# Systems Biology of Mature Human B cells in Health and Illness

By

Cara Ellen Wogsland

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Microbiology and Immunology

June 30, 2017

Nashville, Tennessee

Approved:

Jonathan M. Irish, Ph.D.

Andrew J. Link, Ph.D.

Holly M. Algood, Ph.D.

Leslie J. Crofford, M.D.

R. Stokes Peebles, M.D.

This work is dedicated to my children, Alora, Brittan, Maxwell, and Zara; and my nieces, Lexi, Avery, and Ellie. Remember, perspective is everything. When life seems too hard, adjust your perspective, have a bite to eat and a nap, and carry on. May you find your way in life and be willing to work hard, shed tears, and climb over or tear down walls to get there.

# ACKNOWLEDGEMENTS

There are so many people to thank. I would not be the person I am today without all the people who have been in my life. We shape every life we touch and those lives in turn, shape us. Here I will acknowledge a few of those lives that have shaped mine.

I would like to thank my mentor, Jonathan Irish, for his guidance and support throughout graduate school. Thank you for giving me the freedom to explore and define my own project. I learned more from this freedom than I would have thought possible. Thank you for sticking up for me and sticking by me even when my world felt like it was falling apart. Thank you for never giving up on me. We are not so very different which I think is why we sometimes drive each other crazy. I look forward to future meetings where we can ponder things like what makes a population a population.

I would like to thank my committee chair, Andrew Link, for helping me throughout graduate school and remaining a calming, grounded voice of wisdom. I would like to thank my entire committee including committee members Holly Algood, Leslie Crofford, Stokes Peebles, and Mark Boothby for challenging me and forcing me to think outside the box. Your input and questions helped make my work grounded and balanced.

I would like to thank my fellow lab members and the members of the Rebecca Ihrie lab for their comradery, friendship, and scientific conversations. I need to especially thank Kanutte Huse. We mentored each other and became fast friends of the best kind. We can discuss just about anything. We have both taught each other so much about science and life.

I would like to thank my mother, who knew the world she was raised in was not fair. She was not allowed to take the advanced math class because she was a girl. When she played basketball, girls were only allowed three steps before they could pass

the ball because they might get overexerted if they did more. She knew these limitations were unfair and unnecessary and never imposed limitations on me. She made sure I had a mix of “boy” and “girl” toys available. I had dolls, cowboys and Indians, Barbies, and a tool bench. My mother was one of the toughest people I have ever known. She raised two children while her husband, my father, was slowly dying and then continued to raise my sister and I as a widow and single mother for three years before meeting my stepfather. My mother found out that she had breast cancer shortly after she met my stepfather. He stayed by her side and helped her through surgery and recovery while taking care of my sister and I as well. She beat the cancer and it went away for over 20 years before rearing its ugly head again. When the cancer found her the second time, it was in her bones and I had just started at Vanderbilt the summer of 2011. I got almost another year with her before the cancer went to her brain. It killed her within weeks after that. One day before she died, we were eating in the cafeteria between her cancer treatments and she started talking about my “speriments.” I had no idea what she was talking about. She explained, telling me that when I was little, I was always doing little experiments and “speriments” is what I called them. I treasure this memory. I guess I have been a scientist at heart my whole life. Thank you, Mom. My mother never realized how smart she was. She wondered where I got it from. I do not.

I would like to thank my father, Joel, who was patient and kind and left us far too soon. My father had type 1 diabetes. He was diagnosed when he was 12 years old. Back then, checking insulin was not something that was done on a regular basis or even at home. His diabetes was not well regulated and it was hard to control. That took a toll on his body and by the age of 30, he was very sick. I remember going to dialysis with him because his kidneys no longer functioned. I remember him being in insulin shock and him

lying in bed with his eyes half open but not seeing me. I remember the friendly paramedics that carried him down the stairs on a stretcher late one night. But my most cherished memory is of my dad Joel brushing my hair. My mom was always in a hurry and tended to pull on my tangles and it hurt. My dad Joel never rushed. He was so patient and gentle. He did not always have the energy to brush my hair but when he did, it was wonderful. Having three girls of my own, I still do not know how he managed to be so gentle. I do not have many memories of my dad because he died when I was only five years old. Most of what I know of him comes from the people and things he left behind. He was a creator, a painter, a Star Trek fan, and Lord of the Rings fan. He made his own detailed map of Mordor which hangs on my wall. He made the frame for it too. He was not afraid to try new things and was always starting projects and creating. I thank him for the creativity that I inherited from him. Both of my parents were intelligent, thoughtful, and passionate and they passed those attributes on to me.

I would like to thank my dad Ray, my stepfather, who allowed me to be fearless with computers. My mother remarried when I was eight years old and it was not long before my sister and I were calling Ray, Dad. He was so loving and kind and accepted my sister and I as his own. He worked for IBM for years and so in 1988 when my mom and my dad Ray got married, I got a computer in my house. That was a big deal in 1988. And bigger than that, at 8 years old, I was allowed to use it. It is rare for someone my age, 37 in 2017, to have a parent that knows more about computers than they do but that is the case for me. I will never know as much about computers as my dad Ray. I never had any fear of breaking computers because he could fix anything I did. I thank him for this because it has helped me to learn new computer systems, software, and programming

languages without fear. Computers are such a huge part of my scientific and personal life. I owe him a great deal of thanks.

I need to thank my husband, Bradley Wogsland, for always believing in me and encouraging me. He has the ability to sometimes see me better than I see myself, and for that I am grateful. When I went back to school in 2009, my plan was to learn about food allergies and start with a degree in food science. He encouraged me to do biochemistry because it would open more doors. I was terrified. I was not sure that I could do it. I failed one of first tests that first semester in honors Cell Structure. It was bad, one of those you know you failed when you are leaving the room because you did not even finish it. His belief in me never wavered. I kept pushing through and rallied to get straight As that first semester after being out of school for 7 years. I quickly realized that not only could I successfully complete my degree in biochemistry but that I was exceling at it and loving it. Bradley encouraged me to continue on to get my PhD after my bachelors and was willing to follow me anywhere that might take me. Here I stand about to get my PhD and he never finished his. So, with the power that I have in my own acknowledgements section, I grant you the title of Dr. Wogsland herein. Thank you Dr. Bradley Wogsland for being there for me and supporting me throughout our many years together. Thank you for believing in me even when I do not believe in myself. I am sorry for all the times I thought you were crazy for thinking I could do it.

I would like to thank my four children, Alora (19), Brittan (19), Maxwell (13), and Zara (11), who have suffered through having a parent in school for most of their lives. It is hard to be a kid with a parent who has homework to do in the evenings and on weekends. They remind me why I am doing this and have always been so proud of me. They are the reason I persevered. They thought I was cool even when I was sure I would

not make it through this process. My son, Maxwell, was my inspiration for embarking on this research journey. He has multiple food allergies and I wanted to make a difference in the lives of people with food allergies through research. My project drifted away from allergy research but, throughout graduate school, he listened to me talk about my science. He asked insightful questions from a young age and regularly tried to engage me in scientific discussions, particularly at bedtime.

I would also like to thank Bethany Dale for her support and friendship. She made my days brighter during my last year of graduate school when things seemed bleakest. I will treasure our science and yoga always.

Thank you also to my Aunt Cindy and my husband's Aunt Dell for their support during an especially hard time when I was writing my dissertation. I do not think I would have finished it without your love and support.

TABLE OF CONTENTS

Page

# LIST OF TABLES

Table Page

# LIST OF FIGURES

Figure Page

# LIST OF ABBREVIATIONS

Ab : antibody

ADCC : antibody dependent cellular cytotoxicity

Ag : antigen

AID : activation-induced cytidine deaminase

BCL2 : B-cell lymphoma 2, apoptosis regulator

BCR : B cell receptor

Breg : regulatory B cell

CD : cluster of differentiation, naming system for cell surface molecules

CITRUS : cluster identification, characterization, and regression algorithm

CLL : chronic lymphocytic leukemia

CSR : class switch recombination

CVID: common variable immune deficiency

CXCR4 : C-X-C chemokine receptor type 4, also known as CD184

CXCR5 : C-X-C chemokine receptor type 5, also known as CD185

CyTOF : mass cytometer made by Fluidigm (formerly DVS)

EBF1 : early B cell factor 1 protein

FBS : fetal bovine serum

Fc : fragment crystallizable region of an antibody

FCS : flow cytometry standard data file

FcγR : Fc gamma receptor

FL : Follicular lymphoma

GC : germinal center

HLA : human leukocyte antigen complex, human MHC, eg HLA-DR, -DQ, -DP

HSC : hematopoietic stem cells

Ig : Immunoglobulin

IHC : immunohistochemistry

IMD : integrated mass data file, raw data file generated by CyTOF mass cytometers

MEM: marker enrichment modeling

MHC: major histocompatibility complex (proteins)

PAX5 : paired box 5 protein

PBMC : peripheral blood mononuclear cells

PLC : phospholipase C

PMT : photomultiplier tubes, used in fluorescence cytometry

ROS : reactive oxygen species

RPMI-1640 : Roswell Park Memorial Institute 1640 cell medium

SFK : Src family kinase

SHM : somatic hypermutation

SHP-1 : src-homology-2 domain-containing phsophatase-1 protein

SPADE : spanning-tree progression analysis of density-normalized events

SYK : spleen associated tyrosine kinase

TCR : T cell receptor

Tfh : T follicular helper cells

Treg : regulatory T cells

tSNE/t-SNE : t-distributed stochastic neighbor embedding, specifically the Barnes-Hut implementation known as BH-SNE

V(D)J/VDJ : variable (V), diversity (D) and joining (J) gene recombination in B and T cells

viSNE : a visualization tool for high dimensional single cell data based on tSNE (BH-SNE)

# PREFACE

In this dissertation, I describe the application of mass cytometry to human B cells, germinal center (GC) biology and follicular lymphoma (FL) and its impact on B cell biology research and understanding. I have designed detailed mass cytometry phenotype staining panels for the study of B cell biology and used these to characterize human B cells in healthy peripheral blood, tonsil, and lymphoma tumors.

In-depth characterization of B cells from a healthy immune system was needed before the disease state could be explored. I began with fluorescence flow cytometry to identify peripheral blood mononuclear cell (PBMC) subsets and to phenotype mature human B cells. I then started collecting mass cytometry data and developed a B cell-focused panel that could identify mature B cell populations with room for additional immunophenotyping. This depth of protein phenotyping required novel approaches for data analysis to take advantage of the 30+ markers per cell. I probed the lymphocyte populations of follicular lymphoma tumors at the single cell level. I contributed to a study characterizing mature tonsillar B cell subsets by phenotype and phospho-signaling responses. I optimized computational approaches such as viSNE to maximize the biological veracity of the output viSNE maps. I have contributed tools and approaches that enable unbiased in-depth phenotyping of human B cells and have shown that these tools work in healthy and disease settings. My research has primarily been hypothesis generating in nature. The scope of my data collection with mass cytometry reveals much and is useful for generating new ideas.

In Chapter 1, I introduce the fields of systems biology, human B cell biology, germinal center biology, follicular lymphoma, mass cytometry, and mass cytometry

analysis tools and approaches. The chapter discusses the field when I entered, the changes that have occurred since, and my contributions to the field.

In Chapter 2, I used my knowledge of healthy peripheral blood and tonsil and applied it to mass cytometry phenotyping of a small patient cohort of resected follicular lymphoma (FL) tumors (Wogsland, Greenplate et al. 2017). Using healthy donor tonsils as a comparison point, I thoroughly characterized the lymphoma phenotypes of T cells, malignant B cells, and non-malignant B cells. In FL, a substantial proportion of the B cells are the malignant disease-causing cells and the disease phenotype is pronounced. I identified a potential GC deficit in FL tumors that could contribute to immune deficits by inhibiting new GC reactions. In my characterization, I found that the heterogeneity within tumors was driven strongly by cell to cell variations in expression levels of human leukocyte antigen D related antigen (HLA-DR). Surprisingly I also found that FL malignant B cells which are thought to arise from a germinal center B cell, were phenotypically distinct from GC B cells.

Chapter 3 entitled, “Cutting edge: redox signaling hypersensitivity distinguishes human germinal center B cells” was published in the Journal of Immunology (Polikowsky, Wogsland et al. 2015). Using a high dimensional mass cytometry panel, we identified cell types and characterized signaling features. I developed the first human B cell-focused mass cytometry panel that was the framework for this panel. The published panel contains a backbone of B cell markers that identify the mature B cell subsets of naive, memory, class switched memory, germinal center, and plasmablasts/plasma cells. The B cell panel was adapted for this study to include intracellular signaling readouts. We modified the surface panel by adding phospho-antibodies to study B cell signaling in the germinal center. We explored B cell responses to titrating levels of reactive oxygen

species (ROS; hydrogen peroxide) and found germinal center B cells to be more sensitive than other B cell types found in tonsil.

In Chapter 4, I present the computational experiments I performed while testing the robustness of viSNE with large flow cytometry standard (FCS) file datasets. High dimensional mass cytometry requires more than biaxial plots to get the full use of such a dense data set. It is crucial that the results produced by computational tools reflect and reveal biological truth. I found that decreased cell number or increased iteration number both improved the quality of the viSNE map when the data set is large.

Conclusions, unanswered questions, and future directions are addressed in Chapter 5. I discuss ways to test several unanswered questions related to the follicular lymphoma phenotyping project. I discuss other topics of interest including B cells, allergy, and mass cytometry.

Additional protocols, figures, and co-authored texts are included in the appendices.

# CHAPTER 1 INTRODUCTION

## Systems biology

*Systems biology is the study of systems of biological components, which may be molecules, cells, organisms or entire species. Living systems are dynamic and complex, and their behavior may be hard to predict from the properties of individual parts.*

- Excerpt from the Department of Systems Biology at Harvard Medical School webpage (2017)

Researchers have struggled to define and understand the emerging field of systems biology (Kirschner 2005, 2007, Breitling 2010). The usefulness of systems biology is broadly recognized but understanding and applying this holistic view of a system as opposed to a reductionist view has been met with resistance. The need for systems biology and systems immunology becomes clear when working in a human system where traditional reductionist approaches often used on mouse models or cell lines are not applicable. Systems biology research does not focus on single proteins and their role in a cell or an organism like traditional reductionist research. Although studying multiple proteins and cell types, like I have done here, might seem to be taking on too much, or unmanageable; it is only with this broad systems look that we can really understand certain features of disease or health. Taking a systems biology approach to characterize a disease or biological system allows it to be understood at a high level and can be used to not only answer questions but to also generate new hypotheses. The mass cytometry approach enables a broad yet deep look at the protein features of individual cells. Not all systems biology approaches capture data at the single cell level like mass cytometry. Systems biology studies can also include combinations of transcriptomics, metabolomics,

cytokine studies, or proteomics of tissue, sorted cells, or single cells. The single-cell systems data collected with mass cytometry is invaluable for revealing the nuances of protein expression within and between cell types. In their recent review, Brodin and Davis discuss the value of systems immunology in translating animal model knowledge to humans(Brodin and Davis 2017). They stress the value of systems immunology to understand how perturbations to the immune system change immune cell populations. Specifically, they refer to systems vaccinology, as a way to study the human immune response to vaccines (Brodin and Davis 2017). Tracking the immune response over time to vaccines (Hagan, Nakaya et al. 2015), or to infection at the single cell level (Appendix F) reveals the changes that occur in immune cell populations. This knowledge aids in understanding the impact of the vaccine or infection on the immune system.

In Chapter 2, I took a similar approach, using single cell systems biology but instead of looking over time, I looked at a single pre-treatment time point and compared between the disease state of follicular lymphoma and the healthy state of tonsil. The systems biology approach confirmed FL characteristics reported in the literature, including a lack of IgD co-expression on IgM+ malignant cells (Grier, Al-Quran et al. 2012), low to negative CD44 expression on malignant cells (Detry, Drenou et al. 2004, Eberth, Schneider et al. 2010), and retention of BCR expression on malignant cells (Kuppers 2005, Zuckerman, McCann et al. 2010, Buchner and Müschen 2014). This approach also led to multiple new findings; GC B cells were diminished in FL tumors; FL malignant B cells did not resemble GC B cells phenotypically; the malignant B cells had a high level of intra-tumoral heterogeneity that was driven by HLA-DR in all tumors studied; the T cell subsets were not significantly different between tumor and healthy samples although they trended towards an activated phenotype in FL tumors (Wogsland,

Greenplate et al. 2017). The implications and ideas for future studies are presented in Chapter 5.

## Significance of B cell biology

B cells are an integral part of the adaptive immune system. They have multiple functions including antibody production, antigen presentation, T cell activation, and cytokine production (LeBien and Tedder 2008, Shen and Fillatreau 2015, Hoffman, Lakkis et al. 2016). A healthy B cell repertoire is essential for fighting off infections and developing lasting memory responses to a multitude of pathogens (Conley, Dobbs et al. 2009). This is apparent in B cell diseases, including antibody deficiencies, where patients are far more susceptible to infection than individuals with a normal B cell repertoire (LeBien and Tedder 2008, Conley, Dobbs et al. 2009).

B cells, primarily plasma cells, produce antibodies that circulate throughout the body (Medina, Segundo et al. 2002). Secreted antibodies are integral to fighting and preventing infection as part of the humoral immune response. The B cell heavy chain isotype plays a prominent role in antibody function (Murphy, Travers et al. 2008). The antibody isotypes IgD and IgM are produced by B cells that have not undergone class switch recombination (CSR). IgG(1-4), IgA(1,2), and IgE are found on the B cell surface or are secreted by plasmablasts/plasma cells that have undergone CSR (Klein and Dalla-Favera 2008). Secreted antibodies have an array of circulation patterns with IgM antibodies generally forming pentamers, IgA circulating in monomeric form in blood and dimeric form in the mucosa, and IgG and IgE secreted as monomers (Woof and Kerr 2006, Schroeder and Cavacini 2010). The Fc portion of the different isotypes direct which receptors the secreted antibodies can bind. The most striking of these is IgE's ability to

bind to the high affinity IgE receptor (IgεRI) without first binding antigen (Kinet 1999, Gould, Sutton et al. 2003). The Fc portion of the antibody determines the functional role the antibody. Fc regions have different affinities for different Fc receptors that are located on different cell types and in different tissues. IgG antibodies have a family of six Fc gamma receptors (FcγR) that are found on multiple cell types, have varying affinities for the IgG subtypes, can be activating or inhibitory, and play different roles in antibody dependent cellular cytotoxicity (ADCC) and phagocytosis (Li and Kimberly 2014). The role of IgG antibodies and FcγRs in the context of cancer immunotherapy has been reviewed by Stewart et al. (Stewart, Hammond et al. 2014). B cells can express FcγRIIb (CD32) which is considered an inhibitory receptor that binds IgG1,3,4 but not IgG2 (Stewart, Hammond et al. 2014). Engagement of CD32 has an inhibitory effect on BCR signaling (Karnell, Dimasi et al. 2014). There are six IgG Fc receptors, some activating and others inhibitory, with variable binding affinities for the four IgG subtypes (Nimmerjahn and Ravetch 2008).

B cells are professional antigen presenting cells (APC). They can uptake specific antigen via the BCR and display it on MHC-II complexes to cognate CD4 T cells. This specific antigen display can activate the cognate T cell which in turn activates the presenting B cell via CD40/CD40L binding (Rodríguez-Pinto 2005, Chen and Jensen 2008). B cells can also display non-specific antigen on MHC-II molecules which acts to tolerate CD4 T cells and does not activate the presenting B cell (Chen and Jensen 2008). The human MHC class II complexes HLA-DR, DP, and -DQ present exogenous antigen on the surface of APC including B cells (Janeway, Travers et al. 2001). The mass cytometry panels in this dissertation include staining for HLA-DR. In healthy B cells HLA-DR, DP, and DQ have coordinate expression, whereas in some lymphomas and

leukemias including FL that is not the case and expression patterns may vary (Guy, Krajewski et al. 1986). This is one of the few papers published that specifically measured DR DP and DQ individually. The FL data from Guy et al. was collected by immunohistochemistry (IHC) and there is no single cell information or images in the publication. In Chapter 2, HLA-DR was found to be highly variable on FL malignant cells. All tumors showed that same variability of expression on the malignant cells.

B cells are key players in effective vaccination strategies. Vaccines are dependent on generating a strong T cell-mediated B cell response that leads to the expansion of antigen specific memory B cells, antibody secreting plasma cells, and antigen specific T cells (Siegrist 2008). During the early primary immune response, IgM B cells and antibodies are the primary isotype. B cells will class switch to IgG later in the immune response (Siegrist 2008). This B cell response to vaccination (and infection) is an essential part of a healthy immune system (Clem 2011).

B cells are involved in many diseases including cancer, autoimmunity, and allergy (Gould, Sutton et al. 2003, Browning 2006, Yanaba, Bouaziz et al. 2008, Conley, Dobbs et al. 2009, Pillai, Mattoo et al. 2011, Braza, Chesne et al. 2014, Schwartz, Zhang et al. 2016). B cells may inhibit tumor growth, promote tumor growth, or be the tumor causing cells, as will be explored here in the context of follicular lymphoma (Tsou, Katayama et al. 2016, Yuen, Demissie et al. 2016). In B cell lymphomas, the B cell is the malignant pathogenic cell (Kuppers 2005, Gerstein, Zhou et al. 2015). Directly targeting B cells in these lymphomas has changed the field and the outcome for many patients. Specifically, the introduction of anti-CD20, Rituximab/Rituxan, in 1997 as a therapeutic agent has dramatically improved lifespans for patients with B cell lymphomas (Dotan, Aggarwal et al. 2010) (Evans and Clemmons 2015). A new anti-CD20 drug, Obinutuzumab/Gazyva,

was released in 2013 for use in chronic lymphocytic leukemia (CLL) (Evans and Clemmons 2015). It was approved in 2016 as a second line therapy in FL (2016). Obinutuzumab is thought to have a greater antibody-dependent cell-mediated cytotoxicity (ADCC) effect than Rituximab (Alduaij, Ivanov et al. 2011, Evans and Clemmons 2015). It is too soon to say if the progression free survival and overall survival in FL patients will be better than with Rituximab.

B cells play a pathogenic role in autoimmune diseases by producing autoreactive antibodies (Mietzner, Tsuiji et al. 2008, Bax, Huizinga et al. 2014, Suurmond and Diamond 2015). These self-identifying antibodies lead to cell death and inflammation. Anti-CD20 therapies have more recently been introduced in autoimmune diseases with good results (Perosa, Prete et al. 2010). Defective B cells are the culprit in many immune deficiencies, often due to the inability to produce protective antibodies such as in the chronic variable immune disease (CVID) family of diseases. B cell immune deficiencies are usually monogenic. Patients present with a history of recurring infections (Warnatz, Denz et al. 2002, Conley, Dobbs et al. 2009, Salzer, Santos-Valente et al. 2013).

## Human B cell development and the germinal center reaction

### Primary B cell development

B cell development begins in the bone marrow with hematopoietic stem cells (HSC) (LeBien 2000). The expression of transcription factors EBF1 and PAX5 are required for B cell development (Hagman and Lukin 2006, Somasundaram, Prasad et al. 2015). EBF1 controls pre-BCR expression and the PI3 Kinase signaling pathway (Hagman, Ramirez et al. 2012). PAX5 plays multiple roles in B cell development. It comes

on early in B cell development and represses non-B lineage gene expression while promoting B-lineage protein expression. PAX5 maintains B cell identity throughout B cell development and plays a role in V(D)J recombination (Medvedovic, Ebert et al. 2011)

B cell precursors undergo V(D)J recombination in multiple steps producing a wide array of gene combinations and B cell receptor (BCR) variable regions (Brack, Hirama et al. 1978, LeBien and Tedder 2008). B cell development is subject to the principles of central tolerance. B cells must have a successful B cell receptor gene rearrangement to survive and reach maturity. They require a low level tonic signal through the B cell antigen receptor (BCR) for survival (Lam, Kuhn et al. 1997, Srinivasan, Sasaki et al. 2009). Too strong of a BCR signal can lead to cell death, anergy, or receptor editing (Nemazee 2006, Zou and Diamond 2013). If a B cell receives to strong of a signal through the BCR, it will undergo light chain receptor editing and attempt to produce a modified BCR that is no longer self-reactive (Nemazee 2006). Mature B cells circulate throughout the body in blood and lymph. The naïve B cells that exit the bone marrow express a diverse repertoire of BCR gene rearrangements, thus enabling the detection of a wide range of potential antigens.

**Germinal center reactions and peripheral B cell development**

The interplay between kinase activity and phosphatase regulation is thought to determine the fate of mature B cells undergoing the germinal center (GC) reaction. Chapter 3 explores this with peroxide stimulation of tonsil B cells and find that GC B cells have a more robust phospho-signaling response than other tonsillar B cell subsets. Chapter 2 and Chapter 3 use mass cytometry for an in-depth analysis of GC B cells. Chapter 2 focuses on GC B cells in relation to FL B cells while Chapter 3 focuses on

kinase and phosphatase differences. When a B cell is activated by binding antigen to the B cell receptor (BCR) and receiving additional signals in the form of cytokines or T cell help, the activated B cell travels to a B cell follicle in lymphoid tissue and begins to proliferate and form a germinal center reaction (Klein and Dalla-Favera 2008). The GC reaction is a dynamic process with the result being improved antigen recognition in long lived memory B cells and the production of long lived antibody secreting plasma cells. Surface molecules and signaling cascades play crucial roles in the GC reaction. Deficits in surface molecule expression or signaling pathways can lead to imbalances and deficits in the humoral immune system.

When a B cell encounters antigen with its B cell receptor and receives T cell help via CD40/CD40L, it is activated and can then move to the B cell follicle in the lymphoid tissue. Mature naïve B cells are IgD+, CD20+ and CD19+ . They lack CD27 and CD38 which are expressed on memory B cells, and GC B cells respectively (Yurchenko, Kovalevska et al. 2010, Huse, Bakkebo et al. 2011, Wei, Jung et al. 2011). Activated T cells in the periphery can activate B cells. These T cells can be detected by activation markers such as PD-1, MHC II, ICOS, CD69, and CD38 (Rao, Gurish et al. 2017). Surface markers such as CD62L, CXCR5, and CXCR4 assist expressing cells in homing and entry to lymphoid tissue(Stein and Nombela-Arrieta 2005, Wirth, Badovinac et al. 2009). Activated B cells form a germinal center inside the B cell follicle as they begin to proliferate and modify their B cell receptors during a process known as affinity maturation (Klein and Dalla-Favera 2008).

During affinity maturation, activation-induced cytidine deaminase (AID) is expressed and plays a role DNA modification in somatic hypermutation (SHM) and class switch recombination (CSR) (Muramatsu, Kinoshita et al. 2000) SHM induces point

mutations in the B cell receptor while CSR makes a change in heavy chain isotype. These mutations may produce an antigen specific high affinity BCR, make the B cell receptor less likely to bind antigen, destroy the B cell receptor entirely, or cause mutations to other DNA sites (Klein and Dalla-Favera 2008). Many B cells die at this point due to unsuccessful BCR changes or DNA damage. It is thought that many B cell lymphomas arise from inappropriate mutations during the germinal center reaction (Klein and Dalla-Favera 2008).

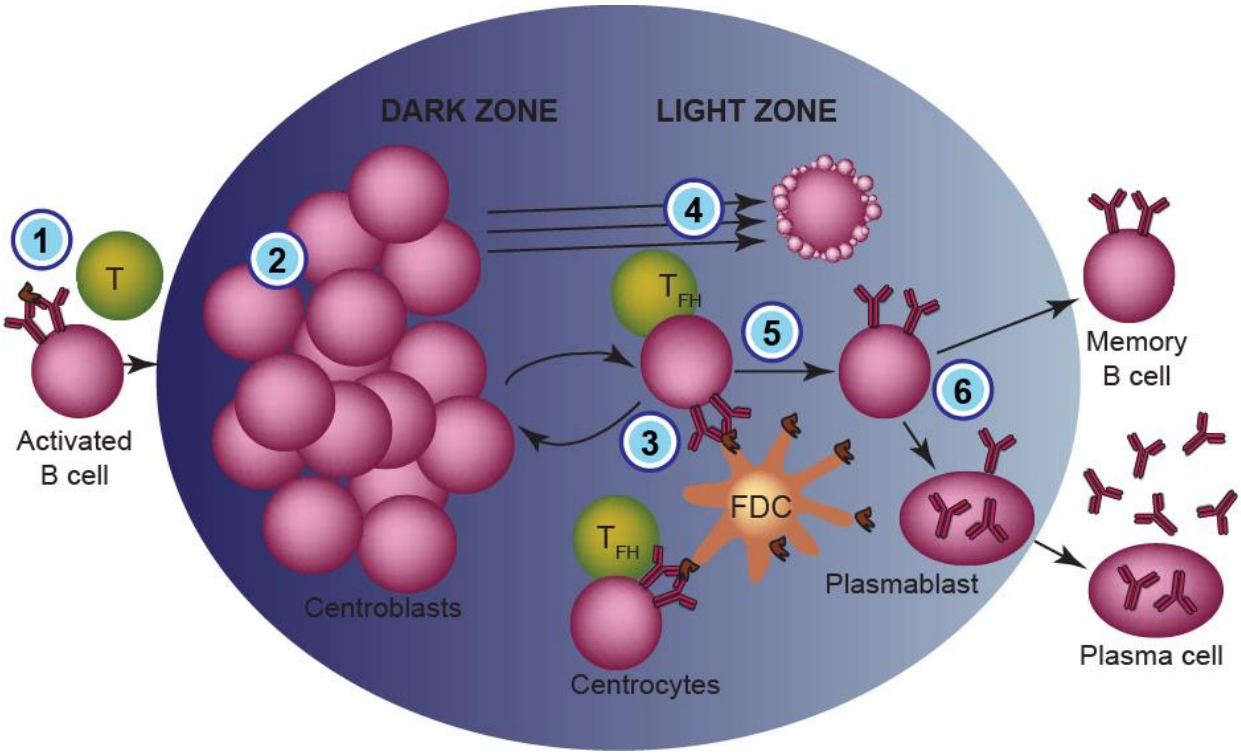



*Figure 1-1 Cartoon of the germinal center reaction.*
*1. An activated mature B cell enters a B cell follicle in lymphoid tissue after receiving T cell help. 2. The B cell undergoes a process called affinity maturation which includes clonal expansion and somatic hypermutation of the BCR in an attempt to optimize antigen specificity. These proliferating cells are located in a region known as the dark zone. 3. Resting B cells with mutated BCRs move to the light zone seeking T cell help. B cells may move between the dark and light zones multiple times until successful BCRs are established or apoptosis occurs. 4. B cells with deleterious mutations that are unable to receive T cell help undergo apoptosis. This is the majority of GC B cells. 5. Successful B cells with enhanced antigen specificity can receive T cell help and are positively selected for survival. 6. Some of these B cells undergo class switch and then differentiate to become a memory B cells or a plasma cells.*
*Wogsland and Huse Adapted from (Kuppers 2005)*

B cells are one of the few cell types that have permission to make DNA changes which makes it more likely that unwanted changes will go unchecked. Both B and T cells undergo V(D)J gene rearrangements (Murphy, Travers et al. 2008). Only B cells make guided purposeful point mutations during SHM and change out the adapter portion of the BCR during isotype class switching (Cerutti 2008). The BCR and its isotype provide information about the developmental path of the B cell. B cells that receive T cell help are more likely to class switch and to have more extensive point mutations due to SHM (Berkowska, Driessen et al. 2011). A minority of GC B cells will successfully mutate a BCR that binds antigen with higher affinity and be selected for survival.

