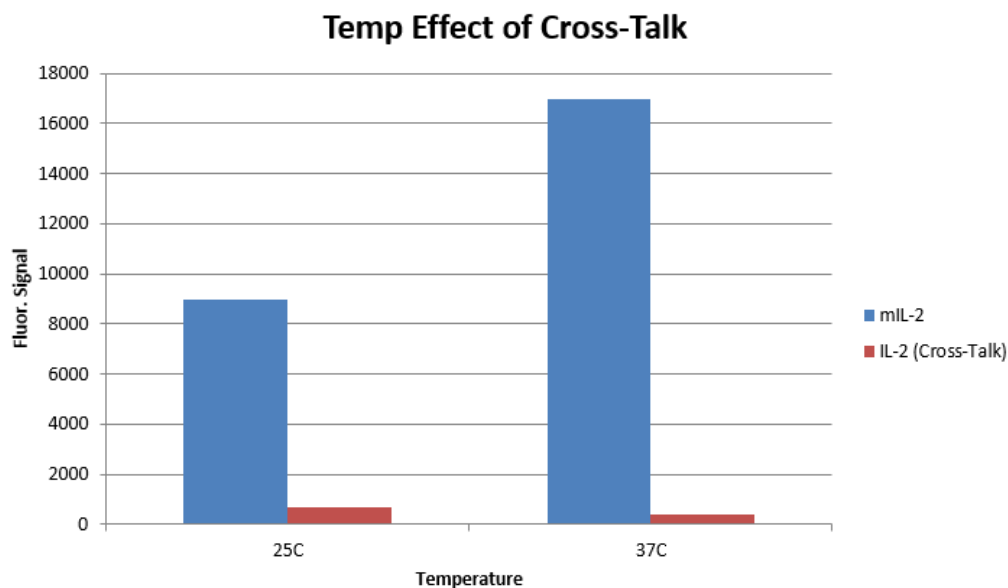
Figure 5: Temperature effect on mIL-2 and IL-2 crosstalk when mIL-2 recombinant was tested and cross-talk from IL-2 was measured.

26

The reproducibility of the test was measured for multiple conditions, which are summarized in Table 1. The CV% between different users was the most variable at 22%. This was well within the range to yield accurate protein measurements [36]. The device reproducibility was good, with CV% for proteins being less than 10%. CV% from different repeats most likely came from the slight difference in DNA spot leading. These variations were reduced when signal was averaged in a well, since the well-to-well variation was lower.

Table 1: Variations in protein measurements under different conditions.

	Average % CV of Cy3 Signal	Average % CV of Protein Signal (p=4 proteins)
Different Repeats	12.28	17.7
Different Wells	6.32	12.9
Different Devices (Different Days, Same User)	3.5	6.7
Different Users (Same Day)		22

2.3.3 - Measuring patient samples

Patient information is listed in Appendix A Table 1. The protein measurement dataset contained n=92 patient samples from 20 unique patients at different time points, and p = 35 proteins. On average, the patient blood samples were collected 24 days after the start of the drug treatment.

2.3.4 - Statistical Learning

2.3.4.1 - Statistically significant proteins

Box plots of select protein levels – gated on tumor Growth vs. No Growth – are shown in Figure 6. A full list of protein significance values are shown in Appendix A Table 2. Because 35 proteins were tested for significance at one time, the Bonferroni Correction was applied and the true significance was set at p = 0.0014 [33]. Using this correction, the only proteins that showed significantly different levels between patients with growth and no growth were TGFb1, HGF, and VEGFR2.
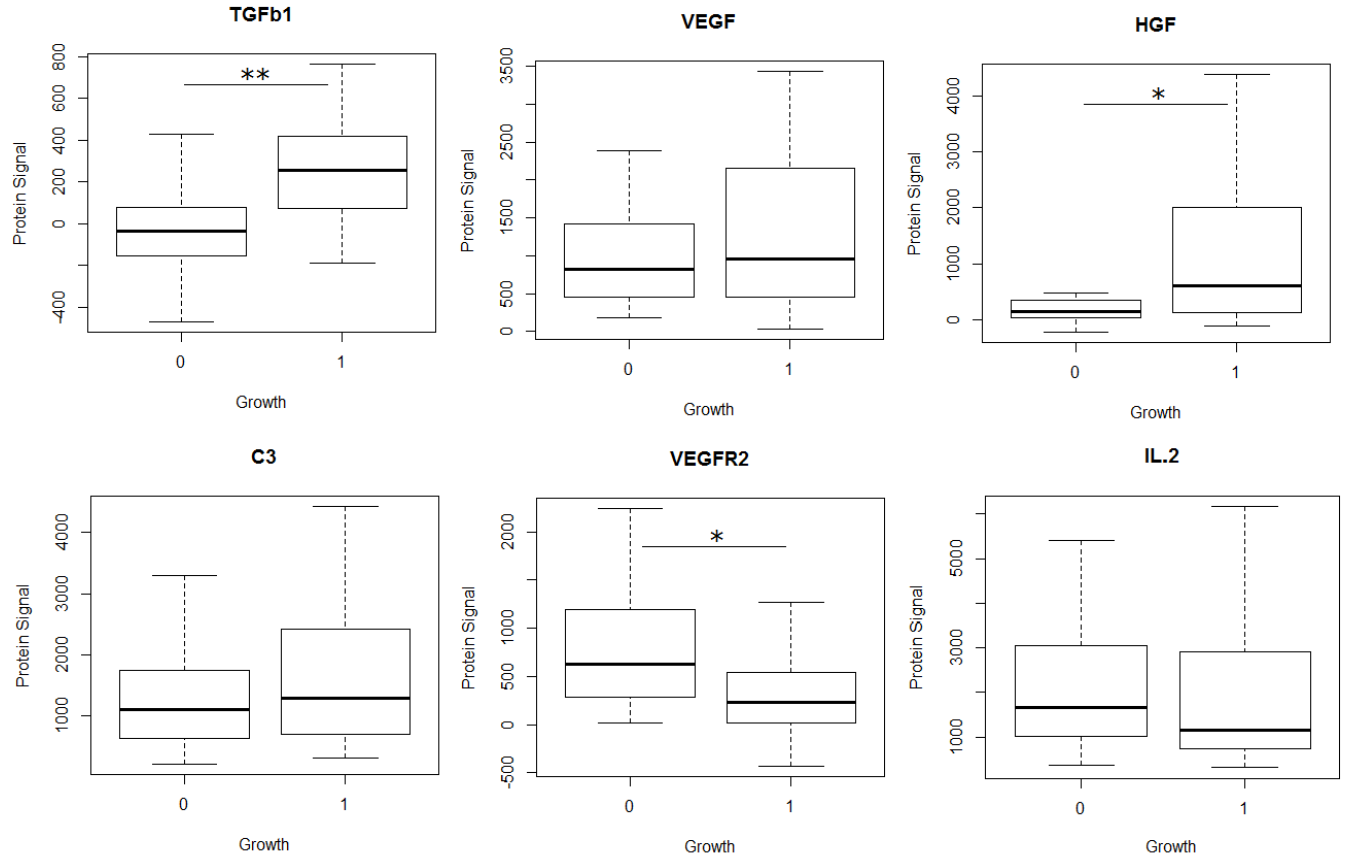
*Figure 6: Box plots of select proteins, gated on Growth (1) or No Growth (0). Data is from the background subtracted dataset. Significance indicators are shown above the box plots. Significance thresholds are lower than normal due to the Bonferroni correction. Significance markers: * = p < 0.001, ** = p < 0.0001.*

2.3.4.2 - Logistic Regression

Before fitting to a logistic regression curve, the correlation matrix was investigated for signs of collinearity. The heatmap of the correlation matrix is shown in Appendix A, Figure 2. There were three instances of collinearity in the data: C3 and EGF, VEGF and IL.12, and MMP2 and IL.1b. In each of these cases, one protein was removed in order to reduce the chance of collinearity affecting standard error measurements.

From developing the logistic regression model, it was found that the bootstrap accuracy of prediction using all 35 proteins was 73.9%. The AUC was 0.721, which is generally considered a "C grade" in a standard school grading scale [37, 38]. The full list of prediction quality values for the complete panel of proteins is listed in Table 2, labeled "LR Model (35 Proteins)." Furthermore, a subset

28

of the patients were removed as outliers because they were above the r-student outlier threshold [34]. Three of 92 patients were removed, resulting in a loss of ~3% of the total patient data. This samples were left out of subsequent analyses.

Table 2: Statistical learning model results for background subtracted dataset.

Model Name (Background Subtracted)	Bagging Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	AUC
LR Model (35 Proteins)	73.9	58.4	72.7	59.1	72.1	0.751
LR Model (5 Proteins)	85.2	75.0	91.9	85.8	84.9	0.901
kNN	72.3	50.0	87.5	72.2	72.8	
Random Forrest	75.3	65.0	81.9	70.2	78.1	
Ensemble Model	86.5	84.8	87.5	80	90.7	

After optimization, five proteins were found to be significant predictors in the fit. These proteins are listed in Table 3, along with their significance values and coefficient weights. The bootstrap accuracy was 85.2% and the AUC was 0.901. The complete list of quality parameters for this LR model are listed in Table 2, under the row label "LR Model (5 Proteins)." This AUC is much higher than with the full protein panel, indicating a better model fit.

Table 3: List of protein predictor coefficients and significance values for LR mode (5 Proteins). Significance markers: * = p < 0.05, . = p < 0.1.

Protein Name	p-value	Significance	Avg. Coeff. Weight
VEGF	0.0231	*	5.21
VEGFR2	0.00946	*	-5.67
TGFb1	0.00707	*	5.33
IL-2	0.0113	*	-3.51
C3	0.101	.	1.98

2.3.4.3 – Ensemble Method

The prediction accuracy of the three ensemble models was 86.5%, 75.3%, and 74.1% for the LR, kNN, and RF models, respectively. The correlation table of their predictions is shown in Table 4, with no correlation exceeding 0.574. These correlations were low enough to perform an ensemble prediction.

29

Table 4: Correlation table of ensemble predictions.

Ensemble Correlations	LR	kNN	RF
LR	1	0.556	0.456
kNN	0.556	1	0.574
RF	0.456	0.574	1

The quality parameters for the ensemble model are shown in Table 1, under the row labeled "Ensemble Model."

2.4 - Discussion

Though the CV% of protein measurements were typically lower than 20% in our experiments, it would be desirable to reduce their values to 10% or lower. According to Reed et al., a CV% of 20% in a protein measurement will be off by a factor of 1.5x ~15% of the time and off by a factor of 2x ~2% of the time [36]. However, if our CV% was reduced to 10%, the chance of being off by a factor of 1.5x or 2x goes to ~0%.

In our study, HGF was found to be upregulated in the patients with tumor growth relative to those with no growth. It has previously been observed in colorectal cancer that HGF increases with tumor growth during anti-VEGF therapy [39]. HGF is responsible for many functions, including cell mobility, angiogenesis, and cell growth [40]. It is secreted from mesenchymal cells, which was interesting in respect to its strong correlation with TGFb1. This is because TGFb1 is secreted primarily by leukocytes, rather than mesenchymal cells [41]. Their correlation was likely indirectly correlated through regulatory mechanisms. TGFb1 has very similar functions in the body as HGF. Therefore, it is possible that general regulatory mechanisms for cell growth, cell mobility, and angiogenesis caused their levels to become correlated.

Our results show promise in predicting whether a GBM patient was responding to Bevacizumab treatment. Both the logistic regression and ensemble models predicting with greater than 80% accuracy. The quality of the predictions obtained were comparable to other protein biomarker diagnostics. In various studies of early detection of breast cancer, sensitivity/specificity ranged from 80 to 100% [42]. The ensemble model did have a significantly higher sensitivity than the individual models, meaning it was better able to pick out true positives (tumor growth) as positive. In our case, this is more important, since these patients are the ones we want to catch early in order to stop treatment. Therefore, in the real world setting, the ensemble model would be chosen, rather than the LR model.

The LR model yielded good results, with an AUC > 0.9, which is equivalent to an "A" grade in a standard grading scale [37]. The validity of this model was supported by the fact that two of the predictors in the model, TGFb1 and VEGFR2, were significantly distinguishable between the Growth and No Growth groups.

Two immune proteins were found to be predictive in the logistic regression model: IL-2 and C3. None of these proteins were expressed to significantly different levels on their own. However, they both had low standard error values in the logistic regression model, indicating they had a relationship with the patient outcome. C3 had a positive coefficient, meaning an increase in its levels predicts a worse outcome, whereas IL-2 had a negative coefficient, meaning an increase predicts a better outcome. Because logistic regression model coefficients give the relationship between a single protein, with all others held constant, we can make hypotheses as to how each predictor weight may be related to patient response [34].

High levels of C3 would indicate the innate immune system is active [43]. This is at its highest during the initiation of an adaptive immune response. Therefore, the positive correlation between C3

levels and tumor growth may be because the patient's immune system is still in its initial stages of launching an immune response towards the tumor, or perhaps is stagnated at that stage.

High levels of IL-2 would indicate T cell expansion during the initiation of a Type I immune response. Because our model showed that higher levels of IL-2 is a predictor of "No Growth," this could imply that adaptive immunity is a positive correlator with tumor reduction in Bevacizumab treatment of GBM. IL-2 has previously been observed to cause tumor reduction in GBM, though most references on this topic come from the 1990s [44, 45].

It was observed that VEGF was a positive predictor in the LR model, which means high levels indicate a progressing tumor. Intuitively, this makes sense: if VEGF is higher, one expects the tumor to have more blood vessels and be growing faster. This observation has been seen in multiple studies predicting patient response via VEGF levels [46, 47].

TGFb1 and VEGFR2 were the most discriminating proteins between growth and no-growth and were best predictors in the LR model. We observed that TGFb1 was overexpressed in the GBM patients with growing tumors, and consequently was a positive predictor in the LR model (increased level indicated the patient was more likely to have a growing tumor). Normally, TGFb1 has many roles, including a cell proliferation regulator, immunosuppressor, extra cellular membrane remodeler, and angiogenesis promoter [48, 49, 50]. It has been shown to promote cancer growth by suppressing NK cells, which enhances tumor cell metastasis [51, 52]. Furthermore, increased levels of TGFb1 have been documented in colorectal, prostate, bladder, liver, and brain cancers like GBM [53, 50]. Therefore, it is logical that TGFb1 is a positive predictor for tumor growth in GBM.

We observed that VEGFR2 was underexpressed in patients with growing tumors and it was a negative predictor in the LR model. This indicates that less VEGFR2 in one's blood stream predicts a growing tumor. Because VEGFR2 is a membrane receptor found on endothelial cells, one would typically

not assume VEGFR2 would be present in the bloodstream unless these cells were being lysed [54]. Therefore, if VEGFR2 is found in the blood stream, this indicates that endothelial cells are being lysed. One hypothesis is that the higher levels of VEGFR2 in "No Growth" patients are coming from the lysed tumor blood vessels. Their destruction would, in turn, starve the tumor, which we believe is why VEGFR2 is a negative predictive of tumor growth.

An alternative to this hypothesis about VEGFR2 would be that we are measuring the soluble form of VEGFR2, referred to as sVEGFR2. The role of sVEGFR2 is hypothesized to be a blocking agent for VEGF, thereby reducing angiogenesis [55]. This makes sense in light of our results, since an increase in sVEGFR2 would imply less angiogenesis and slower tumor growth.

2.5 – Conclusions

Two statistical learning models were developed that predicted whether a GBM patient was responding to Bevacizumab treatment with >80% accuracy. The LR model revealed that TGFb1 and VEGFR2 were key predictors, with TGFb1 being a positive predictor and VEGFR2 being a negative predictor. An ensemble model was developed using LR, kNN, and RF, which showed slightly better accuracy than the LR model alone. These models have demonstrated a powerful means to predict whether Bevacizumab is effective or not for a GBM patient.
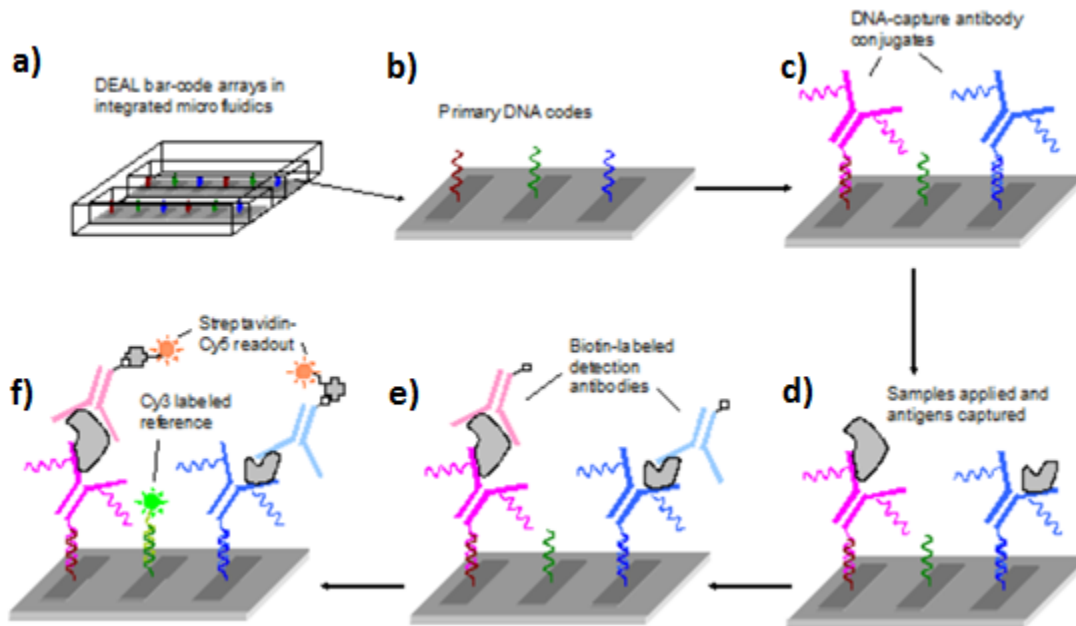
Figure 1: Overview of DEAL method. a) DNA is patterned onto a surface using spotted microarray printer or PDMS template (shown). b) The DNA is then attached to the slide surface. Different sequences of DNA are shown with different colors. c) DEAL conjugates are introduced, which hybridize with the surface bound DNA. This converts the DNA array into an antibody array. d) The sample's proteins are captured by the surface bound antibodies. e-f) Finally, biotinylated secondary antibodies and a streptavidin-dye conjugate are introduced. This allows for the detection of captured protein levels.
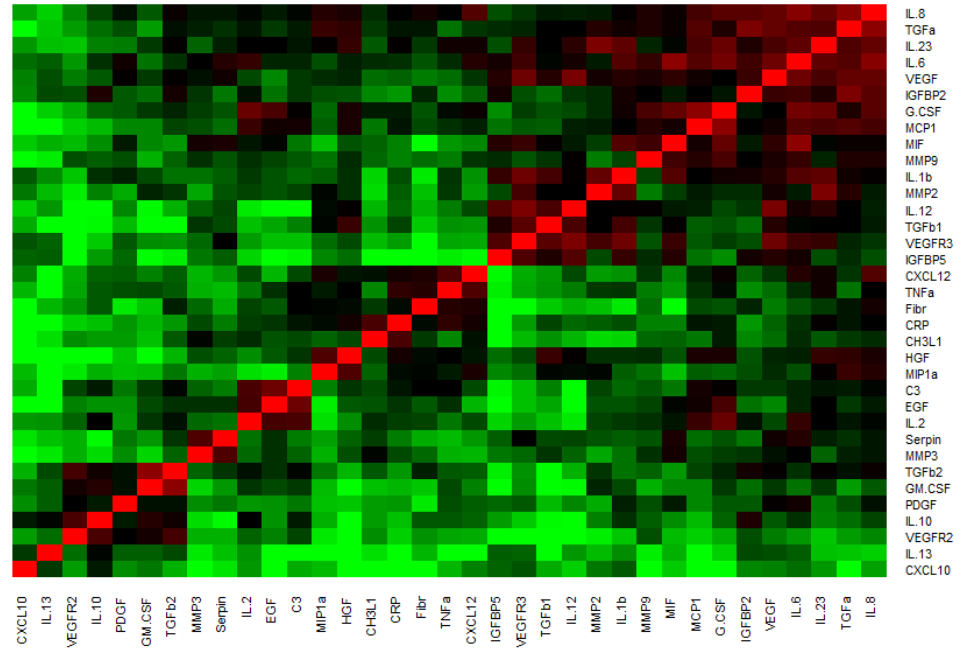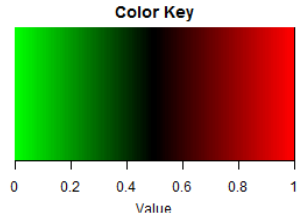
Figure 2: Correlation of protein levels from background subtracted data.

Appendix A Table 1: Patient information of 92 samples measured.

	Age		Gender		Drug		Recurrence				
	Mean	Range	Male	Female	Avastin	Avastin 184	New	1st	2nd	3rd+	N/A
Tumor Growth	55	30-82	16	9	20	5	3	9	3	5	12
No Tumor Growth	54	30-71	10	11	19	2	4	7	2	3	5

Appendix A Table 2: Dataset of Mann-Whitey U tests results of significance when comparing protein levels between Growth and No-Growth patients. Background subtracted dataset was used.

