Highly Informative Analytical Platforms for Rapid, Non-Invasive Diagnosis and Stratification of Patients with Cancer

Thesis by

Ophir Vermesh

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2011

(Defended August 30, 2010)

"Our truest life is when we are in dreams awake."

-Henry David Thoreau

Acknowledgments

I wish to express my deepest gratitude to a number of people who have had an indelible impact on my graduate experience and without whom this dissertation would not have been possible. First of all, I would like to thank my advisor, Professor James Heath, for his vision and guidance. He has created a lab environment that encourages thinking outside the box and coming up with novel, big ideas to solve some of the most seemingly intractable scientific problems. It was truly a privilege to have access to the nearly unlimited resources of his first-rate lab, including a top-of-the-line in-house microfabrication facility. The Nanosystems Biology Cancer Center (NSBCC), which encompasses Caltech, UCLA, and Leroy Hood's Institute for Systems Biology (ISB), and which is headed by Jim, fosters the kind of interdisciplinary, collaborative environment that is needed to enable progress in these projects.

I would also like to thank Professor Paul Mischel (UCLA) and Dr. Tim Cloughesy (UCLA) for their prominent collaborative roles in our clinical trial on glioblastoma patients. We received all of our GBM patient blood samples from Dr. Cloughesy's neuro-oncology clinic. Dr. Cloughesy also provided a large amount of clinical information about each patient, allowing us to investigate differences in many clinical parameters within each of the cohorts he designated. Professor Mischel, with his broad and deep expertise in the glioblastoma field, helped us choose a suitable protein panel for our plasma assays, highlighted those aspects of the patient cohorts that were most interesting to investigate further, and helped us to fully appreciate the significance and potential impact of our results. I would also like to thank him for his generous gift in providing us with a full set of capture and detection antibodies when our first set ran out. Further, I would like to extend my gratitude to Dr. Alyssa Ziman for providing us with blood

samples from healthy individuals for our control experiments. I would also like to thank Tom Bourgeois and his colleagues at the Clinical Immunology Research Laboratory (CIRL) at UCLA for centrifuging and processing the hundreds of patient samples used in our study. Thanks also to Bruz Marzolf at ISB for printing our DNA microarray-spotted slides and for troubleshooting issues related to DNA-loading, spot morphology, spot-to-spot and slide-to-slide consistency, and so forth.

I had the distinct pleasure of working closely with my brother, Udi Vermesh, on the glioblastoma clinical trial. We both appreciated just how rare it is for two brothers to end up in the same graduate school, in the same lab, and working on the same project. Because our experiments and assays were long and arduous, often extending well into the night, it was a blessing to work side-by-side with Udi, with whom I could joke, share entertaining experiences, and take breaks to go to the gym. Our shifted work schedules also matched up quite well. I also had the pleasure of collaborating with Dr. Rong Fan (now an Assistant Professor at Yale), with whom I co-authored the *Nature Biotechnology* paper "Integrated barcode chips for rapid, multiplexed analysis of proteins in microliter quantities of blood." The paper could not have had the impact it did without Rong's impressive clinical experiments and his method for patterning DNA at high-density. I would also like to thank Dr. Brian Yen, who was instrumental in the early development of the blood separation chip. He wrote a program in MATLAB that optimized the lengths and widths of the various microfluidic device channels to allow for a high degree of plasma separation efficiency. His experience with microfluidics was extremely helpful in guiding various aspects of the design and fabrication process. Thank you to Marino DiFranco, our undergraduate SURF student, for writing the various batch files we used to automate our statistical analysis. Thanks also to Alok Srivastava (ISB) for his help in the early DNA, antibody,

and DEAL conjugate validation experiments. I would also like to thank Chao Ma and Kiwook Hwang for their help with the tissue engineering project (not described herein).

Other colleagues and good friends who I would like to thank are Gabe Kwong, Tiffany Huang, and Shawn Sarkaria. My conversations with Gabe greatly enhanced my knowledge of and intensified my interest in biomedical technology development, and inspired me to become innovative in my own right. He also greatly assisted me in the early stages of the tissue engineering project. From Tiffany, I learned an enormous amount about GBM as well as about the various experimental methods used in cellular and molecular biology. Both Tiffany and Shawn (Mischel Lab) were instrumental in identifying a panel of proteins with high biological relevance to glioblastoma and to cancer generally. Our multiplexed assay platform was designed to target the 35 proteins they chose. In addition, they optimized the procedure for conjugating antibodies to DNA and contributed purified conjugates for our experiments.

I would also like to thank Mike McAlpine (now an Assistant Professor at Princeton) and Dr. Slobodan Mitrovic for many exciting and entertaining conversations about science and other topics throughout the years. I would like to thank Diane Robinson for her administrative help, for being a good friend, and for just being fun. I would also like to thank Kevin Kan for ensuring the smooth operation of the clean room facility and, more recently, the biology lab. Thanks to Jackie Barton, Jack Beauchamp, and Nate Lewis for serving on my thesis committee. Finally, I would like to thank my parents for their love and support throughout the years, and helping to shape me into who I am today.

Abstract

As the tissue that contains the largest representation of the human proteome, blood is the most important fluid for clinical diagnostics. However, although changes of plasma protein profiles reflect physiological or pathological conditions associated with many human diseases, only a handful of plasma proteins are routinely used in clinical tests. Reasons for this include the intrinsic complexity of the plasma proteome, the heterogeneity of human diseases and the rapid degradation of proteins in sampled blood. The first part of this thesis reports an integrated microfluidic system, the integrated blood barcode chip (IBBC) that can sensitively sample a large panel of protein biomarkers over broad concentration ranges and within 10 minutes of sample collection. It enables on-chip blood separation and rapid measurement of a panel of plasma proteins from quantities of whole blood as small as those obtained by a finger prick. The device holds potential for inexpensive, noninvasive and informative clinical diagnoses, particularly in point-of-care settings.

Proteomic approaches, on which the IBBC platform is based, have shown great promise in recent years for correctly classifying and diagnosing cancer patients. However, no large antibody-based microarray studies have yet been conducted to evaluate and validate plasma molecular signatures for detection of glioblastoma and monitoring of its response to therapy. In the second part of this thesis, plasma samples from 46 glioblastoma patients (72 total samples) are compared with those of 47 healthy controls with respect to the plasma levels of 35 different proteins known to be generally associated with tumor growth, survival, invasion, migration, and immune regulation. Average-linkage hierarchical clustering of the patient data stratified the two groups effectively, permitting accurate assignment of test samples into either GBM or healthy

control groups with a sensitivity and specificity as high as 90% and 94%, respectively (when test samples within unbiased clusters were removed). The accuracy of these assignments improved (sensitivity and specificity as high as 94% and 96%, respectively) when the cluster analysis was repeated on increasingly trimmed sets of proteins that exhibited the most statistically significant ($p < 0.05$) differential expression. The diagnostic accuracy was also higher for test samples that fell into more homogeneous clusters. Intriguingly, test samples that fell within perfectly homogeneous clusters (all members belonging to the same group) could be diagnosed with 100% accuracy. Using the same 35-protein panel, we then analyzed plasma samples from GBM patients who were treated with the chemotherapeutic drug Avastin (Bevacizumab) in an effort to stratify patients based on treatment-responsiveness. Specifically, we compared 52 samples from (25) patients who exhibited tumor recurrence with 51 samples from (21) patients who did not exhibit recurrence. Again, several proteins were highly differentially expressed and cluster analysis provided effective stratification of patients between these two groups (sensitivity and specificity of 90% and 96%, respectively).

Table of Contents

Table of Figures

List of Tables