The process of selecting high-affinity B cells is dependent upon the ability to present antigen to T follicular helper cells (Tfh) cells (Victora, Schwickert et al. 2010). Antigens are presented on follicular dendritic cells (FDCs) and B cells compete with each other for binding antigen (Gitlin, Shulman et al. 2014). The B cells with high affinity BCRs will pick up the most antigen from the FDCs, process it and present it on MHC-II molecules to Tfh cells. Tfh cells with a cognate T cell receptor (TCR) will bind the MHC-II-peptide complex and provide survival and proliferation signals to the B cells including CD40L/CD40 stimulation (Crotty 2015). B cells that survive, differentiate into either memory B cells or antibody secreting plasmablasts/plasma cells (Shlomchik and Weisel 2012). Both differentiation paths play an important role in adaptive immunity. Many mature plasma cells eventually migrate to the bone marrow and continue to secrete antibodies, potentially for decades (Medina, Segundo et al. 2002, Klein and Dalla-Favera 2008, Jourdan, Caraux et al. 2009). These circulating antibodies become a first line of defense against known pathogens upon subsequent exposure. When subsequent

exposure occurs, the memory B cells will become activated and then undergo a new germinal center reaction that will lead to increased and improved antibody production.

**Mature B cell phenotypes**

B cells that circulate throughout the body and B cell follicles in the secondary lymphoid organs and undergo GC reactions are known as follicular B cells or, B2 B cells (Montecino-Rodriguez and Dorshkind 2012). These B cells are the focus of this dissertation. With mass cytometry phenotyping, I have focused on four mature B cell populations: naïve, memory, GC B cells, and plasmablasts. The cells can be grouped into populations by a series of biaxial gates or by drawing gates directly on a viSNE map. Most of the research presented in Chapters 2 and 3 focuses on naïve, memory, and GC B cell subsets and their characteristics. In Chapter 2, I compare these B cell subsets in healthy tonsil to those in follicular lymphoma tumors. In Chapter 3, the comparison is in the signaling responses (phospho-specific antibodies) to ROS between the B cell subsets.

Naïve B cells have the phenotype CD19+ CD20+ IgD+ CD38- CD27-. The naïve B cells have not undergone class switch recombination (CSR) or somatic hypermutation (SHM). Naïve B cells can express varying levels of IgM. They do not express IgG, IgA, or IgE (Jackson, Wilson et al. 2008). GC B cells have the phenotype CD19+ CD20hi IgD- CD38+ CD44low/-. CD38 is an indicator of GC status in human B cells (Wagner, Hanna et al. 2000, Jackson, Wilson et al. 2008). GC B cells tend to have low surface immunoglobulin (Ig) since they undergo CSR and SHM during a GC reaction. They tend to express low levels of the chemokine receptors CXCR4 and CXCR5 (Allen, Ansel et al. 2004). Memory B cells have the phenotype CD19+ CD20+ IgD- CD38- CD27+. CD27

marks memory B cells in humans. Memory B cells are the result of clonal expansion after antigen encounter in T cell dependent GC reaction or T-independent reactions(Jackson, Wilson et al. 2008). Gating strategies for these B cell populations are located in the figures in Chapter 2. Traditional biaxial gating and on-viSNE gating strategies are shown.

**Dysregulation of B cells in follicular lymphoma**

Follicular lymphoma is the most common indolent non-Hodgkin lymphoma (McLaughlin 2002, Zinzani 2005). Patients survive for decades but relapses are common and transformation to more aggressive lymphoma is always a risk (Dave 2006) Follicular lymphoma is thought to originate from the germinal center reaction (Dave, Wright et al. 2004). Even with the use of the life-extending anti-CD20 drug Rituximab, FL has remained largely incurable and maintains a risk of transformation to the more aggressive diffuse large B cell lymphoma (DLBCL). Follicular lymphoma malignant cells form follicles reminiscent of GCs (Dave, Wright et al. 2004). The malignant cells express GC B cell signature genes although their protein profile may vary from patient to patient with the loss of one or more GC B cell signature protein (Shaffer, Rosenwald et al. 2002). FL is genetically characterized by a genomic translocation that puts BCL2 under the control of the heavy chain promoter. BCL2 represses apoptosis and since the heavy chain is constitutively expressed in B cells, this means that BCL2 is as well (Ngan , Chen-Levy et al. 1988). When BCL2 is expressed, it blocks BAX and BAK from permeabilizing the mitochondrial membrane and releasing ROS and cytochrome C (Hardwick and Soane 2013). Since GC B cells are already undergoing cycles of proliferation, the block on apoptosis can quickly lead to an overexpansion of B cells. I hypothesized that the phenotyping of FL tumors would show the malignant cells to be phenotypically similar to

GC B cells observed in tonsil. This was however not the case. The malignant cells, had high levels of BCR expression, low CD40, low CD38, and low CXCR4 which contrasted with GC B cell phenotypes (Figure 2-10B). The malignant cells reflected GC B cells by being low to negative for CD44 and expressing some CXCR5. These similarities were not enough to map the cells together using the dimensionality reduction tool viSNE.

In addition to malignant FL cells not resembling GC B cells, there was a lack of GC B cells in the resected tumors. Instead, the GC population is almost nonexistent and there is a large malignant B cell population. These results are shown in Chapter 2. I discuss what might be causing this lack of GC population and how it could be tested in the conclusion chapter.

HLA-DR was found to be highly variable in the malignant B cells we phenotyped. The variability implies a complex interaction with the immune system that is not well understood. It is known that HLA-DR is sometimes expressed on cancer cells such as melanoma (Brocker, Suter et al. 1984, Johnson, Estrada et al. 2016). The cancer cells in FL are B cells which normally express HLA-DR in a healthy state whereas melanocytes do not do not express HLA-DR in a healthy state. It is not clear why the HLA-DR expression is so variable within the malignant tumor cells. All the tumor samples showed this same HLA-DR variability. A variance that surpassed that of non-malignant B and T cells (Chapter 2, Figure 5).

This work generated new insights into follicular lymphoma and new questions. What is causing the diminished GC population? What is the role of HLA-DR and the variability of expression among the malignant cells? What else can be determined if we have patient outcome data? I attempt to answer these questions and others in my final conclusions chapter.

## Tools for the study of B cells

### Human B cell research

While mice have long served as a research model for studying the adaptive immune system and B cells (Eibel, Kraus et al. 2014), we have the ability to study human B cells from healthy and diseased patients. Human B cells can be studied using cell lines or primary human tissue. Cell lines are useful for testing experimental design but primary cells are an invaluable tool for understanding human B cell biology. Cell lines have skewed survival signaling that make them immortal and therefore shift their phenotypes and signaling away from that of primary cells (Irish, Czerwinski et al. 2006, Kaur and Dufour 2012).

Human B cell lines and primary B cells have been cultured for decades (Friend, Marovitz et al. 1978, Jacques and Françoise 1991, Karpova, Schoumans et al. 2005). The discovery that CD40 stimulation, initially via anti-CD40 and later with rCD40L, could keep primary B cells alive was important for primary B cell research (Jacques and Françoise 1991). CD40 stimulation enables *ex vivo* culture to study B cell replication, class switch (Tangye, Ferguson et al. 2002), signaling, and antibody production.

#### Peripheral blood

Human peripheral blood is a renewable resource of primary B cells for research. B cells collected from peripheral blood are robust, capable of undergoing freeze/thaw with minimal effect, and can be grown in culture (Jacques and Françoise 1991, Irish, Czerwinski et al. 2006). The ability to keep B cells alive in culture has made it possible to

learn much from human B cells about class switch, differentiation, apoptosis, etc. B cells can be grow in suspension making them easy to work with (Huse, Bakkebo et al. 2011).

Tonsil

Tonsil has been used for decades as a model to study mature human B cells and the GC reaction (Gadol, Peacock et al. 1988, Feuillard, Taylor et al. 1995, Kremmidiotis and Zola 1995, Jackson, Wilson et al. 2008, Perez, Billordo et al. 2014). Pitfalls of tonsil too, not the exact location as tumors, always activated, possibly inflamed.

**Flow cytometry**

Fluorescence flow cytometry

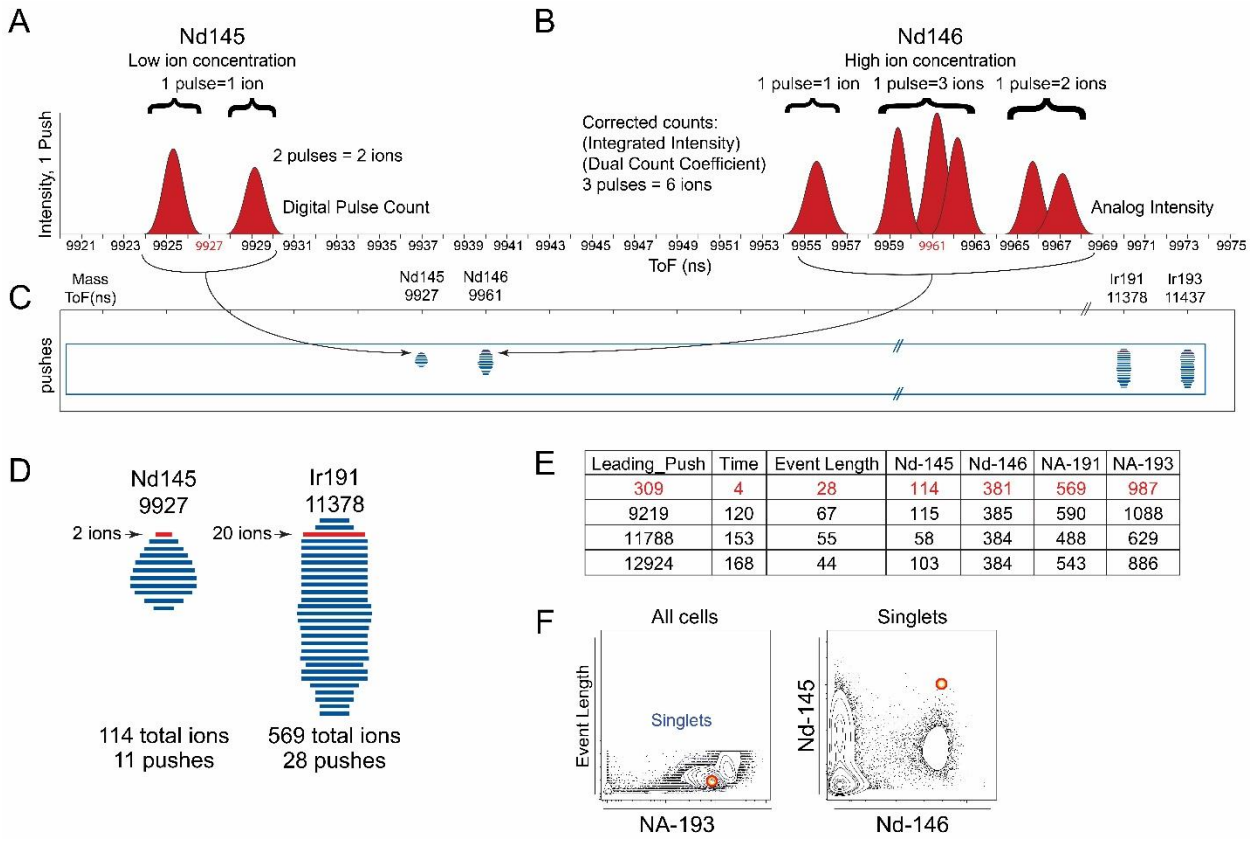
Fluorescence flow cytometry is a single cell analysis tool that uses microfluidics and lasers to collect single cell data. Fluorescence flow cytometry has continually advanced by adding more lasers, detection channels, new fluorescent tags, and improved software (Preffer and Dombkowski 2009). While multiple markers can be measured at once there is still the concern of compensation due to overlapping emissions spectra (Herzenberg, Parks et al. 2002). This overlap limits the number of tags that can be used simultaneously. Due to the nature of fluorescence flow cytometry, values are relative and vary based on voltage settings which limits signal quantitation (Givan 2001). This makes fluorescence flow cytometry less than ideal for quantitative flow cytometry. Fluorescence marker/channel intensities vary depending on the voltages for each collection channel. A negative value can look positive if the voltages are too high. This can cause confusion and it makes quantitative measurements and advanced analysis difficult (Givan 2001).

Mass cytometry

The introduction of mass cytometry with the CyTOF mass cytometer has provided a single cell platform that eliminates the need for compensation and has the capacity to measure more markers simultaneously and quantitatively (Bandura, Baranov et al. 2009, Ornatsky, Bandura et al. 2010). With mass cytometry, zero is zero, it is not relative. The ions detected are the ions reported, there are no voltage settings to change that shift signal levels. This, combined with the ability to build 30+ marker phenotype staining panels, makes it ideal for a systems biology approach at the single cell protein level.

I have contributed tools (Appendix E) and approaches (Chapter 2 and Chapter 4) that enable unbiased in-depth phenotyping of human B cells with mass cytometry. I have applied these tools and approaches in settings of health and disease to uncover new biology and a deeper understanding of existing biology.

Mass cytometry uses the same antibodies as fluorescence flow cytometry but a different chemistry that conjugates a polymer that can chelate metal isotypes used as the detection molecules (Bandura, Baranov et al. 2009, Majonis, Herrera et al. 2010). The metal reporter ions become bound to the cells via antibodies during the staining process. Each antibody gets a different purified metal isotope. The reporter isotopes remain with the cell during ionization and atomization. The reporter ion cloud eventually reaches a time of flight detector and the ion pulses are recorded in an integrated mass data (IMD) file. The path of an ion is presented as a cartooned schematic in Figure 1-2. The CyTOF software parses the IMD file to generate flow cytometry standard (FCS) and text files with cell level information similar to fluorescence flow data although without the forward and side scatter. See Figure 2-6 for mass cytometry gating strategy.

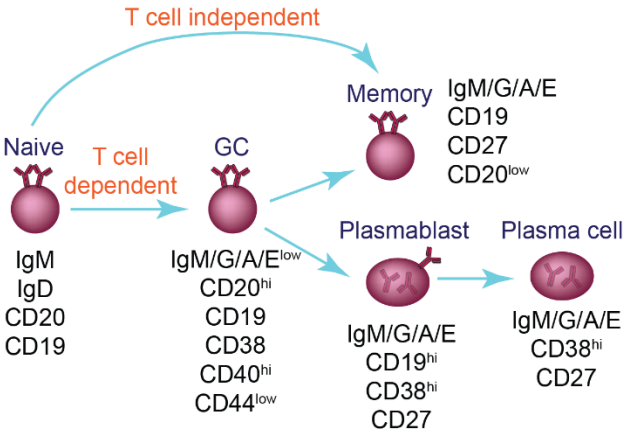



| Leading_Push | Time | Event Length | Nd-145 | Nd-146 | NA-191 | NA-193 |
|---|---|---|---|---|---|---|
| 309 | 4 | 28 | 114 | 381 | 569 | 987 |
| 9219 | 120 | 67 | 115 | 385 | 590 | 1088 |
| 11788 | 153 | 55 | 58 | 384 | 488 | 629 |
| 12924 | 168 | 44 | 103 | 384 | 543 | 886 |

*Figure 1-2 Understanding data generation from isotopes to cells: Depiction of pulses, ions, pushes, and events in mass cytometry.*
*A,B) Cartoon of one push zoomed in to show only two masses (adapted from Fluidigm/DVS). The high and low ion concentration are used together to form a single Dual Signal scale. The x axis is nanoseconds. A) At low metal concentrations, the pulses can be summed to find the correct number of ions, referred to as the digital puklse count. 1 pulse equals 1 ion. B) At higher metal concentrations, pulses start to overlap and summation will not give an accurate ion count. The use of a calibrartion coeficient is necessary to find the accurate number of ions. When the claibration coeficient is multiplies by the anaolg/integrated intensity, the correct number of ions is determined. C) Cartoon of the pushes/time of flight collection window zoomed in to a single event and showing only 4 metals including Ir intercalator. D) Cartoon of the pushes for two metals that are integrated for a single event. E) Simplified view of what the pre-FCS data looks like. Parameters are labeled at the top of each column and each row after the first is an event. F) Example biaxial plots from an FCS file showing where the cell tracked in red is located.*

As an example, the identification of 4 mature B cell subsets and major BCR heavy chain isotypes uses eleven markers/channels (CD19, CD20, IgD, IgM, IgG, IgA, CD27, CD38, CD3-, and CD45, CD44). Eleven markers is on the high end of panel design for fluorescence cytometry whereas for mass cytometry that is less than a third of the panel space. The other channels can then be used to characterize the B cell subsets and

identify and characterize other immune populations (Figure 1-3). See Table 1-1 for an example B cell panel.




| B cells | B cell signaling | B cell subsets | Other Immune Cells | Adhesion & Trafficking |
|---|---|---|---|---|
| IgLambda | CD22 | CD5 | CD3 | CD24 |
| IgKappa | CD40 | CD10 | CD4 | CD49F |
| CD45 | CD23 | CD43 | CD5 | CD62L |
| HLA-DR | CD79B | | CD8 | CXCR4 |
| CD72 | IL4Ra | | CD33 | CXCR5 |
| CD23 | CD32 | | CD36 | |
| | | | CD56 | |
| | | | CD86 | |

35 Marker CyTOF Panel

*Figure 1-3 The value of mass cytometry – getting key markers into one tube*
*Mass cytometry enables the identification and characterization of multiple cell types in a single staining panel. Five mature B cell types can be identified and then further characterized simultaneously. The panel allows for the identification of T cell, NK cell, and Monocytes as well. The trafficking and adhesion markers and other markers such as CD38, CD27, and HLA-DR are helpful in characterizing these non B cell immune subsets.*

The high dimensionality and quantitative nature of mass cytometry lends itself to the use of dimensionality reduction visualization tools. These tools incorporate all of the parameters collected about the cellular events and allow the data to be present in a 2-dimensional (2D) or 3D format that is comprehensible by the human eye. These tools include viSNE, SPADE, PCA, and CITRUS (Qiu, Simonds et al. 2011, Newell, Sigal et al. 2012, Amir el, Davis et al. 2013, Bruggner, Bodenmiller et al. 2014, Chester and Maecker 2015, Newell and Cheng 2016). My research has primarily used viSNE and SPADE. viSNE is excellent for preserving the single cell level data while SPADE is useful for clustering the data based on t-SNE features. When I applied viSNE to the B cells from the lymphoma tumors, the malignant cells were clearly distinguishable from non-malignant tumor B cells and healthy tonsillar B cells with minimal human involvement in the data analysis (Figure 2-1). viSNE was able to separate malignant and non-malignant B cells thereby identifying malignant and non-malignant samples in computational

experiments without the need to look at light chain exclusion. This could be put to use as an automated tool for FL identification, especially if other B cell lymphomas map differently in viSNE.

**Example Human B cell mass cytometry CyTOF staining panel**

| # | Isotope | Antigen | Notes | | Clone | Supplier |
|---|---|---|---|---|---|---|
| 1 | 141Pr | CCR6 (CD196) | Chemokine receptor for CCL20, involved in mucosal immunity, on T cells | surface | G034E3 | Biolegend |
| 2 | 142Nd | CD19 | B cells | surface | HIB19 | Fluidigm |
| 3 | 143Nd | CD5 | T cells | surface | UCHT2 | Fluidigm |
| 4 | 144Nd | CCR5 (CD195) | chemokine receptor, on memory T cells, DCs, and macrophages | surface | NP6G4 | Fluidigm |
| 5 | 145Nd | CD4 | T helper cells | surface | RPA-T4 | Fluidigm |
| 6 | 146Nd | IgD | non-class switched B cells | surface | IA6-2 | Fluidigm |
| 7 | 147Sm | CD20 | B cells | surface | 2H7 | Fluidigm |
| 8 | 148Nd | IgA | Class switched B cells | both | Poly | Fluidigm |
| 9 | 149Sm | CD23 | low affinity IgE receptor, found on B cells | surface | EBVCS-5 | Biolegend |
| 10 | 150Nd | CD43 | B1 B cells, T, mono, gran, may play a role in activation | surface | 84-3C1 | Fluidigm |
| 11 | 151Eu | Ig lambda | B cell light chain | surface | MHL-38 | Fluidigm |
| 12 | 152Sm | CD36 | FAT, scavenger receptor, IgM plasmas, B2 B cells after CD40 stim | surface | 5-271 | Fluidigm |
| 13 | 153Eu | CD62L | L-selectin lymphoid homing | surface | DREG-56 | Fluidigm |
| 14 | 154Sm | CD45 | leukocytes | surface | HI30 | Fluidigm |
| 15 | 155Gd | CD27 | memory B cells | surface | L128 | Fluidigm |
| 16 | 156Gd | CD86 | B7-2, involved in T cell interactions | surface | IT2.2 | Fluidigm |
| 17 | 158Gd | CD33 | mono, myeloid, binds sialic acid, ITIM | surface | WM53 | Fluidigm |
| 18 | 159Tb | CD22 | BCR inhibitory receptor | surface | HIB22 | Fluidigm |
| 19 | 160Gd | Ig kappa | B cell light chain | surface | MHK-49 | Fluidigm |
| 20 | 161Dy | CD32 | in house conjugation | surface | FUN-2 | Biolegend |
| 21 | 162Dy | CD79B | Igβ, BCR signaling subunit | surface | CB3-1 | Fluidigm |
| 22 | 163Dy | IgG-APC* | majority of class switch B cells | both | G18-145 | BD |
| 23 | 164Dy | CD49F | α6 integrin subunit, involved in cell adhesion | surface | G0H3 | Fluidigm |
| 24 | 165Ho | CD40 | B cells, co-stimulatory molecule, interacts with CD40L on T cells | surface | 5C3 | Fluidigm |
| 25 | 166Er | CD44 | hyaluronic acid receptor, negative on GC B cells, positive for Teff, mem | surface | BJ18 | Fluidigm |
| 26 | 167Er | CD38 | GC B cells, plasmablasts/plasma cells | surface | HIT2 | Fluidigm |
| 27 | 168Er | CD8 | CTLs | surface | SK1 | Fluidigm |
| 28 | 169Tm | CD24 | adhesion: B cells | surface | ML5 | Fluidigm |
| 29 | 170Er | CD3 | T cells | surface | UCHT1 | Fluidigm |
| 30 | 171Yb | CXCR5 (CD185) | allows access to lymph node and spleen | surface | J252D4 | Biolegend |
| 31 | 172Yb | IgM | Non-class switched B cells | surface | MHM-88 | Fluidigm |
| 32 | 173Yb | IL4Rα | IL-4 cytokine receptor on B cells and T cells | surface | G077F6 | Biolegend |
| 33 | 174Yb | HLA-DR | antigen presenting cells | surface | L243 | Fluidigm |
| 34 | 175Lu | CXCR4 (CD184) | B cells, binds CXCL12 | surface | 12G5 | Fluidigm |
| 35 | 176Yb | CD56 | NCAM, Leu-19, and NKH1, type I transmembrane glycoprotein, NK cells | surface | CMSSB | Fluidigm |
| 36 | Ir191 | - | Natural Iridium intercalator, cell marker for mass cytometry | IC | - | Fluidigm |
| 37 | Ir193 | - | Natural Iridium intercalator, cell marker for mass cytometry | IC | - | Fluidigm |

*IgG-APC antibody used, with an anti-APC secondary from Fluidigm (APC003)
IC – intracellular
BD - Becton, Dickinson Biosciences
(Kondo, Takata et al. 2007, Barmania and Pepper 2013)

*Table 1-1 Example B cell focused mass cytometry panel*

# CHAPTER 2 MASS CYTOMETRY OF FOLLICULAR LYMPHOMA TUMORS

Full title: Mass cytometry of follicular lymphoma tumors reveals intrinsic heterogeneity in proteins including HLA-DR and a deficit in non-malignant plasmablast and germinal center B cell populations (Wogsland, Greenplate et al. 2017)

Cara Ellen Wogsland, Allison Rae Greenplate, Arne Kolstad, June Helen Myklebust, Jonathan Michael Irish, Kanutte Huse

## Abstract

*Background:* Follicular lymphoma (FL) is an indolent non-Hodgkin lymphoma that has a risk of transformation to more aggressive lymphoma. Relatively little is known about the non-malignant B-cell and T-cell subset composition within the tumor microenvironment and whether altered phenotypes are associated with patterns of lymphoma B-cell heterogeneity.

*Methods:* Two mass cytometry (CyTOF) panels were designed to immunophenotype B and T cells in human FL tumors. Populations of malignant B cells, non-malignant B cells, and T cells from each FL tumor were identified and their phenotypes compared to B and T cells from healthy human tonsillar tissue.

*Results:* Diversity in cellular phenotype between tumors was greater for the malignant B cells than for non-malignant B or T cells. The malignant B-cell population bore little phenotypic similarity to any healthy B-cell subset, and unexpectedly clustered closer to naïve B-cell populations than GC B-cell populations. Among the non-malignant B cells within FL tumors, a significant lack of GC and plasmablast B cells was observed relative

to tonsil controls. In contrast, non-malignant T cells in FL tumors were present at levels similar to their cognate tonsillar T-cell subsets.

*Conclusion*: Mass cytometry revealed that diverse HLA-DR expression on FL cells within individual tumors contributed greatly to tumor heterogeneity. Both malignant and non-malignant B cells in the tumor bore little phenotypic resemblance to healthy GC B cells despite the presence of T follicular helper cells in the tumor. The non-malignant GC B cell population was diminished in the tumors however the Tfh cell population was not diminished. These findings suggest that ongoing signaling interactions between malignant B cells and intra-tumor T cells shape the tumor microenvironment.

## Introduction

Follicular lymphoma (FL) has been observed to be very different between patients in terms of phenotype, progression, transformation, and outcome. Here, I examined the differences between the cells within individual tumors and between tumors. This provided several new insights. The malignant B cells are heterogeneous within the tumor. Heterogeneity in FL exists not only between patients but within the malignant cells of a single tumor. Using dimensionality reduction tools and clustering I identified HLA-DR as highly variable within the tumors and the top contributor to malignant cell phenotypic diversity. All eight tumors examined had variable expression of HLA-DR on the malignant cells. Since HLA-DR presents peptides to immune cells, this variability could be due to a complex interaction with the immune system. The malignant cells might be playing a complex balancing act between receiving T cell help (high-HLA-DR) and becoming less immunogenic (low HLA-DR) (Silva, Silva et al. 2013). I discuss HLA-DR in more detail in

Chapter 5. I also found that GC B cells had a diminished presence among the non-malignant B cells in the tumors. This is contrary to that normally robust GC population that is seen in tonsil.

Patient demographics are shown in Table 2-1. This information was not used in the analysis of the data because the sample set is too small to accurately draw conclusions about phenotype based on the patient demographics. The Sample IDs in the table were used in the figures to label individual tumors. There were 2 male patients and 6 female patients. Patients ages ranged from 31 and 49 years, which is a lower than the average FL patient (>60 years) (Freedman 2015). The therapies and responses to therapies are recorded in Table 2-1 along with overall survival (OS) data. All tumor samples were collected before therapy so the therapies listed have no impact on the data shown here. These patients were treated before the Rituximab era so even if there were enough patients to make an outcome analysis, the findings would not be as relevant since Rituximab is now a key part of first therapy (Freedman 2015). The patient disease presentations were scored using the Follicular Lymphoma International Prognostic Index (FLIPI) at the time of diagnosis. FLIPI uses 5 prognostic factors and a total scale of 0-5 to assess FL in a patient: patient age (>60 is +1), Ann Arbor Stage (>II is +1), hemoglobin (<120 g/L is +1), number of nodal areas (>4 is +1), serum lactate dehydrogenase (LDH) (>normal is +1) (Solal-Celigny, Roy et al. 2004). FLIPI scores of 0-1 fall into the good risk category, 2 is intermediate risk, and 3-5 means the prognosis is poor (Solal-Celigny, Roy et al. 2004). The light chain restriction status determined clinically matched what we saw by mass cytometry.