Protein Name	p-value	Significance (Bonferroni Cutoff: 0.0014)
IL.2	0.203	
MCP1	0.327	
IL.6	0.924	
G.CSF	0.258	
MIF	0.930	
EGF	0.738	
VEGF	0.258	
PDGF	0.493	
TGFa	0.152	
IL.8	0.239	
MMP3	0.220	
HGF	0.000979	*
Cy3	0.787	
CXCL10	0.356	
CXCL12	0.981	
IGFBP2	0.524	
IGFBP5	0.762	
MIP1a	0.00746	.
TGFb1	0.000002	**
CH3L1	0.126	
VEGFR3	0.627	
TNFa	0.627	
C3	0.297	
MMP2	0.054	
IL.10	0.738	
IL.1b	0.981	
IL.12	0.0269	
MMP9	0.054	
TGFb2	0.098	
GM.CSF	0.245	
CRP	0.949	
VEGFR2	0.000801	*
IL.13	0.0274	
IL.23	0.532	
Serpin	0.949	
Fibr	0.768	

Appendix A Table 3: Statistical leaning model results when Raw dataset was used.

Model Name (Raw Data)	Bagging Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	AUC
LR	85	83	86	77	90	0.82
RF	70	65	72	48	84	
kNN	74	73	74	52	88	
Ensemble Model	94	93	94	90	96	

Appendix B (Chapter 2)

Appendix B, Table 1: List of proteins used (Adapted from Udi Vermesh's Thesis, © 2011).

	Company	Capture Antibody (Catalogue #)	Detection Antibody (Catalogue #)
IL2	BD	555051	555040
MCP1	eBioscience	16-7099-85	13-7096-85
IL6	eBioscience	16-7069-85	13-7068-85
G-CSF	R&D systems	MAB214	BAF214
MIF	R&D systems	MAB289	BAF289
EGF	R&D systems	MAB636	BAF236
VEGF	R&D systems	MAB293	BAF293
PDGF	R&D systems	MAB1739	BAF221
TGFa	R&D systems	AF-239-NA	BAF239
IL8	BD	554718	554716
MMP3	R&D systems	AF513	BAF513
HGF	R&D systems	MAB694	BAF294
IP10	R&D systems	MAB266	BAF266
SDF1	R&D systems	MAB350	BAF310
IGFBP2	R&D systems	MAB6741	BAF674
IGFBP5	R&D systems	MAB8751	BAF875
MIP1a	R&D systems	AF-270-NA	BAF270
TGFb1	BD	559119	559119
Ch3L1	R&D systems	DY2599	DY2599
VEGFR3	R&D systems	MAB349	BAM3492
TNFa	eBioscience	16-7348-85	13-7349-85
C3	abcam	ab17455-100	ab14232-50
MMP2	R&D systems	DY1496	DY1496
IL10	eBioscience	16-7108-85	13-7109-85
IL1b	eBioscience	16-7018-85	13-7016-85
IL12	eBioscience	14-7128-82	13-7129-85
MMP9	R&D systems	MAB9092	BAM909
TGFb2	R&D systems	DY302	DY302
GM-CSF	BD	554502	554505
CRP	R&D systems	MAB17071	BAM17072
VEGF R2	R&D systems	MAB3573	BAF357
IL13	eBioscience	16-7139-81	13-7138-81
IL23	eBioscience	14-7238-85	13-7129-85
Serpin E1	R&D systems	MAB1786	BAF1786
Fibrinogen	abcam	ab10066-250	ab14790-200

References (Chapter 2)

[1] E. Petricoin et al., "Proteomic applications for the early detection of cancer.," *Nat. Rev. Cancer,* vol. 3, pp. 267-275, 2003.

[2] K. Blennow et al., "Cerebrospinal fluid and plasma biomarkers in Alzheimer disease," *NATURE REVIEWS NEUROLOGY,* vol. 6, no. 3, pp. 131-144, 2010.

[3] N. Kosaka et al., "Circulating microRNA in body fluid: a new potential biomarker for cancer diagnosis and prognosis," *CANCER SCIENCE,* vol. 101, no. 10, pp. 2087-2092, 2010.

[4] R. Simpson et al., "Exosomes: proteomic insights and diagnostic potential," *EXPERT REVIEW OF PROTEOMICS,* vol. 6, no. 3, pp. 267-283, 2009.

[5] D. Pittawy et al., "Doubling times of human chorionic gonadotropin increase in early viable intrauterine pregnancies," *Am J Obstet Gynecol,* pp. 299-302, 1985.

[6] A. Lokshin et al., "Development of a Multimarker Assay for Early Detection of Ovarian Cancer," *J Clin Onco,* vol. 28, no. 13, pp. 2159-2166, 2010.

[7] R. A. Weinberg, The Biology of Cancer, Garland Science, 2014.

[8] D. Sidransky, "Emerging molecular markers of cancer," *Nature,* vol. 2, pp. 210-219, 2002.

[9] B. Alberts et al., Molecular Biology of the Cell, Garland Science, 2007.

[10] M. Dougados et al., "Diagnostic and predictive value of anti-cyclic citrullinated protein antibodies in rheumatoid arthritis: a systematic literature review," *ANNALS OF THE RHEUMATIC DISEASES,* vol.

65, no. 7, pp. 845-851, 2006.

[11] T. Wyss-Coray et al., "Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins," *Nature Medicine,* vol. 13, pp. 1359 - 1362, 2007.

[12] S. Kingsmore, "Multiplexed protein measurement: technologies and applications of protein and antibody arrays," *Nat. Rev. Drug Discovery,* vol. 5, pp. 310-321, 2006.

[13] R. Fan et al., "Integrated barcode chips for rapid, multiplexed analysis of proteins in microliter quantities of blood," *Nature Biotech.,* vol. 26, pp. 1373-1378, 2008.

[14] M. Hartmann et al., "Protein microarrays for diagnostic assays," *Analytical and Bioanalytical Chemistry,* vol. 393, no. 5, pp. 1407-1416, 2008.

[15] D. Felschow et al., "The SELDI-TOF MS Approach to Proteomics: Protein Profiling and Biomarker Identification," *Biochemical and Biophysical Research Communications,* vol. 292, no. 3, pp. 587-592, 2002.

[16] G. Rice et al., "Performance of a multianalyte test as an aid for the diagnosis of ovarian cancer in symptomatic women," *Journal of Translational Medicine,* vol. 10, no. 45, 2012.

[17] M. Mancini et al., "Assessment of intra-observer variability in measurement of high-grade brain tumors.," *J Neurooncol.,* vol. 108, p. 477–83, 2012.

[18] A. Fauci et al., Harrison's Principles of Internal Medicine, New York: McGraw-Hill, 2005.

[19] S. Kesari et al., "Malignant gliomas in adults," *N Engl J Med.,* vol. 359, p. 492–507, 2008.

[20] J. H. Sherman et al., "The Role of Avastin in the Management of Recurrent Glioblastoma,"

NEUROSURGERY CLINICS OF NORTH AMERICA, vol. 23, no. 2, p. 331, 2012.

[21] T. Cloughesy et al., "Response Assessment Criteria for Glioblastoma: Practical Adaptation and Implementation in Clinical Trials of Antiangiogenic Therapy," *Curr Neurol Neurosci Rep,* vol. 13, pp. 347, 1-11, 2013.

[22] N. Ferrara et al., "VEGF and the quest for tumour angiogenesis factors.," *Nature Rev. Cancer,* vol. 2, p. 795–803, 2002.

[23] J. Folkman, "Tumor angiogenesis: therapeutic implications.," *N. Engl. J. Med. ,* vol. 285, p. 1182–1186, 1971.

[24] J. Loeffler et al., "Lessons from phase III clinical trials on anti-VEGF therapy for cancer," *Nat Clin Pract Oncol.,* vol. 3, no. 1, pp. 24-40, 2006.

[25] M. Arjaans et al., "Bevacizumab-induced normalization of blood vessels in tumors hampers antibody uptake," *Cancer Res.,* vol. 73, no. 11, pp. 3347-55, 2013.

[26] N. Ferrara et al., "Discovery and development of bevacizumab, an anti-VEGF antibody for treating cancer," *NATURE REVIEWS DRUG DISCOVERY,* vol. 3, no. 5, pp. 391-400, 2004.

[27] E. Repasky et al., "Using extracellular biomarkers for monitoring efficacy of therapeutics in cancer patients: an update," *Cancer Immunol Immunother.,* vol. 57, no. 6, pp. 759-775, 2008.

[28] C. C. C. D. T. P. T. M. C. T. D. Schmidt, "Costly Cancer Drugs Trigger Proposals To Modify Clinical Trial Design," *JNCI J Natl Cancer Inst ,* vol. 101, pp. 1662-1664 , 2009 .

[29] R. M. De Waal et al., "Anti-angiogenic treatment of human cancer: Pitfalls and promises," *International Journal of Cancer,* vol. 86, no. 6, p. 870–873, 2000.

[30] C. Zhou et al., "Safety and efficacy of first-line bevacizumab-based therapy in advanced non-squamous non-small-cell lung cancer (SAiL, MO19390): a phase 4 study," *Lancet Oncol.,* vol. 11, no. 8, pp. 733-40, 2010 .

[31] R. Bailey et al., "DNA-Encoded Antibody Libraries: A Unified Platform for Multiplexed Cell Sorting and Detection of Genes and Proteins," *JACS,* vol. 129, no. 7, p. 1959–1967, 2007.

[32] X. Wang et al., "Immobilized probe and glass surface chemistry as variables in microarray fabrication," *BMC Genomics,* vol. 5, no. 53, pp. 1-8, 2004.

[33] C. F. et al., "Multiple correlations and Bonferroni's correction," *Biol Psychiatry,* vol. 44, p. 775–777, 1998.

[34] G. James et al., An Introduction to Statistical Learning with Applications in R, New York: Springer, 2013.

[35] B. Lantz, Machine Learning in R, Packt Publishing, 2013.

[36] G. Reed et al., "Use of Coefficient of Variation in Assessing Variability of Quantitative Assays," *Clin. and Diag. Lab. Imm.,* pp. 1235-1239, 2002.

[37] J. Fan et al., "Understanding receiver operating characteristic (ROC) curves," *CJEM,* vol. 8, no. 1, pp. 19-20, 2006.

[38] "The Area Under an ROC Curve," University of Nebraska Medical Center, [Online]. Available: http://gim.unmc.edu/dxtests/roc3.htm. [Accessed 7 May 2015].

[39] J. Heymach et al., "Phase II Trial of Infusional Fluorouracil, Irinotecan, and Bevacizumab for Metastatic Colorectal Cancer: Efficacy and Circulating Angiogenic Biomarkers Associated With

Therapeutic Resistance," *JCO,* vol. 28, no. 3, pp. 453-459, 2010 .

[40] S. Aaronson et al., "Identification of the hepatocyte growth factor receptor as the c-met proto-oncogene product," *Science 251 (): ,* vol. 251, no. 4995, p. 802–804, 1991.

[41] S. Wahl et al., TGF-β and Related Cytokines in Inflammation, Switzerland: Springer, 2001.

[42] R. Tollenaar et al., "Proteomic serum biomarkers and their potential application in cancer screening programs," *Int J Mol Sci. ,* vol. 11, no. 11, pp. 4175-4193, 2010.

[43] K. Murphy, Janeway's Immunobiology, Garland Science, 2011.

[44] I. Royston et al., "Interleukin-2 gene therapy in a patient with glioblastoma," *Gene Ther. ,* vol. 2, no. 2, pp. 164-167, 1995.

[45] C. Ursaciuc et al., "Interleukin-2 (IL-2) in the treatment of malignant brain tumors (glioblastomas)," *Rom J Neurol Psychiatry. ,* vol. 31, no. 3-4, pp. 195-206, 1993.

[46] E. Winer et al., "VEGF as a marker for outcome among advanced breast cancer patients receiving anti-VEGF therapy with bevacizumab and vinorelbine chemotherapy," *Clin Cancer Res. ,* vol. 14, no. 23, pp. 7871-7877, 2008.

[47] J. Heymach et al., "Baseline vascular endothelial growth factor concentration as a potential predictive marker of benefit from vandetanib in non-small cell lung cancer," *Clin Cancer Res.,* vol. 15, no. 10, pp. 3600-3609, 2009 .

[48] J. Pouyssegur et al., "Transforming growth factor beta 1 (TGF-PI) promotes endothelial cell survival during in vitro angiogenesis via an autocrine mechanism implicating TGF-alpha signaling," *MOLECULAR AND CELLULAR BIOLOGY,* vol. 21, no. 21, pp. 7218-7230, 2001.

[49] H. Lodish et al., "Role of transforming growth factor beta in human disease," *N Engl J Med,* vol. 342, p. 1350–1358, 2000.

[50] D. Fortin et al., "Transforming growth factor-beta and its implication in the malignancy of gliomas," *Targ Oncol,* vol. 10, p. 1–14, 2015.

[51] D. Heo et al., "Elevated TGF-beta1 Secretion and Down-Modulation of NKG2D Underlies Impaired NK Cytotoxicity in Cancer Patients," *The Journal of Immunology,* vol. 172, p. 7335–7340, 2004.

[52] M. Nakajima et al., "Transforming growth factor beta stimulates mammary adenocarcinoma cell invasion and metastatic potential," *Proc Natl Acad Sci,* vol. 87, p. 7678–7682, 1990.

[53] D. Nikolic´-Vukosavljevic et al., "Elevated plasma levels of transforming growth factor-b1 (TGF-b1) in patients with advanced breast cancer: association with disease progression," *European Journal of Cancer,* vol. 39, pp. 454-461, 2003.

[54] M. Cross et al., "Vascular endothelial growth factor receptor-2: Structure, function, intracellular signalling and therapeutic inhibition," *Cellular Signaling,* vol. 19, pp. 2003-2012, 2007.

[55] C. Pequeux et al., "Implication of soluble receptors of VEGF, sVEGFR-1 and sVEGFR-2, in angiogenesis," *American journal of reproductive immunology,* vol. 60, no. 1, 2008.

Chapter 3: Phenotyping Cells in Microfluidic Devices using a Photocleavable Conjugate

3.1 - Introduction

Single cell proteomic technologies have been expanding since in the 1970s, when fluorescence activated cell sorting (FACS) was developed [1]. Since then, FACS has seen great strides in analyzing cell populations in a rapid and multiplexed fashion, with techniques like CyTOF allowing over 30 surface makers to be tested simultaneously [2]. In addition to FACS technologies, microfluidics-based single cell technology have become more prolific in the field. Recently, the Heath Lab measured multiple secreted proteins from T cells in a microfluidic device [3, 4]. C. Love's group investigated the temporal secretion dynamics of T cells under different stimulation conditions [5]. In this study, the authors not only investigated the protein secretion but also the phenotype of each cell. This was accomplished using fluorophore conjugated antibodies that bound specific surface markers. During the experiment, images were taken of each single cell chamber using a fluorescent microscope and the cells phenotyped based on which fluorophores were excited. This type of technology has been adapted in the Heath Lab, as described in Chapter 5.2.6. We phenotyped the cells in a similar fashion to C. Love's group, which allowed us to determine two different phenotypes: $CD8^+$ and CD4 ($CD8^-$) T cells.

Phenotyping with fluorescent markers on a standard 3 color microscope allows for the determination of 8 phenotypes at most.[1] This is miniscule relative to a standard FACS instrument, which can measure > 256 phenotypes. Therefore, there is a need for single cell technologies to catch up to FACS in terms of phenotype detection capabilities.

---

[1] The equation for this is 2^n, where n is the number of fluorophores.

One phenotype detection method developed by Weissleder et al. can detect nearly infinite phenotypes using the detection of amplified DNA. It is performed as follows: 1) antibodies that recognize cell surface markers are linked to unique DNA strands, 2) cells are incubated with these antibody-DNA conjugates and unbound antibodies are washed away, 3) the DNAs are cleaved off the antibodies and amplified via a polymerase chain reaction, and measured to identify which surface markers were bound to cells [6]. This method allows for the measurement of nearly unlimited phenotypes, limited only by the number of different PCR products they can detect. Therefore, the application of this technology to a single cell microfluidic platform would easily match the number of phenotypes detected by FACS.

In the current chapter, we describe a method that adapts Weissleder's phenotyping technique with the Heath Lab's DEAL method in order to detect complex phenotypes on microfluidic devices. The general schematic for this is shown in Figure 1. We use Weissleder's method of binding antibody-DNA conjugates to cells (Figure 1a) and then release the DNA in the microfluidic single cell chamber using photocleavage (Figure 1b). Instead of detecting with PCR, we attach a fluorophore to the DNA. This ssDNA-fluorophore is then captured on the surface bound ssDNA used in the DEAL method (Figure 1c). Each antibody will have a unique DNA and be fluorophore-labeled. This will not only multiplex detection of surface markers but also allow for its quantitation (Figure 1d). Using this technique, we have demonstrated the detection of surface bound streptavidin (SA) on protein coated beads and CD3 on Jurkat T cells in microfluidic chambers.
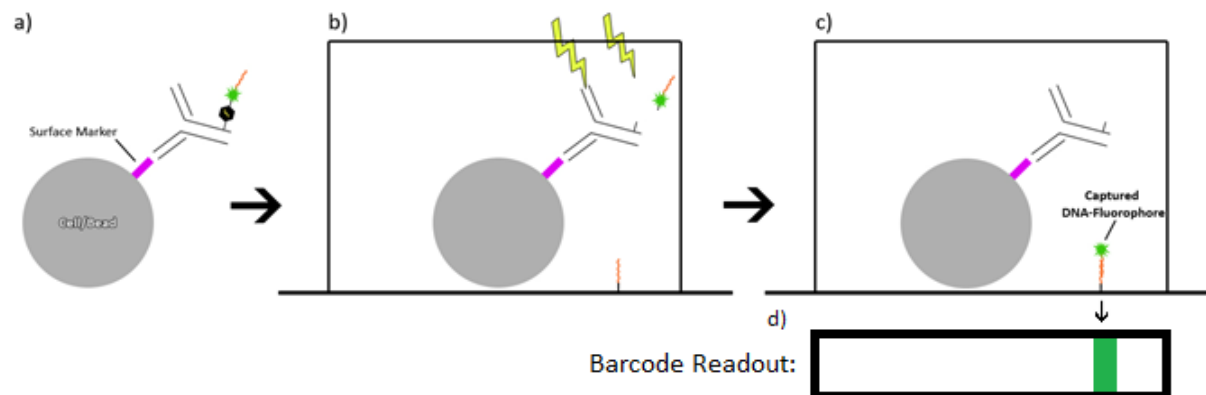
Figure 1: Overview of photocleavage conjugate (PC conjugate) detection of cell surface markers. a) The PC conjugate binds to a surface marker on a cell. b) The cell is then trapped in a microfluidic, single cell chamber and UV is introduced to photocleave a DNA-fluorophore conjugate. c) The DNA-fluorophore then binds to the slide's surface, using the Heath Lab's DEAL method. d) Finally, the slide is scanned to obtain a fluorescence intensity for where the DNA-fluorophore bound.

3.2 – Materials and Methods

3.2.1 - Synthesis of conjugate

The photocleavable (PC) conjugate is synthesized in the same manner as standard DEAL conjugates. This method is described in detail in Chapter 2.2.1. The DNA used, however, is different from the standard ssDNA used to make DEAL conjugates. The base ssDNA is attached to an internal Cy3, an internal photocleavable group, and terminated with an amine group. The DNA is custom made by IDT DNA.

5'-Amine/Internal Photocleavable Spacer/Internal Cy3/Base-DNA-3'

In order to optimize the linkage of the DNA to antibody, multiple conditions were attempted. First, two different amount of N,N-Dimethylformaminde were added to the DNA solution during the synthesis (10 µL vs. 20 µL). This helps solubility when the DNA is hydrophobic. In addition, two concentrations of DNA were tested, 140 µM and 175 uM.

In order to see if the PC conjugate could bind to surface proteins, an anti-CD3 PC conjugate was synthesized and used to stain a J45.01 Jurkat cell line (ATCC). The T cells were spun down for 4 minutes at 500xg to pellet them. Approximately 95% of the solution was removed to obtain a high concentration

of cell. Then 5 µL of PC conjugate at 1 µM was added and the cells were allowed to sit on ice for 30 minutes. Afterwards, the cells were washed twice by diluting up with cold RPMI 1640 media (Life Technologies) +10% Fetal Bovine Serum (FBS, Life Technologies), spinning down the cells, and then removing the media. The cells were then placed on a microscope slide and visualized using an IX81 Fluorescence Microscope (Olympus). The GFP channel was used to visualize the Cy3 present in the PC conjugate. As a positive control, Jurkat cells were stained in the same fashion with anti-CD3 PE (Biolegend). Images were acquired and saved for analysis. Quantification of fluorescence intensity included cropping cells from the full image, isolating the clusters where the antibodies bound, and averaging the fluorescence intensity in that region.

3.2.2 - Photocleavable staining and release experiments

First, the optimum time for UV photocleavage was tested by measuring the amount of ssDNA-fluorophore release vs. UV exposure time. This was accomplished by synthesizing a PC conjugate with an anti-Streptavidin antibody. Streptavidin (SA) coated beads (10 µm diameter, Spherotech) were used in place of cells during control experiments. The beads were diluted to 2×10^6 beads/mL in 1% BSA and allowed to block for 15 minutes. The solution was then spun down at 2000 xg for 2 min and ~95% of the solution was removed to concentrate the beads. 5uL of PC Conjugate was then added to the solution and allowed to incubate on ice for 30 min. The bead solution was then washed twice with 1% BSA. The beads were separated into multiple vials and exposed to either UV-A (365 nm), UV-B (302 nm), or UV-C (254 nm) light in a VWR UV-Crosslinker. Beads were exposed for varying amount of time, ranging from 0 to 15 minutes at a dosage of 150 mW m^{-2}. In addition, negative control beads were used that had BSA coating rather than streptavidin.

After exposure, each sample was spun down at 2000 xg for 2 minutes, and the supernatant was collected and placed into a PDMS microwell. The PDMS microwell's substrate was a DEAL patterned

substrate that had the complementary ssDNA strand bound to the surface in order to capture the released DNA-fluorophore (see Chapter 4 for details on DEAL patterning). The samples were incubated on the DEAL substrate for 1 hr at room temperature to allow capture of photocleaved ssDNA-fluorophore. During this time, the PDMS was covered with aluminum foil to prevent photobleaching. The microwells were then washed with 1% BSA three times and then the slide was washed with 1x PBS, and finally 0.5x PBS. The slide was scanned on a Genepix 4400A fluorescent slide scanner. The fluorescence intensity of all spots in a microwell were averaged to obtain the relative amount of ssDNA-fluorophore that was released from the beads. This test was then repeated for Jurkat T cells using a PC Conjugate specific to the CD3 surface marker.

3.2.3 - Bead Tests in microfluidic device

A proof of principle experiment was run using a similar method as 3.2.2; however, instead of measuring the amount of ssDNA-fluorophore release in the supernatant, the beads were incubated in single cell chambers and the ssDNA-fluorophore was captured therein. The general procedure was shown in Figure 1. Beads were incubated with PC conjugate as described before, washed, and then loaded into a microfluidic device at a density of 1×10^6 beads/mL. The beads were then trapped in single cell chambers and exposed to UV-A light for 6 minutes. The microfluidic device was then incubated for 1 hour in order to allow diffusion of the ssDNA-fluorophore throughout the single cell chambers. During this time, video was captured to obtain bead counts in each chamber. Afterwards, beads were washed out and the slide was washed and scanned as described above.

The data was analyzed by measuring the fluorescence signal levels in each chamber using the Genepix Pro software. This data was then aligned with the cell counts obtained by analyzing the capture video of each chamber. The actual cell count was acquired by counting the number of beads in each chamber from the light microscope image. This was considered the "actual cell count." To determine if

48

the fluorescence signal could positively identify a chamber with beads in it, the mean plus two standard deviations from the zero cell signal level was used as a cutoff. A Mann-Whiteney U test was used to determine whether the fluorescence signal could separate zero and one bead chamber.

3.2.4 – T cell tests

In order to test the PC conjugate method of phenotype detection on cells, CD3 was detected on J45.01 Jurkat cell line. The toxicity of the anti-CD3 PC conjugate was tested with Jurkat T cells by mixing 50 ug/mL with $1x10^6$ cells/mL. The cell viability was measured after 16 hours of incubation. A control sample was tested by adding PBS solution instead of PC conjugate.

In addition to testing toxicity of the PC conjugate, the toxicity of UV-A light was tested on Jurkat cells. This was accomplished by exposing Jurkats to UV-A light (dosage: 150 W m^{-2}) for different amounts of time and measuring the presence of apoptotic markers using Annexin-V staining (Biolegend).

Testing the Jurkats on the microfluidic device was performed in a similar fashion to the bead experiment from Chapter 3.2.3. The Jurkats were concentrated down at 500 xg for 4 minutes and 5 µL of an anti-CD3 PC conjugate was added (stock concentration was 0.5mg/mL). This was incubated on ice for 30 min and then washed twice by spinning down, discarding supernatant, and diluting back up with RPMI 1640 media. The Jurkats were loaded into the microfluidic device at a concentration of $1x10^6$ cells/mL. The device was then exposed to UV-A light for 6 minutes. The remaining procedure and data analysis were conducted as described in Chapter 3.2.3.

3.3 – Results

3.3.1 - Synthesis of conjugate

The initial yield for the PC conjugate was lower than a standard DEAL conjugate by a factor of ~2. After increasing the conjugate DNA from 140 to 175 µM and decreasing the DMF amount from 20 µL

to 10 µL, the FPLC peak was improved by a factor of four (Figure 2). This improvement resulted in a

drastic difference in subsequent experiments. Before the improved method of synthesis, the signal to

noise ratio was approximately 1.6:1. After the improved method of synthesis, the signal to noise ratio
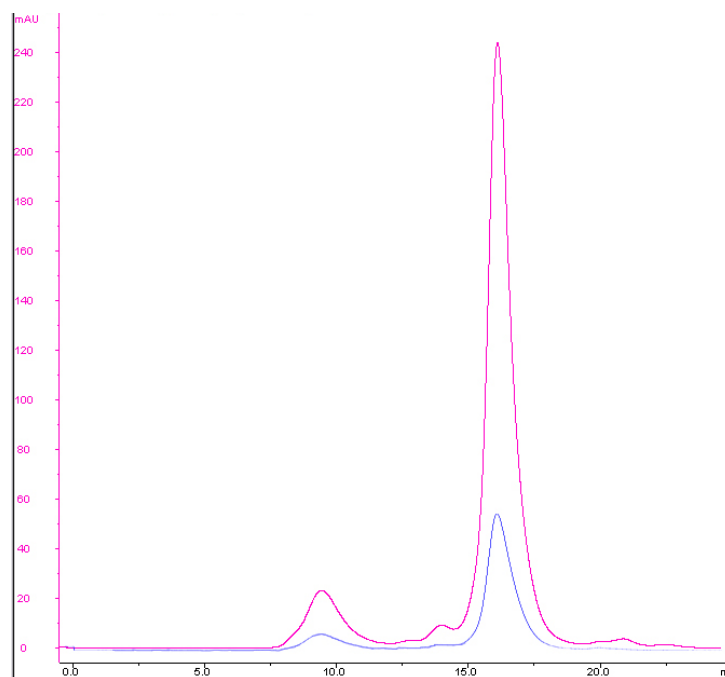
was 2.8:1.

Figure 2: Overlaid FPLC peaks of PC conjugate made with 140uM DNA + 20 µL DMF (Purple Line) or 175 µM DNA + 10 µL DMF (Pink Line). The x-axis is the mL of solution eluted from the FPLC column and the y-axis is the mAU absorbance at 280 nm.
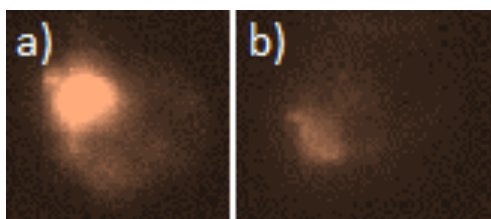

Figure 3: Fluorescence image of Jurkat cells stained with a) anti-CD3 PE or b) anti-CD3 PC Conjugate.

Images of Jurkats stained with a standard fluorophore and PC conjugate are shown in Figure 3. A

standard staining fluorophore (anti-CD3 with PE dye) used in FACS experiments is shown in Figure 3a.

Figure 3b is the PC conjugate. The maximum fluorescence of the PC conjugate was 64% that of the

standard FACS fluorophore. Although this was lower, it was sufficient to detect the CD3 with high signal

to noise.

50

3.3.2 – Photocleavable staining and release experiments

Beads stained with PC conjugate are shown in Figure 4 after 0, 5, and 15 minutes of UV-A light

exposure. The average intensity of the beads vs. UV-A exposure time is shown in Figure 5 with the

microarray captured fluorescence signal overlayed. The graph clearly shows that the fluorescence of the

beads decreases as the ssDNA-fluorophore is photocleaved off. Simultaneously, the ssDNA-fluorophore

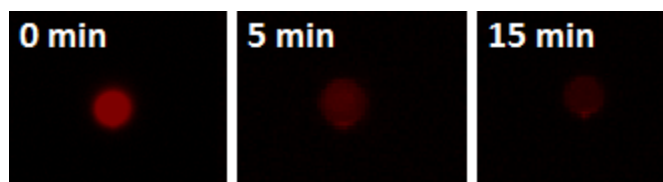captured on the microarray increases.

Figure 4: Fluorescent images of SA coated bead stained with PC Conjugate after 0, 5, and 15 minutes of UV-photocleavage. The intensity of the bead decreases as DNA-fluorophore is cleaved off.
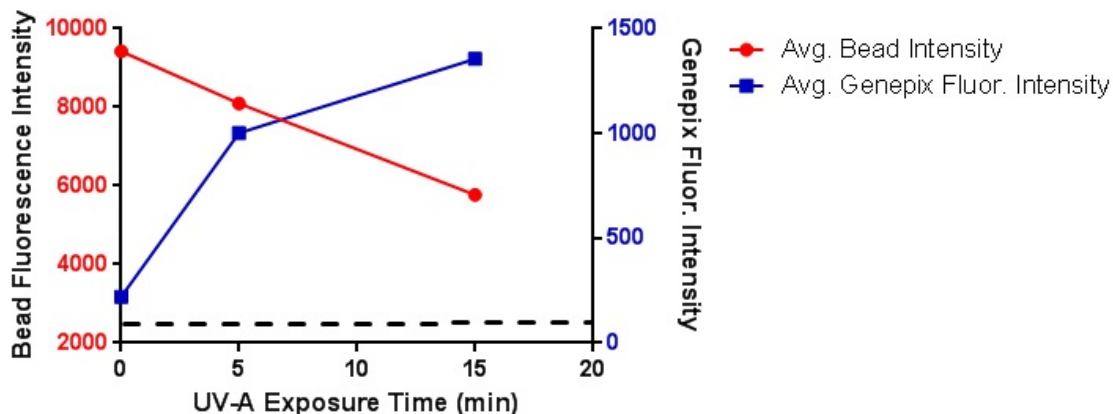
Figure 5: Graph of average bead fluorescence intensity as a function of UV exposure time, as well as the measured Genepix fluorescence intensity measured at the same time. The dashed black line is the background level for the bead fluorescence.

The average fluorescence signal captured by the substrate after 0, 5, and 15 minutes is shown

for all wavelength of UV light in Figure 6. UV-A and UV-B are significantly better at photocleaving the PC-

51

conjugate than UV-C. Furthermore, UV-B has a slight advantage over UV-A, with UV-B photocleaving

23% and 8% more ssDNA-fluorophore at 5 and 15 minutes of exposure time, respectively.
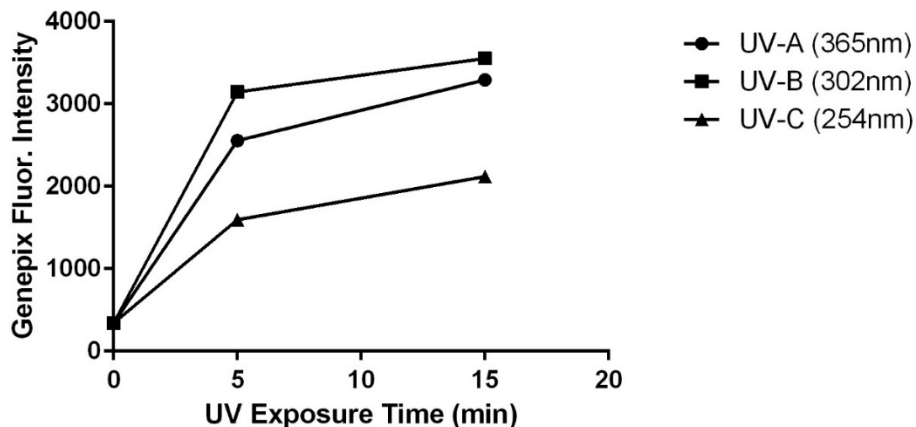

Figure 6: Comparison of UV exposure time and measured Genepix fluorescence signal versus the UV wavelength. Each point is the average intensity of the DNA-fluorophore captured on a DEAL substrate.

The average PC conjugates released over time were then measured with higher frequency to

determine the optimum photocleavage time. The results from this experiment are shown in Figure 7.

The signal to noise was 4.4:1 at 6 minutes and 5.6:1 at 15 minutes.
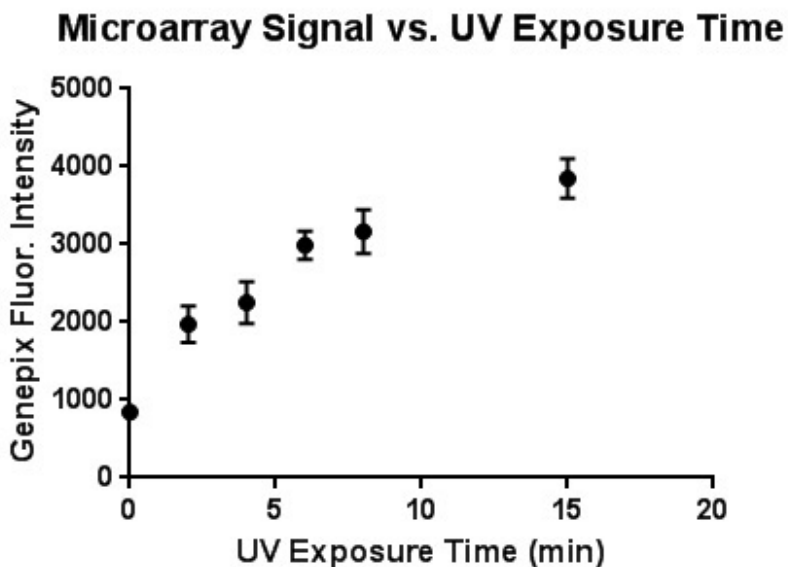

Figure 7: Plot of Genepix Fluorescence intensity and UV exposure time with high frequency measurements. Each data point shows the average and standard deviation of the DNA-fluorophore intensity captured on a DEAL substrate. The DNA-fluorophore was released from beads exposed for the indicated time.

Jurkats were stained and photocleaved in a similar fashion; however, only UV-A was used due to the phototoxic effects of UV-B and UV-C [7]. The average fluorescence signal captured after photocleaving for 0 and 15 minutes was 6,500 and 17,000. This corresponded to a signal to noise ratio of 2.6:1.

3.3.3 - Bead Tests in microfluidic device

Figure 8 shows an example of anti-Streptavidin PC conjugate-labeled beads in microfluidic chambers. In the light microscope image (Figure 8a), the beads show up as small black dots. In the fluorescence image (Figure 8b), they show up as red dots due to the fluorophore on the PC conjugate.
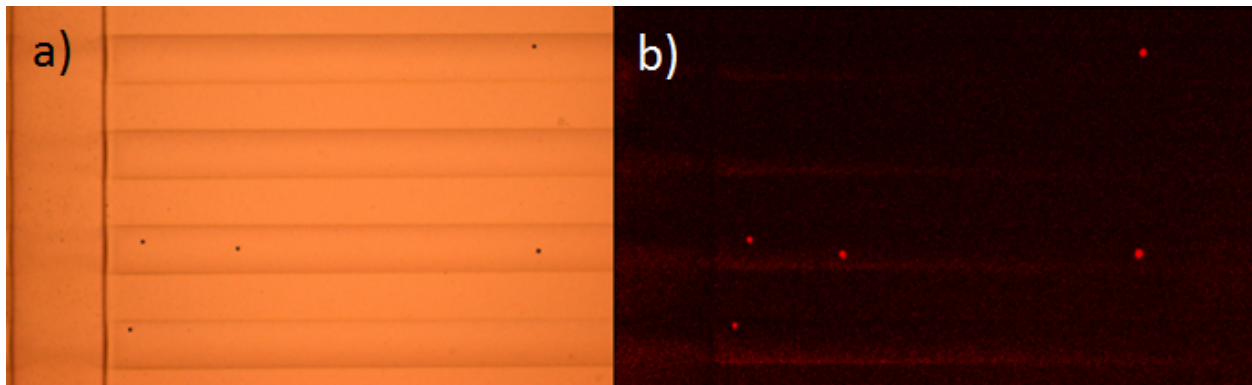
Figure 8: Images of beads in single cell microfluidic device. a) Bright field image of beads in four different single cell chambers. The bead counts in each chamber, counting from the top chamber down, are 1, 0, 3, and 1. b) The same region as the image in part a, taken with a Cy5 filter on a fluorescence microscope.

A representative image of the captured fluorescence signal in single cell chambers is shown in Figure 9. Each chamber's actual bead count is displayed at the top and the white boundary shows the single cell chamber's edge. Visually, the difference in fluorescence can be discerned between chambers with 0, 1, and 2-3 beads per chamber.

Figure 9: Genepix Fluorescence Image of seven single cell chambers. The green fluorescence is the captured ssDNA-fluorophore released during the bead test in the microfluidics device. Each chamber has the number of beads displayed at the top.
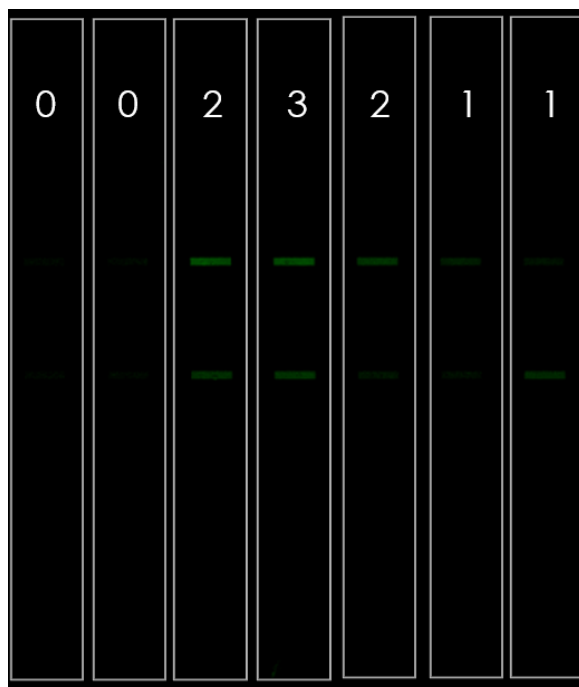
The positive predictive rate of cell detection (that a chamber had at least one cell) was 92% using the photocleavable fluorescence signal. The false positive rate was low as 0.5%.

The distribution of fluorescence signal captured vs. bead count is shown in Figure 10a and a scatter plot in Figure 10b. Bead count distributions were significantly distinguishable using a Mann-Whitney U Test (0 bead, 1 bead, and 2 beads). Most importantly, 91.4% of single cell chambers had a fluorescence signal higher than all zero cells chambers. 97% of two bead chambers had a fluorescence signal higher than all zero bead chambers. This indicates that zero and two bead chambers were very distinguishable from zero bead chambers in terms of their fluorescence signal, which demonstrates a proof of principle for detecting beads in microfluidic single cell chambers using a photocleavable conjugate.
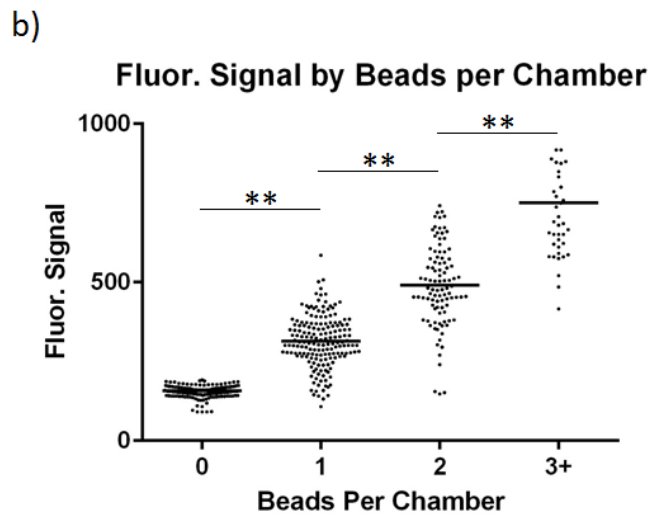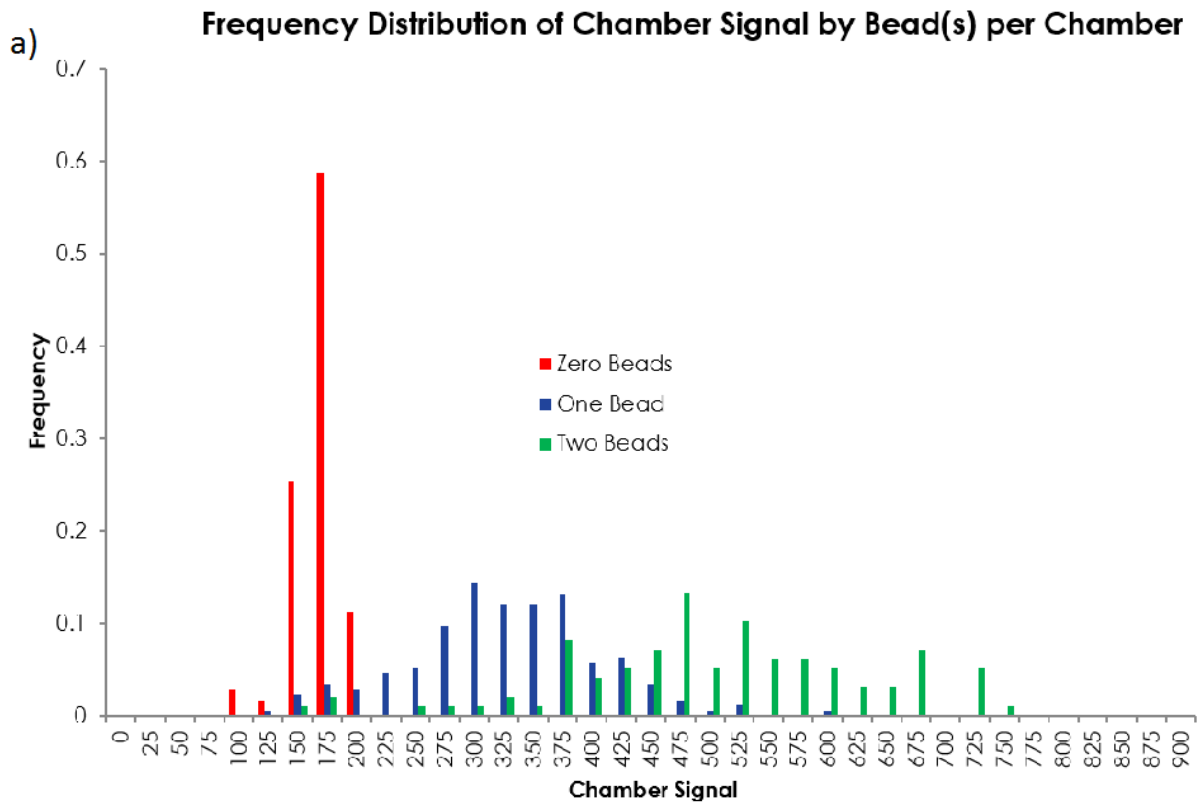
Figure 10: Single cell chamber signal vs. number of beads per chamber. a) Data is represented as a histogram. b) Data is represented as a scatter plot, where each point is a different microfluidic chamber.