I designed the core B cell mass cytometry marker panel (Chapter 1, Table 1-1) and assisted in adapting the panel for this study (Table 2-2). The markers included were

chosen to identify mature B cell subsets, characterize those B cell subsets, and identify other immune cell types, primarily T cells. Although a T cell panel was run for the detailed

| Sample ID | Sex | Age | Response to first therapy | First therapy | Stage | FLIPI | Censor for OS [1 = Dead] | Overall survival relative to date of diagnosis | Immunophenotyping (diagnosis) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | CD37 | Ig kappa | Ig lambda | LC tumor restrictio | CD3 |
| FL3 | male | 49 | CR | CHOPx6 | | 0 | 0 | 23.0 | 60 | 2 | 64 | lambda | 35 |
| FL4 | female | 31 | no therapy | no therapy | IVA | 3 | 0 | 22.4 | 69 | 47 | 9 | kappa | 28 |
| FL7 | male | 44 | PR | CHOPx4 | IIIA | 2 | 0 | 21.7 | 61 | 13 | 43 | lambda | 40 |
| FL10 | female | 49 | CR | Leukeran | IIIA | 2 | 0 | 21.4 | 66 | 53 | 6 | kappa | 35 |
| FL12 | female | 40 | CR | Leukeran | IIIA | 0 | 1 | 3.7 | 68 | 60 | 5 | kappa | 26 |
| FL5 | female | 42 | PR | CHOPx7 | IVA | 1 | 1 | 8.3 | 80 | 1 | >60 | lambda | 15 |
| FL9 | female | 34 | CR | CHOPx8 | IVA | 2 | 1 | 9.1 | >85 | >90 | 1 | kappa | 15 |
| FL14 | female | 48 | CR | CHOPx8 | IVA | 1 | 1 | 3.5 | 75 | 81 | 1 | kappa | 11 |

*Table 2-1 Patient demographics for FL tumors phenotyped by mass cytometry.*
*All tumor samples were taken pre-therapy. Ann Arbor stage is recorded in roman numerals with higher numerical values indicating more advanced disease. The FLIPI scale is a prognisis scale of FL disease. A higher FLIPI score indicates a worse prognosis (Solal-Celigny, Roy et al. 2004). Overall survival (OS) is marked as 0 for alive and 1 for deceased. The cloumn to the right of Censor for OS indicates the years since diagnosis for the OS measure. CD37 is a tetraspannin that is highly expressed on B cells and B cell lymphomas and is thought to play a role in B cell activation. It has been considered as an immunotherapy target in B cell lymphomas (Grosmaire, Hayden-Ledbetter et al. 2014). CD37 was not measured by mass cytometry in this study. The kappa/lambda percentages were measured by clinical fluorescence flow and the tumor light chain restriction was determined and reported here. CD3 positivity was also measured in the tumor as an indicator of T cell presence. These sample were all FL lymph node tumors resected in Oslo, Norway in the last 2 decades of the 20th century.*

analysis of the T cells, it is important to have at least a few T cell markers in the B cell panel in order to definitively mark the T cells as not B cells. T cells are CD3+ and B cells are CD3-. This marking of the T cells as not B cells, enables confident gating out of T cells and a robust viSNE map if the cells were to be analyzed together. I designed the strategy for and performed the data processing, clean up gating, manual population gating, automated cluster generation with SPADE, and the multiple viSNE runs. The markers chosen as input parameters for the B cell viSNE runs were markers that identified mature B cell populations such as CD38, IgD, CD27, CD44, and CD20 and markers that characterized those populations including CD22, CXCR5, CXCR4, and HLA-DR. I chose not to use markers that were previously used to positively gate the B cells such as CD45 and CD19 since those did not add value to the analysis. CD20 was used for pre-gating

however its expression levels are indicative of B cell subset; GC B cells have higher CD20 levels than other mature B cell subsets. CD45 staining was variable between samples stained on different days so including that could have introduced error into the analysis.

There are several pitfalls to note for this study. First off, as I mentioned above, the samples were not all stained on the same day or with the same staining cocktail. Samples were also not run on the CyTOF on the same day. While these variables could introduce problems into the data interpretation, I have looked extensively at the primary data for all the samples and I am confident that the signals are true enough as to not introduce significant errors into the findings. I discussed BCL2 in the introduction chapter however I am showing no data on BCL2. The reason for this is because the BCL2 staining did not work and showed no positivity. I hypothesize that the staining failure is due to the fact that we used an anti-BCL2 antibody that required a secondary antibody for detection. The anti-BCL2 antibody was labeled with PE, which is a very bright marker by fluorescence flow. The problem with using it for mass cytometry with a secondary anti-PE antibody is that PE is a large molecule and does not leave room for the binding of many PE molecules to an antibody. This in turn means that there are few spots for anti-PE to bind so there is little amplification of the signal.

I designed the viSNE analyses, choosing appropriate markers that would address the questions asked. For example, IgK and IgL were not included in the analyses because I knew that the malignant cells would have either high or negative expression depending on tumor status (Table 2-1).

| Isotope | Antigen | B-cell panel | T-cell panel | Order of stain[2] | Clone | Supplier |
|---|---|---|---|---|---|---|
| Pr141 | CCR6[B,T] | + | + | Surface | G034E3 | Fluidigm |
| Nd142 | CD19 | + | + | Surface | HIB19 | Fluidigm |
| Nd144[3] | IgG | + | | IC | HP6017 | Biolegend |
| Nd144[4] | IgG | + | | IC | Poly goat | Invitrogen |
| Nd144 | CCR5[T] | | + | Surface | NP-6G4 | Fluidigm |
| Nd145 | CD4[T] | | + | Surface | RPA-T4 | Fluidigm |
| Nd146 | IgD[B] | + | | Surface | IA6-2 | Fluidigm |
| Nd146 | CD8[T] | | + | Surface | RPA-T8 | Fluidigm |
| Sm147 | CD20[B] | + | | Surface | 2H7 | Fluidigm |
| Sm148 | IgA | + | | IC | Poly goat | Fluidigm |
| Sm149 | CCR4[B,T] | + | + | Surface | 205410 | Fluidigm |
| Nd150 | CD43[B,T] | + | + | Surface | 84-3C1 | Fluidigm |
| Eu151 | Igλ | + | | Surface | MHL-38 | Fluidigm |
| Eu151[3] | ICOS[T] | | + | Surface | C398.43 | Biolegend |
| Sm152 | CD36[B] | + | | Surface | 5-271 | Fluidigm |
| Sm152 | TCRgd[T] | | + | Surface | 11F2 | Fluidigm |
| Eu153 | CD62L[B] | + | | Surface | DREG-56 | Fluidigm |
| Eu153 | CD45RA[T] | | + | Surface | HI100 | Fluidigm |
| Sm154 | CD45 | + | + | Surface | HI30 | Fluidigm |
| Gd156 | CD86[B] | + | | Surface | IT2.2 | Fluidigm |
| Gd156 | CXCR3[T] | | + | Surface | G025H7 | Fluidigm |
| Gd158 | CD33[B] | + | + | Surface | WM53 | Fluidigm |
| Tb159 | CD22[B] | + | | Surface | HIB22 | Fluidigm |
| Tb159 | CCR7[T] | | + | Surface | G043H7 | Fluidigm |
| Gd160 | Igκ | + | | Surface | MHK-49 | Fluidigm |
| Dy162 | CD79B[B] | + | | Surface | CB3-1 | Fluidigm |
| Dy162 | CD69[T] | | + | Surface | FN50 | Fluidigm |
| Ho165 | CD40[B] | + | | Surface | 5C3 | Fluidigm |
| Er166 | CD44[B,T] | + | + | Surface | BJ18 | Fluidigm |
| Er167 | CD27[B,T] | + | + | Surface | O323 | Fluidigm |
| Er168[3] | CD38[B] | + | | Surface | HIT7 | Biolegend |
| Er168[3] | CTLA-4[T] | | + | Surface | L3D10 | Biolegend |
| Tm169 | CD25[T] | | + | Surface | 2A3 | Fluidigm |
| Er170 | CD3[B] | + | + | Surface | UCHT1 | Fluidigm |
| Yb171 | CXCR5 (CD185)[B,T] | + | + | Surface | 51505 | Fluidigm |
| Yb172 | IgM | + | | IC | MHM-88 | Fluidigm |
| Yb172 | CD57[T] | | + | Surface | HCD57 | Fluidigm |
| Yb174 | HLA-DR[B] | + | | Surface | L243 | Fluidigm |
| Yb174[3] | PD-1[T] | | + | Surface | EH12.2H7 | Biolegend |
| Lu175 | CXCR4 (CD184)[B,T] | + | + | Surface | 12G5 | Fluidigm |
| | light_chain[B] | + | | Computational combination of Igκ and Igλ | | |
| Ir191/193 | - | Natural Iridium intercalator (cell identifier) | | IC | - | Fluidigm |

[1] Some antibodies were excluded from analyses due to minor panel mismatch between samples
[2] Indicates if staining was performed prior to permeabilization (surface) or after permeabilization (intracellular, IC).
[3] In-house conjugation performed in the Irish Lab at Vanderbilt
[4] Detected with secondary anti-FITC-Nd144 (FIT-22; Fluidigm)
[T] Denotes the 20 features used in viSNE analysis of samples stained with T-cell panel (Figures 1,2,5).
[B] Denotes the 20 features used in viSNE analysis of samples stained with B-cell panel (Figures 1,3,4,5).

*Table 2-2 List of mass cytometry antibodies used in the FL study.*

What I wanted to do was see if I could separate out the malignant cells on the viSNE map without using light chain restriction. My computational experiment was a

success and viSNE did separate out the malignant cells (Figure 2-1). It is crucial to have markers that are positive and negative for multiple populations. This helps spread the cell populations apart in phenotypic tSNE space. The difference in overall phenotype is larger if a population is positive for one marker and negative for another and another population shows the opposite expression patterns. This is observed most readily in viSNE maps that have altered marker use. I interpreted the results for the viSNE runs and manual gating. Moreover, I designed and created the figures for the body of the text.

FL is a B-cell malignancy and the second most common non-Hodgkin lymphoma. Although patient overall survival is now measured in decades, FL is considered incurable and multiple relapses are common (Dave 2006). FL is named for its follicle-like appearance and is thought to arise from mature germinal center (GC) B cells (Dave, Wright et al. 2004). Malignant FL B cells are characterized by light-chain restriction and the t(14;18) translocation that leads to overexpression of B cell lymphoma 2 (BCL2) (Ngan , Chen-Levy et al. 1988). As the disease evolves, more mutations are acquired, leading to genetic heterogeneity in the tumor (Eide, Liestøl et al. 2010, Green, Gentles et al. 2013). The high rate of relapse and the identification of negative prognostic cells present at high levels within some FL tumors (Irish, Myklebust et al. 2010) has highlighted the need to understand the biology of FL tumor cell heterogeneity and interaction with the microenvironment.

Much of the past FL research has focused on genetics (Dave, Wright et al. 2004, Dave 2006, Green, Gentles et al. 2013) and flow cytometric analysis with a limited number of markers measured per single cell (Mantei 2009, Irish, Myklebust et al. 2010). The

introduction of mass cytometry and the associated field of computational high-dimensional analysis have paved the way for in-depth analysis of single cell phenotype (Greenplate, Johnson et al. 2016, Saeys, Gassen et al. 2016). Here, malignant and non-malignant tumor cells were characterized at the single cell protein level and compared to the lymphocytes from healthy tonsils. Non-malignant healthy donor tonsil is an accessible lymphoid tissue that contains mature B cells that are pre-, during-, and post-germinal center reaction (Allen, Okada et al. 2007) and that has been studied previously by mass cytometry (Sen, Mukherjee et al. 2014, Polikowsky, Wogsland et al. 2015, Wong, Chen et al. 2015). While differences exist in the phenotype and organization of the cells in human secondary lymphoid organs including spleen, lymph nodes, and tonsils (Brachtel, Washiyama et al. 1996, Vidal-Rubio, Sanchez-Carril et al. 2001, Allen, Okada et al. 2007), both tonsils and reactive lymph nodes are valuable comparison points for studies of tumor-involved lymph nodes and FL tumors.

In this exploratory study, mass cytometry was used to study malignant B cells and non-malignant B and T cells in the tumor microenvironment with a focus on intra- and inter-tumor heterogeneity, and on changes in the composition of immune cells within tumors. Mass cytometry was selected due to the ability to characterize >35 features of individual cells and to reveal unexpected malignant and non-malignant cell subsets with unusual phenotypes (Becher, Schlitzer et al. 2014, Bendall, Davis et al. 2014, Irish and Doxie 2014, Greenplate, Johnson et al. 2016, Leelatian, Doxie et al. 2016).

## Materials and methods

### Tissue sample collection

FL samples were selected from the biobank at The Norwegian Radium Hospital, Oslo University Hospital. Tonsils were obtained from patients undergoing tonsillectomy at Agroklinikken, Asker, Norway. All samples were obtained with patient consent in accordance with the Declaration of Helsinki. The study was approved by the regional committee for research ethics, Oslo Norway, and the Vanderbilt institutional review board (IRB).

### Mass cytometry

Cryopreserved single cells were thawed, pelleted, and rested in RPMI-1640 with 10% FBS for 45 minutes before staining. Viability was determined by trypan blue staining and manual cell counting. Samples with less than 50% viability were excluded from the study. At least one million live cells per sample were stained with surface antibodies (listed in Table 2-2) for 30 minutes, then washed and fixed in 1.6% PFA for 5 minutes followed by cell membrane permeabilization with >90% cold methanol. Cells were stored in methanol at -80°C for up to two weeks. Samples were washed twice and stained with intracellular antibodies (listed in Table 2-2) for 30 minutes. Cells were then incubated in iridium cell tracker (Fluidigm) at the recommended concentration for 20 minutes. Stained and nucleic acid marked (iridium intercalated) samples were collected on a CyTOF 1 (Fluidigm) at the Vanderbilt flow cytometry core.

**Data analysis**

The cloud-based flow cytometry platform, Cytobank (Kotecha, Krutzik et al. 2010), was used for file storage and data analysis including biaxial data display and gating, viSNE (Amir el, Davis et al. 2013), SPADE (Qiu, Simonds et al. 2011), histograms, heat maps, and statistics consistent with standard computational analysis workflow (Diggins, Ferrell et al. 2015). The statistical language R with RStudio was used for hierarchical clustering (RStudio Team (2015) URL http://www.rstudio.com/, R Core Team (2013) URL https://www.R-project.org).

A computational light-chain channel was added to the B-cell panel files post data collection. The channel was named “light_chain” and was created by selecting the value of either IgL or IgK, whichever was higher, for each cell event. This new light chain channel reported the light chain level for each B cell regardless of isotype and was used in computational analysis of phenotype.

The dimensionality reduction similarity mapping tool viSNE (Amir el, Davis et al. 2013) was used to create two-dimensional t-SNE visualizations of multidimensional cellular phenotypes (Diggins, Ferrell et al. 2015, Newell and Cheng 2016, Saeys, Gassen et al. 2016). Cells that had similar phenotypes were placed close together on viSNE maps created based on analysis of 20 cellular measurements. T cells from all tonsil and tumor samples were analyzed together by viSNE using 20 markers from the T-cell panel (Figure 2-1 and in depth in Figure 2-5). A total of three B cell viSNE analyses were performed each using the same 20 B-cell panel markers. All B cells were initially analyzed together (Figure 2-1), and that map was used to gate for malignant and non-malignant B cells. In the two subsequent B cell viSNE analyses, the non-malignant B cells from FL samples and all tonsillar B cells were analyzed together in a “non-malignant B cell” viSNE (Figure

2-9), and the malignant B cells were analyzed together in a “malignant B cell” viSNE (Figure 2-10). SPADE was applied to the t-SNE axes created by viSNE (Figure 2-11) to cluster cells into cell subsets (Becher, Schlitzer et al. 2014, Shekhar, Brodin et al. 2014, Diggins, Ferrell et al. 2015). Variance in marker expression between clusters was calculated for nodes containing at least 100 cells using transformed median values (arcsinh scale, cofactor of 15, as in (Irish, Myklebust et al. 2010)).

## Results

### FL malignant B cells were phenotypically distinct from tonsillar B cells.

Eight FL lymph node tumor samples were phenotyped by mass cytometry. Each sample was stained with two different panels, one focused on T-cell makers and one on B-cell markers (Table 2-2). Three non-malignant healthy donor tonsils were stained with the same panels to serve as healthy controls (Figure 2-1). After initial gating to identify singlet T cells or B cells (Figure 2-2A), viSNE analyses were performed. CD45+ CD3+ CD19- cells from the 11 T-cell panel files were analyzed together in the “T cells” viSNE, (Figure 2-1A; tSNEs_T cells). The panel was designed to identify and characterize T-cell subsets and included for viSNE analysis the following 20 markers: CCR6, CCR5, CD4, CD8, CCR4, CD43, ICOS, TCRγδ, CD45RA, CXCR3, CCR7, CD69, CD44, CD27, CTLA4, CD25, CXCR5, CD57, PD1, CXCR4. These markers were not used to gate for T cells. CD45+ CD19+ and/or CD20+ CD3- B cells were analyzed together in a similar fashion for the “all B cells” viSNE analysis (Figure 2-1A; tSNEs_all B cells). The panel was designed to identify and characterize B cell subsets, and included for viSNE analysis

the following 20 markers: CCR6, IgD, CD20, CCR4, CD43, CD36, CD62L, CD86, CD33, CD22, CD79B, CD40, CD44, CD27, CD38, CD3, CXCR5, HLA-DR, CXCR4, light_chain.

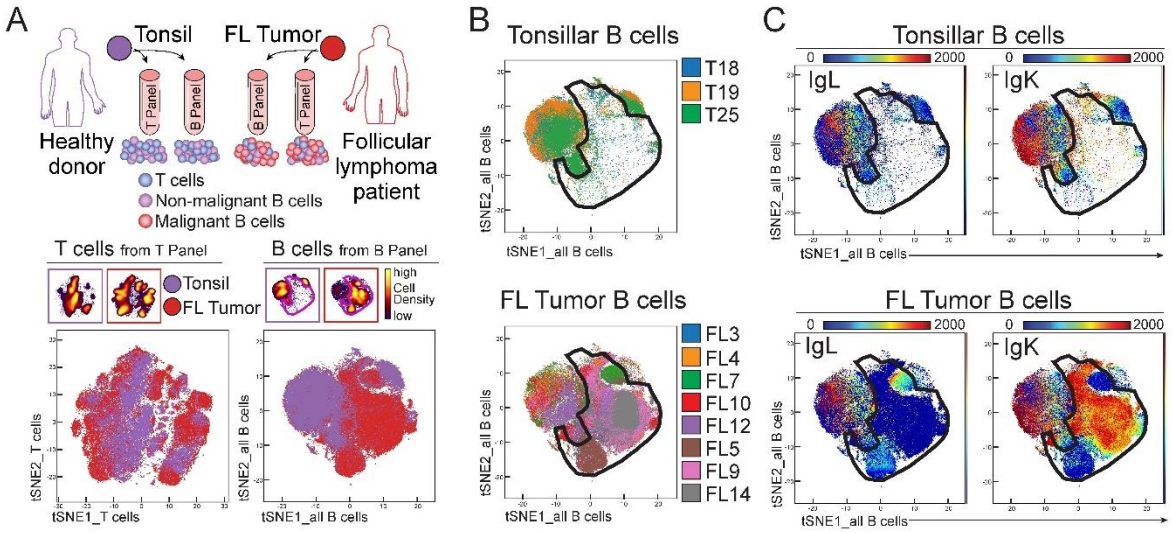



*Figure 2-1 High dimensional phenotyping of lymphoma B cells, non-malignant B cells, and tumor-infiltrating T cells. A) FL tumor samples and tonsils from healthy donors were studied by mass cytometry. Each sample was split in two and analyzed by two antibody panels, one focused on B cells and the other on T cells, to identify three main populations of cells in FL tumor samples (malignant B cells, non-malignant B cells, and T cells) as shown in the cartoon. CD3+ T cells (bottom left) and CD19+ B cells (bottom right) from tonsils and FL samples were analyzed together in a viSNE based on 20 markers from the T-cell panel and 20 markers from the B-cell panel, respectively. B-C) Malignant B cells were gated in an area of the viSNE map where tonsillar B cells were mostly absent (B) and the cells were light-chain restricted (C). Top row shows all tonsil samples combined, bottom row shows all FL samples combined.*

The tonsillar B cells mapped almost entirely to the periphery of the all B cells viSNE map, leaving a large space in the center of the map where the majority of B cells from the eight FL tumors fell (Figure 2-1, Figure 2-3). The central area on the viSNE map where the tonsillar B cells were primarily absent was gated as malignant B cells (Figure 2-1B,C). The viSNE map was made without the use of IgLambda (IgL) and IgKappa (IgK) to guide the separation of malignant cells; instead, a computational light-chain channel was used to indicate positivity of total light chain. IgL and IgK expression were used afterward to confirm the isotype exclusion of the malignant B cells for each tumor sample (Figure 2-4). The area outside the malignant gate, the non-malignant B-cell area, for tonsil and FL tumors showed a mixture of light-chain isotypes, as expected (Figure 2-1C). Two IgL

lymphomas (FL5 and FL7) were clearly identifiable on the viSNE map as IgK- and IgL+ (bottom, Figure 2-1C). The third lambda positive sample, FL3, does not follow this pattern so I do not think this is a lambda specific feature, merely a tumor specific occurrence. The intensities of IgL and IgK showed that the malignant area contained light-chain-restricted cells whereas the non-malignant area contained a mix of IgK and IgL cells, similar to the tonsillar B cells (Figure 2-1C). A key finding is the fact that light-chain-restricted cells fell into the malignant area regardless of light-chain isotype. The FL samples contained cells that mapped in the non-malignant area, but to varying degrees. These cells were interpreted as non-malignant B cells present within the tumors.

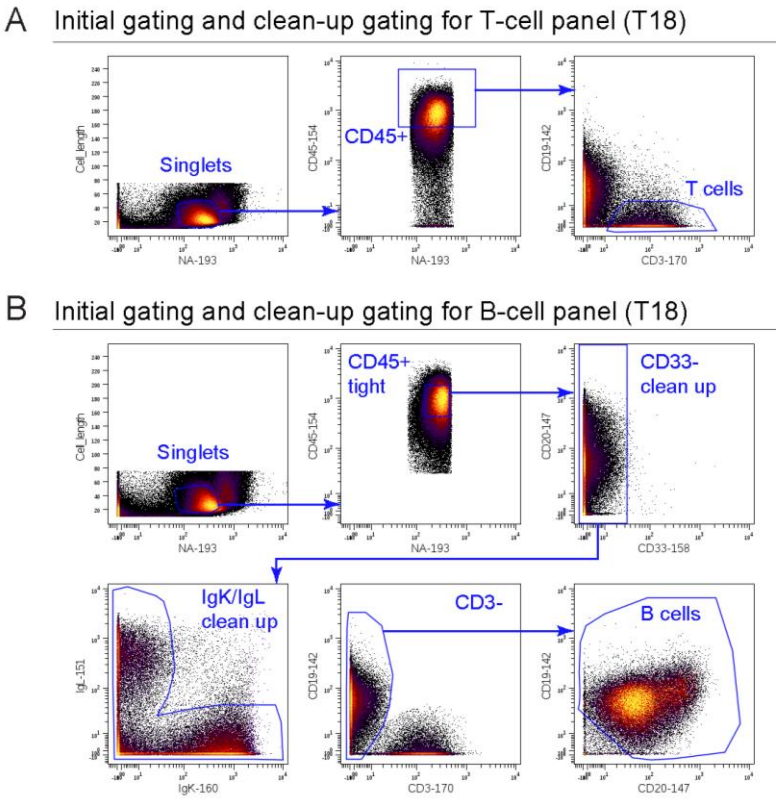



*Figure 2-2 Gating strategy used for initial gating before viSNE or traditional biaxial gating comparison. A) T cell pre-gating. B) B cell pre-gating and clean-up gating*

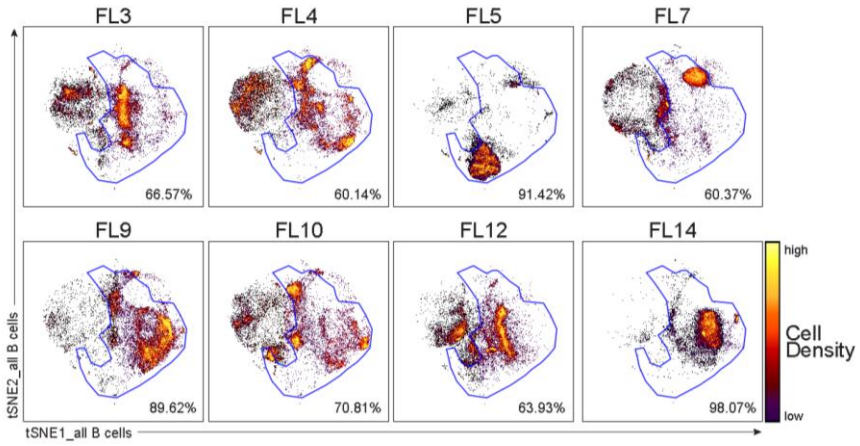



*Figure 2-3 viSNE plots of all B cells for each tumor.*
*The percent shown is the percentage of cells inside the malignant gate. FL5, FL9, and FL14 were excluded from the nonmalignant B-cell analysis.*

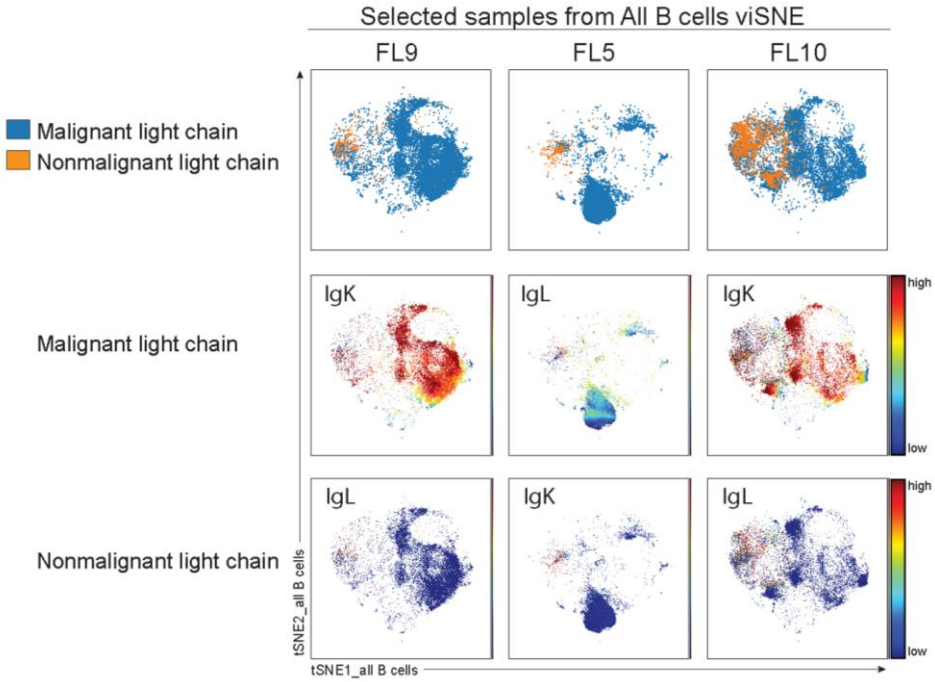



*Figure 2-4 Backgating of light-chain biaxial gating*
*(viSNE view of traditional gating of malignant/nonmalignant cells by light-chain exclusion)*

### T-cell distribution trended toward an activated state

Key T-cell markers and the T cells viSNE map (Figure 2-5A) were used to subset the T-cell panel samples into five populations (Figure 2-5B). For comparison, traditional

biaxial gating was performed in parallel (Figure 2-6A,B). The eight FL tumors and the three tonsils had a high degree of overlap in the T cell viSNE map (Figure 2-1D-E), suggesting a similar phenotype of tumor-associated T cells and healthy tonsillar T cells. There were no significant differences between T-cell subset distributions in FL tumors and tonsils (Figure 2-5C, Figure 2-7A). T cells from most samples were distributed throughout the viSNE map with the exception of FL14 (Figure 2-5D-E). FL14 had a restricted distribution and contained predominantly CD8+ T cells.

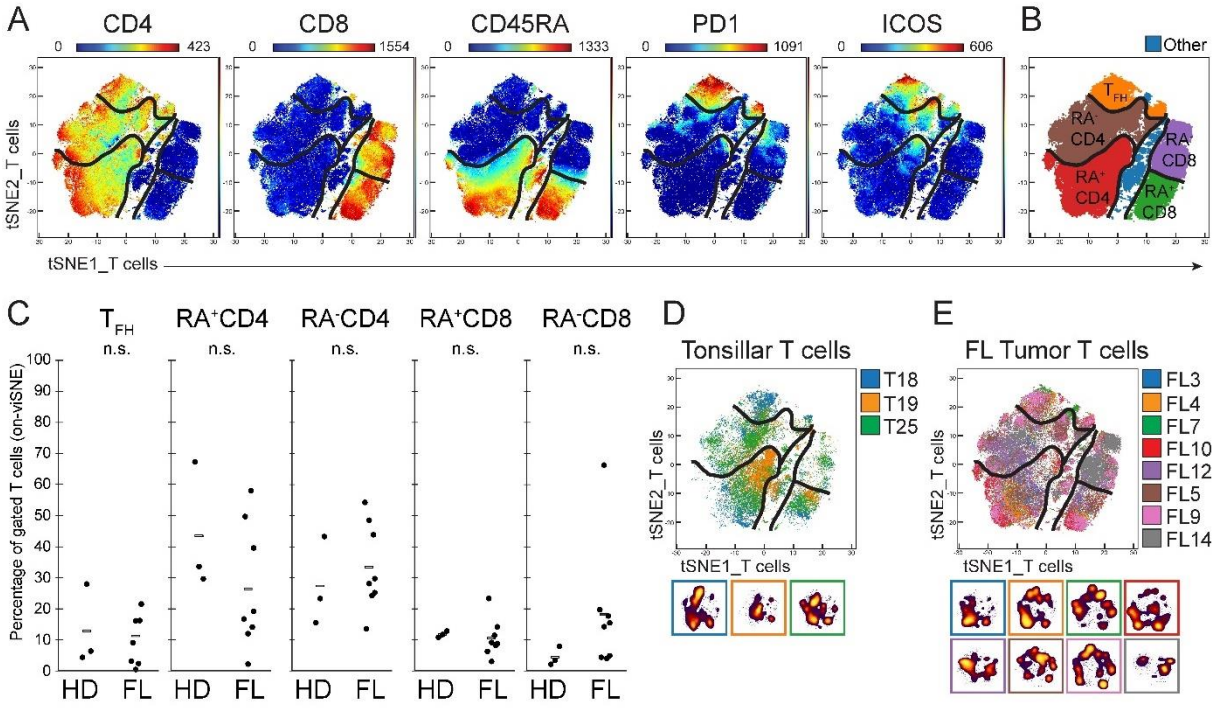



*Figure 2-5 The abundance and phenotypes of tumor infiltrating T-cell subsets are comparable to those of healthy tonsillar T cells.*
*A) Expression of measured proteins is shown as a heat plot on t-SNE axes (tSNEs_T cells). Markers shown were used to make the expert-gated populations shown in (B). Populations are denoted by black lines. B) T-cell populations were gated in viSNE based on markers shown in A. C) Distribution of T-cell subsets across samples. Figure shows all healthy donor (HD) tonsil samples (D) and all FL samples (E) overlaid with individual cell-density plots beneath.*

Although there were no significant differences between tonsillar T cell and FL tumor T cell-subset proportions, FL tumors tended to have fewer CD45RA+ or more CD45RA- T cells, suggesting fewer naïve cells were present (Figure 2-5C). It is not uncommon for T

cell populations to shift from naïve (CD45RA+) to memory (CD45RA-) in disease (Ezawa, Yamamura et al. 1997, Larbi and Fulop 2014). FL tumors also displayed a skewing towards CD8+ T cells (Figure 2-8). I suspect that the Tfh cells are interacting with the malignant B cells.