3.3.4 - T cell tests in microfluidic device

A graph of Jurkat viability vs. treatment is shown in Figure 11. The y-axis is the percent of

apoptotic T cells found after exposing to UV-A light or to the stimulant phorbol 12-myristate 13-

acetate/Ionomycin (PMA/I). The percent of apoptotic cells found after 15 minutes was significantly

higher than 5 minutes. Furthermore, the percent of apoptotic T cells found at 5 minutes was less than

the PMA/I stimulation. Therefore, using this information, and the fact that the amount of DNA-

fluorophore released between 6-15 minutes did not change significantly, a UV-A exposure time of 6

minutes was subsequently used for Jurkat experiments.
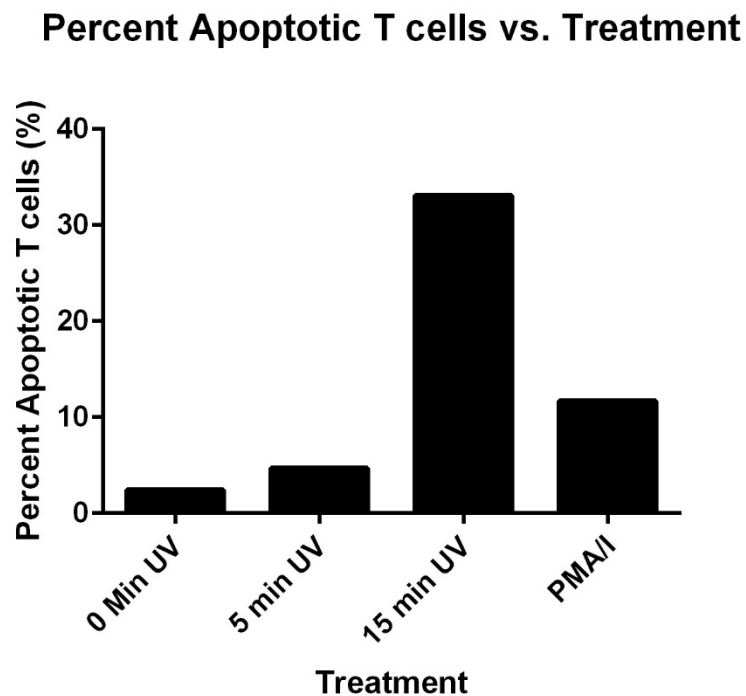
Figure 11: Cell apoptosis percentage vs. treatment type. Jurkat T cells were measured for the presence of Annexin V after 16 hours of each treatment.

A representative image of a single cell microfluidic test performed using Jurkats is shown in

Figure 12. The cells in the bright field image are the white circular objects with dark borders.
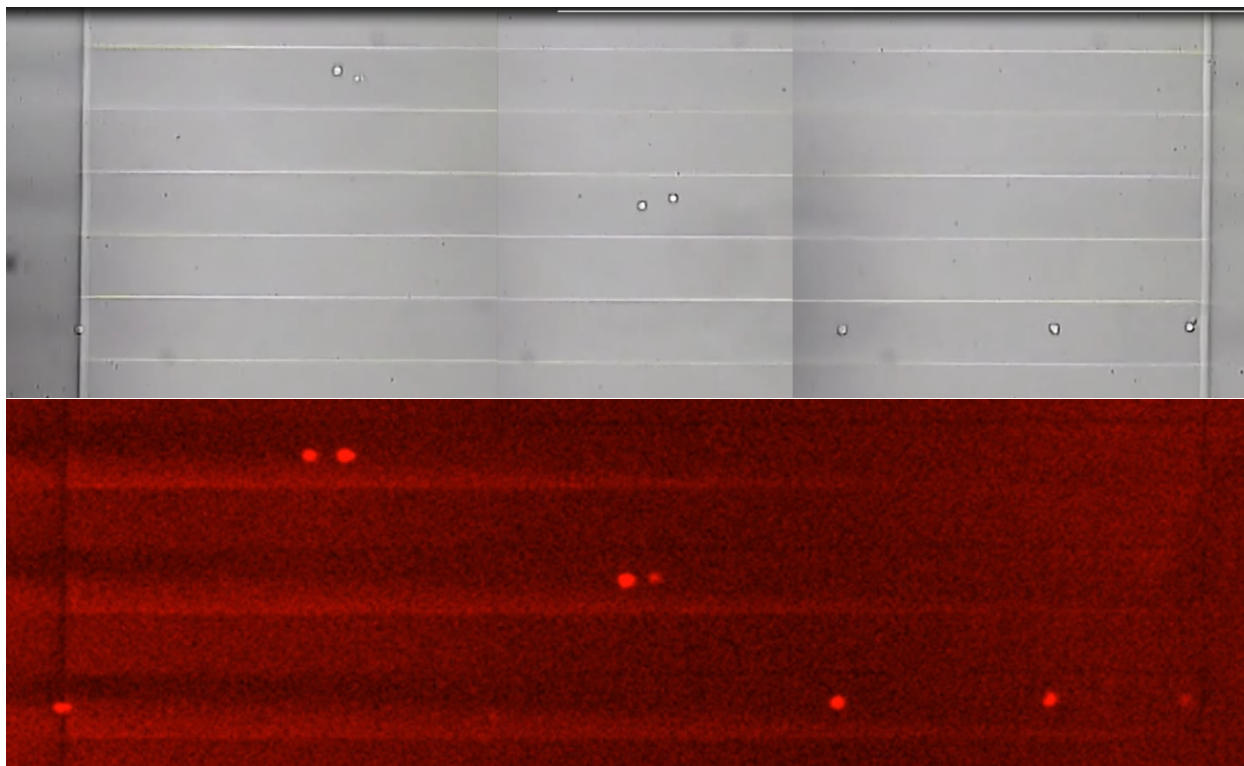
Figure 12: Images of single cell microfluidic chambers with Jurkat T cells. a) Bright field image of chambers. b) Fluorescence image of chambers.

Using the fluorescence signal from the release PC conjugate, the positive predictive rate was

49% with a false positive rate of 1.4%. Figure 13 shows a scatter plot of fluorescence signal vs. cell count

per chamber.
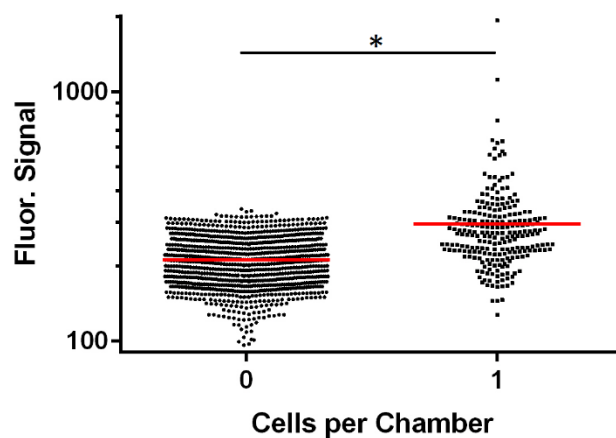
*Figure 13: Scatter plot of single chamber signals vs. number of Jurkat in chamber. The significance indicator is shown above the geoms. Significance markers: * = p < 0.05.*

3.4 – Discussion

Comparing the fluorescence signal of T cells with anti-CD3 PE and the PC conjugate, we observed a significantly higher signal in the case of anti-CD3 PE. This is likely because of the higher extinction coefficient of PE relative to the Cy3 fluorophore used in the PC conjugate (max extinction coefficients for PE and Cy3 are $2x10^6$ and $1.4x10^5 M^{-1}cm^{-1}$, respectively) [8, 9]. In spite of its lower extincition coefficient, Cy3 was used because of its ready integration as an "internal" group in the photocleavable DNA-fluorophore synthesis. If a stronger fluorophore was able to be integrated as an internal group, this would likely improve the detection sensitivity of the photocleavable marker.

The detection sensitivity of the Jurkats was significantly lower than that of the streptavidin coated beads: the rate of distinguishing one cell chambers from zero cell was 49% for Jurkats and 92% for the beads. This demonstrates a significant gap between these two systems. The Jurkat experiment's lower positive predictive value demonstrates there are significant differences between the beads and cells. One key difference is that the surface coverage of streptavidin on beads is much higher than CD3 on Jurkats. Human T cells regularly have ~$6x10^4$ CD3 per cell, whereas beads have $1.8x10^7$ SA proteins per bead.[2] [10] Another difference was that cell membrane proteins like CD3 can be internalized and degraded, which would reduce the number of PC conjugates detected [11]. In order to improve the lower sensitivity of detecting CD3 on Jurkats, one could 1) use stronger fluorophores (as discussed in the first paragraph), 2) increase the loading of fluorophores on the PC conjugate, 3) amplify the detected signal, or 4) reduce the internalization of CD3 with Filipin. The first three ideas would be the most facile changes, as they would not interfere with the biology, as the fourth idea would.

One problem with the PC conjugate was the stability of the construct itself. Batches would come out working well, whereas other batches would have nearly no activity. Having reagents that

---

[2] SA bead density was calculated using the following approximations: 1) SA is a sphere with radius 2.25nm [13], 2) SA arranged in a monolayer, hexagonal layout of SA, and 3) the bead is a sphere is 10 μm in diameter.

perform with varying degrees of success would not be acceptable in the long term, and therefore, time would need to be spent debugging this issue. One likely cause of this is light exposure photocleaving the DNA-fluorophore off prematurely. These reagents were made in a communal lab, so there was a chance they would inadvertently be exposed to light. In this project, the reagent was protected from fluorescent lights by coating anything it was placed into with aluminum foil (e.g., FPLC column, microfluidic device) and storing it in black-colored vials. However, it will inevitably be exposed to UV during the FPLC run, where it used 280 nm wavelength light to determine the absorbance of each fraction. Finding an alternative method to purify the PC conjugate would likely improve its stability.

The viability of the Jurkats limited the amount of UV-A dose we could apply towards photocleaving the PC conjugate. As shown in Figure 11, the viability decreased rapidly between 5 min and 15 minutes of UV-A exposure. Malanado et al. showed that 5000 J m^{-2} is the maximum dosage that thymocyes could receive before DNA fragmentation occurred [7]. In our experiment, 6 minutes of UV-A was equal to 5400 J m^{-2}. Therefore, 6 minutes of UV-A in our experiment is just at the safe limit for Jurkats. The downside to such a short UV-A exposure time is the PC conjugates may not be fully cleaved after 6 minutes. This was demonstrated in Figure 5 when measuring the reduction in fluorescence as a function of UV exposure time: the bead intensity did not reach the background level, even after 15 minutes of UV-A. Therefore, it would be advantageous to either 1) improve the photocleaving efficiency on the PC conjugate or 2) find an alternative cleavage method. One such alternative cleavage method would be to use DNA as a linker and cleave it using a restriction enzyme.

It is worth noting that this method is not designed to be a cell counting mechanism since 1, 2, and 3 cell chambers were not perfectly separable, even in the bead test. The PC conjugate would best be applied in situations where the cell count is determined automatically (see Chapter 1.2.1) or a cell trap is used in the microfluidics. A cell trap would guarantee that each chamber would only have one cell

59

(or zero cells), allowing the user to trust that the phenotype data obtained from the PC conjugate readout is from a single cell.

With the resolution of the issues raised in this Discussion section, this method holds great promise. In a single cell experiment where three fluorescent dyes are tested on four separate microfluidic devices, the method saves a great deal of time relative to manual analysis. Furthermore, the potential ability to quantitate the data is a further advantage.

3.5 – Conclusions

A photocleavable conjugate was developed in order to detect surface markers and phenotype cells in a single cell microfluidic device. Two proof of principle experiments were conducted with beads and Jurkat T cells. Surface streptavidin on beads was readily detected with a 92% positive predictive rate. Jurkat T cells were detected with a 49% positive predictive rate and zero cell chambers were distinguishable from one cell chambers via a Mann-Whiteny U test. These experiments show promise in determining cell phenotype on-chip via PC conjugates.

References (Chapter 3)

[1]   L. Herzenberg et al., "The history of the fluorescence activated cell sorter and flow cytometry: a view from Stanford.," *Clin Chem,* vol. 48, pp. 1819-1827, 2002.

[2]   G. Nolan et al., "Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum," *Science,* vol. 332, pp. 687-696, 2011.

[3]   J. Heath et al., "A clinical microchip for evaluation of single immune cells reveals high functional

heterogeneity in phenotypically similar T cells," *Nature Medicine,* vol. 17, p. 738–743, 2011.

[4] J. Heath et al., "Multifunctional T-cell Analyses to Study Response and Progression in Adoptive Cell Transfer Immunotherapy," *Cancer Discovery,* vol. 3, pp. 418-429, 2013.

[5] C. Love et al., "Polyfunctional responses by human T cells result from sequential release of cytokines," *Proc. Natl. Acad. Sci.,* vol. 109, pp. 1605-1612, 2012.

[6] S. Weissleder et al., "Photocleavable DNA Barcode–Antibody Conjugates Allow Sensitive and Multiplexed Protein Analysis in Single Cells," *J. Am. Chem. Soc.,* vol. 134, no. 45, p. 18499–18502, 2012.

[7] C. e. a. Maldonado, "Ultraviolet exposure of thymoctes: selective inhibition of apoptosis," *Int. J. Radiat. Biol.,* vol. 90, no. 6, pp. 445-450, 2004.

[8] "EXTINCTION COEFFICIENTS AND FLUORESCENCE DATA," Glen Research, [Online]. Available: http://www.glenresearch.com/Technical/Extinctions.html. [Accessed 5 May 2015].

[9] L. Stryer et al., "Fluorescent phycobiliprotein conjugates for analyses of cells and molecules," *J Cell Biol.,* vol. 93, no. 3, p. 981–986, 1982.

[10] C. Lambert et al., "Variability of CD3 membrane expression and T cell activation capacity," *Cytometry Part B ,* vol. 78B, 2010.

[11] W. Mitch et al., "Protein Degradation by the Ubiquitin–Proteasome Pathway in Normal and Disease States," *JASN,* vol. 17, no. 7, pp. 1807-1819, 2006.

[12] R. Fan et al., "High-throughput Secretomic Analysis of Single Cells to Assess Functional Cellular Heterogeneity," *Analytical Chemistry,* vol. 85, no. 4, p. 2548–2556, 2013.

[13] M. Rao et al., "Immobilization of streptavidin on 4H–SiC for biosensor development," *Applied Surface Science,* vol. 258, p. 6056–6063, 2012.

Chapter 4: Predicting Adoptive Cell Transfer Therapy Success with T cell Polyfunctionality using Single Cell Microfluidic Assays

4.1 – Introduction

Immunotherapy has shown great promise in treating cancer. There exist multiple methods of using the immune system to tumor cells, with the most prevalent being cancer vaccines, checkpoint inhibitors, and adoptive cell transfers [1, 2, 3]. Their unifying mechanism is harnessing the immune system to dispose of tumor cells, rather than using a direct acting exogenous agent.

Adoptive cell transfer (ACT) has been developed over the past 100 years, reaching a sprint in the late 20th century. In 1922, it was suggested that tumor infiltrating lymphocytes (TILs) were associated with longer patient survival [4]. After the discovery that TILs grown from melanoma tumors could lyse autologous tumor cells, a clinical trial was run using TILs to treat metastatic melanoma [5, 6, 7]. Since then, this therapy has shown objective responses in 40-70% of Stage IV melanoma patients, and complete response rates in 10-40% [8, 9]. However, the pre-selection process requires the patient to have resectable tumor with TILs present, which excludes 10-15% of candidates from taking part in this promising therapy [10].

An alternative form of ACT uses genetically engineered, autologous T cells, which does away with the need for TILs. The method is outlined in Figure 1. First, a patient undergoes a leukapheresis to harvest autologous T cells. A viral vector is then inserted into the T cells. This vector encodes for a T cell receptor (TCR) that recognizes a specific tumor MHC peptide [11]. Next, the T cells are expanded until they reach billions to trillions in number. Finally, the patient undergoes lymphodepletion treatment and the T cells are infused. These T cells then go on to kill the tumor cells.
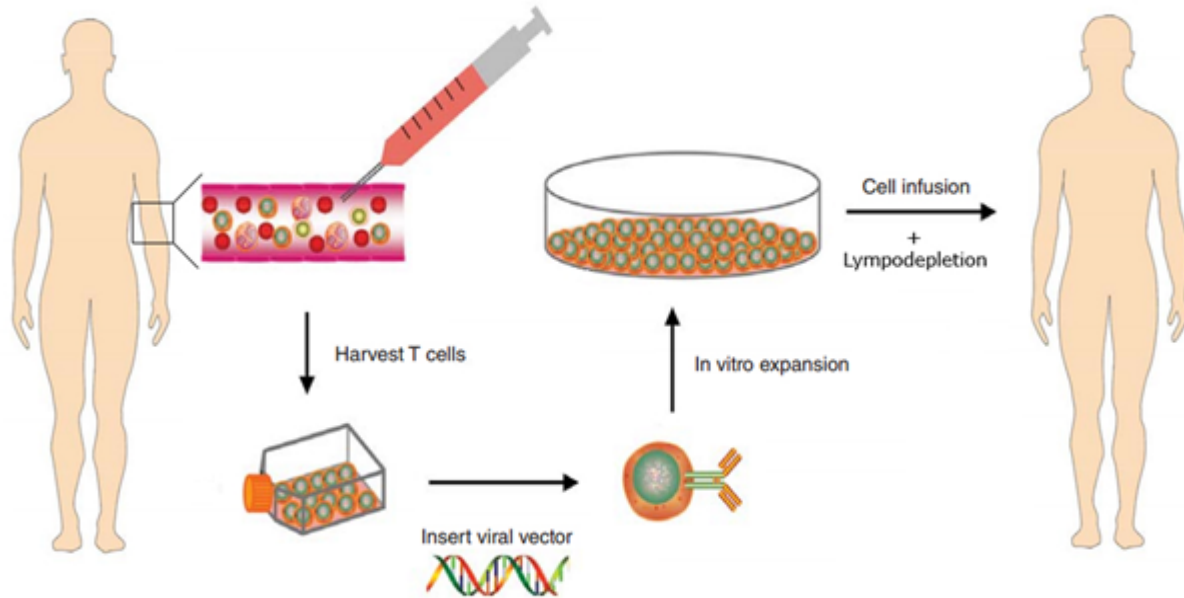
Figure 1: Overview of genetic engineered ACT. Adapted from [8].

Though genetically engineered ACT is promising in allowing more patients to undergo immunotherapy, the response rates are lower than that of the tradition TIL ACT, ranging from 10-30% for an objective response and complete response rates of < 20% [12]. Therefore, it would be desirable to understand why a small percent of patients do well, and find an avenue to increase the success rate.

Previously, Ma et al. from the Heath Lab investigated the single cell protein secretion of T cells from patients in a genetic engineered ACT trial [13]. Measurements were taken every ~10 days post-T cell infusion and showed a strong correlation between T cells polyfunctionality, or multiple protein secretions from the same single cell, and tumor lysis. Polyfunctional T cells are known to be advantageous in immune response, for instance, in the case of HIV-1 vs. HIV-2. HIV-2 is considered a less pathogenic disease, and has higher numbers of polyfunctional T cells, as compared to HIV-1 [14] [15]. Therefore, we hypothesized that T cell polyfunctionality was an important factor in the success of genetically engineered ACT.

Using this hypothesis, we set out to determine if T cell polyfunctionality can be used to predict whether a patient will succeed or fail a genetic engineered ACT therapy. The prediction was based on the extent of polyfunctionlaity present in the initial T cells infusion product (referred to as "Pre-Infusion" T cells). Like Ma et al. from the Heath Lab, single cell microfluidics were employed to measure the functional protein secretion of T cells from 15 patients in two different genetic engineered ACT trials. Multiple data analysis methods were used to relate this single cell data with treatment success.

4.2 – Materials and Methods

4.2.1 – DNA Barcodes

DNA barcodes were patterned using a PDMS template. The PDMS was made using Sylgard 184 (Dow Corning) in a 10 to 1 ratio of Part A to Part B. The PDMS was mixed for 2 minutes by hand using a glass-stir rod and then 10 minutes in a Unicyclone (Fancort Industries, Inc.). The barcode masters were treated with chlorotrimethylsilane (CTMS) for 10 minutes before the mixed PDMS was poured over top. The PDMS was degassed for 10 minutes under vacuum and then baked for 2 hours at 80 degC. Holes were punched into the PDMS using a 0.5mm diameter punch. The PDMS was cleaned and adhered to a poly-lysine coated slide (Thermo Fisher). This was baked for another 2 hours at 80 °C to yield a completed barcode.

The barcode was then used to flow pattern ssDNA onto the surface of the poly-lysine slide. The night before flowing the ssDNA, 4uL of 0.1% poly-lysine (Sigma) was flowed into each chamber at 4psi. This was left to flow until the channels were dry, which was typically 12-16 hours. Next, 300 µM ssDNA (IDT DNA) in 40% DMSO and PBS is mixed 1:1 with 2mM BS3 in PBS. This is repeated for every different sequence of ssDNA. These are listed in Chapter 2 Appendix A. This is left to react at room temperature for 1 hour. Afterwards, 5uL of each ssDNA is flowed into the device for 2 hours. The devices are kept in a 100% humid chamber for another 2 hours before the PDMS is peeled from the slide and then washed

with 0.02% SDS once and water three times. The slides are stored in a dessicator (humidity < 20%) until use.

4.2.1 – PDMS Devices

PDMS was originally made using RTV 615 PDMS (Momentive) but was switched to Sylgard 184. The Sylgard Part A and Part B were mixed in a ratio of 10 to 1 for the control (thick) layer and 20 to 1 for the flow (thin) layer, mixing for 2 minutes by hand using a glass-stir rod, then mixing on a Unicyclone for 10 minutes. During this time, the silicon wafers with the single cell microfluidic design patterns were placed in a chamber with 5-10 drops of CTMS for at least 2 minutes. The control layer PDMS was poured onto the silicon wafer masters to a thickness of ~6mm and then placed in a degassing chamber for 45-60 minutes. The thin layer PDMS was placed in the degassing chamber while still in the stirring cup. After degassing the thin layer PDMS, it was poured over the silicon flow master, and spun at 500 rpm for 5 seconds (100 rpm/sec ramp), followed by 2000 rpm for 60 seconds (300 rpm/sec ramp). After spinning, the wafer was left on a flat surface for 5 minutes to allow the PDMS to settle.

Both the thick and thin layer PDMS were placed in an 80 °C oven for 15-20 minutes, and then allowed to cool for 15 minutes. Holes were then punched in the control ports in the thick layer. The two layers were aligned and adhered, and the presence of dust between the layers was then searched for using a light microscope.

The PDMS was then baked at 80 °C for another 1.5 hours to bond the two layers together. The PDMS was cooled for 15 minutes and then the flow channel ports were punched using a 1mm punch. The PDMS was then bonded to a pre-patterned barcode slide and baked for 2 hours. The device could be used for a single cell experiment within one week of its creation.

4.2.2 - Sorting T cells

T cells were thawed overnight after washing/diluting in RPMI 1640 media + 10% FBS. The T cells were spun down at 500 xg for 4 minutes to pellet the cells. The following staining dyes were added per the manufactutor's instructions: 7-AAD (Biolegend), CD3 (Biolegend, OKT3), CD4 (Biolegend, OKT-4), CD8 (Biolegend, HIT8a), and MART-1 tetramer (MBL Int.). The pellets were then broken up and placed on ice for 30 minutes. The cells were then washed with cold media and diluted up to ~10^7 cells/mL.

T cells were sorted using a BD Aria II or Aria III machine. Cells were gated on FSC-A, FSC-H, SSC-A (singlet, lymphocyte identification), 7-AAD, CD3, CD4, CD8, and MART-1 tetramer. The machine was calibrated with CompBeads (BD Bioscience) prior to sorting. The T cells were sorted into two groups, 7-AAD$^-$CD3$^+$CD4$^+$CD8$^-$MART-1$^+$ and 7-AAD$^-$CD3$^+$CD4$^-$CD8$^+$MART-1$^+$, which are henceforth referred to as CD4 and CD8, respectively.

4.2.3 - Microfluidic Test

The microfluidic test was started concurrently with the sorting, so that when the cells were done sorting, the microfluidic device was prepared.

First, a 3% BSA in PBS solution was flowed in the microfluidic device and dead-end filling was performed to remove bubbles in the flow channels. The BSA was then flowed for 1 hour to block the surface. Conjugates were then diluted to 50 µM in 3% BSA and 200 µL was flowed in the device for 1 hour. At this point, the sorted T cells were washed three times with RPMI 1640 + 10% FBS, diluted up to a concentration of 10^6 cells/mL and then stimulated. The stimulation included MART-1 tetramer alone (8 µL per 3x10^6 cells), and MART-1 plus CD28 (CD28 was diluted 1:250 from a stock 0.5 mg/mL concentration). The T cells were allowed to sit with the stimulant for 10 minutes before loading into the microfluidic device.

The T cells were then flowed into the device and the pneumatic chambers were actuated to sequester them into the microfluidic chambers. Videos were then taken of the cells and then the device was placed in a 37 °C incubator with 5% CO_2 for 14 hours. After this incubated, the microfluidic device was removed from the incubator and the cells were washed out. The secondary antiboides were diluted to 50 µM in 3% BSA, along with Streptavidin-Cy5 (0.5mg/mL, eBioscience) and a reference ssDNA-Cy3 (10uM), both diluted at a ratio of 1:100 in BSA. 200 µL of this solution was flowed over 1 hour, followed by 45 minutes of 3% BSA to wash the device out. The PDMS was then stripped off of the slide and the slide was dunked in 0.5x PBS three times. After drying the slide on a slide-spinner (VWR), it was scanned on a Genepix 4400A. The gain and power were set at 450/15% for Cy3 and 600/80% for Cy5.

4.2.4 - Patient Study

The T cells studied were collected from metastatic melanoma patients enrolled in a clinical study performed at UCLA [UCLA, Los Angeles, CA; IRB (08-02-020 and 10-001212) under an IND filed with the U.S. Food and Drug Administration (IND 13859)] [11]. Patients were chosen based on the presence of MART-1 marker on tumor immunohistochemistry assays and HLA-A*0201[+]. The Response Evaluation Criteria in Solid Tumors (RECIST) metric was used to determine tumor burden throughout the study. This trial is henceforth referred to as the F5 trial. A second trial performed by S. Rosenberg targeted NY-ESO-1 in synovial sarcoma and melanoma patients [16]. This trial will henceforth be referred to as the NIH trial.

The T cells measured in this study were frozen aliquots of T cells that most closely resembled the "pre-infusion product" that were to be injected into the patients: these T cells had been transduced with the MART-1 or NY-ESO-1 TCR viral vector and expanded previously.

4.2.5 - Data Analysis

The cell count was acquired by counting the number of cells in each chamber using the videos taken prior to incubating the cells overnight. The protein signals were acquired by using the "Block Aligning" method in the Genepix Pro 7 software (Molecular Devices). A 20 µm spot size was aligned onto the barcode, and the mean signal was used as the protein's fluorescent signal. These datasets were then aligned using custom written Matlab code. This program also scanned for possible cross-talk and calculated local background levels, as detailed in Chapter 1.2.2.2.

The percent of T cells secreting a certain protein was calculated in both a relative and absolute fashion:

$$Absolute\ Secretion\ Rate_j = \frac{\sum_{i=1}^{n} \begin{cases} 1\ if\ Protein_{ij} > Background\ Threshold \\ 0\ if\ Protein_{ij} < Background\ Threshold \end{cases}}{n} \times 100\%$$

Where *Protein$_{ij}$* is the protein level of the jth protein for the ith cell and *n* is the total number of single cells for that patient.

$$Relative\ Secretion\ Rate_j = \frac{Absolute\ Percent\ T\ cells\ Secreting_j}{\sum_{j=1}^{p} Absolute\ Percent\ T\ cells\ Secreting_j} \times 100\%$$

As an initial test, protein signal levels were correlated with patient outcome. This was done by performing a Pearson's and Spearsman's correlation and recording the p value. A Bonferroni Correction was used to account for the large number of correlation hypotheses tested.

To predict patient outcome, a partial least squares regression model was generated. This method is useful in fitting multivariate data to response outcomes, as it uses algorithmically determined principal components to fit the data (rather than standard principal components like in Principal Component Regression). Feature selection was performed using a custom written Genetic Algorithm,

which used Q^2 score as the quality metric for gene survival. Q^2 is a measure of the PLSR model's "goodness of fit" from cross-validations, similar to R^2 in linear regression. It is defined as:

$$Q^2 = 1 - \frac{PRESS}{TSS}$$

where PRESS is the Predictive Residual Sum of Squares and TSS is the total sum of squares.

Finally, the polyfunctional strength index (pSI) was fit to the response data to determine if it was predictive. This was done by using a linear discriminant analysis estimation. The equation for calculating the pSI is shown below [13]:

$$pSI = \% \text{ polyfunctional T cells within T cells} \times \sum_{i=1}^{P} \text{MFI of cytokine } i \text{ of the polyfunctional T cells}$$

where p is the number of proteins in the "function panel." The function panels are proteins that share the same role in the body, determined in reference [13]. For a T cell to be considered as polyfunctional, it was required to be secreting at least 3 of the total 10 proteins measured. The average protein levels of T cells that expressed proteins for a specific function were then calculated and summed over all proteins in that function. This is multiplied by the percent of T cells that had that function to obtain the pSI.

4.3 – Results

4.3.1 - Patient and T cell Statistics

The patient statistics are detailed in Appendix A, Table 2.

Cell viability of the patient T cells was 15±9% and 7±7% for the F5 and NIH trial, respectively. In order to rule out the cell viability accounting for differences in the T cell measurements, it was correlated with individual protein levels and average pSI magnitude. This yielded R^2 values < 0.01, signifying no correlation between the viability and T cell functionality.

The T cell statistics showed large differences between patient T cell functionality. A scatter plot of the protein copy numbers per T cell is shown in Figure 2. One can see strong variation in protein secretion level between patients. This was especially true in the case of CCL4, where patient F5-2's T cells secreted nearly 10x higher on average than all the other patients. CCL4's average signal CV% between patients was 109%, whereas the average of the remaining proteins was 34.5%. Far fewer copies of TNFa were secreted on average; however, this was not the case when PMA/I stimulation was used to stimulate the T cells (not shown). An additional noteworthy aspect displayed in Figure 2 is the right skew of CCL4, TNFa, and IFNg levels, due to large outliers. For instance, the average CV% of CCL4's protein levels from the same patient was 305%.

The absolute and relative percent of T cells secreting a specific protein is shown for the F5 trial in Figure 3a and 3b, respectively. Across all proteins, the CV% of T cell secretion rate was 52.5%±16 and 38.8%±14 for the absolute and relative cases, respectively. This shows that the distribution of the relative T cell secretion rate is fairly similar across patients, as can be seen visually in Figure 3b. Much of the difference in this comes from patient F5-7. When this patient was removed, the CV% for dropped to 33%.
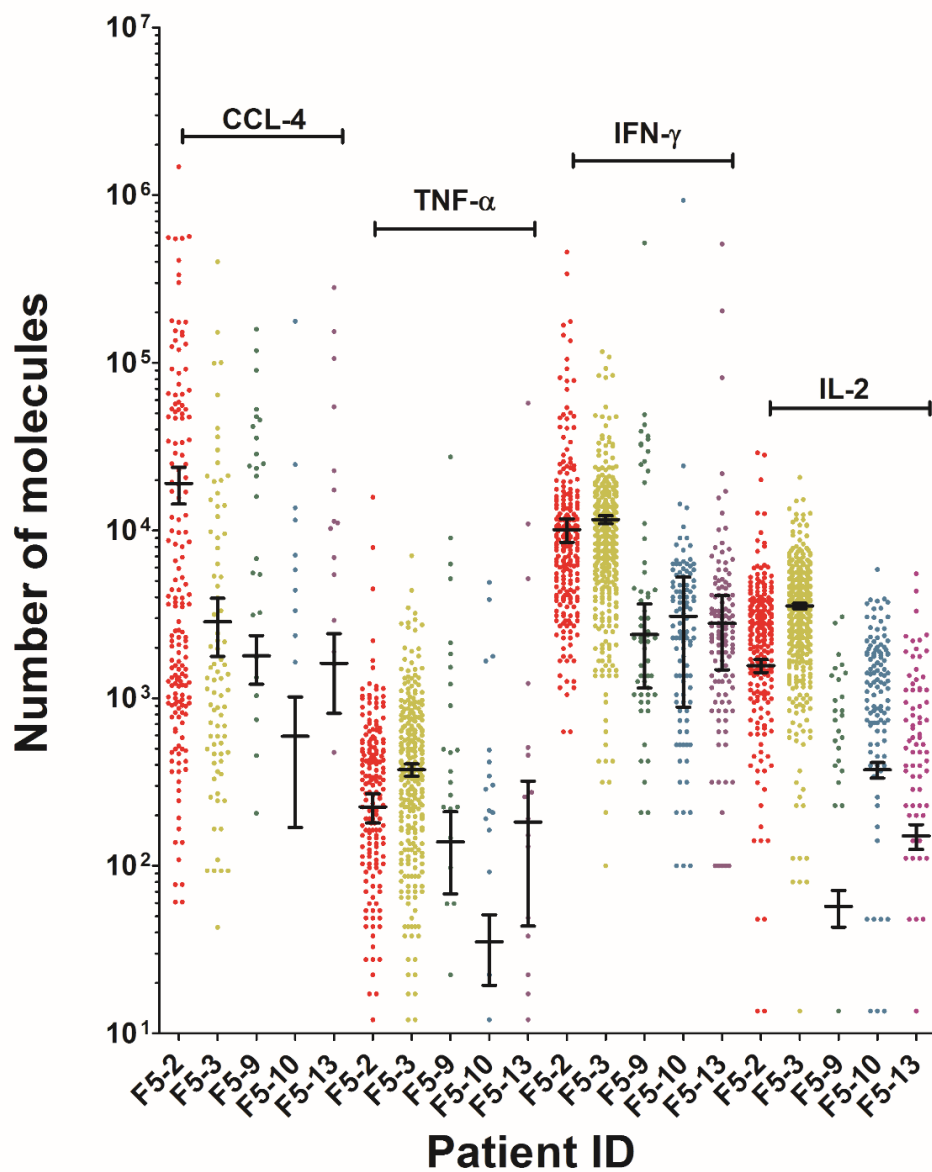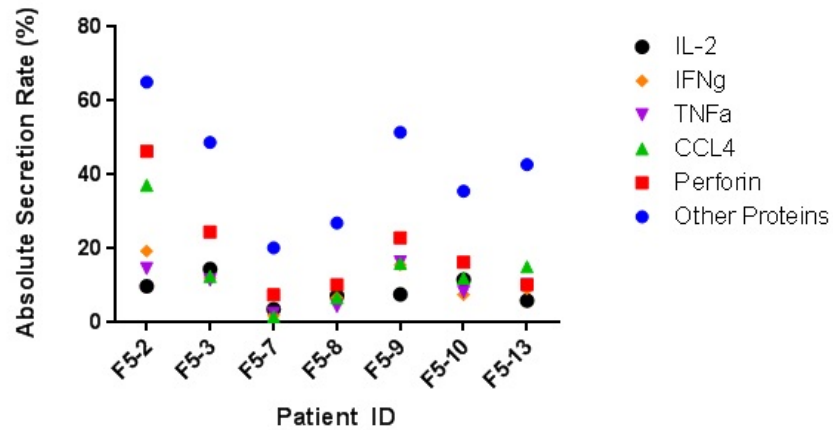
Figure 2: Scatter plot of single T cell protein secretion levels from five F5 trial patients. Each point represents a measured single cell chamber readout. Abbreviations: CCL, chemokine C-C motif ligand; IFN, interferon;TNF, tumor necrosis factor; IL, Interleukin.

a)

Protein Secretion Rate of T cells (Absolute Percentages)

b)

Secretion Rate of T cells (Relative Percentage)
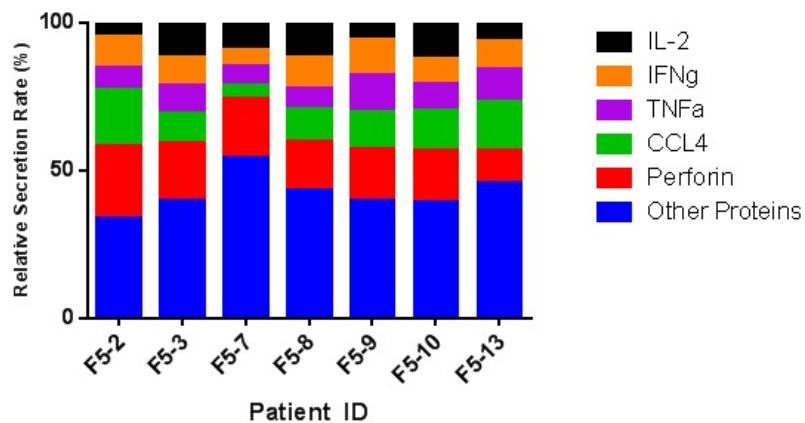
Figure 3: Secretion rates of T cells from F5 trial patients. a) The absolute percent of all individual T cells that were considered positive secretors for each protein. b) The relative rate of T cell secretion rate. This was obtained using the equation in Chapter 4.2.5.

From correlating measured parameters with the progression free survival (PFS) of each patient, no significant correlations were found. This was the case both with Pearson's and Spearsman's correlation. The full list of p values is summarized in Appendix A Table 3. Only the standard deviation of IL-2 showed good correlation (Figure 4). However, because of the large number of hypotheses tested, its significance level was not adequate to reject the null hypothesis.
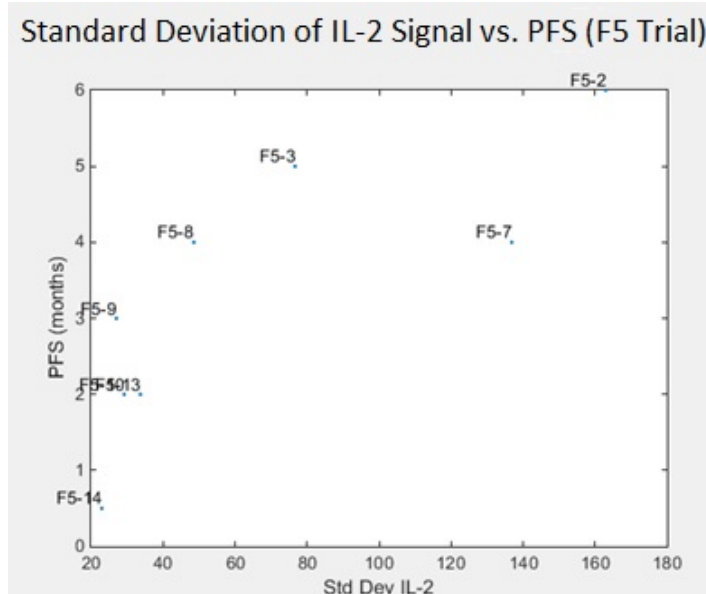
Figure 4: Correlation between standard deviation of IL-2 and progression free survival in the F5 Trial. The p-value was 0.0069, which was close to the significance threshold determined by the Bonferroni correction (p < 0.0014).

The percent of cells with a certain number of functions is shown in Figure 5. This shows that patient F5-2 has significantly higher polyfunctional cells, regardless of the number of functions. All other patients exhibited a similar trend.
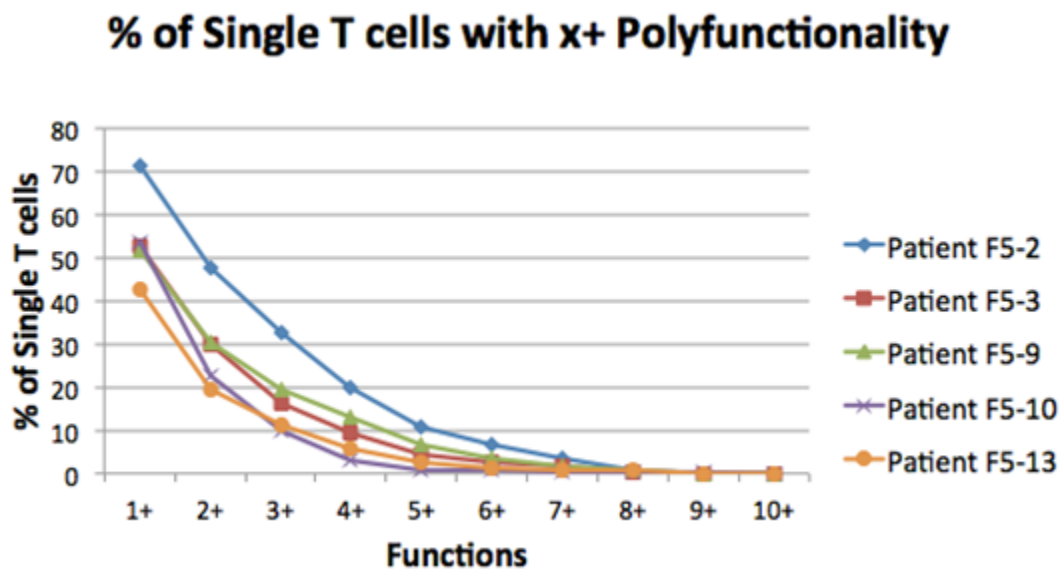
Figure 5: Percent of T cells with a certain number of functions for patients in the F5 trial. Each data point gives the percent of total single T cells that had that number of functions or more.

4.3.2 - PLSR Prediction

There were many well predicting PLSR models, with Q^2 values of >0.9. However, because of the limited number of patients, cross-validation was not possible. Therefore, no definite conclusions could be made regarding the best model to predict PFS using single cell T cells data.