**Diminished GC and plasmablast populations among non-malignant B cells in FL tumors**

Non-malignant tumor-infiltrating B cells were next investigated by analyzing them in a new viSNE together with the tonsillar B cells (tSNEs_non-mal. B) using the same markers as the initial B cell viSNE (Figure 2-9). In Figure 2-1, B cells positioned outside the malignant gate were considered non-malignant due to their overlapping with tonsillar B cells in the viSNE map and the mixture of IgK and IgL expressing cells. Three of the FL tumors contained less than 10% non-malignant B cells and were excluded from this analysis (Figure 2-3). Key B-cell markers were used to draw expert gates on the non-mal. B viSNE map to divide the B cells into four mature B-cell subsets; naïve, memory, germinal center (GC), and plasmablasts (PB) (Figure 2-9A-B, Figure 2-6C,D). The FL tumor non-malignant B cells and the tonsillar B cells showed a high degree of overlap in the B cell viSNE map (Figure 2-9D-E), suggesting a similar phenotype of non-malignant B cells and healthy tonsillar B cells. There were significant differences in B cell-subset distributions between tonsil and FL tumor for GC B cells and plasmablasts (Figure 2-9C, Figure 2-7B) with average percentages of 15% and 2.8% for GC B cells and 2% and 0.2% for plasmablasts, respectively.

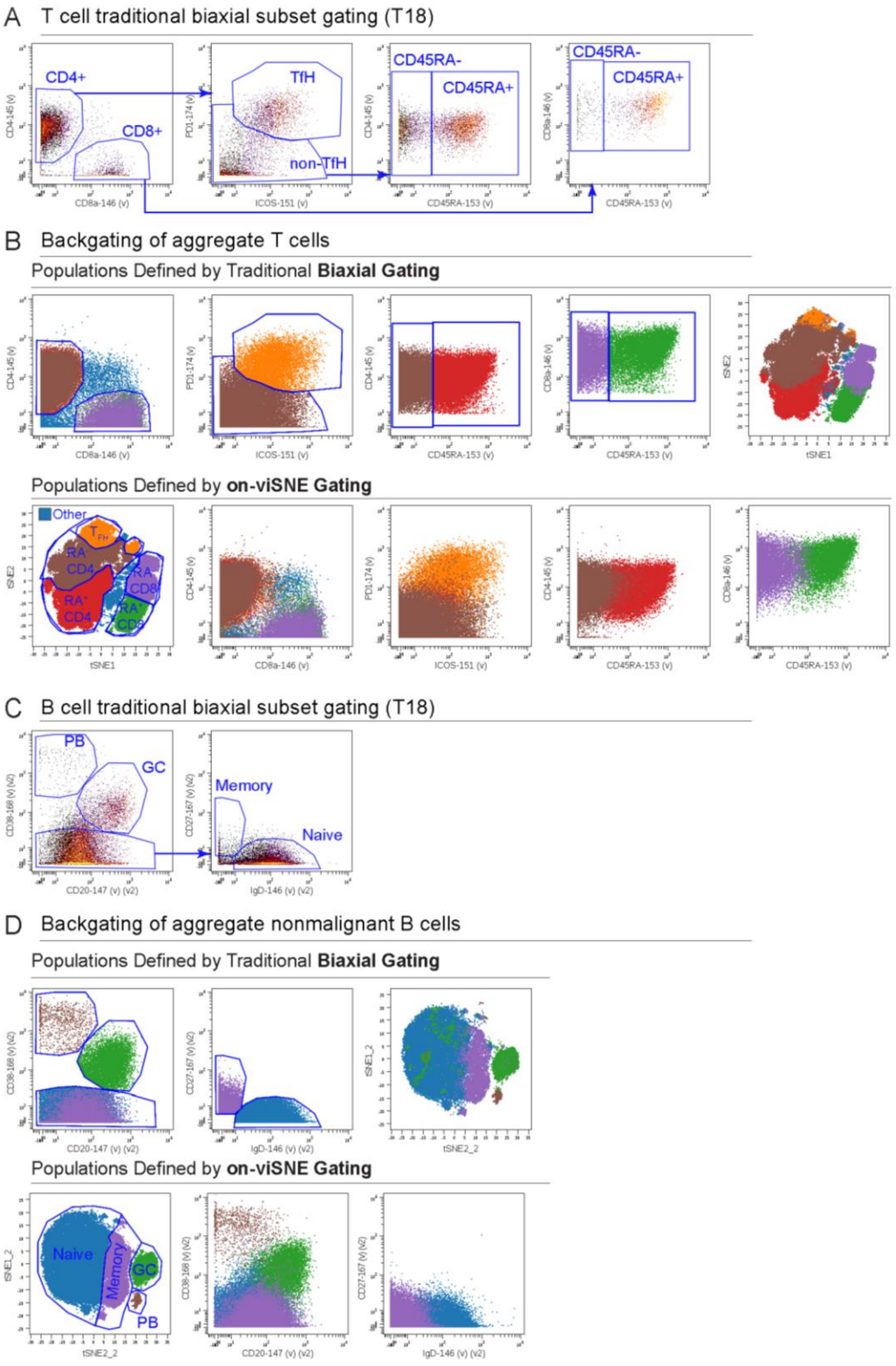



*Figure 2-6 Comparison of biaxial gating and gating on viSNE maps.*
A) *Traditional biaxial gating of T-cell subsets. B) Backgating of biaxial-gated T-cell populations onto viSNE map (top), and backgating of on-viSNE-gated populations onto biaxial plots (bottom). D) Traditional biaxial gating of B-cell subsets. B) Backgating of biaxial-gated B-cell populations onto viSNE map (top), and backgating of on-viSNE-gated populations onto biaxial plots (bottom).*

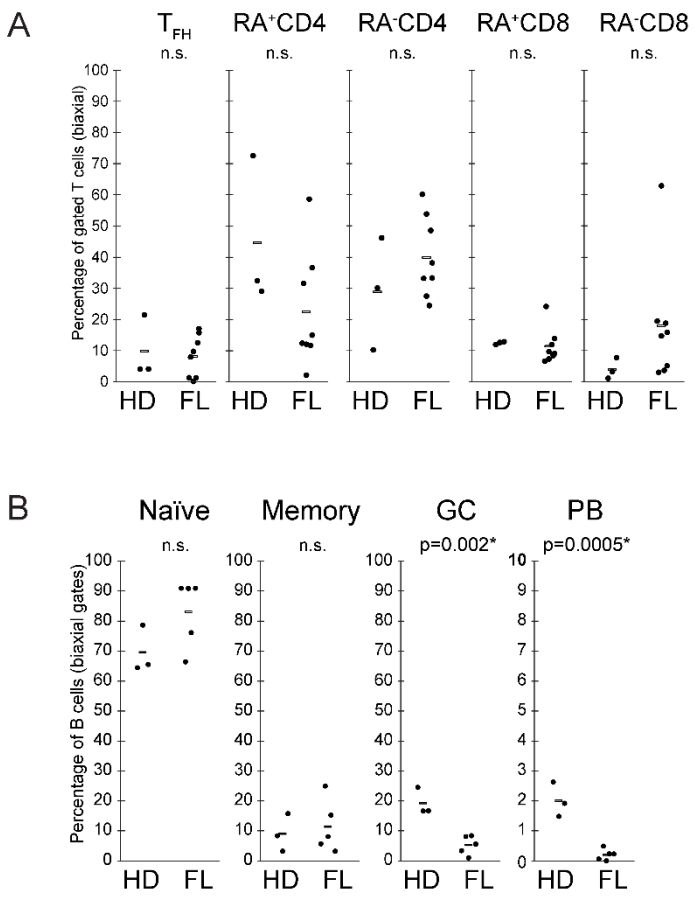



*Figure 2-7 Distribution of B & T cell subsets identified by traditional biaxial gating.*
*Figure shows all healthy donor (HD) tonsil samples and all FL samples. A) T-cell subsets as gated in Figure 2-6A. B) B-cell subsets as gated in Figure 2-6C.*

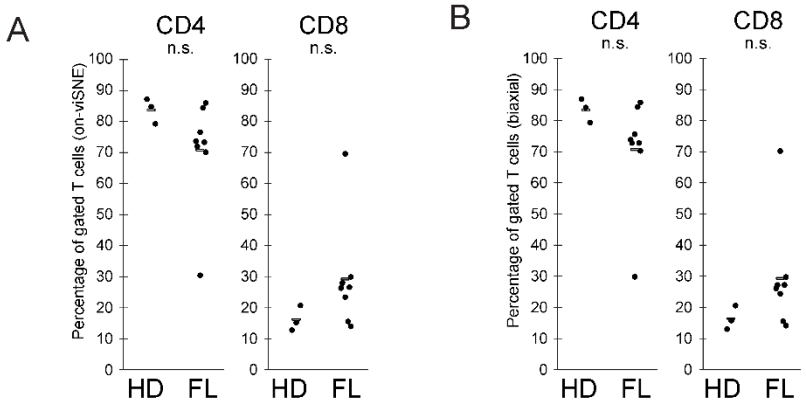



*Figure 2-8 T cell percentages plotted as total CD4 and total CD8 T cells for healthy donors and FL tumors.*
*A) On-viSNE gating values. B) Traditional biaxial gating values*

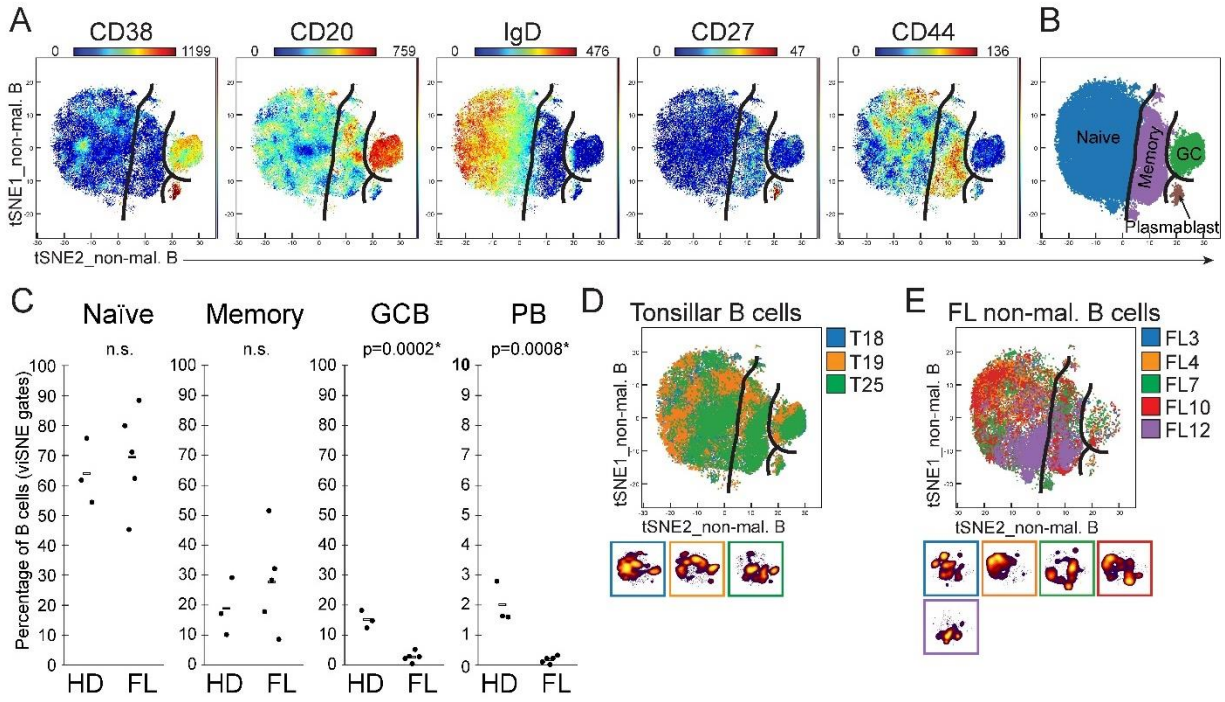



*Figure 2-9 Disruption of germinal center and plasmablast populations was observed in tumor-associated non-malignant B cells.*
*Non-malignant B cells, identified in Figure 2-1, were analyzed in a new viSNE (tSNEs_non-mal. B). FL samples with 90% or more malignant cells were excluded from the analysis. A) Expression of measured proteins is shown as a heat plot on t-SNE axes. B) Markers shown in (A) were used to gate established B-cell populations. C) Distribution of B-cell subsets across samples. Figure shows all healthy donor (HD) tonsil samples (D) and all FL samples (E) overlaid and with color coded individual cell density plots beneath.*

## FL malignant B cells were not germinal-center like in phenotype

The (non-malignant) tonsillar B-cell populations for each tonsil (4 populations from 3 tonsils) from Figure 2-9 were compared to the malignant B-cell population from each FL tumor sample using median intensity of the same 20 markers (Table 2-2) used in the B cells viSNE analyses. The populations were hierarchically clustered by median marker intensities, thereby creating a dendrogram that showed how phenotypically similar the populations were. The malignant B cells primarily clustered together but had a phenotype more similar to naïve and memory B cells than GC and plasmablasts (Figure 2-10A). There was no single marker that separated out the malignant B cells (Figure 2-10B).

Malignant B cells had higher levels of CD79B and IgM than tonsillar B-cell subsets suggesting strong dependence of BCR signaling. However, in contrast to naive B cells, all malignant B cells were IgD-. The malignant B cells also showed downregulation of the co-receptor CD40.

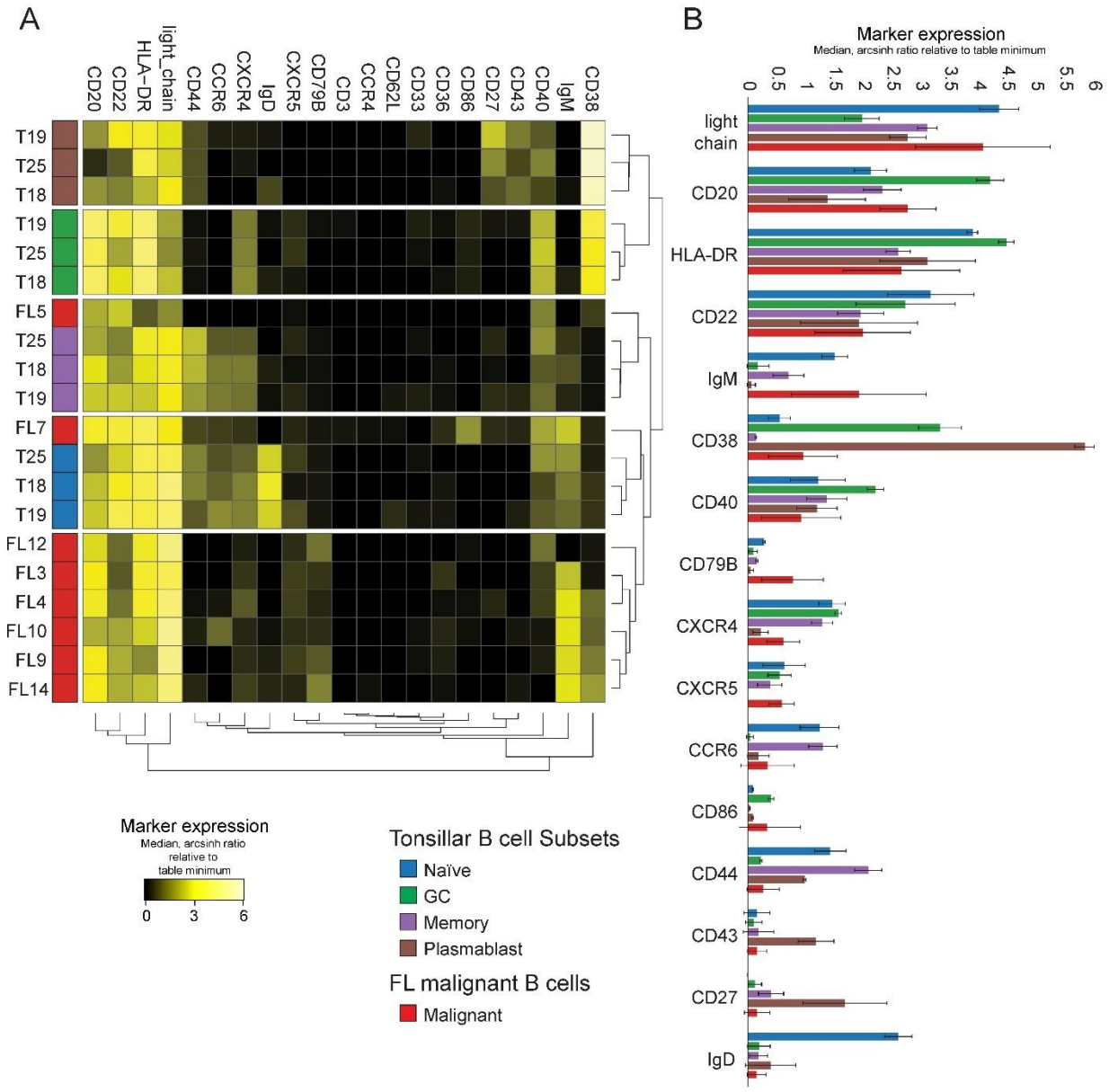



*Figure 2-10 Hierarchical clustering grouped most malignant B-cell populations and revealed malignant B cells are more phenotypically similar to naive and memory B cells than to GC B cells.*
*A) Dendrogram shows hierarchical clustering of healthy tonsillar B cell subsets and FL malignant B cells (as gated in Figure 2-9 and Figure 2-1 respectively) based on median marker expression. B) Average median marker expression of healthy tonsillar B cell subsets and FL malignant B cells. Markers are ordered by FL marker median mass intensity (MMI) from top to bottom.*

**Mass cytometry characterized high variability of intra-tumor phenotypic heterogeneity in malignant B cells**

The malignant B cells from FL samples were analyzed in a separate viSNE analysis (tSNEs_mal. B cells). In contrast to the non-malignant B cells and T cells, malignant B cell viSNE maps varied between the samples with cells occupying different areas of the map (Figure 2-11A). To quantify this variation, unsupervised clustering using SPADE was utilized to make 10 unbiased clusters of cells for each of the viSNE maps (T cells, malignant B, and non-malignant B; Figure 2-11B). Cell distribution between the nodes could then be studied within a sample and across samples (Figure 2-11C). There was some variation in the distribution throughout SPADE nodes for all the three cell types, but with greater variability for the malignant B cells than the non-malignant B and T cells (Figure 2-11C). For the malignant B cells, some FL samples were dominated by a few SPADE nodes (FL5, FL7, and FL14) whereas others spread out in the viSNE map (FL3, FL4, FL12). The markers that varied the most within tumors were HLA-DR, light_chain, and CD38 (Figure 2-11D-E). These results demonstrated variable phenotypic heterogeneity among the malignant B cells within a single tumor.

## Discussion

Minimally biased, automated computational analysis using viSNE accurately separated malignant B cells from non-malignant B cells without using BCL2 expression or light-chain restriction when distinguishing these cell subsets. Light-chain isotype was found to be restricted within the computationally defined malignant cell area, as expected for FL (Davidson, Risberg et al. 1999, Irish, Czerwinski et al. 2006, Craig and Foon 2008, Horna, Olteanu et al. 2011).

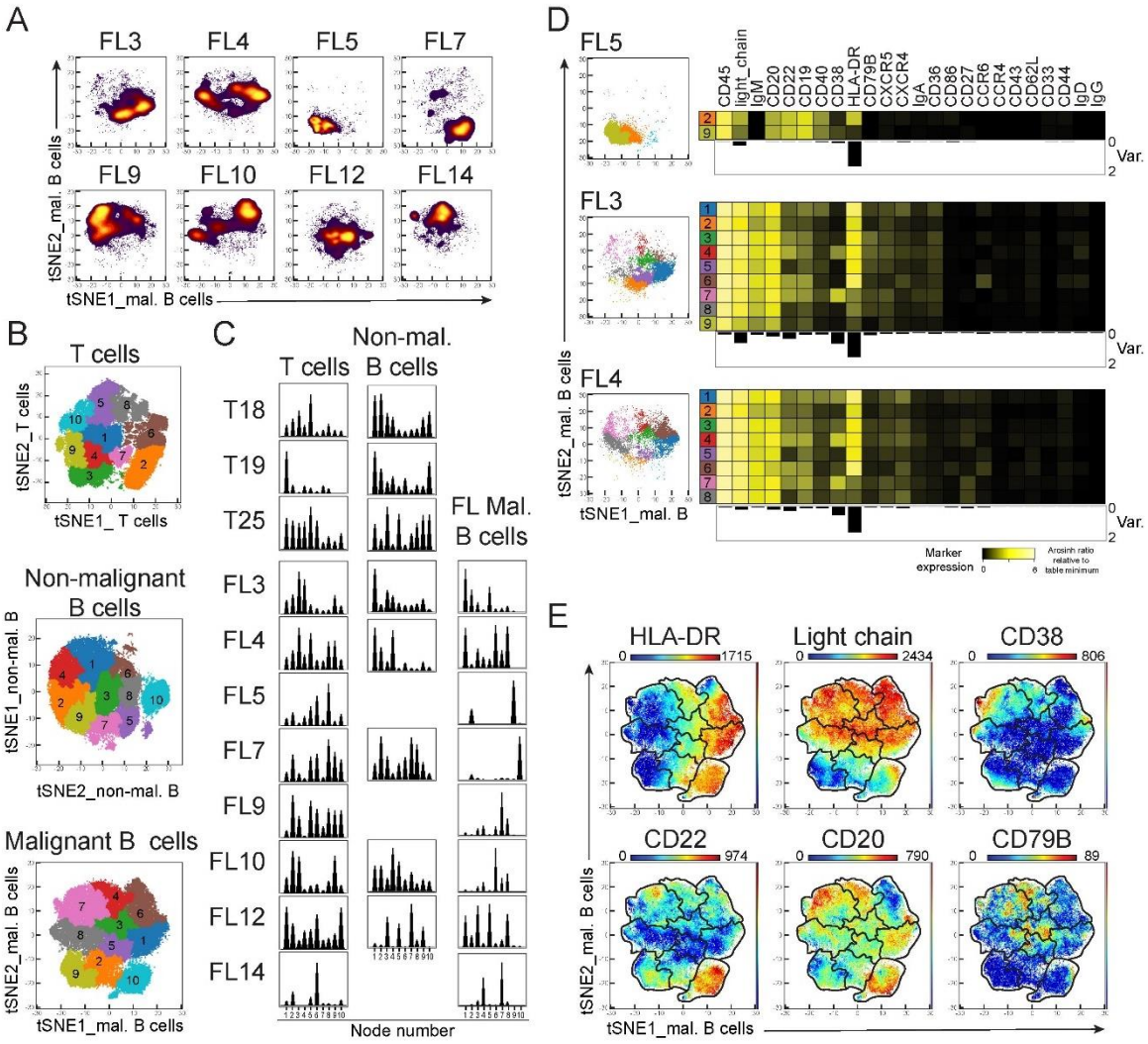



*Figure 2-11 Intratumor heterogeneity differed between the FL samples and HLA-DR was the most variable marker across samples.*
*A) Malignant B cells, gated in Figure 2-1, were analyzed in a new viSNE. Individual density plots are shown. B) SPADE analysis with 10 nodes were analyzed on the viSNE of T cells (top), non-malignant B cells (middle), and malignant B cells (bottom). Figure shows each numbered node as a separate color on a viSNE map of aggregated files. C) Histograms display number of events in each SPADE node. D) Heat map of median marker expression for each SPADE node with more than 100 cells for the FL samples. Below each heat map, the variance in each marker between the SPADE nodes are displayed as bar graphs. E) Expression of measured proteins is shown as a heat plot on t-SNE axes (malignant B cells viSNE) for some of the markers with high variance across several samples. Gates represent SPADE nodes.*

Thus, automated computational analysis was effective at identifying malignant and non-malignant cells despite the significant heterogeneity of lymphoma tumors. viSNE also successfully separated populations of non-malignant B cells and T cells. Malignant cells

are identified clinically using fluorescence flow cytometry however deep phenotyping of malignant cells is not possible with fluorescence flow cytometry and the non-malignant B cells are not characterized (Freedman 2015). viSNE mapping reveals cellular relationships in an easy to understand manner. The in-depth and quantitative nature of mass cytometry combines well with viSNE to reveal the biology of the data.

Multiple viSNE analyses were performed with and without the combined light-chain channel and the malignant cells plotted separately from the tonsillar B cells with both strategies. This result suggested that marker expression other than light-chain restriction can be used to distinguish FL malignant B cells from healthy B cells. While skewing of the light chain ratio is useful in identifying B cell malignancies, it does not allow for the separate analysis of the malignant cells because non-malignant B cells of the matching light chain will be present. This non-exclusion of the non-malignant B cells will skew any kind of bulk analysis of the malignant cells. In addition to this, identifying other markers that have differential expression between malignant and non-malignant cells could lead to the development of new therapeutic targets. CD79B, the signaling subunit of the B-cell receptor, was expressed at higher levels in malignant B cells than in all non-malignant B-cell populations (Figure 2-10B). Similarly, malignant B cells typically displayed high per-cell expression of Ig light chain and IgM (Figure 2-10B). This high level of BCR protein expression on the lymphoma B cells is consistent with a continued dependence on BCR signaling in FL (Irish, Myklebust et al. 2010). Activation of BCR signaling by auto-antigen might be one of the initial driving forces in FL, and several auto-antigens have been identified (Coelho, Krysov et al. 2010, Sachen, Strohman et al. 2012, Cha, Qin et al. 2013). Furthermore, high BCR expression, potentially followed by BCR-induced activation, aligns with the observed skewing of the non-malignant B-cell populations and

suggests that the malignant B cells outcompete the non-malignant B cells that are most dependent on support from the microenvironment for survival and selection, such as GC B cells and plasmablasts (Victora and Nussenzweig 2012). Together with the survival advantage provided by overexpression of anti-apoptotic BCL2, this may explain the deficit in non-malignant GC and plasma B cell populations observed here in FL tumors.

FL is traditionally considered to be a germinal center malignancy. FL malignant B cells contain evidence of SHM in their immunoglobulin genes (Zhu, McCarthy et al. 2002), and exhibit a GC-like gene expression pattern (Alizadeh, Eisen et al. 2000, Elenitoba-Johnson, Jenson et al. 2003). However, key signaling receptors differed between GC and malignant B cells and the protein expression pattern of malignant B cells was not GC-like (Figure 2-10). Critically, higher per-cell expression of BCR subunits CD79B, light-chain, and heavy-chain provided a clear distinction between malignant B cells and GC B cells, which expressed less surface BCR subunit proteins than naïve healthy B cells and malignant B cells. The malignant B cells were lower than GC B cells for other markers, including CD20 and CD38, as shown previously (Mantei 2009). However, malignant B cells did not phenotypically match naïve B cells or any other subset of non-malignant B cells. For example, both malignant B cells and healthy tonsillar GC B cells expressed lower levels of CD44 than naïve B cells (Figure 2-10). Thus, the expression profile of signaling receptors on malignant B cells distinguished them from all subsets of non-malignant B cells, including GC B cells.

Prior studies of clonal evolution in FL have revealed genetic heterogeneity (Eide, Liestøl et al. 2010, Green, Gentles et al. 2013, Green, Kihira et al. 2015). Single cell analysis of BCR signaling and patterns of protein expression with fluorescence flow cytometry revealed lymphoma negative prognostic (LNP) cells that exists at diagnosis in

patients with poor overall survival (Irish, Myklebust et al. 2010). Here, we observed similar patterns of CD20 expression as in these prior studies and identified additional proteins that are highly variable among lymphoma cells from the same tumor. In particular, mass cytometry revealed HLA-DR expression as one of the most variable features of FL. Green *et al.* have previously found *CREBBP* to be commonly mutated in FL, a mutation that is associated with decreased antigen presentation and expression of HLA-DR on FL B cells (Green, Kihira et al. 2015). Furthermore, this intratumor as well as intertumoral variation in HLA-DR is significant as HLA-DR expression has been previously reported in other cancers to be associated with a positive response to anti-PD1 checkpoint inhibitor therapy (Johnson, Estrada et al. 2016).

In conclusion, the use of mass cytometry to obtain deep profiling of cell subsets enabled identification of biologically important features, such as tumor heterogeneity and loss of non-malignant B-cell subsets.

# CHAPTER 3 REDOX SIGNALING HYPERSENSITIVITY DISTINGUISHES HUMAN GERMINAL CENTER B CELLS

Full title: Cutting Edge: Redox signaling hypersensitivity distinguishes human germinal center B cells (Polikowsky, Wogsland et al. 2015)

Hannah G. Polikowsky, Cara E. Wogsland, Kirsten E. Diggins, Kanutte Huse, and Jonathan M. Irish

## Abstract

Differences in the quality of BCR signaling control key steps of B cell maturation and differentiation. Endogenously produced $H_2O_2$ is thought to fine tune the level of BCR signaling by reversibly inhibiting phosphatases. However, relatively little is known about how B cells at different stages sense and respond to such redox cues. In this study, we used phospho-specific flow cytometry and high-dimensional mass cytometry (CyTOF) to compare BCR signaling responses in mature human tonsillar B cells undergoing germinal center (GC) reactions. GC B cells, in contrast to mature naive B cells, memory B cells, and plasmablasts, were hypersensitive to a range of $H_2O_2$ concentrations and responded by phosphorylating SYK and other membrane-proximal BCR effectors in the absence of BCR engagement. These findings reveal that stage-specific redox responses distinguish human GC B cells.

## Introduction

In this study, we found that ROS caused increased signaling through the B cell receptor (BCR) pathway in germinal center (GC) B cells. Additionally, we saw increased phosphatase (SHP-1) expression in a subset of GC B cells that we suspect are the light zone B cells. Understanding healthy GC B cells is key to making advancements in B cell

lymphomas. In Chapter 2, I performed a deep phenotyping of eight follicular lymphoma (FL) tumors with tonsillar cells as the healthy counterpart. The work done here helped pave the way for incorporating disease states into the mass cytometry analysis performed in Chapter 2. Here I focused on the markers that identify mature healthy B cell subsets and the markers that characterize those populations. This study paved the way for the FL study by providing me with insight into mass cytometry deep phenotyping of mature B cell subsets. This depth of phenotyping had not been done previously in FL. Fluorescence flow FL studies have not had large enough staining panels to perform this depth of phenotyping.