4.3.3 - pSI Prediction

Pie charts showing the pSI are shown for the best, average, and worst responders in Figure 6. The diameter of the pie chart is proportional to the average pSI value, with the pie slices representing each function's contribution. Across all patient, the MIP1b (CCL4) function dominated the contribution, with the Killing function coming second. The correlations between each function and the PFS time for patients was measured in Appendix A Table 3. There were no strong correlations with the current datasets. However, the patients with the highest average pSI were the two top responders in the F5 trial. This can be seen in Figures 7a. The relationship between average pSI and PFS in the NIH trial is shown in Figure 7b. In both cases, the blue line represents the estimated linear discriminant analysis boundary for what was considered a "non-responder" and "responder," according to the leading doctors of the study. This was the red line demarcates the idea separation based on pSI to predict "non-responder" from "responders." In both cases, the accuracy was 57%, with a PPV of 67% and a NPV of 50%.
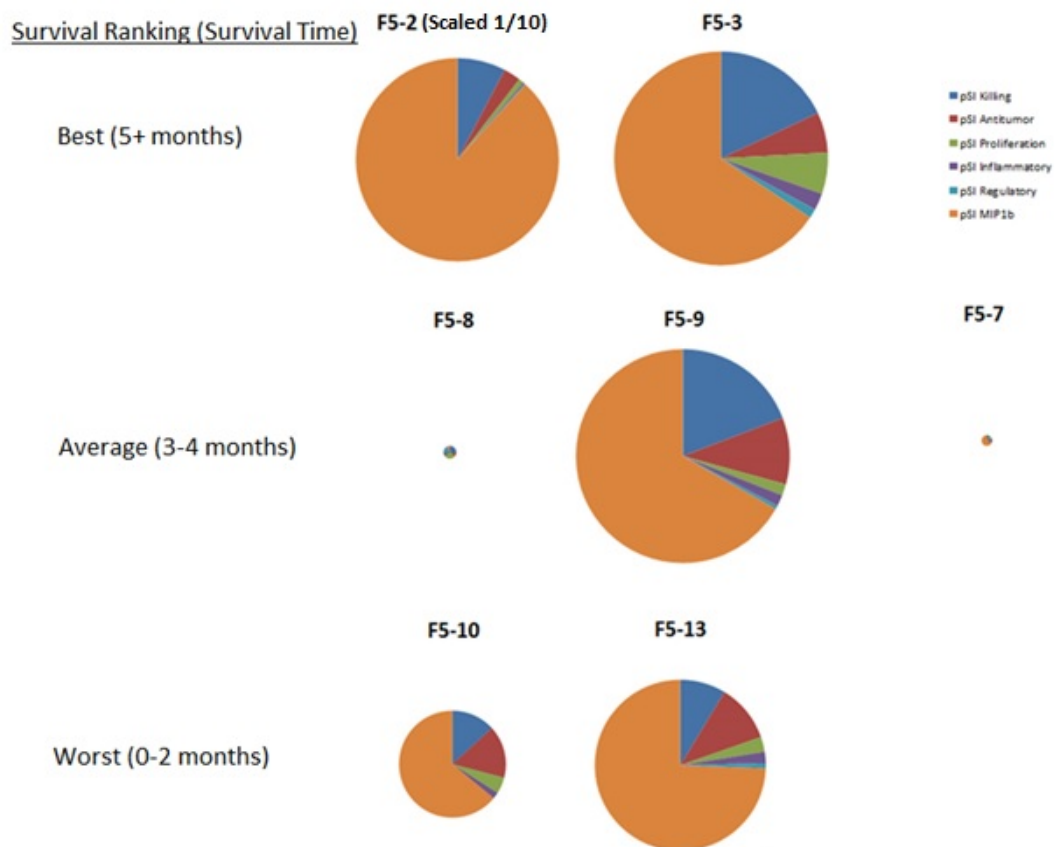
Figure 6: pSI pie charts for the best, average, and worst responders in the F5 trial. Pie charts are scaled in proportion to the average pSI value. Each slice represents a pSI function.
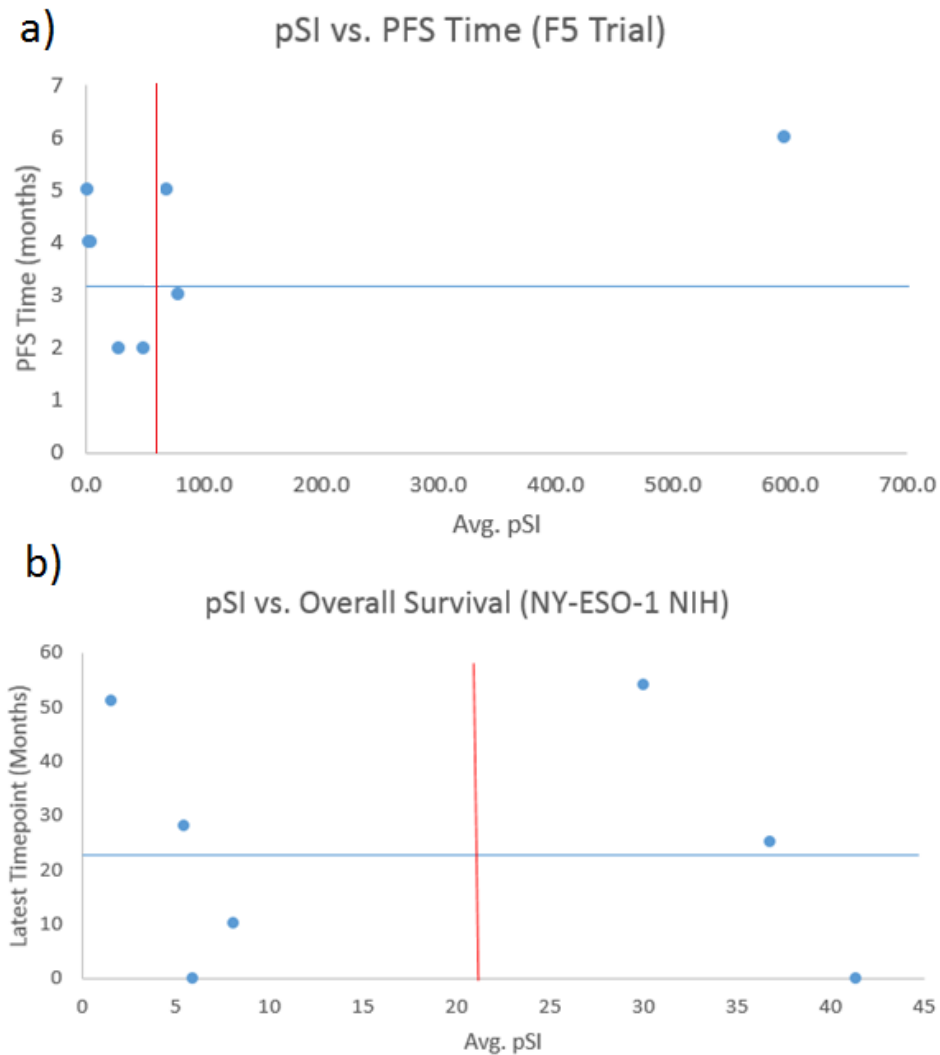
Figure 7: Scatter plot of average pSI vs. progression free survival (PFS) for the a) F5 trial and b) NIH trial. The blue line shows the general boundary for responders and non-responders. Below the line is a non-responder, and above is a responder. The red line represents the boundary line for the best separation of responders/non-responders based on average pSI.

4.4 – Discussion

Though we did not observe a correlation between pSI and patient PFS, this does not negate our ability to predict response. Firstly, we require more data before making a solid conclusion about polyfunctionality's relationship to patient success in genetic engineered ACT. Secondly, our method successfully captured the top two responding patients in the F5 trial by using average pSI as a prediction metric. This shows promise that our method could be used as a predictive model for only the patients with the highest polyfunctional T cells, without making predictions for patients that had average pSI

values below a certain threshold. Therefore, it would be interesting to further investigate this phenomenon with more statistics.

One possible reason our study did not show a strong correlation between polyfunctionality and patient response was that it could not take epitope spreading into account. Epitope spreading is when protein fragments are released from a lysed tumor and then captured by DCs. These DCs go on to prime naïve T cells in the draining lymph nodes. This generates a second generation of T cells with a diverse set of epitope targets, hence the name epitope spreading. This phenomenon has been shown to correlate with successful immunotherapy trials, in the case of Type I immune responses [17]. Therefore, because we were not able to measure the potential magnitude of epitope spreading, our predications may be missing a key component to determining immunotherapy success.

Another possible reasons we did see a strong correlation between polyfunctionality and patient response was that the T cell viability was low. We opted to test the T cells directly after thawing rather than culturing them, because this was closer to the actual T cell infusion product that the patients received. Culturing would mature/differentiate the T cells and may not reflect the actual population phenotype they would have received upon infusion [18]. This unfortunately meant that we had to sort out a significant amount of dead cells, and use the remainder for our tests.

The PLSR model showed promise in obtaining a relationship between single cell T cell data and patient response; however, there was not enough data to train and test a full model. Therefore, we believe that with additional patients, it is still possible to relate single cell T cell information to patient response.

4.5 – Conclusions

Single T cell protein measurements were conducted on the Pre-Infusion product of 15 metastatic melanoma patients undergoing genetically engineered ACT. In one of the two trials, the top

two patients were predicted as good responders based on their polyfunctional strength index. Further

data is necessary to conclude if polyfunctionality is predictive for ACT success.

Appendix A (Chapter 4)

Table 1: Proteins used and ssDNA sequences.

Protein Name	Vendor	Capture Antibody	Detection Antibody	DNA ID	ssDNA Sequence
IL-2	R&D Systems	MAB602	BAF202	B	5'-AAA AAA AAA AAA AGC CTC ATT GAA TCA TGC CTA -3'
IL-17	R&D Systems	MAB317	BAF317	C	5'- AAA AAA AAA AAA AGC ACT CGT CTA CTA TCG CTA -3'
IFNg	R&D Systems	MAB2852	BAF285	D	5'-AAA AAA AAA AAA AAT GGT CGA GAT GTC AGA GTA -3'
TNFa	Biolegend	502802	502904	E	5'-AAA AAA AAA AAA AAT GTG AAG TGG CAG TAT CTA -3'
CCL4 (MIP1b)	R&D Systems	MAB271	BAF271	F	5'-AAA AAA AAA AAA AAT CAG GTA AGG TTC ACG GTA -3'
Perforin				H	5'-AAA AAA AAA AAT TGA CCA AAC TGC GGT GCG-3'
Granzyme B	R&D Systems	EL2906	EL2906	I	5'-AAA AAA AAA ATG CCC TAT TGT TGC GTC GGA-3'
IL-4	Biolegend	500701	500804	K	5'-AAA AAA AAA ATA ATC TAA TTC TGG TCG CGG-3'
IL-10	R&D Systems	MAB2172	BAF217	L	5'-AAA AAA AAA AGT GAT TAA GTC TGC TTC GGC-3'
IL-6	R&D Systems	MAB206	BAF206	M	5'-Cy3-AAA AAA AAA AGT CGA GGA TTC TGA ACC TGT-3'

Table 2: Patient information from F5 Trial.

ID	Sex	Ethnicity	Age	Prior Treatments	Active Metastatsis Sites	IL-2 Doses	No. F5 TCR transgeneic cells (x10^9)	% MART-1 Tetramer Pos Cells	IFN-g with K562-A2-MART-1 (pg/mL/10^6)
F5-2	F	White	46	MKC prime-boost vaccine, IL-2	Skin, LN, Bone	6	1	74.1	15.9 x 10^5
F5-3	M	White	61	-	Lung, Liver	13	1	74.6	3.4 x 10^5
F5-7	M	White	48	IL-2	SC, Bone	9	1	63.7	4.2 x 10^5
F5-8	M	White	44	-	LN, Liver	11	1	66.7	7.8 x 10^5
F5-9	F	White	46	-	Skin, LN	11	1	69.7	11 x 10^5
F5-10	F	White	47	-	Liver, Adrenal, SC, LN, Orbit	14	5	55.9	4.7 x 10^5
F5-13	M	White	60	Left axillary lymphadnectormy, Adjuvant interferon, radiation, surgical resection, IL-2, ipilimumab	Lung, Abdomen, SC	4	4.4	49.7	8.7 x 10^5

Table 3: P values of correlation data from measured parameters and progression free survival. Both Spearsman's and Pearson's correlations are given.

	IL-6 % secreting Protein of total Cells	IL-10 % secreting Protein of total Cells	IL-4 % secreting Protein of total Cells	GB % secreting Protein of total Cells	Perforin % secreting Protein of total Cells	CCL4 % secreting Protein of total Cells	TNFa % secreting Protein of total Cells	IFNg % secreting Protein of total Cells	IL-17 % secreting Protein of total Cells

Spearsman	0.546	0.8512	0.3579	0.9444	0.8004	0.5841	0.9444	0.8004	0.4647
Pearsons	0.542	0.8548	0.3779	0.7962	0.3768	0.2282	0.9668	0.6122	0.4211

	IL-2 % secreting Protein of total Cells	Relative IL-6 % secreting Protein of total Cells	Relative IL-10 % secreting Protein of total Cells	Relative IL-4 % secreting Protein of total Cells	Relative GB % secreting Protein of total Cells	Relative Perforin % secreting Protein of total Cells	Relative CCL4 % secreting Protein of total Cells	Relative TNFa % secreting Protein of total Cells	Relative IFNg % secreting Protein of total Cells
Spearsman	0.9444	0.1556	1	0.8988	0.546	0.8004	0.3706	0.2948	0.8512
Pearsons	0.911	0.4944	0.9805	0.7418	0.9666	0.9397	0.1713	0.1521	0.6252

	Relative IL-17 % secreting Protein of total Cells	Relative IL-2 % secreting Protein of total Cells	Avg Signal IL-6	Avg Signal IL-10	Avg Signal IL-4	Avg Signal GB	Avg Signal Perforin	Avg Signal CCL4	Avg Signal TNFa
Spearsman	0.8512	0.2948	0.129	0.3401	0.406	0.148	0.1937	0.7349	0.9885
Pearsons	0.7089	0.4671	0.0554	0.2719	0.6312	0.1147	0.139	0.3095	0.7457

	Avg Signal IFNg	Avg Signal IL-17	Avg Signal IL-2	Std Dev IL-6	Std Dev IL-10	Std Dev IL-4	Std Dev GB	Std Dev Perforin	Std Dev CCL4
Spearsman	0.9885	0.406	0.3873	0.4647	0.1024	0.8988	0.148	0.0361	0.9444
Pearsons	0.8971	0.4743	0.334	0.4372	0.0298	0.9494	0.1565	0.0119	0.5075

	Std Dev TNFa	Std Dev IFNg	Std Dev IL-17	Std Dev IL-2	pSI Killing	pSI Antitumor	pSI Proliferation	pSI Inflammatory	pSI Regulatory
Spearsman	0.3873	0.3706	0.2405	0.0183	0.8004	0.7956	0.6417	0.6417	0.8952
Pearsons	0.1882	0.2629	0.2311	0.0125	0.2245	0.5374	0.1845	0.5569	0.7619

	pSI MIP1b	% Viable Cells of total	% CD4+ MART-1+	% CD8 MART-1+	CD4:CD8 Ratio	%MART-1 Cells from total CD3	IL-2 Doses	Peak % of MART-1	Num. TCR transduced (*10^9)
Spearsman	0.7147	0.2663	0.3254	0.9135	0.6262	0.4988	0.6667	0.9885	0.1548
Pearsons	0.1512	0.434	0.1379	0.8556	0.7081	0.2761	0.6645	0.9577	0.1399

References (Chapter 4)

[1] D. WHITE, "TREATMENT OF 283 CONSECUTIVE PATIENTS WITH METASTATIC MELANOMA OR RENAL-CELL CANCER USING HIGH-DOSE BOLUS INTERLEUKIN-2," *JAMA,* vol. 271, no. 12, pp. 907-913, 1994.

[2] S. Rosenberg et al., "Cancer Regression and Autoimmunity in Patients After Clonal Repopulation with Antitumor Lympocytes," *Science,* vol. 298, pp. 850-854, 2002.

[3] C. June et al., "Chimeric Antigen Receptor-Modified Cells in Chronic Lympoid Leukemia," *NEJM,* vol. 365, no. 8, pp. 725-733, 2011.

[4] W. Maccarty, "Factors which influence longevity in cancer: a study of 293 cases," *Ann Surg,* vol. 76, no. 1, pp. 9-12, 1922.

[5] S. Topalian, et al., "Immunotherapy of patients with advanced cancer using tumor-infiltrating lympocytes and recombinant interluekin-2: a pilot study," *J Clin Oncol,* vol. 6, no. 5, pp. 839-853, 1988.

[6] L. Muul, "Identification of specific cytolytic immune responses against autologous tumor in humans bearing malignant melanoma," *J Immunol,* vol. 138, no. 3, pp. 989-995, 1987.

[7] S. Rosenberg, et al., "Identification of specific cytolytic immune responses against autologous tumor in humans bearing malignant melanoma," *J Immunol.,* vol. 138, no. 3, pp. 989-995, 1987.

[8] S. Rosenberg, et al., "Adoptive Cell Transfer for Patients With Metastatic Melanoma: The Potential and Promise of Cancer Immunotherapy," *Cancer Control,* vol. 20, no. 4, pp. 289-297, 2013.

[9]   S. Rosenberg et al., "Durable complete responses in heavily pretreated patients with metastatic melanoma using T-cell transfer immunotherapy," *Clin Cancer Res,* vol. 17, no. 13, pp. 4550-4557, 2011.

[10] S. Goff, et al., "Tumor Infiltrating lymphocyte therapy for metastatic melanom: analysis of tumors resected for TIL," *J Immunother.,* vol. 33, no. 8, pp. 840-847, 2010.

[11] A. Ribas et al., "Adoptive transfer of MART-1 T cell receptor transgenic lymphocytes and dendritic cell vaccination in patients with metastatic melanoma," *Clinical Cancer Research,* vol. 20, pp. 2457-2465, 2014.

[12] L. Johnson et al., "Gene therapy with human and mouse T-cell receptors mediates cancer regression and targets normal tissues expressing cognate antigen," *Blood,* vol. 113, no. 3, pp. 535-546, 2009.

[13] C. Ma et al., "Multifunctional T-cell Analyses to Study Response and Progression in Adoptive Cell Transfer Immunotherapy," *CANCER DISCOVERY,* vol. 3, no. 4, pp. 418-429, 2013.

[14] D. Macallan et al., "Comparing HIV-1 and HIV-2 infection: Lessons for viral immunopathogenesis," *Rev Med Virol,* vol. 23, no. 4, pp. 221-240, 2013.

[15] M. Duvall et al., "Polyfunctional T cell responses are a hallmark of HIV-2 infection," *European Journal of Immunology,* vol. 38, no. 2, pp. 350-363, 2008.

[16] S. Rosenberg et al., "Tumor regression in patients with metastatic synovial cell sarcoma and melanoma using genetically engineered lymphocytes reactive with NY-ESO-1," *J Clin Oncol.,* vol. 29, no. 7, pp. 917-924, 2011.

[17] M. Disis, "Immunologic biomarkers as correlates of clinical response to cancer immunotherapy,"

Chapter 5: T cell-T cell Contact and its Effect on Protein Secretion in Microfluidic Cell Assays

5.1 – Introduction

Leukocyte communication is a critical aspect of the immune response that drives adaptive immunity [1]. In one arm of the adaptive immune response, naive T (T_N) cells communicate with dendritic cells in the lymph nodes, where they are primed to become T memory (T_M) or effector (T_E) cells [1]. The determination of a T_N's fate to become either a memory or effector T cell is still not fully understood; however, much effort has been put towards elucidating this mechanism [2, 3]. This is critical to understand, because it can determine whether a patient will become immunized towards a certain foreign species or not. This is critical in vaccines and the developing field of immunotherapy.

One such immunotherapy, Adoptive T cell immunotherapy (ACT), has shown great promise in treating late stage cancer patients [4]. As mentioned in Chapter 4, ACT rates of success vary drastically from patient to patient. One metric for predicting its long term success is the persistence of T cells beyond one months [5]. This indicates a memory pool of T cells is necessary for ACT success. Therefore, applications like immunotherapy would be much more viable, if T cell priming could be manipulated to increasing the number of memory cells produced.

A recent theory in T cell memory formation is that T cell-T cell interactions promote long-term memory. These interactions are similar to that of T cells communicating with dendritic cells (DCs). M. Krummel et al. demonstrated that after priming from DCs, T cells will cluster together and form synapses [6]. These synapses focus protein secretion towards the other T cell, enhancing the capture efficiency of paracrine signaling [7]. Disrupting this T cell-T cell, or homotypic, interaction slowed down immune activation rate. Furthermore, inhibiting T-T contact reduces the number of T_M cells produced

after an infection [8, 9]. This has been demonstrated by inhibiting T cell adhesion molecules (e.g., LFA-1 and ICAM-1) prior to infection and measuring the numbers of persisting T_M cells post-infection. Additionally, J. Zhou has demonstrated that the strength of T cell response is increased when T cells are incubated together [Unpublished] Using mice with OT-1+ tumors, she performed an ACT with T cells that were either directly injected into mice or incubated together before injecting. The mice that received T cells incubated together showed significantly stronger tumor responses. Therefore, there is much promise in improving immunotherapies by better understanding the mechanism behind T-T communication.

The topic of this chapter is understanding the functional effect of T-T communication. We wish to better understand how T cell communication influences the strength of the T cell response, specifically, its functional protein secretion. As in the previous chapter, this investigation was accomplished using single cell proteomic technology. Additional software was developed in order to automate the analysis of parameters related to two cell data measured in this project (e.g., distance between cells)

5.2 – Methods

5.2.1 - Microfluidic Device

The microfluidic device used for this study was the same design as that used in Chapter 4. The process of making this device is detailed in Chapter 4.2.1. One difference in the process was that the PDMS device was cured in an aluminum block to attain a flat-top surface, as detailed in Chapter 1.2.1.1. This allowed for the automated imaging of the cells through the PDMS by an Olympus IX81 microscope.

5.2.2 - Automated programs

The general function of the cell counting programs used to count the cells and analyze the secreted protein levels is detailed in Chapter 1.2.1 and Chapter 1.2.2, respectively. In addition, two

functions were added to the automated cell counting program: T cell-T cell distance calculation and fluorescence detection.

For the distance calculation, the pixel-value distance between two cells was measured by the program and converted to microns, using reference dimensions on the microfluidic device. After all two cell chamber distances had been calculated, the user manually inspects each two cell chamber. They then input whether the T cells are contacting or non-contacting, or reject it as a faulty chamber.