Previously, Irish et al. showed that FL malignant B cells had more robust BCR pathway signaling responses than non-malignant B cells in the same tumors as determined by BCR expression (Irish, Czerwinski et al. 2006). That study did not have the ability to characterize the four main B cell populations that were characterized in Chapter 2 and will be characterized by signaling responses in this chapter. It did however provide evidence that FL B cells responded more robustly to $H_2O_2$ stimulation much in the same way that GC B cells do here. So, although we found FL B cells to be phenotypically distinct from GC B cells (Figure 2-10), they share heightened signaling responses to reactive oxygen species (ROS).

I designed the core B cell panel (Table 1-1) used here and helped adapt the core panel for this study (Table 3-1). The adaptation included the intracellular readouts p-PLCγ and SHP-1. Phospho-antibody and phosphatase channels need to be considered in viSNE analysis design. Whether or not to include one of these signaling molecules depends on the question being asked. If we want to know which cell subsets express a

| Marker | Clone | Mass | Type |
|---|---|---|---|
| CD19 | HIB19 | 142 | δ |
| CD5 | UCHT2 | 143 | δ |
| *IgG | G18-145 | 144 | δ |
| CD4 | RPA-T4 | 145 | β |
| IgD | IA6-2 | 146 | δ |
| CD20 | 2H7 | 147 | δ |
| CD16 | 3G8 | 148 | β |
| IgA | polyclonal | 148 | α |
| Igλ | MHL-38 | 151 | δ |
| CD45 | HI30 | 154 | δ |
| CD27 | L128 | 155 | β |
| *p-PLCγ | K86-6 89.37 | 156 | α |
| CD33 | MW53 | 158 | β |
| CD22 | HIB22 | 159 | δ |
| Igκ | MHK-49 | 160 | δ |
| CD79B | CB3-1 | 162 | β |
| CD40 | 5C3 | 165 | δ |
| CD44 | BJ18 | 166 | β |
| CD38 | HIT2 | 167 | δ |
| CD8 | SK1 | 168 | δ |
| CD3 | UCHT1 | 170 | δ |
| IgM | MHM-88 | 172 | δ |
| HLA-DR | L243 | 174 | δ |
| **SHP-1 | C14H6 | 175 | β |
| Iridium | N/A | 193 | δ |

Marker Type Key:
α- Mass cytometry panel 1 only
β-Mass cytometry panel 2 only
δ-Both panels

All antibodies were purchased from Fluidigm, unless otherwise specified
*BD Biosciences antibody
**Cell signaling technology antibody

*Table 3-1 Two mass cytometry panels with a large degree of overlap were used in this study. Phospho-PLCγ was stained for in panel 1 only and total SHP-1 was stained for in panel 2 only.*

signaling molecule, then we would leave it out of the viSNE analysis as in Figure 3-5 so that its expression is not directing the shape of the map. I assisted in the interpretation of the data and figure design. Although the study was not powered to truly differentiate light and dark zone GC B cells, based on the phenotyping that we performed and the literature on GC B cells, I suspect that the cells higher for the phosphatase SHP-1 are likely light

zone GC B cells that have recently received T cell help based on the fact that they expressed higher levels of CD79B, CD40, HLA-DR, CD22, and CD86 (Figure 3-5A). The mass cytometry panel could be improved by adding CXCR4 and CXCR5 to help differentiate light and dark zone B cells. One of the major pitfalls is that p-PLCγ and SHP-1 are not on the same mass cytometry panel so we cannot see directly how phosphorylation of PLCγ relates to total SHP-1 expression at the single cell level.

In addition to BCR signaling, secondary messengers control the signaling context and help determine functional outcomes in B cells. $H_2O_2$ is the primary reactive oxygen species (ROS) produced by B cells. $H_2O_2$ amplifies BCR signaling by transiently inhibiting BCR-associated protein tyrosine phosphatases (Reth 2002). $H_2O_2$ is also produced as part of innate immune responses to wounds and infection (Nathan and Cunningham-Bussel 2013). However, it is not known what impact $H_2O_2$ has on healthy human B cell signaling responses and whether B cells undergoing GC reactions respond differently to $H_2O_2$.

Seconds after BCR crosslinking, a network of signaling molecules becomes activated through posttranslational modifications. As signaling directs B cells down differentiation pathways, B cells adopt well-characterized signatures defined primarily by protein expression (Maecker, McCoy et al. 2012). Naive B cells in humans are defined by expression of CD19, CD20, and IgD. GC B cells are defined as CD19+, CD20hi, CD38+, IgD− B cells. Memory B cells, alternatively, express CD19, CD20, and CD27. Furthermore, human plasmablasts are defined as CD38hi, CD20lo cells that are in the process of downregulating surface BCR and most other surface antigen.

The GC is a highly active environment vital for proper functioning of the adaptive immune system. GC B cells undergo affinity maturation, which involves iterative cycles of clonal expansion, somatic hypermutation, and selection that result in class-switched memory B cells and Ab-secreting plasma cells (Victora, Schwickert et al. 2010, Gitlin, Shulman et al. 2014). How high-affinity B cells are selected in the GC is not entirely clear. Increased Ag capture and presentation lead to increased rates of cell division (Gitlin, Shulman et al. 2014, Shulman, Gitlin et al. 2014). It is also possible that actively proliferating GC B cells produce unique signals that promote their survival and proliferation. Additionally, GC B cell signaling is regulated by protein tyrosine phosphatases (Klein, Tu et al. 2003, Khalil, Cambier et al. 2012). For example, cell surface CD22 can recruit phosphatases, such as SHP-1, to attenuate BCR signaling (Singh, Kumar et al. 2005, Khalil, Cambier et al. 2012). Opposing this activity are NADPH oxidases, such as DUOX1, which produce $H_2O_2$ and lower BCR signaling thresholds by reversibly inhibiting phosphatases (Nathan and Cunningham-Bussel 2013). The environment surrounding the BCR simulates NADPH oxidase, which produces endogenous ROS (Woo, Yim et al. 2010). In turn, ROS oxidize the extracellular compartment and activate the BCR signaling pathway, creating a positive feedback loop. BCR signaling governs B cell functions, and activation and termination of BCR signaling are finely tuned by multiple levels of regulation in healthy cells.

Although the biochemistry of BCR signaling is well understood in model systems, little is known about the quality of in vivo BCR signaling in mature, healthy human B cells. Addressing this gap by mapping the influence of ROS on healthy B cell signaling is important for placing into context the extreme BCR signaling and $H_2O_2$ responses observed in B cell diseases and disorders (Irish, Myklebust et al. 2010). In this study, we

used high-dimensional mass cytometry, phospho-specific flow cytometry, and novel computational data analysis tools (Amir el, Davis et al. 2013, Becher, Schlitzer et al. 2014, Bendall, Davis et al. 2014) to better understand how ROS regulate BCR signaling within subsets of primary human tonsillar B cells.

## Materials and Methods

### Human samples

Tonsils were obtained from children undergoing routine tonsillectomies in accordance with the Declaration of Helsinki following protocols approved by Vanderbilt University Medical Center Institutional Review Board (IRB# 121328). Specific patient demographics are unknown since the samples were collected with an IRB exemption and classified as non-human subjects. Single-cell suspensions were prepared and stored in liquid nitrogen.

### Antibodies

Fluorescent Abs for CD20, IgD, CD38, CD3, CD27, p-SRC, p-SYK, p–phospholipase C (PLC)γ, and p–NF-κB were conjugated to BV421, PerCP-Cy5.5, FITC, PE-Cy7, BUV395, BV570, BV605, PE, and Alexa Fluor 647 (BD Biosciences, Invitrogen, or BioLegend). Mass cytometry Abs are listed in Table 3-1.

### Fluorescence cytometry

Aliquots of cryopreserved single-cell tonsillar samples were thawed into 10 ml warm media (RPMI 1640 [Mediatech, Manassas, VA] plus 10% FBS [Life Technologies, Grand Island, NY]), pelleted by centrifugation at 200 × g, washed with warm media, and pelleted

again at 200 × g before resuspension in flow cytometry tubes. Resuspended samples rested for 15 min in a 5% $CO_2$ incubator at 37°C. Each rested sample was either left unstimulated or stimulated with $H_2O_2$ (Fisher Scientific, Fair Lawn, NJ) for 2 min or CD40L plus enhancer (Enzo Life Sciences, Farmingdale, NY) for 15 min. CD40L and enhancer were prepared per the manufacturer's recommendation. Cells were fixed with 1.6% paraformaldehyde (Electron Microscopy Services, Fort Washington, PA) for 5 min at room temperature following stimulation, washed with PBS (HyClone Laboratories, Logan, UT), pelleted at 800 × g, and permeabilized by 100% ice-cold methanol (Fisher Scientific) in a −20°C freezer overnight. Cells were washed once with PBS and once with cell staining media composed of PBS plus 1% BSA (Fisher Scientific). For each condition, $1 \times 10^6$ tonsillar cells were stained in 100 µl cell staining media. Samples were analyzed using a five-laser BD LSR II (Becton Dickinson, Franklin Lakes, NJ) at the Vanderbilt Flow Cytometry Shared Resource and evaluated using Cytobank software.

**Mass cytometry**

Single-cell tonsillar samples were thawed the same way as samples prepared for fluorescent cytometry. For mass cytometry panel 1, one tonsil sample was left unstimulated and one sample was stimulated with $H_2O_2$ (Fisher Scientific) for 2 min. Cells were fixed (paraformaldehyde) and stained for extracellular targets (Table 3-1). After methanol permeabilization, cells were stained for IgG, IgM, IgA, and p–PLCγ-PE in cell staining media for 15 min at room temperature, then stained with 250 nM iridium intercalator and anti-PE (Fluidigm) for 30 min at room temperature. Cells were washed once in PBS, once in double distilled $H_2O$, suspended in double distilled $H_2O$, and collected on a CyTOF 1.0 at the Vanderbilt Flow Cytometry Shared Resource. Cells

stained with panel 2 (Table 3-1) were stained immediately after thawing, except for SHP-1, which was stained after permeabilization. Mass cytometry data files were evaluated using manual gating and viSNE (Amir el, Davis et al. 2013) in Cytobank. Data were transformed using an arcsinh scale (cofactors of 15, except for SHP-1, which had a cofactor of 5). viSNE maps were generated using the following markers: SHP-1, CD40, IgD, CD3, CD3, CD19, CD20, CD86, CD22, CD44, CD38, CD27, CD79B, and HLA-DR.

## Results and Discussion

### A subset of B cells responded robustly to $H_2O_2$ stimulation

A subset of human tonsillar B cells was initially observed to respond to a 2-min stimulation by 3.3 mM $H_2O_2$ by phosphorylating upstream members of the BCR signaling pathway, including SYK and Src family kinases (SFKs) (Irish, Czerwinski et al. 2006). This $H_2O_2$-sensitive population varied in abundance from 7.3 to 33.24% of CD3− cells (Figure 3-1) and generally expressed higher levels of CD20 compared with other tonsillar B cells (Figure 3-2).

In previous reports, naive B cells in peripheral blood did not respond to 3.3 mM $H_2O_2$ (Irish, Czerwinski et al. 2006). The $H_2O_2$ response of the CD20hi CD3- B cells distinguished these cells from other tonsillar cells and contrasted with the B cell response to other stimuli, such as CD40L, which showed no significant signaling differences across the full range of CD20 expression levels (Figure 3-1). Thus, a novel $H_2O_2$ signaling response distinguished a CD20hi subset of tonsillar B cells.

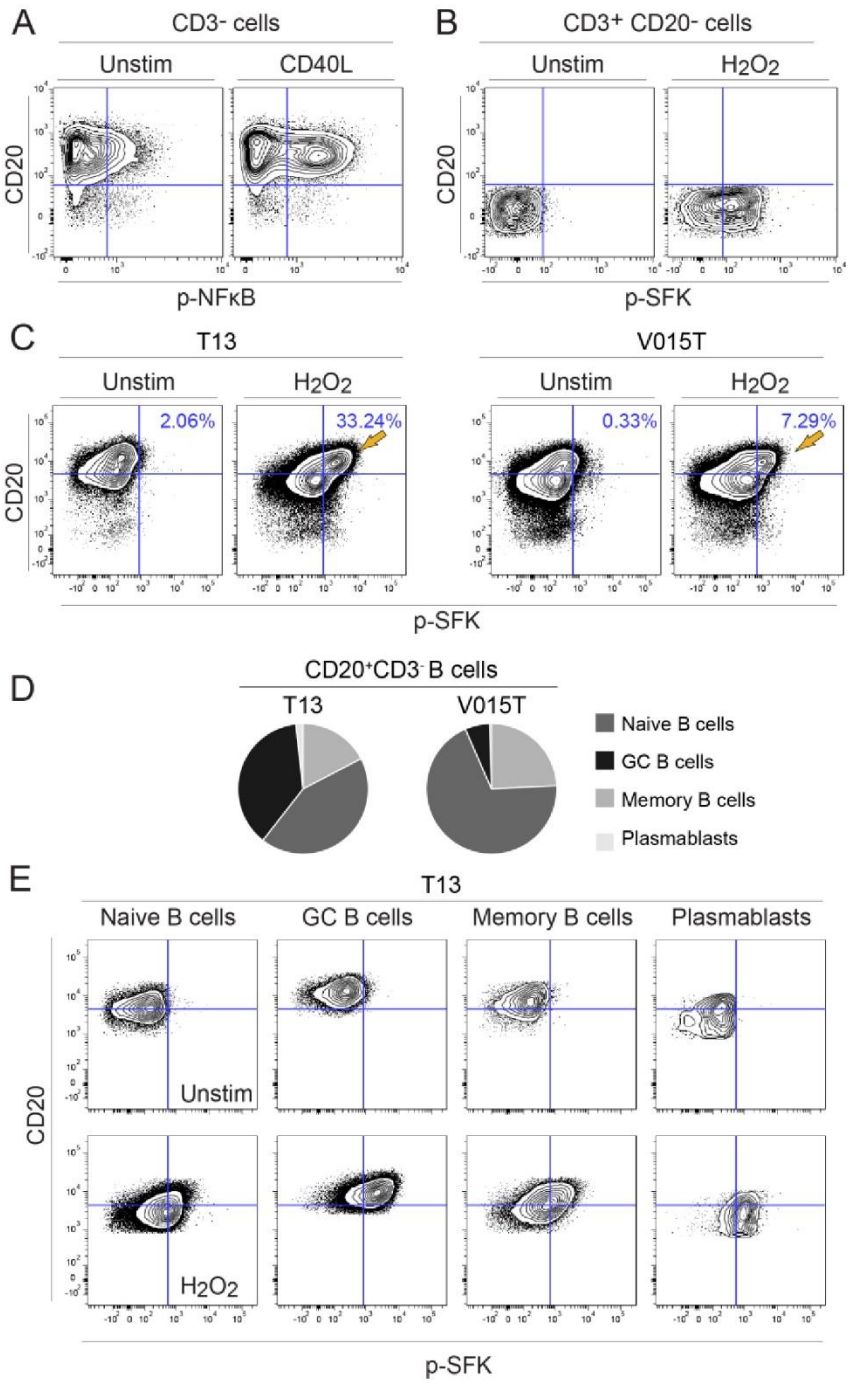



*Figure 3-1 A subset of CD20hi B cells responded robustly to peroxide via phosphorylation of SFK.*
*Populations (CD3- cells and CD3+CD20- cells) defined in figure 3-2 for A-C and gating schematic for D and E (Naïve B cells, GC B cells, Memory B cells, plasmablasts) shown in figure 3-4. (A) Contour plots show p-NFκB in unstimulated cells and cells stimulated by CD40L for 15 minutes in CD3- cells. (B-D) Cells were either left unstimulated or stimulated by 3.3mM of $H_2O_2$ for 2 minutes. (B) Contour plots show p-SFK (i.e. p-LCK) in CD20- CD3+ tonsillar T cells (C) Contour plots show two other CD3- healthy tonsil specimens (T13 left-two plots, V015T right-two plots) and their response to H2O2 stimulation. Sensitivity to $H_2O_2$ in a CD20hi B cell population is indicated (gray arrows). (D) Population portions of identified B cell subsets for two tonsils is shown. (E) Contour plots show raw data for B cell population responses to $H_2O_2$ (bottom row). Data for one representative tonsil (T13, n=3) is shown.*

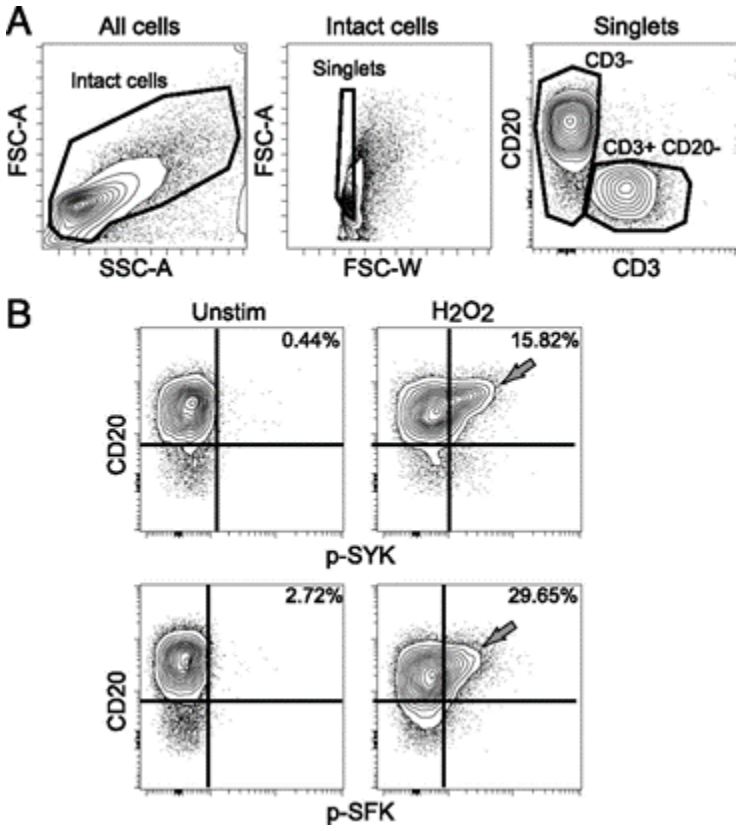



*Figure 3-2 CD20hi B cells in human tonsil were sensitive to $H_2O_2$.*
*(A) Contour plots show gating for CD3− cells and CD3+CD20− T cells in human tonsil. (B) Contour plots show p-SYK and p-SFK in CD3− tonsillar B cells left unstimulated or stimulated by 3.3 mM $H_2O_2$ for 2 min. Sensitivity to $H_2O_2$ in a CD20hi B cell population is indicated (gray arrows). Plots are representative of three tonsils.*

### Comprehensive characterization of $H_2O_2$-responsive B cells by mass cytometry

To determine the identity of the $H_2O_2$-responsive cells, a high-dimensional mass cytometry panel designed to characterize mature B cells was developed (Table 3-1). The $H_2O_2$-sensitive cell population was gated and labeled as “responder” (R) cells, and the signature of protein expression was contrasted with cells labeled as “non-responders” (NR) or CD3+ T cells (Figure 3-3A,B). The $H_2O_2$-sensitive responder cells were characterized by a CD20hi, CD38+, IgD− phenotype that contrasted with the other evaluated populations of non-responder cells and CD3+ cells (Figure 3-3C). This observed responder cell phenotype suggested a GC B cell identity (Jackson, Wilson et

al. 2008, Maecker, McCoy et al. 2012). In agreement with this, a strong relationship was seen between the fraction of $H_2O_2$-sensitive responding cells and the abundance of GC B cells in each tonsil (Figure 3-1).

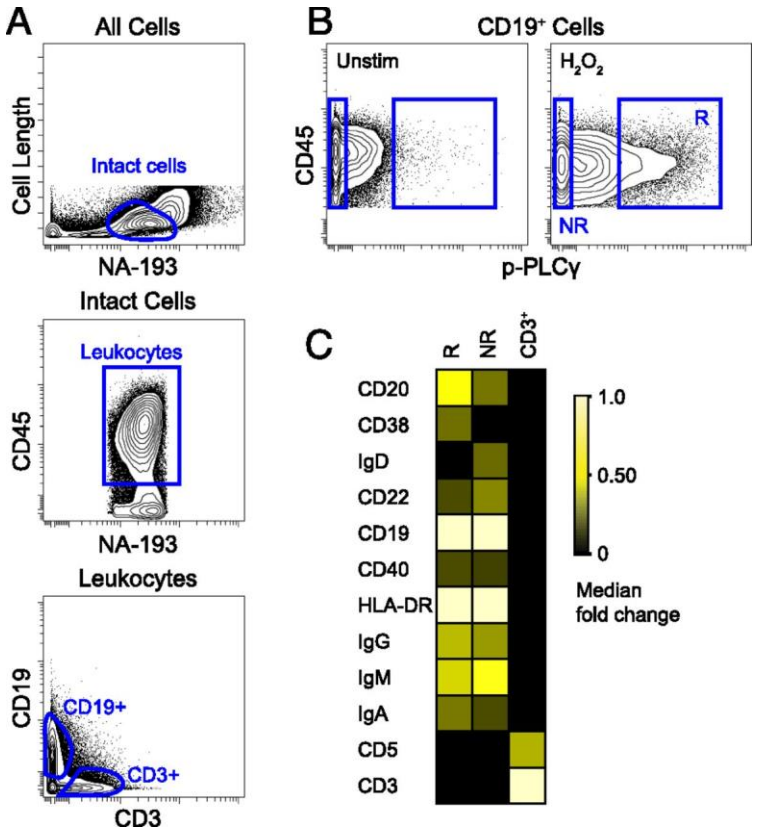



*Figure 3-3 Mass cytometry revealed $H_2O_2$ responder population as CD20hi, CD38+, and IgD−.*
*(A) Contour plots show gating for CD19+ B cells and CD3+ T cells in human tonsil. (B) Contour plots show p-PLCγ in CD19+ tonsil B cells left unstimulated or stimulated by 3.3 mM H2O2 for 4 min. Cells that were sensitive to $H_2O_2$ stimulation were labeled responder (R) cells and contrasted with nonresponder cells (NR). (C) A heat map shows the median fold change in CD3+ T cells, R, and NR subsets.*

### GC B cells were hypersensitive to $H_2O_2$ stimulation

BCR signaling normally triggers a complex, interconnected network of effector signaling pathways (Irish, Czerwinski et al. 2006), and it is currently not known how the quality, magnitude, and duration of BCR signaling "programs" a B cell for contrasting functional outcomes ranging from cell death to proliferation. Phospho-proteins in the BCR signaling network that are rapidly phosphorylated following $H_2O_2$ stimulation might act as effectors

of secondary messenger signaling. To identify $H_2O_2$ signaling effectors and better delineate the $H_2O_2$ sensitivity of B cell populations, a fluorescent panel was developed and cells from three human tonsils were stimulated with varying doses of $H_2O_2$ for 2 min (Figure 3-4). Naive, GC, memory, and plasmablast B cell subsets were distinguished using canonical markers CD3, CD20, CD38, CD27, and IgD (Figure 3-4A). Observed B cell subsets responded to $H_2O_2$ in a dose-dependent manner seen through the phosphorylation of SFK, PLCγ, and SYK; however, GC B cells were the most sensitive to $H_2O_2$ at all concentrations (Figure 3-4B,C).

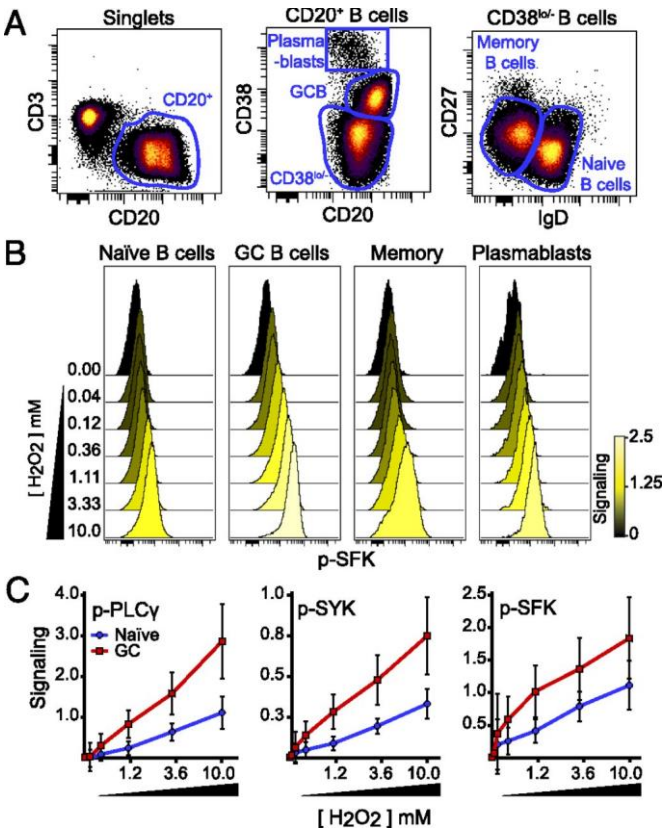



*Figure 3-4 GC B cells were hypersensitive to $H_2O_2$.*
*(A) Density dot plots show gating for identification of plasmablasts, GC B cells, memory B cells, and naive B cells in human tonsils. (B) Histogram overlays show p-SFK in each B cell population [shown in (A)] following 2 min of 3.3 mM $H_2O_2$ (n = 3, representative data shown). Color denotes median fold change in p-SFK expression compared with unstimulated (0 mM $H_2O_2$). (C) Plots illustrate the median fold change in p-PLCγ, p-SYK, and p-SRC in $H_2O_2$-stimulated conditions compared with the unstimulated condition (arcsinh scale). Each point represents the average of three individual tonsil specimens (n = 3) stimulated for 2 min with the indicated concentration of $H_2O_2$, except for the 0.04*

*mM and 0.12 mM $H_2O_2$ stimulated conditions (where n = 2). Red squares represent GC B cells and blue circles represent naive B cells. Error bars denote the SD for each point.*

$H_2O_2$ sensitivity may be an intrinsic characteristic of GC B cells that is necessary for BCR regulation within an active GC. GC B cells may use endogenously produced $H_2O_2$ as a modulator of BCR signaling, whereas BCRs undergo iterative modification. In fact, loss of BCR signaling in healthy B cells reduces B cell survival, and sustained BCR signaling capability is essential for B cell development and survival (Kraus, Alimzhanov et al. 2004). Observed $H_2O_2$ hypersensitivity of GC B cells (Figure 3-4) may be an important feature of accelerating the GC reaction; alternatively, this redox sensitivity may help to cull B cells that do not appropriately execute the delicate process of somatic hypermutation. These results help to place in context the observation that lymphoma B cells are especially sensitive to ROS (Irish, Czerwinski et al. 2006). Prior studies revealed that lymphoma B cells undergo rapid, ROS-mediated apoptosis when glutathione is depleted and that stimulation of lymphoma B cells using anti-BCR F(ab′)2 and $H_2O_2$ negates suppression of BCR signaling that distinguishes clinically relevant lymphoma negative prognostic cells in follicular lymphoma (Irish, Myklebust et al. 2010).

**Heterogeneous SHP-1 expression across B cell populations**

Previous data from GCs generated within transgenic mice reported that GC B cells do not robustly respond to Ag or anti-IgM stimulation compared with non-GC B cells due to colocalization of SHP-1 with the BCR (Khalil, Cambier et al. 2012). To study this relationship in humans, a single-cell approach was used to measure total SHP-1 levels within human tonsillar B cell subsets and quantify any correlation between total SHP-1 protein expression and B cell population identity. Furthermore, an unsupervised computational approach was used to characterize GC B cells and determine whether

additional heterogeneity might exist within this or other B cell populations (Amir el, Davis et al. 2013, Becher, Schlitzer et al. 2014). Elevated phosphatase levels of GC B cells compared with other B cell subsets might explain why GC B cells were hypersensitive to $H_2O_2$ stimulation. To evaluate this hypothesis, an Ab for SHP-1 was added to the mass cytometry panels (Table 3-1). B cell subsets were identified by viSNE analysis using the same key markers as in fluorescent experiments (Figure 3-5). viSNE revealed heterogeneous expression of SHP-1 within naive, GC, and memory B cell populations. Each of these B cell populations contained both high and low SHP-1–expressing cells. In contrast, plasmablasts expressed a consistent, low level of SHP-1. SHP-1 expression contrasted strongly with canonical subset marker expression patterns, which were enriched in subset-specific ways, such as CD20 and CD38 (Figure 3-5B). In the present study, SHP-1 expression was uncorrelated with $H_2O_2$ sensitivity across the B cell stages. Plasmablasts and naive B cells expressed contrasting levels of SHP-1 and had comparable $H_2O_2$ sensitivity, whereas GC and naive B cells had contrasting $H_2O_2$ sensitivity despite similar median levels and per-cell distributions of SHP-1 expression (Figures 3-4, 3-5).

Because SHP-1 expression did not correlate with B cell subset, it is possible that the observed heterogeneity of SHP-1 expression is due to transient differences within B cell subsets that are not reflective of stage, but rather recent stimulation experience. A recent study demonstrated that a subpopulation of light zone GC B cells had more robust BCR signaling compared with all GC B cells (Mueller, Matloubian et al. 2015). Our study was not powered to look at light zone/dark zone differences, but the data suggested that light zone GC B cells may be the GC B cells that are higher for SHP-1. SHP-1 has

previously been found to be higher on light zone GC B cells by IHC (Kossev, Raghunath et al. 2001). Within the GC B cell subset, the cells on the viSNE map that expressed

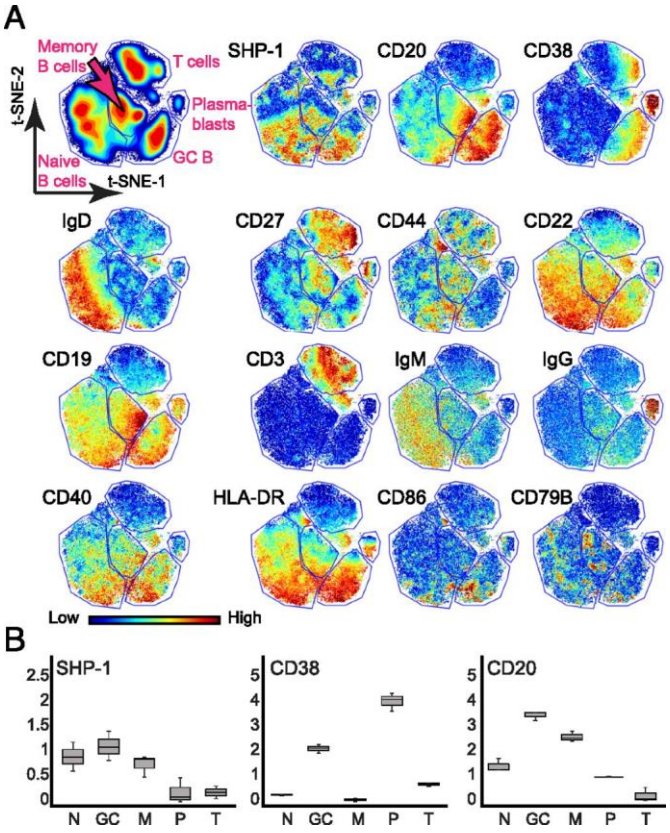



*Figure 3-5 SHP-1 expression was heterogeneous within B cell populations.*
*viSNE maps show CD45+ leukocytes arranged based on marker expression profiles (see gating in Figure 3-3A). Color denotes protein expression, as indicated. (A) Gates were drawn around the main populations identified by viSNE, using protein expression to identify each population. CD19+ B cells were subdivided into naive B cells (CD38−CD27−IgD+), GC B cells (CD20hiCD38+), memory B cells (CD38−CD27+IgD−), and plasmablasts (CD20−CD38hi) and compared with CD3+ T cells. One representative tonsil of four analyzed is shown. (B) Box-and-whisker plots illustrate expression of SHP-1, CD38, and CD20 proteins across three tonsil specimens. Median of each marker is indicated by a black line. Bars denote the minimum and maximum observed mean fluorescence intensity of each marker. GC, GC B cells; M, memory B cells; N, naive B cells; P, plasmablasts; T, T cells.*

higher levels of SHP-1 also expressed higher levels of CD40, HLA-DR, CD22, and CD86 (Figure 3-5A). These proteins relate to T cell signaling interactions and suggest a shift in the signaling relationship between T follicular helper cells and GC B cells.

These results provide new information regarding redox-sensitive signaling in B cell networks that may act to control the outcomes of GC reactions. Precisely how ROS

regulate BCR signaling within GCs remains to be seen; however, the findings in the present study indicate that redox cues specifically impact human GC B cell signaling. These results revealed unknown human GC B cell signaling responses to ROS that can be used as a reference point for studies of diseases originating in cells with GC characteristics, such as B cell lymphomas.

# Chapter 4 CAPTURING BIOLOGICAL TRUTH WITH DIMENSIONALITY REDUCTION TOOLS

Full Title: Accuracy in viSNE analysis is improved by simultaneously increasing iterations with analyzed cell count (unpublished)

Cara E. Wogsland and Jonathan M. Irish

**Abstract**

*Background*: Computational tools like viSNE are effectively revealing known and novel cell subsets in 30+ dimensional cytometry studies of healthy tissue, cancer, and immune disorders (Amir el, Davis et al. 2013, Diggins, Ferrell et al. 2015). However, some viSNE runs fail to effectively separate established cell populations that are normally well-distinguished, such as $CD4^+$ and $CD8^+$ T cells. This study explored the underlying reasons and solutions to address this in the context of canonical human blood cell types.

*Methods*: Human peripheral blood mononuclear cells (PBMCs) from 4 healthy donors were analyzed by 34-dimensional mass cytometry. 422,833 intact CD45+ cells were gated by into 5 populations with the established phenotypes: B cells (CD19+, CD3-), monocytes (CD33+ CD19-, CD3-), CD4 T cells (CD3+, CD4+, CD19-, CD8-), CD8 T cells (CD3+, CD8+, CD19-, CD4-), and NK cells (CD56+, CD19-, CD3-). To assess separation, viSNE maps created in Cytobank or Rstudio with Rtsne were split evenly into 50 x 50 square bins and assessed for average percent purity in bins containing at least 1% of cells. A result of 90% average purity indicated that the most abundant expert-gated population in each bin on average comprised 90% of the cells in that binned region of the

viSNE map. For example, if all the B cells were mapped together, then all the B cell containing bins should have a high percentage of B cells and a low percentage of other cells, giving a high average percent purity score. Bins containing less than 1% of B cells were not counted toward the score. This provided a simple way to rank viSNE maps according to effective separation of canonical cell subsets. Input arguments for viSNE/BH-SNE, including cell number and iterations were varied systematically to test impact on separation of known cell types.

*Results*: viSNE maps made with over 100,000 events and standard input arguments (iterations=1000, perplexity=30) received lower average purity scores indicating a decrease in biological relevance and map integrity.

*Conclusions*: When running viSNE on datasets of over 100,000 events, it is important to increase the number of iterations in order to capture biologically important details and maintain the integrity of the map. The viSNE input arguments should continue to be evaluated as the algorithm and usage evolve.

## Introduction

I recognized the problem of the viSNE map losing integrity with many cells (>100,000) being analyzed simultaneously (Figure 4-1). Integrity refers to the shape of the map and the biological information accurately displayed on the map. If different cell populations such as B and T cells are occupying the same space on the viSNE map, it has lost integrity. Since viSNE is a visualization tool and a large number of cells caused a breakdown in visualization of biologically accurate data, this was a problem that needed

to be addressed. I designed a series of tests to perturb the input parameters to determine which parameters should be changed by the user before a viSNE run and how to allow for the analysis of a large number of cells in a single viSNE run. The work shown here is the qualitative and quantitative summary of that project. I measured integrity with a (percent average) purity score, where the higher the purity, the more pure and biologically true the cell populations are in the map. This means that with high purity, B cells and T cell will not overlap in the map, nor will T cell subsets such as CD4 and CD8.

The purpose of this project was to find the appropriate computational input parameters for identifying biological truth in viSNE analyses of mass cytometry data. In the context of viSNE, identifying biological truth would mean that things an expert knows are “B cells” would be close to each other on a map. Similarly, T cells would be close to T cells and not overlap with B cells. The initial viSNE paper used a maximum of 100,000 events in a single viSNE run (Amir el, Davis et al. 2013). As computational resources have advanced and more cells could be run simultaneously, I have observed in my data and that of others that the viSNE maps lose integrity with more cells added and they often show disparate cells close together.

## Methods

The cellular data points defined by t-SNE coordinates were assigned to a bin in a 50x50 grid (2,500 bins) of the viSNE map. The bins were of equal distance intervals in t-SNE space with the total PBMC t-SNE x,y ranges serving as the outer edges of the grid for each of the five expert gated populations. This ensures that changing the visualization of the map to zoom in or out does not change the bins.

The percentage of expert gated cells in a single bin was calculated for all 2,500 bins, separately for each of the 5 expert gated populations. For every bin that contained at least 1% of an expert population, a purity score was calculated, (expert gated cells)/(total PBMC)*100. Percentages were averaged for each of the five populations to get an average purity score.

## Results and discussion

**Decreasing cell number improved the qualitative look of the viSNE map and the average purity score.**

Figure 4-1 illustrates the problem of the cells clumping together and failing to spread out on the map when cell number increased when using the original parameters of 1000 iterations (I) and perplexity (P) of 30. Perplexity was not tested in this study due to minimal change observed in initial testing and the additional computational demands required.

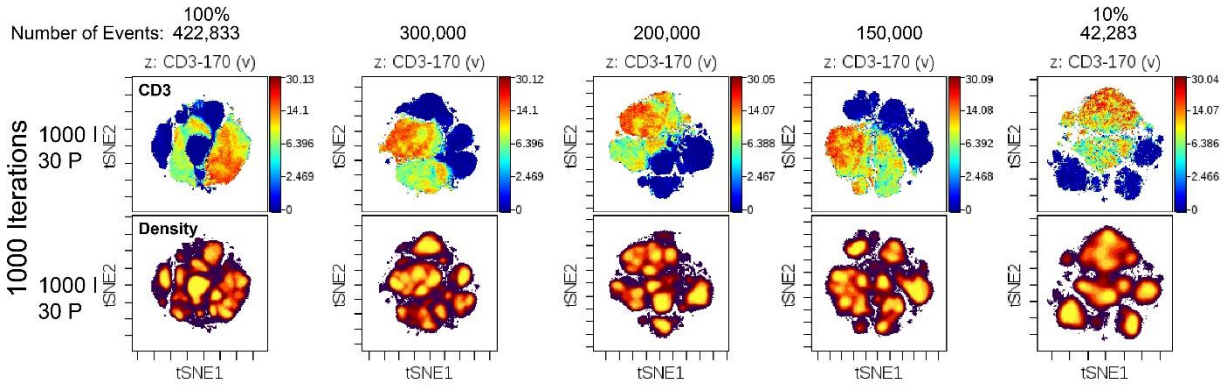



*Figure 4-1 High cell numbers inhibited cell separation in the viSNE map*
*1000 iterations (I) were run for each of five different numbers of cells from the same experimental PBMC dataset. The perplexity (P) was set at the recommended value of 30. The top row shows the heat on CD3 identifying the T cells. The bottom row shows the density of the cell distrubution on the map.*

The large PBMC dataset was subsampled and new viSNEs were run to see if the maps visually improved the separation of the cell population. The maps looked

qualitatively superior as the cell number dropped. When the purity binning matrix was applied and the scores were calculated, it was observed that the average percent purity increased with fewer cells and decreased with more cells (Figure 4-2). The viSNE maps look good and the average purity score is highest when only 100,000 events were analyzed.

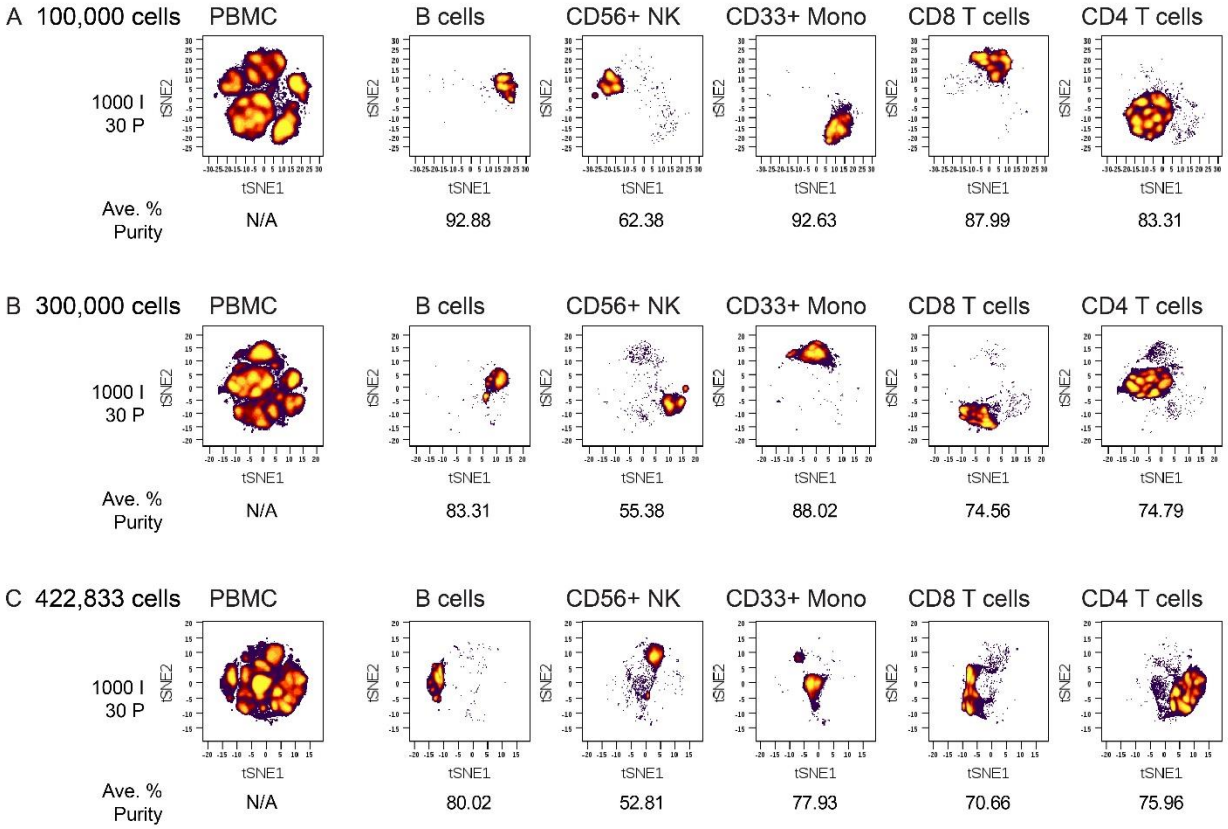



*Figure 4-2 Average percent purity scores dropped as cell number increased with iterations held constant. A) 100,000 cells, B) 300,000 cells, and C) 422,833 cells were run in separate viSNE analyses and average purity scores were calculate for each.The coloumns show the viSNE maps of the different cell populations and the average percent purity scores for each expert gated cell type. Rotation in the viSNE map is normal and does not influence map integrity. Every new viSNE run starts cells at random points (Amir el, Davis et al. 2013).*

Since ignoring a large portion of a dataset is not an ideal solution, increasing the number of iterations was tested to see if that improved the quality and the average purity of the viSNE map.

**Increasing the number of iterations improved the visualization of the map more than it altered the average purity score.**

300,000 events were chosen to be run at 1000, 5000, and 10,000 iterations. The full data set was not chosen because the viSNE run often canceled out when analyzing the whole dataset at 10,000 iterations. The increase in iterations improved the visualization of the map integrity (Figure 4-3). This also makes on-viSNE gating easier since population cut points are more clearly delineated by an absence of cells (white space). The purity scores do not show a consistent increase at 10,000 iterations and were actually lower or the same at 5000 iterations.

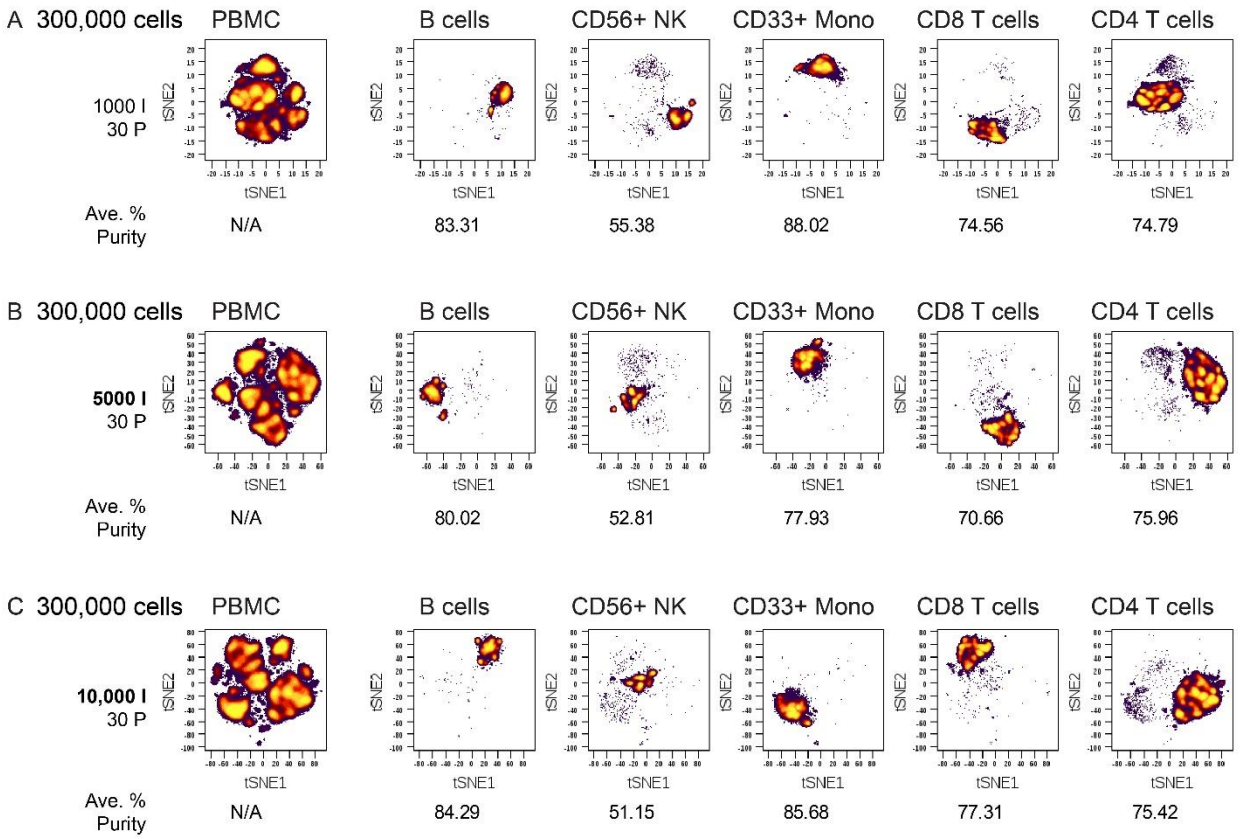



*Figure 4-3 The visualization of the map improved with increased iterations.*
*The same 300,000 events were analyzed by three separate viSNE with varying numbers of iterations, A) 1000 I, B) 5000 I, and C) 10,000 I. The average purity scores were calculated for each expert gated population. A) is the same data as in Figure 4-2B).*

At 10,000 iterations, the total PBMC map reveals clear spacing between the five expert gated populations. Although, the purity score was not consistently increased at 10,000 iterations, that map is superior in its visualization of the cell populations. Visualization to the human eye should be considered in addition to the purity score of map integrity.

There is a balance to be struck when analyzing cells with viSNE. Too many cells and the purity will be poor with cellular populations overlapping and the cell distribution on the map will not be biologically informative. Increasing the number of iterations will compensate and improve the visualization of the viSNE map. More iterations do not appear to negatively impact the data, although it takes more time and is computationally intensive and may result in a failed run. If running a large number of events is not necessary, fewer events will provide a better map. Although there is always the worry the too much down sampling of the dataset will not provide accurate results. viSNE analyses should be run in duplicate or triplicate and several input parameter (cell number and iteration number) settings should be tested for each data set. This will help insure that findings are reproducible and that the correct input parameters are being used for the dataset. The findings from this computational study were applied to the viSNE analysis performed on the FL data in Chapter 2. It is important to always be aware of the total number of cells per viSNE run and the number iterations. I recommend at least 1000 iterations for every 100,000 cells.

# CHAPTER 5 CONCLUSION

## Follicular lymphoma

This mass cytometry study extends knowledge of intra- and inter-tumor heterogeneity that has been observed in prior studies of follicular lymphoma cell genetics, phenotypes, and functional capabilities. We have now characterized the non-malignant B cells in FL at the single cell level. This type of characterization had not been done previously. Understanding the non-malignant cells present in the tumor could impact treatment choices especially immunotherapy choices dependent on forming an adaptive immune response. It remains to be seen if the diminished GC B cell population can be activated to fight the malignant cells. I found viSNE to be a reliable tool to separate malignant and non-malignant cells. Because it takes in so many parameters, 20 here, it is more efficient at differentially mapping malignant and non-malignant B cells than it would be to perform successive manual gating on biaxial plots to pull out the non-malignant cells. I attempted this approach with limited success before the gating strategy became too complex and too many cells were lost. This unbiased stratification of cells performed by viSNE, allowed me to include the non-malignant cells of the same light chain as the malignant cells into the analysis. More importantly, it enabled me to exclude non-malignant cells of the same light chain from the malignant cell analysis. This is a problem in many past studies that cells of the malignant light chain were included in the malignant cell analysis. This can bias the results of what proteins malignant cells actually express, such as IgD (Grier, Al-Quran et al. 2012). Grier et al. used the malignant restricted light chain to mark FL malignant cells for flow cytometry analysis. They found FL malignant cells to include a low level of IgD in some samples. I would argue that these IgD positive cells were non-malignant cells of the same isotype that were grouped in with

the malignant cells for lack of a way to exclude them. With viSNE, we no longer have to settle for light chain restriction as the means to identify malignant cells. viSNE incorporates more information and can plot the malignant and non-malignant cells apart from each other in tSNE space allowing for continued phenotypic analysis as was performed in Chapter 2.

Due to the high dimensional nature of the data, additional computational experiments can be performed without collecting new data. For starters, the HLA-DR variability among the malignant B cells (Figure 2-11) can be further analyzed by gating the HLA-DR positive and negative cells and performing a phenotypic comparison of the two groups such as in Figure 3-3. The gates could be drawn to include all the cells or only the highest and lowest expressers depending on the question asked. Marker enrichment modeling (MEM) (Appendix E) could be added in addition to median shown in Figure 3-3 to display the expression differences of the other markers. The high and low populations could be compared to each other or to a reference population such as healthy GC B cells. MEM works well because it mutes the similarities between the populations and allow the differences to show as in Figure 5-3B. Here the MEM heatmap shows no change for CD45 since all populations express CD45.

I recommend a FL follow up study with a larger patient cohort (including a training and testing set) that includes patient outcomes and patient paired uninvolved lymph nodes. In the case that uninvolved lymph nodes could not be acquired, matched peripheral blood should be collected. The high level of heterogeneity observed at the single cell level hints at malignant B cell biomarkers that could be used to predict treatment response. The tumors studied in Chapter 2 were excised pretreatment.

Phenotyping of tumor tissue over the course of treatment could provide valuable data about which tumors and patients will respond best to different treatment options.

The lack of germinal center B cells in the follicular lymphoma lymph node tumors should be explored. Understanding or reversing the diminished GC B cell population could lead to new treatments in lymphoma. Activating the GC B cells in the tumor could lead to an anti-tumor B cell and antibody pool that could help combat the disease. The GC B cells may be minimally represented due to, 1) the malignant cells using up necessary resources for the healthy GC reaction, 2) the malignant cells stunting the GC reaction through cell-cell contact or 3) secretion of inhibiting molecules or proteins, or 4) little to no GC activity was present prior to the arrival of the malignant cells and the malignant cells are not causing a localized adaptive immune response.

The hypotheses that the malignant cells may be using up resources or inhibiting GC development can be tested *ex vivo* with several different co-culture experiments. B cells can be cultured and clusters of blasting B cells can be observed with the naked eye. The culture of healthy primary B cells from tonsil or peripheral blood has been well established (Tangye, Ferguson et al. 2002, Huse, Bakkebo et al. 2011). T cells from FL tumors and reactive lymph nodes have been successfully co-cultured. Researchers found that the gene expression of reactive lymph node T cells could be changed to reflect that of tumor-associated T cells when cultured with FL cells (Scott and Gascoyne 2014). This means that malignant B cells can survive long enough in culture to have an effect on the cells they are co-cultured with. Malignant B cells could be cultured and then the supernatants collected and applied to GC B cells. Readouts could include, mass cytometry phenotyping, phospho-flow (as in Chapter 3), viability staining via trypan blue for light microscopic analysis or a fluorescent membrane exclusion dye for fluorescence

flow cytometry. A co-culture with membrane separation may be needed to determine if any inhibition in GC reactions was caused by secretions or cell-cell contact. If there truly is no inhibition and the GC B cells just are not being activated by the malignant B cells, then stimulating the bulk tumor cells should reveal a GC response in the non-malignant B cells. Then the problem becomes, can these B cells be activated to act against the malignant cells. In a cancer-immunity cycle model, B cells act as antigen presenting cells (APCs) in presenting antigen to CD4 T cells. This would ideally lead to B cell activation and potentially anti-cancer antibodies. However, the role of B cells in cancer is contentious with some studies showing B cells to promote tumor growth and other studies showing the opposite(Wang, Zhang et al. 2016).

Depending on the results, strategies could be employed to bolster the non-malignant GC B cell response. This may include targeting the T cells or innate immune cells. The key would be to understand the mechanism behind the diminished GC compartment.

I hypothesized that viSNE analysis would spatially separate malignant B cells from non-malignant B cells. Chapter 2 supported this model. The same was not true for T cells although I hypothesized that it would be. The T cells were characterized with their own T cell focused panel. However, this phenotyping was not enough to distinguish T cells from tumor and tonsil. Other researchers have observed T cell differences in signaling which this T cell panel was not designed to capture (Myklebust, Irish et al. 2013). The automated detection of malignant B cells could be further explored to develop a machine gating algorithm that would take the guess work out of human analysis and provide prompt unbiased diagnostic information.

## Redox and hypoxia in B cells

Chapter 3 showed that redox signaling through the BCR pathway was increased for some GC B cells. Additionally we saw increased phosphatase expression in what we suspect were light zone GC B cells (Polikowsky, Wogsland et al. 2015). Recent papers from other groups have shown that light zone GC B cells have higher levels of hypoxia (Cho, Raybuck et al. 2016). There is ample data on redox and hypoxia in cancer research that might be applicable to B cell germinal center responses. ROS can stabilize HIF1a in hypoxic conditions which improves cell survival (Jung, Yang et al. 2008). Hypoxia and redox responses might all be tied to antigen presentation and the ability to get T cell help in the germinal center reaction. It might be that follicular dendritic cells (FDCs) are creating the hypoxic environment due to their high $O_2$ need. GC B cells are subject to poor survival unless they receive T cell help. B cells can produce their own $H_2O_2$ (Reth 2002) which has been shown to have chemo-attractant properties to recruit other cells such as T cells (Rojkind, Domínguez-Rosales et al. 2002, Hara-Chikuma, Chikuma et al. 2012). Maybe B cell survival is not just about making a good antigen receptor (BCR) but also about being able to physically draw a T cell close via $H_2O_2$ chemotaxis. The $H_2O_2$ attractant properties might also play a role in keeping the B cell in the light zone since a successful B cell needs to bind to a T cell for a measurable length of time (Okada, Miller et al. 2005). If a B cell needs a functional BCR in order to produce $H_2O_2$ as Reth 2002 suggests, then that would provide another controlling factor in affinity maturation.

## Ever advancing technologies

In Chapter 4, I explored the input parameters iterations and cell number. I found that increasing iterations or decreasing cell number have the ability to improve the map

integrity when working with a large data set. When viSNE was first introduced, it was only available through MATLAB and had a cell input maximum of 100,000 (Amir el, Davis et al. 2013). At the time, 1000 iterations were enough to maintain map integrity. When viSNE started to be applied to multiple platforms and the cell cap was removed, problems started to arise with balling up of the data with large data sets. I began to formally address that issue in Chapter 4. It is important to continually reevaluate computational tools as their usage changes to ensure that the results are biologically relevant.

The mass cytometry studies discussed herein shed light on multiple new findings and bring about even more questions. The systems biology approach is good for hypothesis generation due to the breadth of molecules studied. The beauty of mass cytometry data is that there is always more to discover within the same data set. Often, I have gone back to reanalyze my data after reading a literature and hearing a relevant lecture. During my time as a graduate student, the mass cytometry field has grown from being unknown, perceived as complicated, and even thought of as unnecessary, to being a sought-after tool for answering questions in a broad range of biological disciplines. This technology and quantitative single cell technologies like it are the future of systems biology. Along with these tools, comes the endless need for algorithms to assist in analyzing this data. Visualization algorithms are key to making the data visually accessible to the researchers collecting it. I look forward to working with the next generation CyTOF, the Helios, that has a larger mass range to detect more parameters. Fluidigm continues to release antibodies conjugated to new metals as the chemistry evolves. The introduction of metal conjugated RNA probes allows for the simultaneous detection of RNA and protein product from a single gene at the single cell level (Frei, Bava et al. 2016). This will help answer many questions about the nature of gene

regulation such as timing and RNA to protein ratios. I suspect the ratios will be different for different genes and vary across cell types and probably during the cell cycle. Fortunately, there are mass cytometry tools for cell cycle study and replication studies (Behbehani, Bendall et al. 2012).

With the introduction of imaging mass cytometry, cellular data with 30+ parameters can be measured while maintaining cell positional information (Giesen, Wang et al. 2014). Imaging of peripheral lymphocytes could be interesting, especially if the cells have been stimulated to form immune synapses, however the real power here is in solid tissue. The ability to achieve single micron resolution of tissue with over 30 antibody markers has the potential to reveal much about protein expression in tissue compartments and tumors. This could be highly valuable for inspecting FL resected tumors. The variations in HLA-DR expression on the malignant B cells might be due to a malignant GC light zone/ dark zone-type localization or could be due to local necrosis. How are the non-malignant B cells oriented? Are they spread throughout the tumor or pushed to the outside away from the bulk of the malignant B cells? I suspect that they do not have their own niche or the GC population would not be so diminished.

## Allergic disease

My project started out by asking what differences existed between the mature B cell compartment of allergic and non-allergic individuals. Literature searches did not turn up any in depth phenotype comparisons between the peripheral blood of non-allergic and allergic individuals so I decided that would be a good place to start. I hypothesized that there would be an obvious B cell phenotype in atopic patients and with a 34-marker panel I was well posed to find it. What I did not count of was the complexity of characterizing

non-atopic normal cells at the level of phenotyping. With so many markers and so many cells, it quickly became clear that normal needed to be characterized and the tools to do so needed to first be optimized in order to effectively apply mass cytometry to the study of B cells in allergic disease. This led to a pause in allergy research while I tested and optimized the dimensionality reduction tools viSNE and SPADE on control cells. Tool building compounded with a substantial decrease in clinical pediatric blood draws, led my dissertation research away from allergic disease.

Although I have no significant findings to report from my pilot study, I would like to outline my hypotheses and thoughts that may be useful to others down the road. B cells have proven to be a useful target in other diseases and I believe the same is true in allergic disease. Anti-CD20 and rapamycin are blunt tools but may work to provide relief in especially severely allergic patients. Diminishing the B cell pool or disrupting B cell function along with the removal of preexisting IgE with anti-IgE (D'Amato, Salzillo et al. 2007) could go a long way to mitigate life threatening allergic reactions.

I hypothesize that a defect in B cells that might be causing them to class switch to IgE more robustly could be detected by perturbing the B cells from allergic and non-allergic individuals with factors that push B cells towards (IL-4, IL-13, IGF-1) an IgE cell fate and factors that pull B cells away (IFNα, IFNγ, PGE2, IL-12, CpG, TGFβ, IL-10) from an IgE cell fate (Pene, Rousset et al. 1988, Gascan, Gauchat et al. 1991, Gascan, Gauchat et al. 1991, Punnonen, Aversa et al. 1993, Punnonen, Yssel et al. 1997, Jeannin, Lecoanet et al. 1998, Koh, Park et al. 2000, Gould, Sutton et al. 2003, Mangan, Fallon et al. 2004, Kumar, Verma et al. 2012). I suggest perturbing the system with stimulating factors such as IL-4 (increase class switch to IgE isotype) and IL-21 (pushes B cells towards an antibody production fate) or inhibiting factors such as TGFβ which inhibits

class switch to IgE isotype. If B cells in allergic individuals possess a defect, it may become apparent when stimulating and observing the readouts of phosphorylation in both intensity and duration, proliferation, cell survival, and ultimately class switch.