The fluorescence detection was accomplished by overlaying the bright-field microscope image of the microfluidic device with the fluorescence image of the device and using the cellular regions of interest from the bright-field image to crop the fluorescence image. The average fluorescence intensity was used to determine if a cell had "positive" fluorescence or not. In the fluorescence detection, the GFP channel was used to visualize the CD8 (FITC) on T cells. If the program detected significant fluorescence from a T cell for that channel, it would be regarded as a CD8 T cell. Otherwise, it was regarded as a CD4 T cell, since the population of T cells was pre-sorted for CD4 or CD8 single positives.

5.2.3 - FACS Cell Sorting

The T cells used were from the F5 or NIH immunotherapy trial, as described in Chapter 4.2.

T cells were sorted using a BD Aria II or Aria III machine. Cells were gated on FSC-A, FSC-H, SSC-A (singlet, lymphocyte identification), 7-AAD (viability), CD4 (Biolegend, OKT-4), CD8 (Biolegend, HIT8a), and MART-1 tetramer (MBL Int.). The machine was calibrated with CompBeads (BD Biosci.) prior to sorting. The T cells were sorted into two groups, CD4+CD8-MART-1+ and CD4-CD8+MART-1+, which are henceforth referred to as CD4 and CD8, respectively.

T cells were sorted with specific colors, in order to allow for fluorescence imaging on an Olympus IX81 microscope. CD4 (fluorohpore: APC/Cy7) and 7-AAD (PerCP) were used because these

fluorophores were undetected in the IX81 microscope. Because CD8 was to be detected, FITC was selected as a fluorophore as it was detectable on the IX81 microscope.

5.2.4 - Two Cell Secretome Experiment

T cell experiments were run in a similar fashion to the experiments described in 5.2.3 with a minor change to the protocol: the devices were imaged on an Olympus IX81 microscope to obtain a full-view, stitched image of the cells in the microfluidic device. The resulting data was analyzed using the automated programs described in Chapter 1. Cells were counted in every microfluidic chamber and the two cells chambers were further analyzed for T cell-T cell distance, as well as if the T cells were contacting.

The percent of T cells secreting a specific protein was gated by "contacting" and "non-contacting," and then plotted by protein. This was accomplished by first determining the threshold for what is considered to be a "positive-secreting" T cells. The cutoff used was the 95 percentile of the protein level in zero cell chambers ($Mean_{zero\ cell}$ + 2 * $StandardDeviation_{zero\ cell}$). The percent of T cells considered positive secretors was then calculated for both contacting two cell chambers and non-contacting two cell chambers.

In addition, the theoretical percent of two T cells that should have been secreting was calculated using the single cell data. This equation was used for calculating the probability for at least one of two independent events (X and Y) to occur [10].

Pr(X|Y) = Pr(X) + Pr(Y) − Pr(X & Y)

where X is the event that one of the two single cells in a chamber secretes a certain protein and Y is the probability that the other cell secretes the protein. In our case, because X and Y have the same probability value, we can simplify using: P(X) = P(Y) and P(X & Y) = $P(X)^2$

$$Pr(X|Y) = 2Pr(X) - Pr(X)^2 \qquad\qquad \textbf{Equation (1)}$$

5.2.5 - Nature of T-T contact

To determine if TCR clustering occurred in the T-T contact synapse, MART-1+ Jurkat T cells were obtained from the Baltimore Lab, courtesy of Michael Bethune. These T cells were stained with PE-MART-1 Tetramer (MBL Int.) and imaged on an Olympus IX81 microscope both under bright-field and TxRed fluorescence. The images were analyzed to get the percent of contacting T cells that have tetramer bridging their synapse.

To investigate the effect of blocking synapse formation, different molecules were added to observe how they altered the degree of contact. To 10^6 cells/mL in 100 µL, each of following molecules were added separately: anti-CD3 (eBioscience, 5uL of 0.5mg/mL), MART-1 Tetramer (MBL Int, 5 µL), and 1x PBS (5 µL). The cells were allowed to incubate overnight for 16 hours. They were then imaged and the percent of cells in contact was counted by hand.

5.2.6 - Two Cell Phenotype Experiment

This experiment was run in a similar fashion to Chapter 4.2.4, with the addition of fluorescent images being acquired during testing to detect CD8 and CD4 phenotypes. Prior to loading the T cells into the microfluidic device, the T cells were stained with CD8-FITC for 30 min to improve coverage of the fluorophore.

Data analysis involved acquiring cell count, two cell distances, and determining if each cell was CD8 positive. CD8 detection was ascertained using a method similar to that of the normal cell detection: the fluorescence image was split into RGB layers and the user selects which layer should be used. For example, the GFP fluorophore would be the green layer. Next, the chambers are cropped out according to the chamber boundaries defined in the normal cell counting program (Chapter 1.2.1.3). Each chamber is then binarized using a user set threshold value and then features are found using the cell detection

method used in Chapter 1.2.1. In the final output, each chamber has an additional value added: the number of fluorescent cells.

The theoretical values for the contacting CD4-CD4, CD4-CD8, and CD8-CD8 were calculated from the percent of CD4 and CD8s sorted in the FACS. Additionally, the theoretical percent of two cell chambers secreting a certain protein was calculated similar to Equation (1). However, this was calculated separately for chambers with CD4-CD4, CD4-CD8, and CD8-CD8 phenotype cells. These theoretical percentages of T cells secreting were then used to see if the two cells in contact, with that same phenotype, were secreting more than the theoretical value. A chamber was considered to be secreting significantly higher when its levels were two standard deviations higher than the mean of the theoretical value calculated from the single cell chambers.

5.3 – Results

5.3.1 - Two Cell Secretome Experiment

The distribution of cell-cell distances is plotted for four different patients in Figure 1. Note, the distance value (x-axis) are binned in intervals that range 400 µm, with an additional bin at 0 with a single value of 0 µm (contact). All patients showed a non-statistical amount of cells at zero distance (contacting). The percent contacting T cells varied between patients, ranging from 5% to 27%, with a mean of 17.3% ± 7.3. This was not affected by stimulation type, as the mean contact percentage for a stronger stimulation of tetrameter+CD28 was 15.6% ± 7.1.
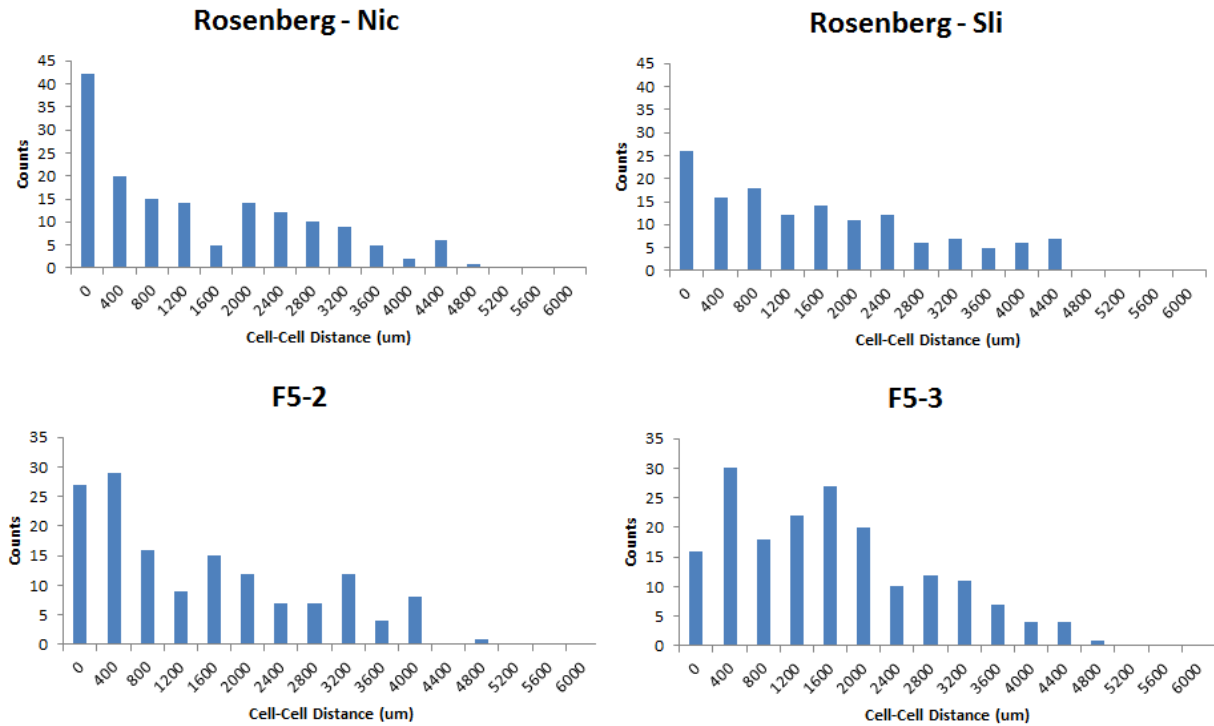
Figure 1: Distance distributions of two cell chambers from four patients in the F5 and NIH trial. The cell-cell distance was measured between from the bright field image of the microfluidic device. These datapoints are only from two cell chambers.

The average protein level of CCL4, a protein related to T cell chemotaxis, was plotted against T cell-T cell distance for four patients in Figure 2. These demonstrated an enhancement of protein signal when cells are contacting. It was also investigated whether close proximity (< 100um) but not contact enhanced protein levels. Only 4.7% of these "close proximity" cases secreted either CCL4, TNFa, GB, or IFNg to a significant level, compared to 52.4% of cells in contact. Because there was no trend in protein signal vs. distance, subsequent analyses binned the data based on if two cells were at 0 distance (contact) or >0 distance (non-contact).

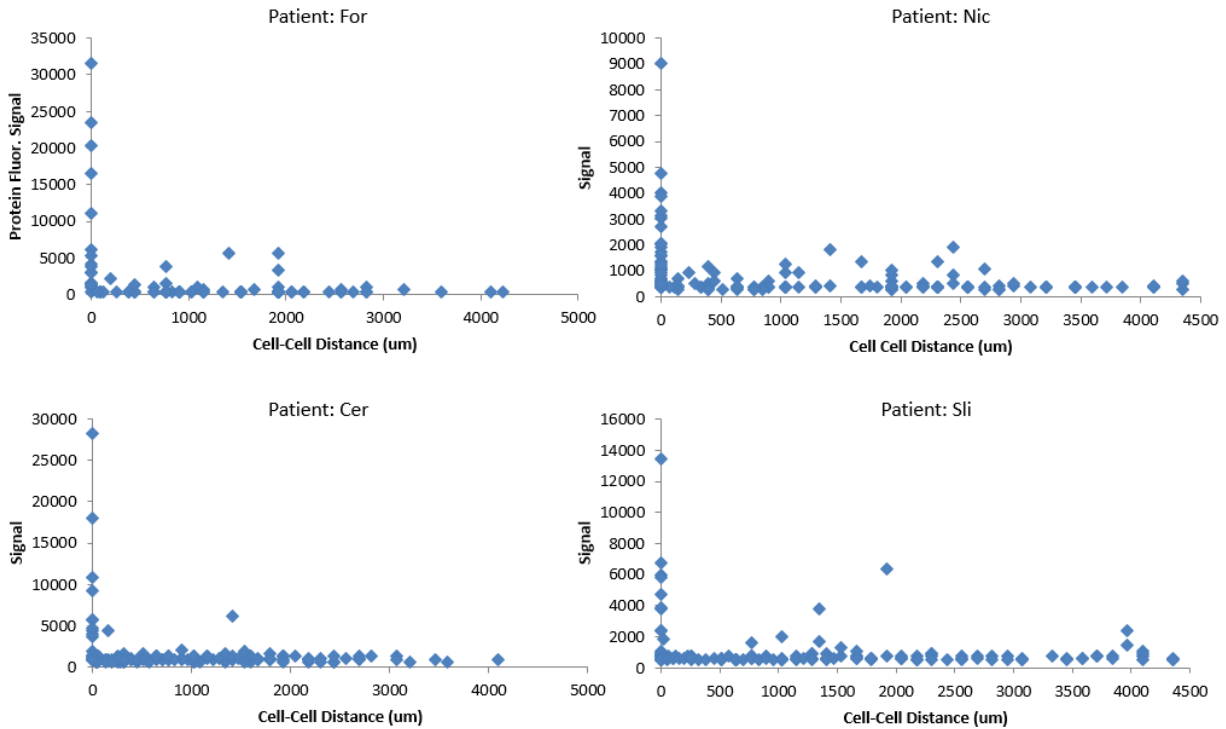# Distance Dependence on MIP1b signal in Two Cell Chambers

Figure 2: CCL4 (MIP1b) signal vs. T cell-T cell distance for four patients from the NIH trial. Each data point represents a measurement from a two cell chamber.

The average protein levels of contacting and non-contacting two cell chambers were plotted in Figure 3. The data was pooled from T cells from four patients in the NIH trial. T cells with protein levels less than the background level are not shown in the plot but were included in significance calculations. The five proteins displayed in the figure were the most statistically discriminating proteins, and the remaining proteins tested (IL-6, IL-4, IL-17, IL-2) were not significantly changed upon contact. A Mann-Whitney U Test showed that TNFa, CCL4 (MIP1b), and GB were statistically distinguishable between the contact and non-contact group.
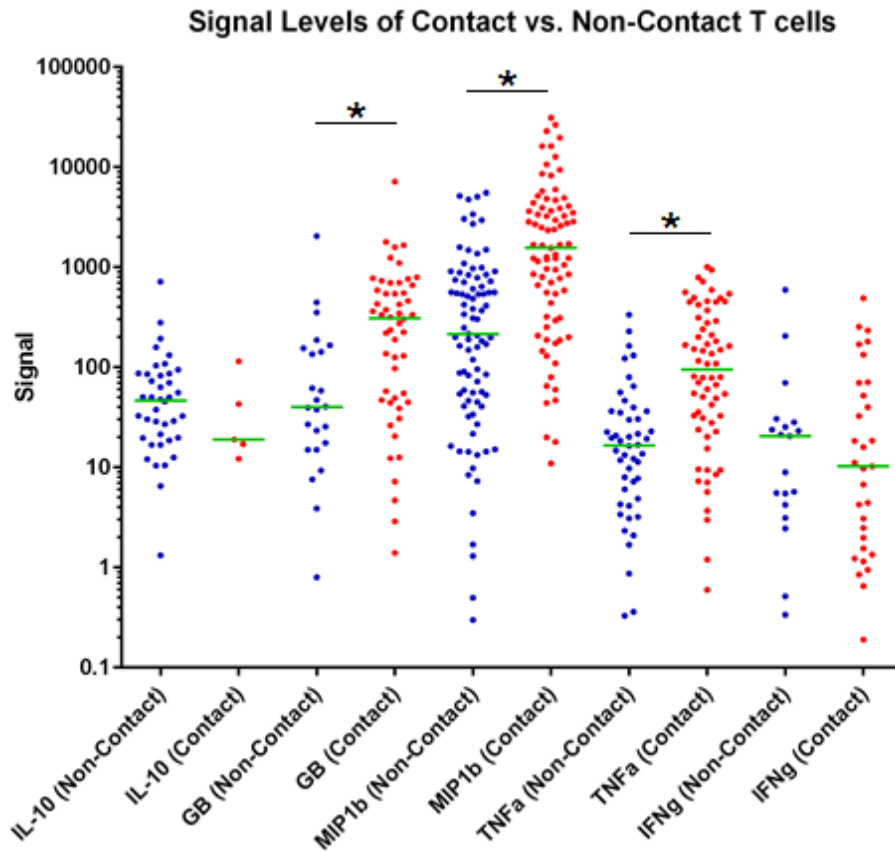
*Figure 3: Scatter plot of protein levels from two cell chambers. This data was gated on if the T cells were in contact or not. Each point represents a two cell chamber measurement of a specific protein. Points that were below what was considered a positive secretion for that protein were excluded from this plot. Significance indicators are shown above proteins that were statistically distinguishable. Significance marker: * = p<0.05.*

The percent of T cells considered positive secretors was plotted for each protein, and gated on contact, and non-contact (Figure 4a). Additionally, the theoretical percent of two T cells secreting a protein based on single cell data was plotted, as calculated from Equation 1. All patient samples showed that contacting T cells enhanced their ability to secrete certain proteins. This occurred for a subset of proteins, including Granzyme B (GB), Perforin, CCL4 (MIP1b), TNFa, and IFNg. The protein expression frequency was averaged over all patients, as shown in Figure 4b. This shows a general trend of upregulation for the subset of proteins mentioned above. Finally, the theoretical value protein secretion frequency matched the measured non-contact value closer than the contact value in 90% of cases (n=8).
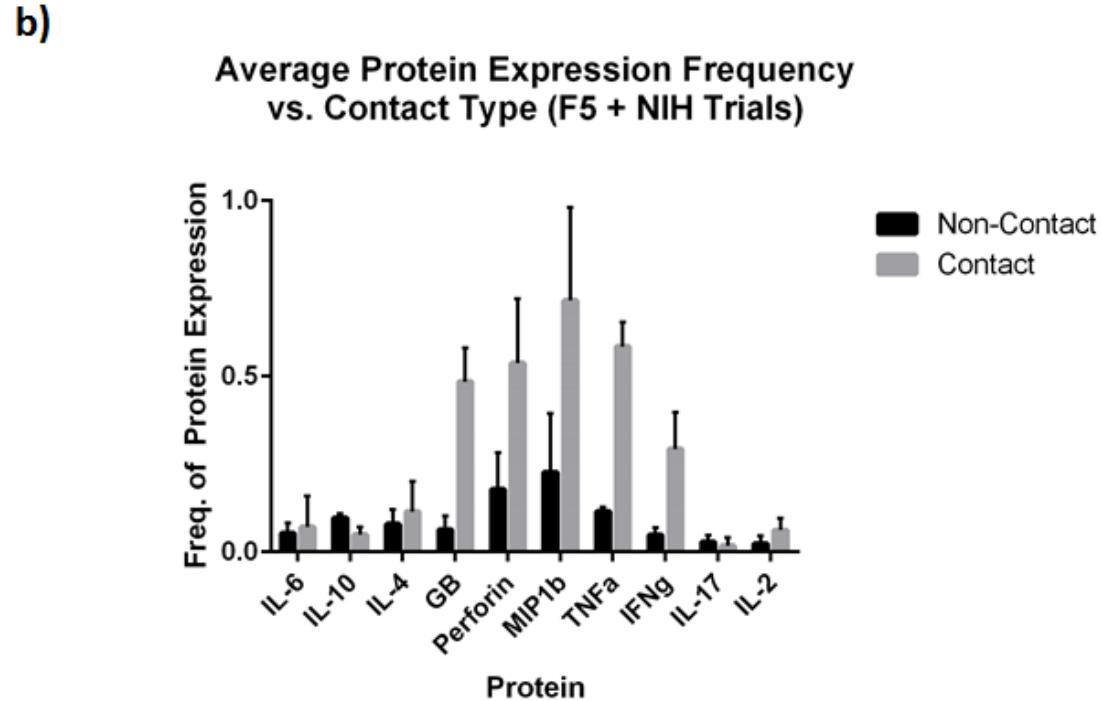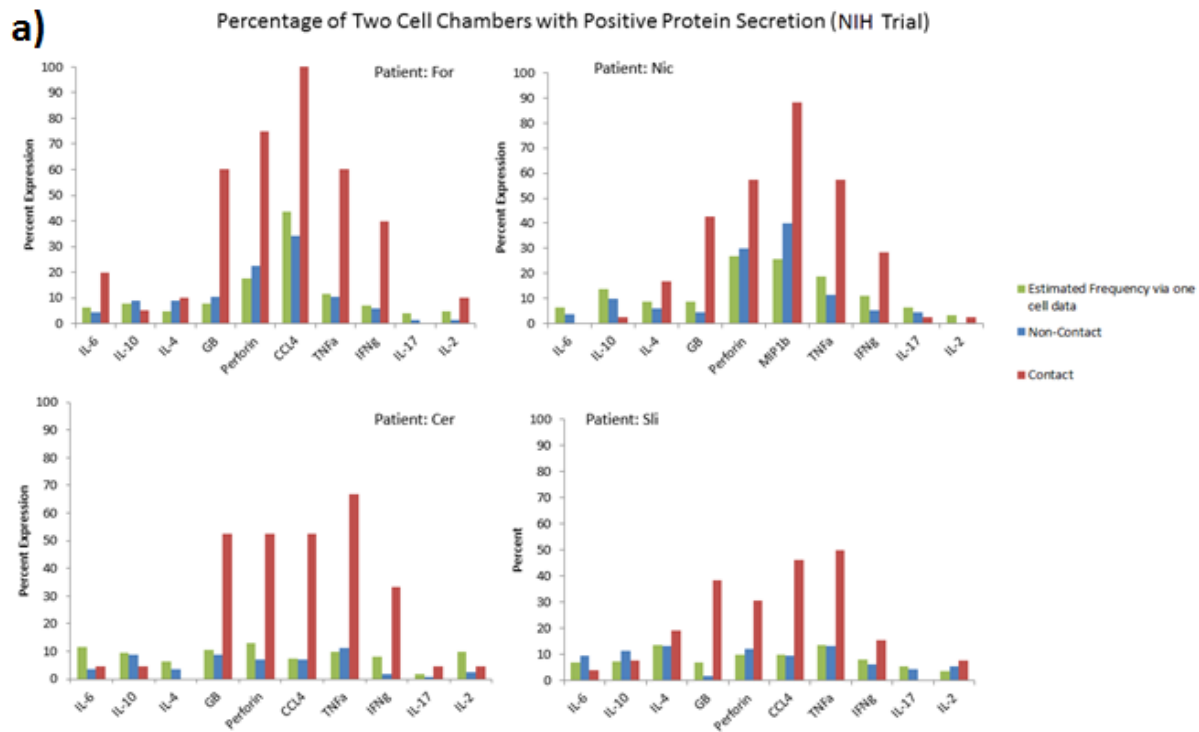
Figure 4: Percent of two cell chambers with positive protein secretion, gated based on if the T cells were in contact or not. a) The percent of T cells secreting a specific protein is shown for four different patients, and is gated on contact, non-contact, or estimated values from single cell data. b) The averaged frequency of protein expressed, gated on contact, and non-contact.