It is entirely possible that there is no B cell defect in allergic individuals. It may be other cell types or micro environmental influences that determine the existence of pathogenic IgE. Finding the root cause is important for understanding allergies. Regardless of the cause though, B cells play a role as they are the cell types that eventually produce the pathogenic and often life threatening IgE. Understanding B cell biology and being able to fine tune and direct class switch would be invaluable for stop allergic pathogenesis by blocking the class switch to IgE or inhibiting the survival of IgE positive cells.

### Autoimmune diseases

B cell mass phenotyping in autoimmune diseases would be an excellent avenue to explore for new drug targets. We know that B cells are involved in autoimmune diseases. B cells become plasma cells that secrete autoantibodies. Regulatory B cells (Bregs) are also involved in disease and should be included in analysis. With the backbone of the B cell panel in place, swapping markers in and out makes the B cell panel flexible for studying Bregs and other B cell types. The flexibility of the panel also allows for tissue specific markers to be added. It is important that the panel can identify B cells in multiple activation states (chemokine receptors and activation markers) and distinguish and positively identify non-B cell populations, such as CD34 on HSCs in the bone marrow.

## Summary

The systems biology approach using mass cytometry to study B cell biology allows for simultaneous identification and characterization of at least 4 major B cell populations in health and disease settings. This in-depth simultaneous approach was not possible before mass cytometry. Mass cytometry enabled the identification and characterization of naïve, GC, memory, plasmablasts, and malignant B cells simultaneously. In Chapter 2, I was able to ask multiple questions with a single data set. I found that 1) the GC B cells were diminished in FL tumors, 2) malignant B cells did not phenotypically resemble GC B cells, 3) the malignant B cells had a highly variable phenotype, and 4) the malignant cell phenotypic variability was driven by diverse expression of HLA-DR surface molecules within the tumors. The most significant contribution that came out of the FL study was the in-depth phenotypic characterization of the non-malignant cells at the single cell level. The non-malignant B cells had not been studied so carefully before.

In Chapter 3, we used mass cytometry to study the kinase and phosphatase activity of the 4 main populations. That study started out using fluorescence flow and looking at phospho-responses from ROS stimulation. We noticed a subset of cells that were responding robustly to the ROS stimulation and had to move platforms to mass cytometry in order to add enough markers to identify the cells of interest. The move to mass cytometry enabled us to identify the GC B cells as the robust responders.

In Chapter 4 and throughout this dissertation, I used and optimized the analysis tool viSNE to display complex data in a comprehensible way. It would be a shame to not use the data from the 30+ parameters but it would be incomprehensible to try and view it all at once. viSNE enabled rapid visualization of cell phenotypes as defined by the tSNEs. This allowed for large differences to be easily spotted and followed up on, such as the

malignant cells in Figure 2-1 and the lack of GC B cells in the FL tumors (Figure 2-9). This work contributed not only biological findings but computational insight and approach to our scientific knowledgebase.

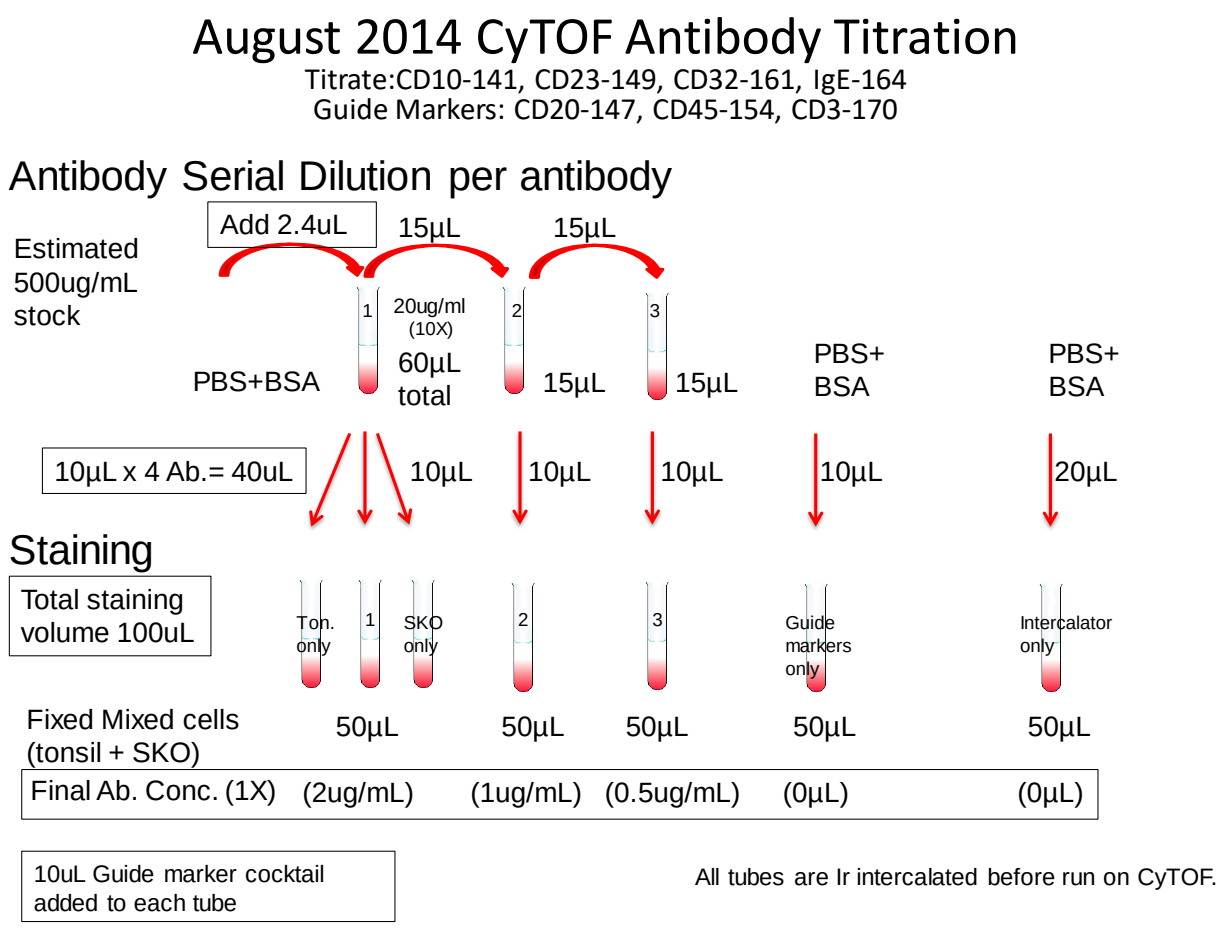



*Figure 5-1 Example schematic of serial dillutions for antibody titration.*
*It is important to always have a negative and positive cell population so that the appropriate antibody concentration can be determined. Subsets of B cells will be positive for CD10, CD23, and/or CD32 and T cells will be negative for these markers. CD20 positively identifies B cells. CD3 positvely identifies T cells. SKO cells will be positive for IgE and PBMC will be negative for IgE. PBMC can be discriminated from SKO cells by their higher level of CD45 expression.*

# APPENDIX B : WHOLE BLOOD PROCESSING PROTOCAL

Written by Cara Wogsland 2014-11-18

**Prepare BUFFY COAT and Plasma**
Put PBS, RPMI+FBS, FBS, and Freeze media at RT
In Green top tubes *Spin down RT, 600xg, 20 min, NO BRAKE*
Pipette plasma to new 15 mL conical down to about 1 cm above buffy coat

**Ficoll Gradient Separation:**
Pipette buffy coat and top layer of RBC into 50 mL conical
Dilute by 1⁄4 with PBS
Add 14 mL of Ficoll to new 50 mL conical
Slowly add the diluted buffy coat on top of Ficoll
*Spin down RT, 400xg, 30 min, LOW BRAKE*
Plasma: *Spin down RT, 1600xg, 10 min, Brake on* to remove platelets
Store plasma in 0.5-1 mL aliquots at -20°C, be careful to leave platelets behind
Ficoll prep after spin: Aspirate down to 3 cm above PBMC layer

| **PBMC layer** | **Granulocyte layer** |
|---|---|
| Pipette the PBMC layer located **above** the Ficoll layer into a labeled 15 mL conical. | Pipette granular cells located **beneath** Ficoll layer into a labeled 15 mL conical. |
| Fill tube with RPMI+FBS to wash | ACK lyse with 10 mL ACK lysis buffer for 4-5 minutes |
| | Top off tube with RPMI+FBS to wash |
| *Spin down RT, 200xg, 10 min, Brake on* | *Spin down RT, 200xg, 10 min, Brake on* |
| Aspirate | Aspirate |
| Wash again in 15 mL RPMI+FBS | ACK lyse again for 1 min with 5 mL lysis buffer |
| | Fill tube with RPMI+FBS to wash |
| *Spin down RT, 200xg, 10 min, Brake on* | *Spin down RT, 200xg, 10 min, Brake on* |
| Aspirate, resuspend in 1 mL FBS to count | Aspirate, resuspend in 1 mL FBS to count |

(Both tubes can be spun together)

**PBMC and Granulocytes Count and Freeze:**
Pipette 10 µL cells into 90uL PBS for counting, for both cell types
Bring cells up in enough FBS for 500 µL per tube and 5-6 million PBMC per tube, 10 million gran. per tube
Add 500 µL 2X freeze media to the bottom of each vial
Aliquot 500 µL of cells to separate cryovials for PBMC and Granulocytes
Freeze in Mr. Frosties or similar at -80°C. Move to $LN_2$ within a week.

**Remaining PBMC in FBS (Optional)**
Bring up in 1 mL PBS
Fix with 100 µL 16% PFA for 5 minutes
Wash with 1 mL PBS
*Spin down RT, 1000xg, 5 min, Brake on*
Decant, Vortex
Bring up in 1 mL cold MeOH
Vortex vigorously
Cap and store at -80°C in 15 mL conical
To be used for fluorescent phenotyping

# APPENDIX C : FLUORESCENCE FLOW STAINING PROTOCOL

## Fluorescent phospho-flow primary samples thaw to staining protocol

Written by Cara Wogsland, last updated 2015-11-12

**Media: warm RPMI+10%FBS**
**Staining Media: cold PBS+1%BSA or RT BV Stain Buffer***

**Staining cocktail(s):**
Make enough cocktail for the number of samples, plus add an extra 10% staining media for slop. Make separate staining cocktail for secondary antibodies. Transfer cells to new tubes for each staining cocktail so that volumes are precise. This is especially important for comparing between samples since staining is concentration dependent. (Staining cocktail calculations can be done in an Excel spreadsheet)

**Thaw:**
Warm cryovial of PBMCs in warm water until thawed
Add 1 mL warm RPMI+FBS to thawed cells. Dump into 15 mL conical containing ~8 mL warm media.
Rinse cryovial and dump into 15 mL conical x2
*Spin down RT, 200xg, 5 min*
Resuspend and count as needed
Keep cells in incubator while counting

*Rest cells in incubator if performing live cell assays such as phospho-flow*

**Stimulation:**
Bring up cells in enough warm media for all conditions. (200 µL per condition)
Aliquot 200 µL of cells per well into plate based on the number of conditions.
Rest in incubator for 10-15 minutes.
Reverse time course based on chosen conditions and time points. (Start with longest time point and count down. Fix all cells together.)
Put cells back in incubator between each time point. All tubes should move in and out together to minimize variables.
All conditions should be fixed at time 0 when the timer goes off.
Remember: two unstim conditions are better than one.

**Stims are used at 50X, add 4 µL of stim to 200 µL of cells:**

**Example time course:** 5 conditions, N=3, 15 wells total

| | Stim1 (15min) | Stim2 (15min) | Stim2 (10min) | Unstim1 | Unstim2 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sample1 | | | | | | | | | | | | |
| Sample2 | | | | | | | | | | | | |
| Sample3 | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |

*Figure 5-2 Example layout of a 96 well plate for stimming cells for a phospho-flow experiment.*
*White empty spaces have cells. Grayed out spaces are considered unused for the purpose of this figure.*

**Fix:**
Add 20 µL 16% PFA to 200 µL of cells (media will turn yellow)
Mix by gentle pipetting
Incubate for 10 min at RT.
Wash with 50 µL PBS
*Spin down RT, 900xg, 5 min* (dead cells are always spun faster than live cells since they do not pellet as well)
Decant, pat on paper towels
Wash with 200 µL PBS, mix by pipetting
*Spin down RT, 900xg, 5 min*

**Permeabilize:**
Decant, pat on paper towels
Resuspend the cells in the residual volume left after decanting by vortexing gently on side of plate.
Add 200 µL of cold methanol (-20ºC) with multi-channel pipet and immediately mix repeatedly by pipetting
Continue to pipet as needed to break up clumps
Incubate cells at -20ºC for at least 10 minutes.
Cells can be left overnight at -20ºC or for weeks at -80ºC. Cover plate so methanol does not evaporate. Wrap in parafilm if stored for longer time periods.

**Intracellular Stain (60 µL total staining volume):**
*Use BV Stain Buffer as staining media if using 2 or more Brilliant Violet dyes in one panel.
Add 70 µL PBS to wash
*Spin down RT, 900xg, 5 min*
Decant, pat on paper towels
Wash with 200 µL staining media, mix by pipetting
*Spin down RT, 900xg, 5 min*
Decant, pat on paper towels
Add 30 µL staining media to each well, mix by pipetting
Transfer 25 µL of cells to new wells for staining
Add 35 µL antibody staining cocktail, mix by pipetting
Stain in dark for 15-30 min (make comp tubes during this time)
Add 200 µL BV stain buffer, transfer wells to labeled FACS tube
Wash each well with 200 µL PBS+BSA and transfer to appropriate FACS tube
Wash cells in FACS tubes with additional 1 mL PBS+BSA
*Spin down RT, 900xg, 5 min*
(Repeat staining as needed for secondary antibodies.)

**Running on Fluorescent Flow Cytometer**
Decant
Bring up in 110 µL PBS.
Store in fridge until ready to run. (Can be run next day but should ideally be run the same day.)

**Compensation tubes are made for each antibody and 1 unstained tube using comp beads**
In labeled FACS tubes add:
100 µL PBS
10 µL positive-negative bead mix (vortex beads well before adding)
Add ¼ the antibody used per 100 µL test.

The positive comp beads in the Irish lab are anti-**mouse** kappa chain. If you stained with antibodies from other animals, you need to find an antibody produced in mouse with matching fluorophore to add to the comp beads. Comp fluorophores should be the actual antibody when using a tandem fluorophore since those can break down over time and show differences between sources.

Comp Beads do not get washed or spun down. They are ready to run through the cytometer as soon as they are made. They should be kept in the dark along with the samples once they are made. They can be stored at 4°C for at least a week.

# APPENDIX D : CYTOF MASS CYTOMETRY STAINING PROTOCOL

CyTOF Primary samples thaw to CyTOF staining protocol
2014-08-26

**Media: warm RPMI+10%FBS**
**Staining Media: cold PBS+1%BSA**

**Thaw**
Warm cryovial of PBMCs in hands or warm water until thawed
Add 1 mL warm RPMI+FBS to thawed cells. Dump into 15 mL conical containing ~8 mL warm media. Rinse cryovial and dump into 15 mL conical
Wash cryovial with warm media.
*Spin down RT, 200xg, 5 min*
Resuspend and count as needed

**Staining cocktail(s)**
Make enough cocktail for the number of samples, plus add an extra 10% PBS+BSA for slop. Make separate staining cocktails for surface and intracellular markers. Make separate staining cocktails for secondary antibodies. You may have up to four staining panels. Transfer cells to new tubes for each staining cocktail so that volumes are precise. This is especially important for comparing between samples since staining is concentration dependent.

**Live Surface Stain**
50 µL stain (double volumes for 100 µL test size)
Staining media: cold PBS+1%BSA.
Cells and antibodies should be in staining media.
Transfer 30 µL of cells to a new FACS tube.
Add 20 µL of premade staining cocktail. (Titrate antibodies ahead of time. In general antibodies can be used at half the recommended concentration. E.g. 0.25 µL into one 50 µL test)
Vortex
Stain RT for 15-30 minutes.
Wash with 1-2 mL PBS+BSA
*Spin down RT, 200xg, 5 min*
(Repeat staining as needed for secondary antibodies.)

**Fix**
Decant
Bring cells up to 200 µL with PBS.
Fix with 20 µL 16% PFA (for a final concentration of 1.6% PFA)
Vortex.
Incubate for 10 min at RT.
Wash with 1-2 mL PBS
*Spin down RT, 900xg, 5 min*

**Permeabilize**
Decant
Resuspend the cells in the residual volume left after decanting by vortexing vigorously.
Add 1 mL of cold methanol (-20ºC)
Vortex immediately.

Pipet as needed to break up clumps
(If a lot of clumping is observed, add 100 µL PBS to next tube before adding methanol. Then bring up in 2 mL cold methanol.)
Incubate cells at -20ºC for at least 10 minutes.
Cells can be left overnight at -20°C or for weeks at -80°C. Cover tubes for longer times so methanol does not evaporate.

**Methanol Removal**
Wash the cells that are in methanol with 2mL PBS
Vortex
*Spin down RT, 900xg, 5 min*
Decant
Repeat wash 1 more time
*Spin down RT, 900xg, 5 min*

**Intracellular Stain (if applicable)**
Resuspend cells in ~40 µL in staining media.
Transfer 30 µL of cells to a new FACS tube.
Add 20 µL of premade staining cocktail.
Vortex
Stain for 15-30 minutes.
Wash with 2 mL PBS+BSA
*Spin down RT, 900xg, 5 min*
(Repeat staining as needed for secondary antibodies.)

**Nucleic Acid Stain**
Decant
Bring cells up in 200 µL PBS
Add 4 µL 50X Iridium nucleic acid intercalator.
Vortex
Incubate 15 minutes at RT. Cells can be left in fridge for several hours.
Wash with 2 mL PBS right before running on the CyTOF.
*Spin down RT, 900xg, 5 min*
Add another 2 mL PBS
Vortex

**Filter cells**
Use a P1000 to facilitate filtering cells. Push tip against filter of a filter top FACS tube and push the cell suspension through. Remove caps when done.
*Spin down RT, 900xg, 5 min*

**Running on CyTOF**
Decant
Vortex void volume
Bring up in 1X normalization beads
Volume varies: A good place to start is 600 µL for a small pellet, 1 mL for a large pellet.
Run 200 µL first to test concentration. Adjust as needed. Streaking and FIFO errors may mean that cells should be diluted further. If you have over 10 million cells, make a separate 1:5 dilution to run first to test the concentration.
**Never run all of your sample at once.**

# APPENDIX E : CHARACTERIZING CELL SUBSETS IN HETEROGENEOUS TISSUES USING MARKER ENRICHMENT MODELING (MEM)

Full title: Characterizing cell subsets in heterogeneous tissues using marker enrichment modeling (Diggins, Greenplate et al. 2017)

Kirsten E Diggins, Allison R Greenplate, Nalin Leelatian, Cara E Wogsland & Jonathan M Irish

## Abstract

Learning cell identity from high-content single-cell data presently relies on human experts. We present *Marker Enrichment Modeling* (MEM), an algorithm that objectively describes cells by quantifying contextual feature enrichment and reporting a human and machine-readable text label. MEM outperforms traditional metrics in describing immune and cancer cell subsets from fluorescence and mass cytometry. MEM provides a quantitative language to communicate characteristics of new and established cytotypes observed in complex tissues.

## Introduction

Marker enrichment modeling (MEM) scores are a useful augmentation to median values that cut through the noise and erroneous results to show where population phenotypes truly differ from each other. MEM labels are an unbiased characterization of cell population phenotypes made using the MEM scores. MEM scores are valuable in visualized form in a heatmap alone or in conjunction with median values. MEM labels help to identify and characterize pre-existing cell populations. Any approach can be used to create the cell populations that will be input into the MEM algorithm including clusters, biaxial gating, and on-viSNE gating. This makes MEM useful as part of a workflow for

data analysis. I have been intellectually involved with MEM since its inception. I was the first to challenge it and argue with the concept and then eventually became one of its strongest supporters as I realized its power in data analysis and started running it on my own data.

I helped design experiments, discussed data visualization, and contributed intellectually to the manuscript. I beta tested the software implementation of MEM and provided feedback about what files and file formats should be generated when MEM is run. I proposed the -10 to +10 scale used for reporting MEM values to make the output values standardized and easily interpretable. I contributed to the R programming language code base implementation of MEM. I assisted with manuscript revisions.

MEM is a powerful tool to assist in the understanding of data however it is very important understand how the algorithm works and how it has been implement in each case. Since the MEM values are often reported in a heatmap, there is a risk of misunderstanding the results and assuming that MEM is like median or mean. MEM is different in that it shows the differences between the populations not the values within a population. For example, CD45 will look negative on leukocyte populations when looking at MEM values whereas it will be very high when looking at median values (Figure 5-3B).

There are several different ways to implement the MEM algorithm. The reference population is a crucial part of the MEM score. The MEM scores and labels are meaningless without knowing the reference population. The most common, and my preferred, implementation is to set the reference population as the bulk non-population cells. That means that all populations in the analysis other than the scored population will be used as the reference point for each population. Every population will be scored so this means that each population's MEM scores will be generated using a different

reference population. (This is all done automatically in the software code.) In practice, the populations are generally gated from a single sample, but that is not necessary. The populations could consist of cell lines or cancer and non-cancer for example. Using the other populations as the reference will provide the best characterization of each population within a given sample. Another way to use MEM is to have all populations run against the same reference population, such as hematopoietic stem cells (HSC).

MEM can be used to compare populations where data has been collected on different platforms, however, I stress again the importance of understanding how the algorithm works to prevent misinterpretation of the results.

**Main text**

Quantitative cytometry workflows have developed diverse approaches to grouping cells into populations and visualizing results in graphs that arrange populations based on phenotype (Diggins, Ferrell et al. 2015, Saeys, Gassen et al. 2016). Important features of populations are typically assumed to be those most highly or differentially expressed. This approach works well when feature variability is low and cells match established types, but computational analysis of single cell data routinely reveals novel cells with non-canonical phenotypes (Becher, Schlitzer et al. 2014, Patel, Tirosh et al. 2014, Greenplate, Johnson et al. 2016). This is especially common in diseases where abnormal expression profiles and signaling responses distinguish clinically significant cell subsets (Irish, Hovland et al. 2004, Irish, Myklebust et al. 2010, Gaudilliere, Fragiadakis et al. 2014, Levine, Simonds et al. 2015, Greenplate, Johnson et al. 2016). Existing statistical approaches can be used to characterize a population's degree of difference from a reference, but may be limited

to a normal distribution or may not account for intra- and inter-population variability in a single metric.

The MEM equation (Eq. 1) produces a signed value for each population feature by quantifying positive and negative, population-specific, contextual feature enrichment relative to a reference cell population (Supplementary Note 1).

$$\text{MEM score} = |MAG_{POP} - MAG_{REF}| + \left(\frac{IQR_{REF}}{IQR_{POP}}\right) - 1,\ (MAG_{POP}\text{-}MAG_{REF}) < 0 \rightarrow \text{MEM} = \text{-MEM}$$

(Eq. 1)

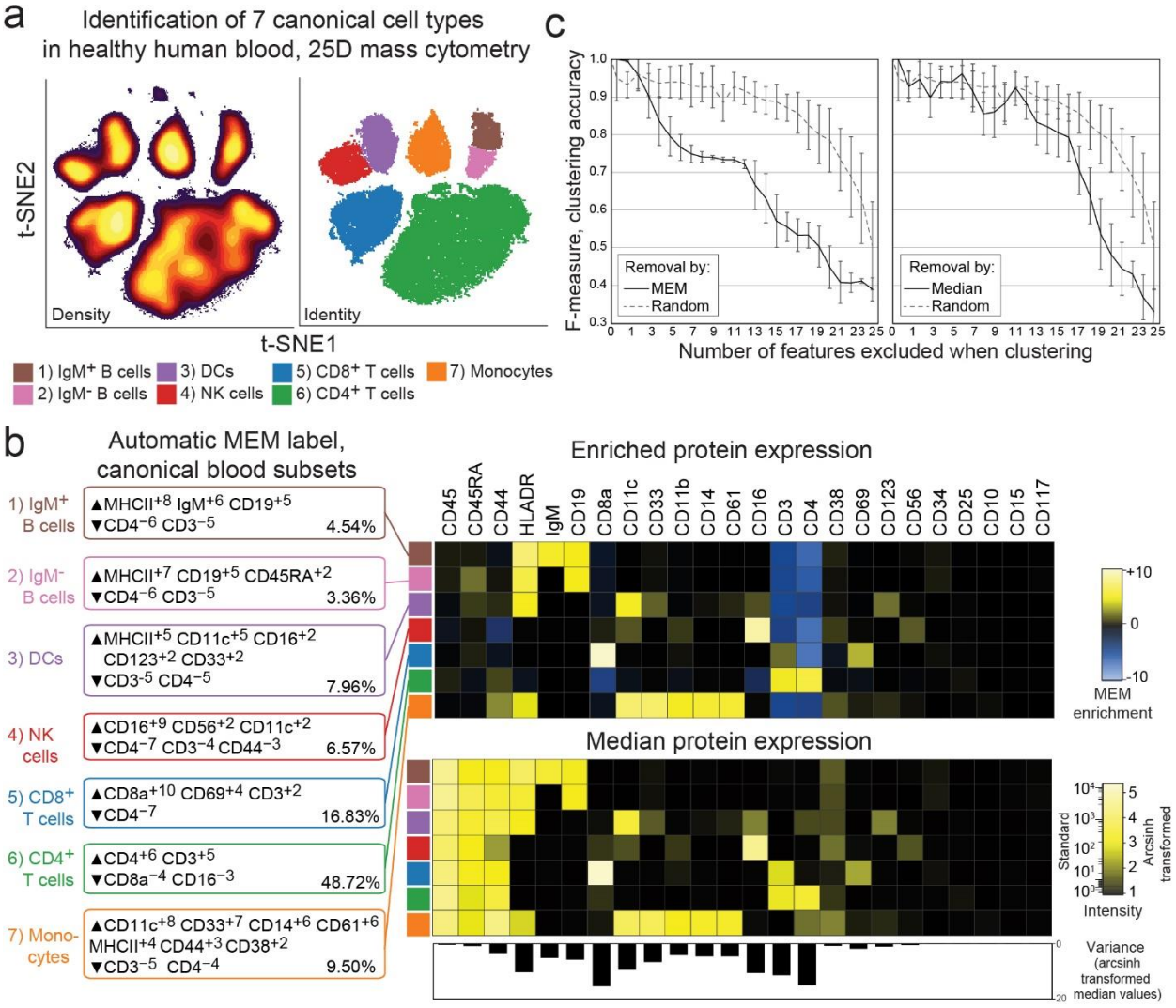
In Eq. 1, POP denotes the population of interest, REF denotes the reference population to which POP will be compared, MAG is feature magnitude (here, median protein expression detected by mass or fluorescence flow cytometry), and IQR indicates the interquartile range. A reference population (REF) is chosen based on a biological comparison of interest (Supplementary Note 1, Supplementary Table 1, Supplementary Fig. 1). MEM was designed to quantify enrichment, whereas other metrics used in cytometry, such as Kolmogorov-Smirnov (K-S) (Young 1977), area under the ROC curve (AUC) (Kim, Donnenberg et al. 2016), and Earth Mover's Distance (EMD) (Orlova, Zimmerman et al. 2016), capture other differences between frequency distributions (Supplementary Note 1). In datasets including healthy human blood, bone marrow, and tonsil, murine tissues, and human tumors, MEM identified key proteins used by experts to distinguish rare and novel cell subsets.

Four cytometry studies, Dataset A (Leelatian, Diggins et al. 2015), Dataset B (Bendall, Simonds et al. 2011), Dataset C (Becher, Schlitzer et al. 2014), and Dataset D, collected as described by Leelatian and Doxie, et al. (Leelatian, Doxie et al. 2016), were

used to evaluate the ability of MEM to identify biological features of expert and machine identified cell subsets. For datasets A, B, and C, populations had been previously identified by experts and by computational tools including viSNE (Amir el, Davis et al. 2013) and SPADE (Qiu, Simonds et al. 2011), which are used in mass cytometry for dimensionality reduction and cell clustering (Diggins, Ferrell et al. 2015), respectively.

Dataset A was mass cytometry data quantifying expression of 25 proteins on healthy human peripheral blood mononuclear cells (PBMC) (Leelatian, Diggins et al. 2015). This dataset was chosen for two reasons: 1) the 7 cell subsets present are well-established, phenotypically distinct populations that served as a gold standard of biological 'truth' and 2) the cells in each of the 7 subsets were characterized for 25 proteins that displayed varying homogeneous and heterogeneous expression patterns. Populations were expert gated following viSNE analysis and each population was compared to the other cells in the sample (Fig. 1, Supplementary Table 2). MEM returned labels that matched prior expert analysis (Leelatian, Diggins et al. 2015) and correctly assigned high positive enrichment values to canonical protein features of each subset (Fig. 1b), including CD4 on $CD4^+$ T cells (▲$CD4^{+6}$ $CD3^{+5}$ ▼$CD8a^{-4}$ $CD16^{-3}$), IgM on $IgM^+$ B cells (▲MHC $II^{+8}$ $IgM^{+6}$ $CD19^{+5}$ ▼$CD4^{-6}$ $CD3^{-5}$), CD11c and MHC II on monocytes (▲$CD11c^{+8}$ $CD33^{+7}$ $CD14^{+6}$ $CD61^{+6}$ MHC $II^{+4}$ $CD44^{+3}$ ▼$CD3^{-5}$ $CD4^{-4}$), and CD16 on NK cells (▲$CD16^{+9}$ $CD56^{+2}$ $CD11c^{+2}$ ▼$CD4^{-7}$ $CD3^{-4}$ $CD44^{-3}$). Proteins that were not significantly enriched on any of the 7 subsets of mature human blood mononuclear cells were correctly assigned near-zero MEM scores (e.g. CD34 and CD117 proteins expressed on hematopoietic stem cells, Fig. 1b). Similarly, proteins with little variability across cell subsets were assigned low, near-zero MEM scores, even for highly expressed proteins (e.g. CD45 on all subsets, CD45RA on non-T cells, Fig. 1b).

Incorporating information about feature variability allowed MEM to capture negative enrichment that was not reflected in magnitude difference ($MAG_{DIFF}$, Supplementary Note 2). Highly enriched proteins were more important to accurate population identification than proteins characterized by high median expression alone (Fig. 1c; Supplementary Fig. 2; Supplementary Fig. 3).