5.3.2 - Nature of T-T contact

The percent of contacting T cells that had TCR localized at the intercell synapse was measured. Only a small percentage of contacting T cells (< 10%) show co-localization of their MART-1 receptor, as shown in Figure 5. The majority of cells had TCR clusters spread non-specifically over the plasma membrane.

A summary of the percent contacting cell vs. stimulation are shown in Table 1. There was a significant reduction in T cell contact when anti-CD3 was added to the T cells (p=0.0139). However, relative to the control T cells, there was no enhancement of T cell-T cell contact when the tetramer was present.

Table 1: Results from contact experiment. T cells were incubated with different molecules and the percent of contacting T cells was measured.

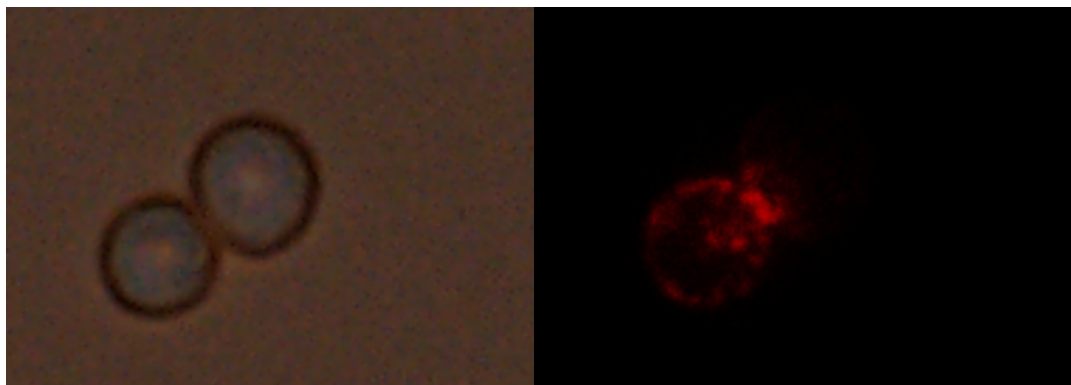| Stimulation	% Contacting T cells	Total pairs of T cells	Total Two Cell Chambers
Control	5.8	22	379
Anti-CD3	2.0	6	299
MART-1 Tetramer	4.41	3	68

Figure 5: Image of Jurkat T cells stained with fluorescent tetramer that bridge the intercellular synapse. a) Bright field image of two contacting Jurkats. b) Fluorescence image of the same contacting Jurkats. Note the TCR is localized at the junction between the two cells.

5.3.3 - Two Cell Phenotype Experiment

T cells behaved in a similar fashion to previous experiments, with the contacting two cell having enhanced protein signals for IFNg, CCL4, and GB (Figure 6).
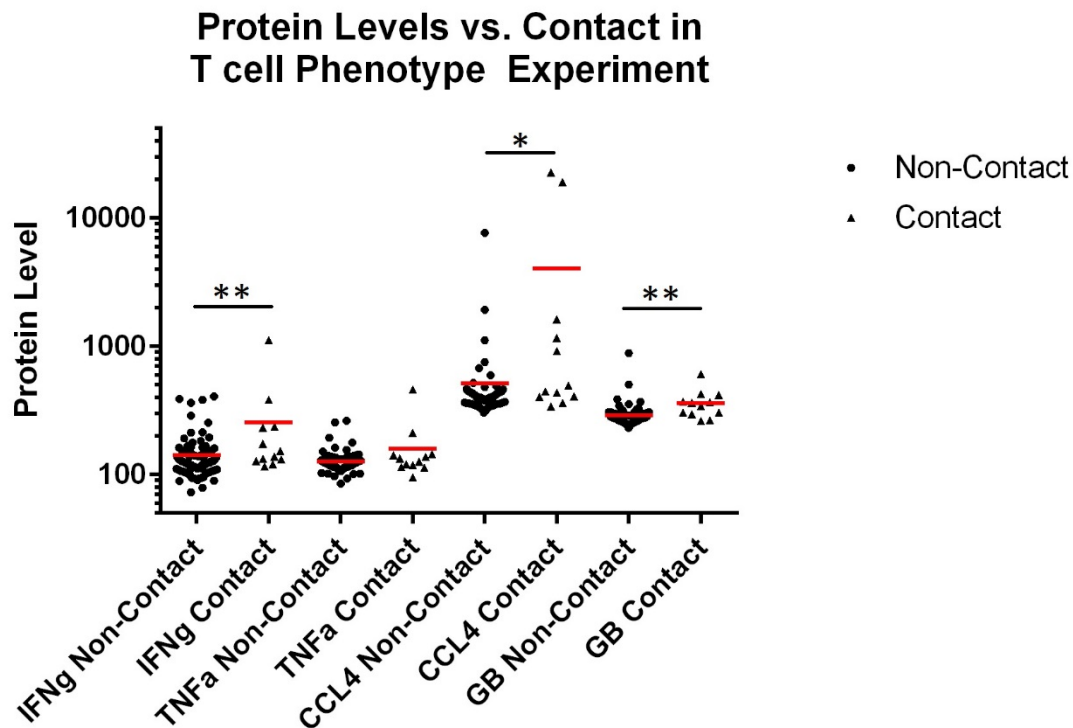

*Figure 6: Protein expression levels for phenotype experiment, gated on contact or non-contact. Each point represents a two cell chamber measurement of a specific protein level. Significance indictors are displayed above the geoms. Significance markers: * = p <0.05, ** = p <0.001.*

Figure 7: Fluorescence and bright field image of contacting T cells with CD8-CD8, CD4-CD4, and CD8-CD4 contacting. On the left is the fluorescence image with the type of contact pair. One the right side is the same region, taken as a bright field image.
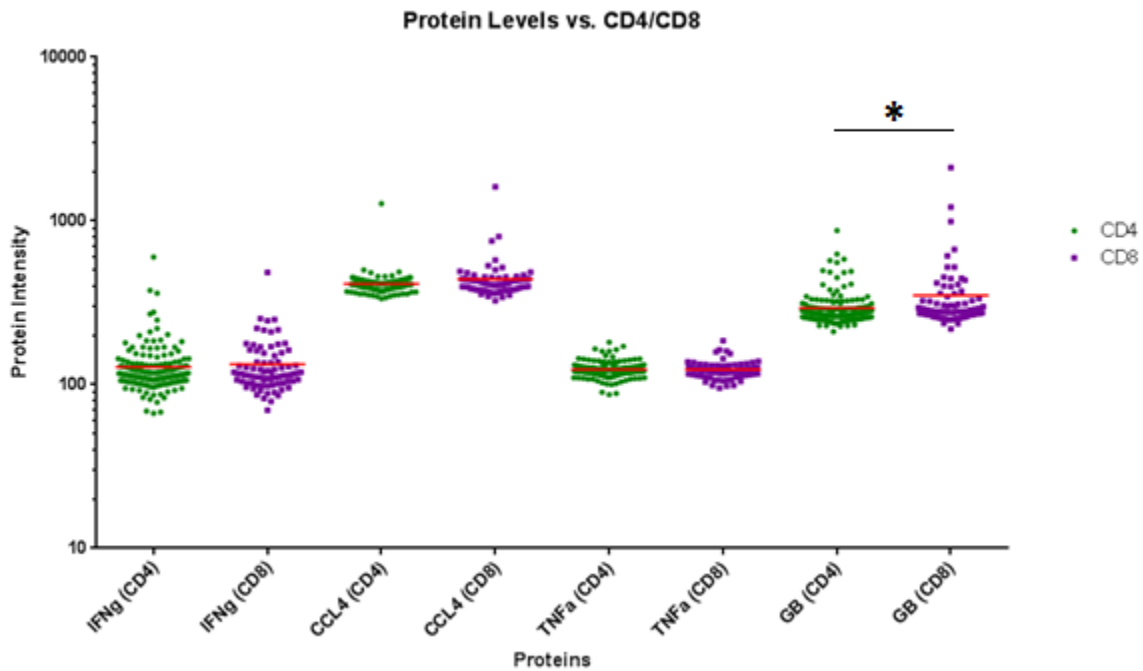
*Figure 8: Plot of protein levels for in single cell chambers from the phenotype experiment, gated on T cell phenotype. Each point represents a single T cell measurement for a specific protein. The red line shows the average secretion level of each protein. Significance plots are shown above the geoms. Significance marker: *= p<0.05.*

T cells were successfully loaded on the chip and phenotyped, as visualized in Figure 7. The figure

shows the three different combinations of T cell sub-phenotypes: CD8-CD8, CD4-CD4, and CD8-CD4. The

contact percentages are listed in Table 2, with their theoretical values shown in parentheses. The CD4-CD8 contact percentage was much lower than the theoretical value, with the CD8-CD8 contact being higher than anticipated.

Table 2: Measured and theoretical percentages of two T cells in contact, gated on phenotype. The measured phenotype is shown in each box, with the theoretical value in parentheses next to it. The theoretical value was calculated from the percent of each phenotype sorted by FACS.

	CD4	CD8
CD4	16 % (10%)	16% (42%)
CD8		68% (48%)

A plot showing the difference in protein secretion level between individual CD4 and CD8 cells is shown in Figure 8. The only protein that discriminated between CD4 and CD8s was Granzyme B.

For each of the combination of T cell subtypes, the significantly enhanced protein levels are shown in Table 3. This demonstrated that CD4s in contact with CD4s will secrete IFNg at a rate that was significantly higher than that of two individual CD4s secreting in a chamber. CD4-CD8 had two proteins that were secreted at a significantly higher rate: IFNg and CCL4. Finally, CD8-CD8 chambers secreted GB and CCL4 at a higher rate.

Table 3: Enhanced proteins in two cell, contacting chamber, gated by phenotype of T cells.

	CD4	CD8
CD4	IFNg	IFNg, CCL4
CD8		GB, CCL4

5.4 – Discussion

Two cells communication has received attention by many groups, including Krummel and Xiang [7, 11]. They have demonstrated that T cells can form synapses with each other and exchange cytokines in a targeted fashion [7]. Furthermore, Xiang demonstrated that preventing T cell communication in the priming stage of an immune response would significantly reduce the number of antigen specific memory T cells found post-infection [11]. Here, we show that T cells in contact tend to be higher producers of select proteins. The fact that we are detecting higher levels of proteins implies that either the cells are producing more absolute value of proteins, or that contacting reduces the number of proteins uptaken into the cells (thereby increasing the amount our devices can detect). Because Krummel et al. have shown that synapse formation also centers receptors to the proteins in the synapse, this implies that T cells in contact tend to secrete towards each other [7]. Therefore, in our case, it is unlikely that we are seeing high protein levels because of reduced uptake/degradation, and so the former hypothesis of protein upregulation is more likely. This could be confirmed by performing single cell RNA-seq to determine if copy numbers increase.

In elucidating the nature of the T cell-T cell contact, we demonstrated that the addition of anti-CD3 significantly reduced the number of contacting T cells. We do not believe, that CD3 is a driving force of contact however. Our hypothesis was that anti-CD3 creates a steric hindrance that impedes synapse formation. What we believe to be the driving force of the contact is LFA-1 and ICAM-1, which have been shown to be critical in homotypic T cell-T cell contact [8, 12]. We have tested this at a bulk level, and showed that T cells do not contact when anti-LFA-1 is added at 10 ug/mL [13]. It would be interesting to detect which proteins localize at the T cell-T cell synapse, as this would give us an idea as to which receptors are responsible for the protein upregulation observed.

Contacting two cells did not upregulate all proteins equally. Many proteins were not upregulated at all. The proteins that were enhanced were related to either cell killing (GB, Perforin), T cell mobility (CCL4), or general immune activation (IFNg, TNFa). In contrast, IL-10 – an anti-inflammatory protein – was not enhanced. This may be the mechanism behind why co-culturing T cells together prior to ACT improves tumor killing ability [Unpublished] There are two possible explanations for the selective upregulation: 1) the exchange of information through cytokines and/or receptors only triggered certain protein synthesis pathways in the cells or 2) the single cell test was not sensitive enough to detect changes in protein expression for certain proteins.

For the first possibility, there is ample evidence that cytokine exchange can selectively enhance protein secretion. For example, pretreatment of T cells with IL-10 can enhance TNFa, and IFNg levels in later secretion, without enhancing IL-2 or IL-4 levels [14]. In addition, it is well known that different stimulants cause selective protein expression upregulation. For instance, PMA/I turns on TNFa expression in leukocyotes [15]. Therefore, it is reasonable to assume that proteins exchanged by the T cell, or receptors activated, enhanced the expression of select proteins. To test this, one could perform selective knock-outs of pathways downstream of relevant cytokine/receptors exchanged by the T cells. Performing blocking experiments, rather than knock-outs, would not be useful in this case, for two reasons: one, antibodies for blocking cytokines would not likely enter into the T cell synapse, and two, antibodies against receptors would likely cause steric hindrance in the synapse formation, as shown in the Nature of T-T contact (Chapter 5.3.3).

For the second possibility, it was possible our single cell test did not pick up smaller changes in select proteins, especially those outside of the linear region of detection [16]. These regions show small changes in fluorescence signal, even with significant protein concentration changes. To test this, one would have to develop higher sensitivity technologies, such as nanowire-based detection of proteins or resonator cavity detection, and integrate these into single cells experiments [17, 18].

In the CD4/CD8 phenotype experiment, it was observed that there were less CD4-CD8s in contact than the theoretical value anticipated (-62%). Consequently, there were more CD8-CD8 contacts than anticipated (+48%). One possible explanation for this would be the detection of false positive CD8 phenotypes. However, every chamber with two cells was checked by hand, so there was no chance for a false positive detection. Another possibility would be that the CD8 T cells were more "sticky" than the CD4s. However, this was also not the case, since the CD4-CD4 contact was actually higher than anticipated. One last possibility, which we believe was most likely, was that our statistics were too low. This could be remedied by increasing the chamber count of the device or engineering cell traps in the device that retain contacting cells rather than having it be purely stochastic [19].

There are numerous further experiments to perform using this experimental platform. First and foremost, it would be interesting to test younger T cells. The T cell samples tested came from metastatic melanoma patients whose T cells were expanded significantly. This expansion can "age" or differentiate the T cells [20]. Krummel et al. showed that this communication exists in younger T cells and it would be interesting to see if enhanced protein expression exists in the younger T cells. This would require detecting additional phenotypes in the microfluidic device. Adding phenotypes for T cell age would be a natural extension for this experiment. The addition of CCR7 and CD45RA would be sufficient for a simple panel for T cell differentiation state (or age) [2, 21]. If further surface markers were detected, one could use the photocleavable construct, described in Chapter 3.

One could also extend this technology towards studying heterotypic contact between other immune cells; for instance, T cells and dendritic cells during the priming stage. This is a critical step in the adaptive immune response, as poor contact between T cells and dendritic cells can reduce the formation of T cell memory [22, 1]. Using our microfluidic technology would give researchers a platform to study isolated interactions between different immune cells.

5.5 – Conclusions

Two cell protein assays conduced with T cells demonstrated that cell-cell contact can selectively enhance protein expression. T cells in close proximity (< 100 μm) did not exhibit the same enhancement, and therefore, direct contact is key. The proteins that were upregulated upon contact were related to cell killing, mobility, and immune activation, demonstrating a potential mechanism behind why T cell-T cell contact can improve the effectiveness of ACT. Therefore, this method has demonstrated that T cells communicate with each other and enhance their cytokine release upon contact.

References (Chapter 5)

[1]   K. Murphy, Janeway's Immunobiology, Garland Science, 2011.

[2]   N. Restifo et al., "Lineage relationship of effector and memory T cells," *Curr. Opin in Imm.,* vol. 25, pp. 556-563, 2013.

[3]   E. Lugli et al., "The who's who of T-cell differentiation: Human memory T-cell subsets," *Eur. J. Immunol.,* vol. 43, pp. 2797-2809, 2013.

[4]   S. Rosenberg et al., "Adoptive Cell Tranfer for Patients with Metastatic Melanoma: The Potential and Promise of Cancer Immunotherapy," *Cancer Control,* vol. 20, no. 4, pp. 289-297, 2013.

[5]   M. Dudley et al., "Adoptive cell transfer: a clinical path to effective cancer immunotherapy," *Nature Reviews Cancer,* vol. 8, pp. 299-308, 2008.

[6] M. Krummel et al., "Secondary T cell–T cell synaptic interactions drive the differentiation of protective CD8+ T cells," *Nature Immunology,* vol. 14, pp. 356-363, 2013.

[7] M. Krummel et al., "A Synaptic Basis for Paracrine Interleukin-2 Signaling during Homotypic T Cell Interaction," *Immunity,* vol. 29, no. 2, pp. 238-248, 2008.

[8] A. Zajac et al., "ICAM-1-dependent tuning of meory CD8 T-cell responses following acute infection," *PNAS,* vol. 110, no. 4, pp. 1416-1421, 2013.

[9] Y. Shimizu et al., "ICAM-1 dependent homotypic aggregates regulate CD8 T cell effector function and differentiation during T cell activation," *J Immunol.,* vol. 191, no. 7, p. 3681–3693, 2013 .

[10] J. Tsitsiklis et al., Introduction to Probability, Belmont, Massachusetts: Athena Scientific, 2008.

[11] J. Xiang et al., "Direct in vivo evidence of CD4+ T cell requirement for CTL response and memory via pMHC-I Targeting and CD40L signaling," *J. Leuk. Bio,* vol. 92, 2012.

[12] Y. Shimizu et al., "ICAM-1-Dependent Homotypic Aggregation Regulate CD8 T Cell Effector Function and Differentiation during T cell Activation," *J Immunol,* vol. 191, pp. 3681-3693, 2013.

[13] B. Henz et al., "Human leukaemic (HMC-1) and normal skin mast cells express beta 2-integrins: characterization of beta 2-integrins and ICAM-1 on HMC-1 cells," *Scand. J. Immunol.,* vol. 45, pp. 471-481, 1997.

[14] J. Lectron et al., "Preincubation of Human Resting T cell clones with Interleukin 10 Strongly Enhances their ability to produces Cytokiens after Stimulation," *Cytokine,* vol. 10, no. 11, pp. 831-840, 1998.

[15] C. RAETZ et al., "PROCESSING AND SECRETION OF TUMOR-NECROSIS-FACTOR-ALPHA IN

ENDOTOXIN-TREATED MONO MAC-6 CELLS ARE DEPENDENT ON PHORBOL-MYRISTATE ACETATE,"
JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 267, no. 32, pp. 23261-23268, 1992.

[16] D. Dziuda, Data Mining for Genomics and Proteomics: Analysis of Gene and Protein Expression Data, Wiley-Interscience, 2010.

[17] C. Lieber et al., "Nanowire biosensors for label-free, real-time, ultrasensitive protein detection," *Methods Mol Biol.,* vol. 790, pp. 223-237, 2011.

[18] X. Fan et al., "Characterization of sensing capability of optofluidic ring resonator biosensors," *Appl. Phys. Lett.,* vol. 97, 2010.

[19] L. Lee et al., "Single-cell enzyme concentrations, kinetics, and inhibition analysis using high density hydrodynamic cell isolation arrays," *Analytical Chemistry,* vol. 78, no. 14, pp. 4925-4930, 2006.

[20] F. Sallusto et al., "Progressive differentiation and selection of the fittest in the immune response," *Nature Rev. Immunol.,* vol. 2, pp. 982-987, 2002.

[21] M. Roederer et al., "T-cell quality in memory and protection: implications for vaccine design," *Nature Rev. Immunol.,* vol. 8, pp. 247-258, 2008.

[22] A. Zajac et al., "Shaping Successful and Unsuccessful CD8 T cell Responses Following Infection," *J. Biomed. and Biotech.,* vol. ID: 159152, pp. 1-12, 2010.

[23] P. Robbins et al., "Cutting edge: Persistence of transferred lymphocyte clonotypes correlates with cancer regression in patients receiving cell transfer therapy," *JOURNAL OF IMMUNOLOGY,* vol. 173, no. 12, pp. 7125-7130, 2004.

[24] J. Xiang et al., "Th Cells Promote CTL survival and Memory via Acquired pMHC-I and Endogenous IL-

2 and CD40L Signaling and by Modulating Apoptosis-Controlling Pathways," *PLOS ONE,* vol. 8, no. 6,

2013.

[25] C. June et al., "Chimeric Antigen Receptor-Modified Cells in Chronic Lympoid Leukemia," *NEJM,* vol.

365, no. 8, pp. 725-733, 2011.

[26] S. Rosenberg et al., "Cancer Regression and Autoimmunity in Patients After Clonal Repopulation

with Antitumor Lympocytes," *Science,* vol. 298, pp. 850-854, 2002.

[27] S. Rosenberg et al., "Tumor regression in patients with metastatic synovial cell sarcoma and

melanoma using genetically engineered lymphocytes reactive with NY-ESO-1," *J Clin Oncol,* vol. 29,

no. 7, pp. 917-924, 2011.

[28] A. Zajac et al., "ICAM-1-dependent tuning of memory CD8 T-cell responses following acute

infection," *Proc Natl Acad Sci,* vol. 110, no. 4, pp. 1416-1421, 2013 .