*5-3 Figure 1: Marker enrichment modeling (MEM) automatically labels human blood cell populations in Dataset A. a) Cells from normal human blood grouped into 7 canonical populations using viSNE analysis and expert review of 25D mass cytometry data (Leelatian, Diggins et al. 2015). b) MEM labels computationally generated for each canonical cell subset. Heatmaps show protein enrichment values used to generate MEM labels and the median protein expression values for each protein on each cell subset. Variability in protein expression across the 7 canonical cell populations is shown below to highlight proteins that were expressed homogeneously (low variability, e.g. CD45) and those that were expressed heterogeneously (high variability, e.g. CD8a, CD4). c) Graphs show decreasing f-measure (clustering accuracy) as markers were excluded from k-means cluster analysis based on high to low absolute MEM or median values, compared to random exclusion.*

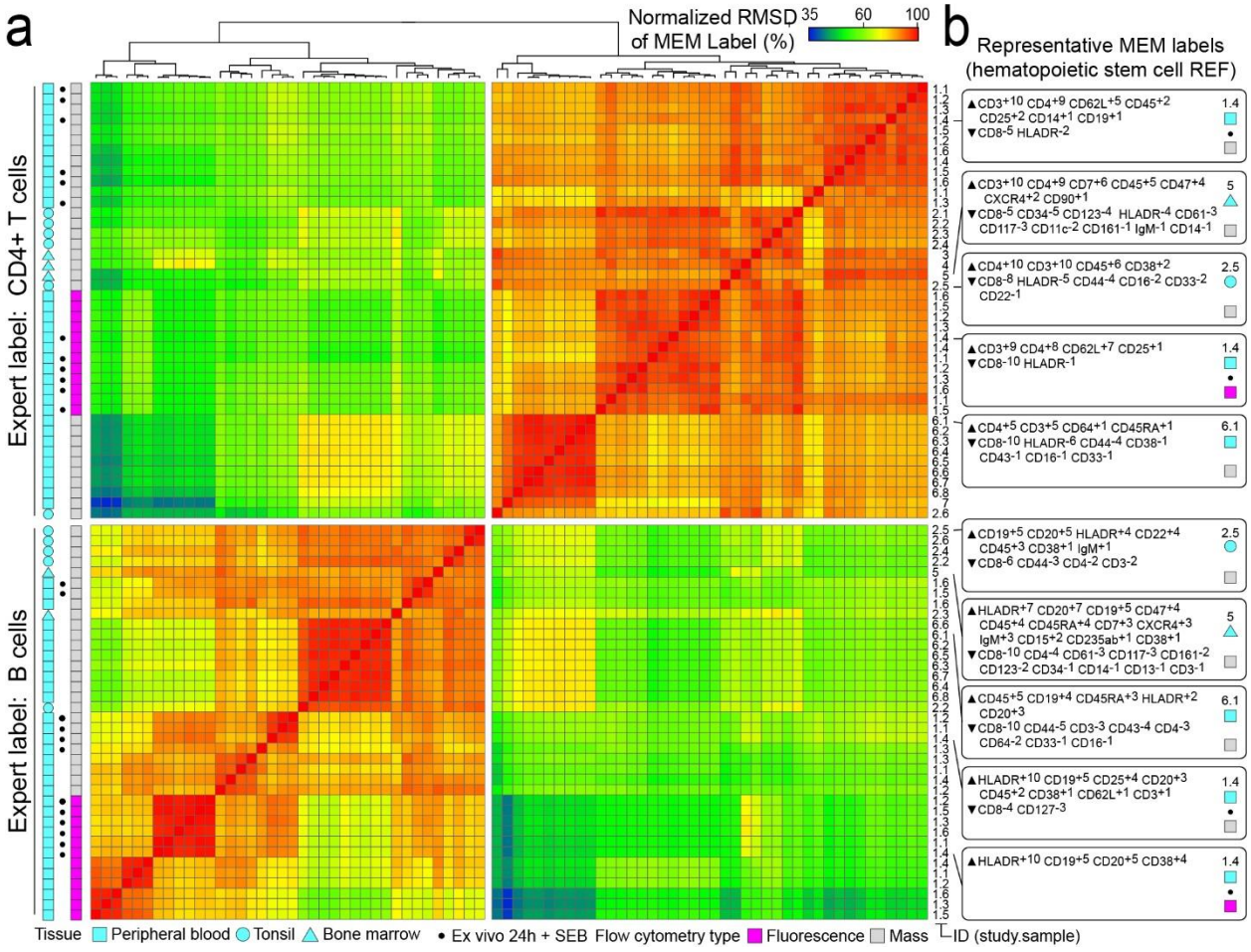
To test the hypothesis that features with high MEM scores would be important for computational cluster formation, the 25 proteins measured in Dataset A (Figure 1b) were sorted in six ways: 1) high to low MEM score, 2) high to low median value, 3) high to low $MAG_{DIFF}$, 4) high to low z-score, 5) high to low K-S statistic, and 6) randomly (Supplementary Table 3). Z-score and K-S statistic values are shown in Supplementary Table 4. The proteins were then sequentially, cumulatively excluded from use in k-means clustering and f-measure was calculated to measure clustering accuracy (Fig. 1c and Supplementary Fig. 2). The order in which markers were excluded is shown in Supplementary Table 3. Random exclusion was performed 15 times and the average result is shown (Fig. 1c). Clustering accuracy was most impacted by excluding proteins based on MEM score. F-measure dropped to 0.75 after removing the proteins with the top 6 MEM scores, whereas a comparable F-measure decrease was only observed after removing the 14 highest markers based on $MAG_{DIFF}$, the 13 highest markers based on z-score, and the 12 highest markers based on K-S statistic values (Supplementary Fig. 2). Removing markers based on median was not significantly different from removing markers randomly until the 15 markers with the highest median signal intensity were excluded (Supplementary Fig. 2). The same analysis was performed with viSNE in place of k-means clustering to visualize loss of population resolution (Supplementary Fig. 3c). In this case, loss of accuracy was reflected in the viSNE map as a loss of separation between “islands” of cells. These results indicated that MEM enrichment scores captured markers that were important to cell identity better than traditional comparisons based solely on median protein expression.

Dataset B was mass cytometry data quantifying expression of 31 proteins on healthy human bone marrow (Bendall, Simonds et al. 2011). Computational and expert analysis had previously identified 23 populations of cells that were analyzed here by MEM (Supplementary Note 3). For example, the cell subset labeled as HSCs was highly enriched for CD34 ($CD34^{+6}$) and negatively enriched for CD45 ($CD45^{-5}$). Dataset B also illustrated the general rule that MEM scores will approach median values as feature variability within populations decreases (Supplementary Fig. 4). MEM captured feature enrichment and heterogeneity better than median in diverse populations, as in Fig. 1c.

Dataset C was mass cytometry data quantifying expression of 38 proteins on murine cells from eight tissues (Becher, Schlitzer et al. 2014) (Supplementary Note 4). In this dataset, "cluster 28" was a novel population identified as $CD11b^{int}$ NK cells. The MEM label for cluster 28 within ILCs was ▲$CD11b^{+5}$ $CD62L^{+3}$ ▼$CD4^{-7}$ $CD103^{-4}$ $Terr119^{-3}$ (Supplementary Note 4 and Supplementary Fig. 5). This MEM label captured the key feature of this novel innate lymphoid cell subset ($CD11b^{int}$) and highlighted additional features that can be used to match this subset to cells identified by others (i.e., to cytotype the population). These results indicate that MEM labels complement unbiased population discovery and effectively characterize *cyto incognito* (Irish 2014) by providing unbiased descriptions that correctly capture key features of novel cell types.

An important aspect of MEM is generation of machine-readable quantitative labels that can be used to register population identities across samples and studies. A MEM label for a newly discovered population can be compared quantitatively against a reference set of established MEM labels or a MEM label reported in a paper. To illustrate this idea, the pairwise, normalized root-mean-squared distance (RMSD) of MEM scores was calculated as a measure of similarity between 80 populations of cells from 7 different

studies including healthy CD4$^{+}$ T cell and B cell (Fig. 2). Cells had highly similar MEM scores within each major cell type, regardless of platform (mass or fluorescence flow cytometry), study, or tissue source. For example, T cells run on mass cytometry from different blood donors were 97% ± 1.3 similar to each other, 85% ± 1.9 similar to T cells from blood run on fluorescence flow cytometry, and 87% ± 2.1 similar to T cells from tonsil run on mass cytometry (Fig. 2, Supplementary Table 5). However, these cells were 66.9% ± 13 similar to any B cell population. This indicates that MEM scores provide a way to communicate cell identity and to quantify similarities of cell types from the text label alone.

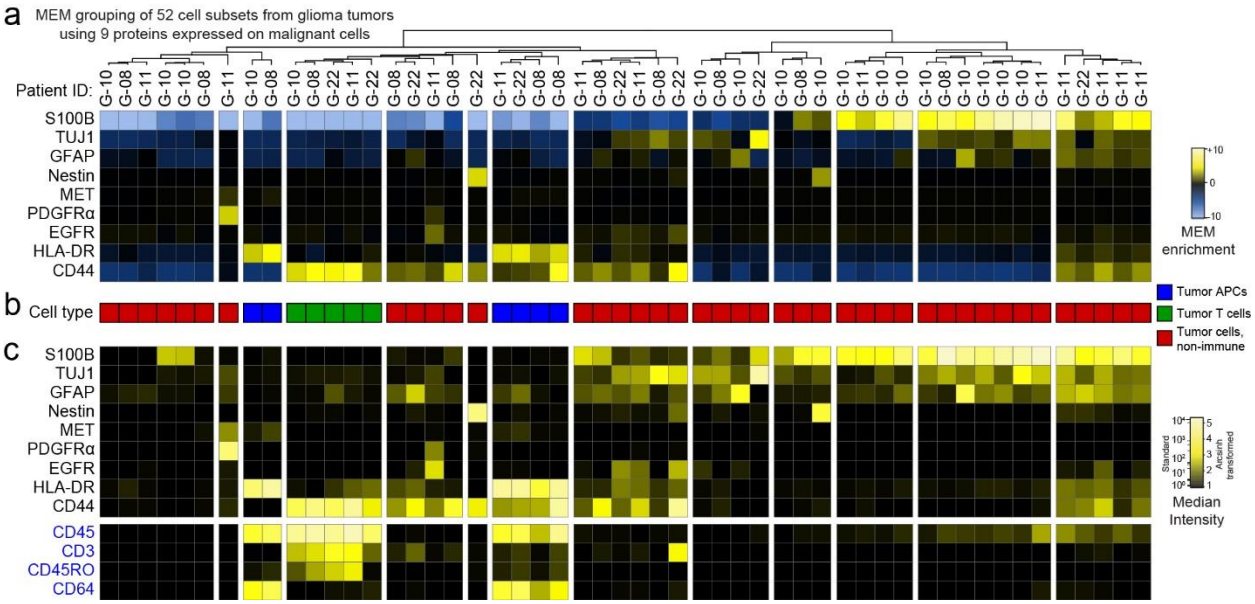



*5-4 Figure 2. Hierarchical clustering based solely on MEM label groups T cells and B cells measured in diverse studies using different cytometry platforms.*

*A) MEM label values were compared for each of 80 populations (CD4+ T cells and B cells) from 3 human tissues representing 6 mass cytometry studies and 1 fluorescence flow cytometry study. Populations are shown clustered according to MEM label percent similarity. Tissue type, source study (numbered 1-7 and referenced in online methods), and individual sample IDs are indicated to the right. *indicates samples stimulated by bacterial superantigen*

*Staphlococcus enterotoxin B (SEB). B) Representative MEM labels for CD4+ T cells (top) and B cells (bottom) from SEB-stimulated normal human blood (1.4, top, mass cytometry), normal human bone marrow (5, mass cytometry), normal human tonsil (2.5, mass cytometry), SEB-stimulated normal human blood (1.4, bottom, fluorescence flow cytometry), and normal human blood (6.1, mass cytometry).*

Dataset D included 52 populations of tumor infiltrating APCs, tumor infiltrating T cells, and non-immune malignant tumor cells identified in human glioma tumors (Leelatian, Doxie et al. 2016). To obtain these populations, each tumor was analyzed by viSNE and cell subsets were expert gated solely on t-SNE cluster density (Supplementary Fig. 6). To determine whether MEM could distinguish immune cell subsets from other tumor cell types with limited information, MEM scores were calculated using only 9 markers that were expected to be expressed on cancer cells (S100B, TUJ1, GFAP, Nestin, MET, PDGFRα, EGFR, HLA-DR, and CD44, Fig. 3a). The 52 populations were grouped into 13 major cell types based on MEM enrichment of 9 analyzed proteins, and these groups were interpreted as tumor infiltrating APCs (Fig. 3b, blue), tumor infiltrating T cells (Fig. 3b, green), or non-immune tumor cells (Fig. 3b, red). To confirm cell identity, four protein features that had been excluded from MEM analysis were assessed (Fig. 3c, CD45, CD3, CD45RO, and CD64). CD45 and CD3 were used to confirm T cell identity and CD45 and CD64 were used to confirm APC identity. MEM correctly identified both immune cell subsets from all tumor types without using key immune lineage markers and without using healthy populations (e.g. APCs from blood or tonsil) to guide the clustering. Thus, MEM labels distinguished populations of cells based on non-traditional features and in a disease context.




*5-5 Figure 3. MEM correctly grouped immune and cancer cell populations from glioma tumors using nine proteins expressed on cancer cells in Dataset D.*
*(A) A heatmap of MEM enrichment scores is shown for 52 populations of cells identified in tumors from 4 glioblastoma patients (G-08, G-10, G-11, G22) in an unsupervised manner using viSNE. (B) Each population was annotated for a cell type based on review of the MEM label and classified as tumor infiltrating APCs (blue), tumor infiltrating T cells (green), or non-immune tumor cells (red). (C) A heatmap of median intensity values is shown for the 13 measured proteins from each of the 52 tumor cell populations.*

MEM labels provided a quantitative language to objectively communicate characteristics of new and established cell types observed in complex tissue microenvironments. Algorithmic comparison of MEM labels correctly identified 80 cell populations from 7 studies of 3 human tissues measured using different instrumentation and distinguished tumor-infiltrating immune cell subsets and malignant cell populations from human glioma tumors. Following additional validation in other cell types, tissues, and instrumentation platforms, it may be possible for machines and humans to use MEM labels to learn and clearly communicate cell identity (cytotype). Given widespread adoption and reporting, MEM labels could be used to communicate cytotypes in a manner analogous to cluster of differentiation (CD) naming of antigen targets of antibodies (1984). MEM can compare populations against a common reference (Supplementary Note 5) and guide feature selection for computational and experimental analysis. MEM can also be

used to monitor changes in tissues over time during treatment. Deviation from a stable MEM score for peripheral blood cell subsets would be expected in the case of emerging malignant cells (Greenplate, Johnson et al. 2016), and lack of change towards a healthy set of MEM scores for blood or bone marrow cell subsets might indicate a lack of response to chemotherapy for a leukemia patient. MEM is expected to assist in machine learning applications by providing quantitative text descriptions of cytotype that can be algorithmically parsed and used to classify newly identified cell subpopulations.

## Data availability statement

The normal human PBMC dataset (Figure 1) were generated by CyTOF analysis as described by Leelatian, et al. (Leelatian, Diggins et al. 2015) and is available as an FCS file in Flow Repository (https://flowrepository.org/experiments/1043).

The normal human bone marrow data set from Bendall and Simonds, et al (Bendall, Simonds et al. 2011) (Dataset B, Supplementary Note 3) was downloaded from Cytobank (Kotecha, Krutzik et al. 2010) as FCS files that included the cell population IDs defined by Bendall and Simonds, et al. (Bendall, Simonds et al. 2011) (https://reports.cytobank.org/1/v1). MEM enrichment scores from Dataset B were compared to the authors' analysis and prior studies of proteins marking stem cells, progenitor cells, and mature cells (Civin, Strauss et al. 1984, Doulatov, Notta et al. 2012)

The murine myeloid CyTOF dataset from Becher, et al (Becher, Schlitzer et al. 2014) (Dataset C, Supplementary Note 4) was downloaded from Cytobank as FCS files that contained gated cell events and cluster IDs as designated by automated analysis conducted by Becher et al (Becher, Schlitzer et al. 2014). MEM enrichment scores from

Dataset C were compared to the authors' analysis and prior studies of neutrophils (Basit, Reutershan et al. 2006, Furze and Rankin 2008).

Datasets for Figure 2 were generated in 7 separate fluorescence and mass cytometry studies by 1) Nicholas et al. (Nicholas, Greenplate et al. 2016), 2) Polikowsky et al. (Polikowsky, Wogsland et al. 2015), 3) Ferrell et al. (Ferrell, Diggins et al. 2016), 4) Amir et al. (Amir el, Davis et al. 2013), 5) Bendall and Simonds et al. (Bendall, Simonds et al. 2011), 6) Greenplate et al., previously unpublished data, and 7) Leelatian et al. (Leelatian, Diggins et al. 2015).

The phospho-flow AML data set generated by Irish et al. (Irish, Hovland et al. 2004) (Supplementary Note 5-Fig.2) was downloaded from Cytobank as FCS files.

The human GBM mass cytometry dataset (Fig. 3) was generated and analyzed as described by Leelatian and Doxie et al. (Leelatian, Doxie et al. 2016) and are available on Flow Repository as text files (https://flowrepository.org/experiments/1044/).

## Online methods

### Code availability

Software for generating MEM scores is available as Supplementary Software.

### CyTOF data pre-processing and analysis

Data analysis was performed using the online analysis platform Cytobank (Kotecha, Krutzik et al. 2010) and the statistical programming environment R. Raw median intensity (MI) values were transformed to a hyperbolic arcsine scale. A cofactor of 15 was used for the PBMC dataset (Fig. 1), and 5 was used for the normal human bone

marrow data set and for the murine myeloid data set. Single, intact cells were gated based on cell length (30-60) and nucleic acid intercalator (iridium). Major PBMC subsets were gated based on CD45 expression (leukocytes) and on canonical lineage marker expression to identify major blood cell subsets.

FCS files were exported from Cytobank as FCS or tab-delimited text files that were parsed for expression intensity information using the R package flowCore (Hahne, LeMeur et al. 2009). MEM was calculated using the arcsinh transformed MI values, as described above. Heatmaps were generated using the heatmap.2 function in the gplots R package (Gregory R. Warnes 2015).

**Fluorescence phospho-flow AML data analysis**

Data were downloaded from Cytobank as FCS files and processed in R as described above. MFI values were transformed to a log normal scale. For each AML patient, a median value and an IQR value was calculated for each marker in the unstimulated condition and for the stimulated conditions. The unstimulated median values were subtracted from the stimulated median values, and likewise for the IQR values. MEM was then calculated by comparing each patient's subtracted median and IQR values to those of the other patients. This enabled a comparison of fold change signaling values rather than raw values.

**Marker enrichment modeling (MEM)**

MEM analysis begins after populations have been identified and aims to provide a simple way to compare findings from experts working with different platforms or performing analysis using different computational tools for population discovery (Lo K

2009, Qiu, Simonds et al. 2011, Bruggner, Bodenmiller et al. 2014, Mosmann, Naim et al. 2014, Shekhar, Brodin et al. 2014) and graphical visualization (Irish, Hovland et al. 2004, Bendall, Simonds et al. 2011, Bendall, Davis et al. 2014, Levine, Simonds et al. 2015, Spitzer, Gherardini et al. 2015). These tools have differing strengths that depend greatly on the structure of the datasets and controls, the biological goals of the study, and the quality of the existing knowledge in the field (Chattopadhyay, Gierahn et al. 2014, Diggins, Ferrell et al. 2015, Saeys, Gassen et al. 2016).

MEM equation

The MEM equation is implemented as an R package (Supplementary Software). Currently, MEM uses medians as the magnitude value; however, depending on the data type, mean may be a more appropriate magnitude statistic and mean could be substituted for median in the equation. Similarly, other statistics, such as variance, might be substituted for IQR. The MEM equation was developed with the intention of capturing and quantifying population-specific feature enrichment in a simple equation that avoids over-fitting or unnecessary computation. The primary goal of this equation is to scale magnitude differences depending on distribution spread. While other distribution features such as skew or shape could be informative, incorporating only two pieces of information – magnitude and spread – into the equation captured enough information to be useful in quantifying both positive and negative population-specific feature enrichment.

MEM output and score scaling

The MEM R script outputs a heatmap of MEM values with a text label summary of feature enrichment as the population (row) names. The + or - value provided along with the marker name is converted to a -10 to +10 scale and rounded to the nearest integer.

As implemented here, the maximum of the scale was set using the highest absolute value MEM score observed across all markers and populations. All values in the matrix are divided by this maximum value and multiplied by 10 to achieve the -10 to +10 scaling. After scaling, the original sign value is reapplied to each MEM score. Scaling the output this way is intended to generate MEM values and labels that are intuitive to human readers and to facilitate comparison of feature enrichment across experiments, samples, batches, time points, and data types.

IQR Threshold

Because MEM uses a ratio of IQR values, near zero values in the denominator, $IQR_{POP}$, will greatly increase MEM scores. For each measurement type, it is important to identify a minimum significant IQR value so that small IQR values below the platform's ability to distinguish signal from noise do not inappropriately increase MEM scores. To automatically determine a minimum threshold for $IQR_{POP}$, the algorithm here calculated the average of the IQR values that were associated with the lowest quartile of population and reference medians. For the mass and fluorescence cytometry datasets used, the automatically calculated IQR threshold was on average 0.5 ± X and so the IQR threshold for all studies here was set to 0.5. The default IQR threshold in the algorithm is also set to 0.5. To have the IQR threshold re-calculated, investigators should specify the "auto" option for the IQR.thresh argument in the MEM function. It is recommended that investigators applying MEM to datasets from different instruments or who are testing MEM for the first time determine whether a change in the IQR threshold is needed.

Reference population selection

MEM scores are contextual; a population's MEM score depends on the reference population(s) to which it is compared. Selection of a reference population should be made deliberately depending on the biological question being addressed. When populations in a MEM analysis arise from different experimental sources, it may be necessary in some cases to normalize measurements prior to MEM analysis to avoid artifacts from experimental variation.

**PBMC processing and mass cytometry**

PBMC were isolated and cryopreserved as described by Greenplate, et al. (Greenplate, Johnson et al. 2016). PBMC were stained with metal conjugated antibodies and prepared for the mass cytometry as previously described (Greenplate, Johnson et al. 2016). The following antibodies were used in the staining panel: CD19-142, CCR5-144, CD4-145, CD64-146, CD20-147, CCR4-149, CD43-150, CD14-151, TCRγδ-152, CD45RA-153, CD45-154, CXCR3-156, CD33-158, CCR7-159, CD28-169, CD29-162, CD45RO-164, CD16-165, CD44-166, CD27-167, CD8-168, CD25-169, CD3-170, CD57-172, PD-L1-175, and CD56-176 (Fluidigm Sciences). In addition, the following purified antibodies from Biolegend were labeled using MaxPar DN3 kits (Fluidigm Sciences), stored at 4°C in antibody stabilization buffer (Candor Bioscience GmbH) and used in the same panel: ICOS-141, TIM-143, CD38-148, CD32-161, HLA-DR-163, CXCR5-171, and PD-1-174.

**Cell subpopulation MEM Score Similarity Calculations**

Comparison of $CD4^+$ T cells to B cells in Figure 2

In order to assess the robustness of MEM across tissue sample types, donors, experimental runs, and flow cytometry platforms (fluorescence and mass cytometry), MEM scores were calculated for cell subsets from 7 different experiments that included 3 healthy human bone marrow samples (Bendall, Simonds et al. 2011, Amir el, Davis et al. 2013, Ferrell, Diggins et al. 2016), 9 healthy human PBMC samples (Leelatian, Diggins et al. 2015, Nicholas, Greenplate et al. 2016), and 6 healthy human tonsil samples (Polikowsky, Wogsland et al. 2015). MEM scores were calculated for each population using as the reference population a combination of hematopoietic stem cells gated as $CD34^+$ $CD38^{lo/-}$ from two studies of healthy human bone marrow (Bendall, Simonds et al. 2011, Ferrell, Diggins et al. 2016). Population similarity was calculated using root mean squared distance (RMSD) calculated on all population MEM scores in a pairwise fashion. MEM scores were calculated using all markers in common between each dataset and the HSC reference (Supplementary Table 5).

RMSD was calculated here as the square root of the average in squared distance between all MEM values in common for each pair of populations (Supplementary Table 5) and then converted into percent maximum possible RMSD. Given the -10 to 10 MEM scale, an RMSD of 20 was the maximum possible difference and corresponded to 0% similarity, whereas an RMSD of 0 between MEM labels indicated 100% similarity. This approach emphasized differences in marker expression when comparing populations. Calculated statistics for $CD4^+$ T cell comparisons included average MEM value +/- standard deviation and p-value calculated using an unpaired, two-tailed Student's t-test.

Human Glioma and Normal Immune Cell MEM Analysis

Glioblastoma data (G-08, G-10, G-11, and G-22) were collected following a published protocol (Leelatian, Doxie et al. 2016). Cells were stained with isotope-tagged antibodies to detect surface and intracellular targets following established protocols (Leelatian, Diggins et al. 2015, Leelatian, Doxie et al. 2016). MEM analysis of glioblastoma patient samples was performed with 9 markers (S100B, TUJ1, GFAP, Nestin, MET, PDGFRα, EGFR, HLA-DR, and CD44), using arcsinh transformation of original median intensity values with a cofactor of 5. Each cell subset was the POP, and the remaining cell subsets were the REF in the analysis.

**Z-score and K-S statistic calculations**

Z-score was calculated between POP and REF as (MEANpop-MEANref)/STDEVref for each marker.

The K-S statistic (Young 1977, Cox, Reeder et al. 1988) was calculated comparing the distribution for each marker on POP and REF using the function ks.test() in R.

**F-measure Analysis**

PBMC populations were defined by expert human gating on canonical markers. For f-measure analysis (Fig. 1c and Supplementary Fig. 2), the 25 measured markers from the CyTOF analysis of healthy PBMC were sorted based on absolute MEM scores, median values, median difference, z-score, and K-S statistic (shown in Supplementary Fig. 2), or randomly across all PBMC populations and the 25 measured proteins. The 5x25 matrix was converted into an ordered vector (length 25X5) and then sorted by absolute value. The first occurrence of each marker in the list was kept and subsequent occurrences of that marker in the list (i.e. that marker's scores on other populations) were

discarded. The order of markers excluded by MEM, median, median difference, z-score, and K-S statistic are shown in Supplementary Table 3. Markers were then sequentially, cumulatively excluded from k-means clustering of cells from high to low absolute for each statistic or score. F-measure was calculated as:

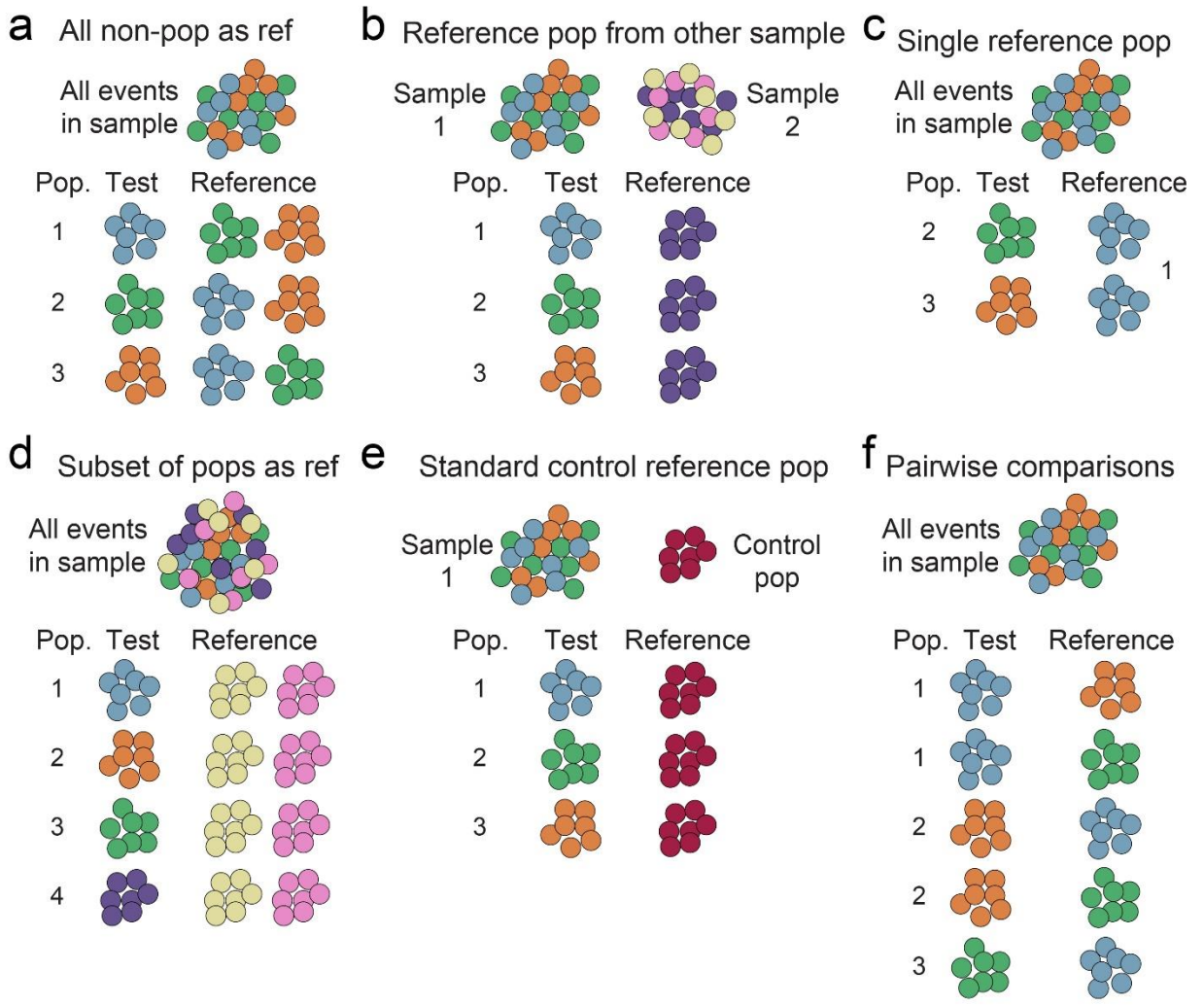
Sensitivity = True Positives/ (True Positives + False Negatives)

Specificity = True Negatives/ (True Negatives + False Positives)

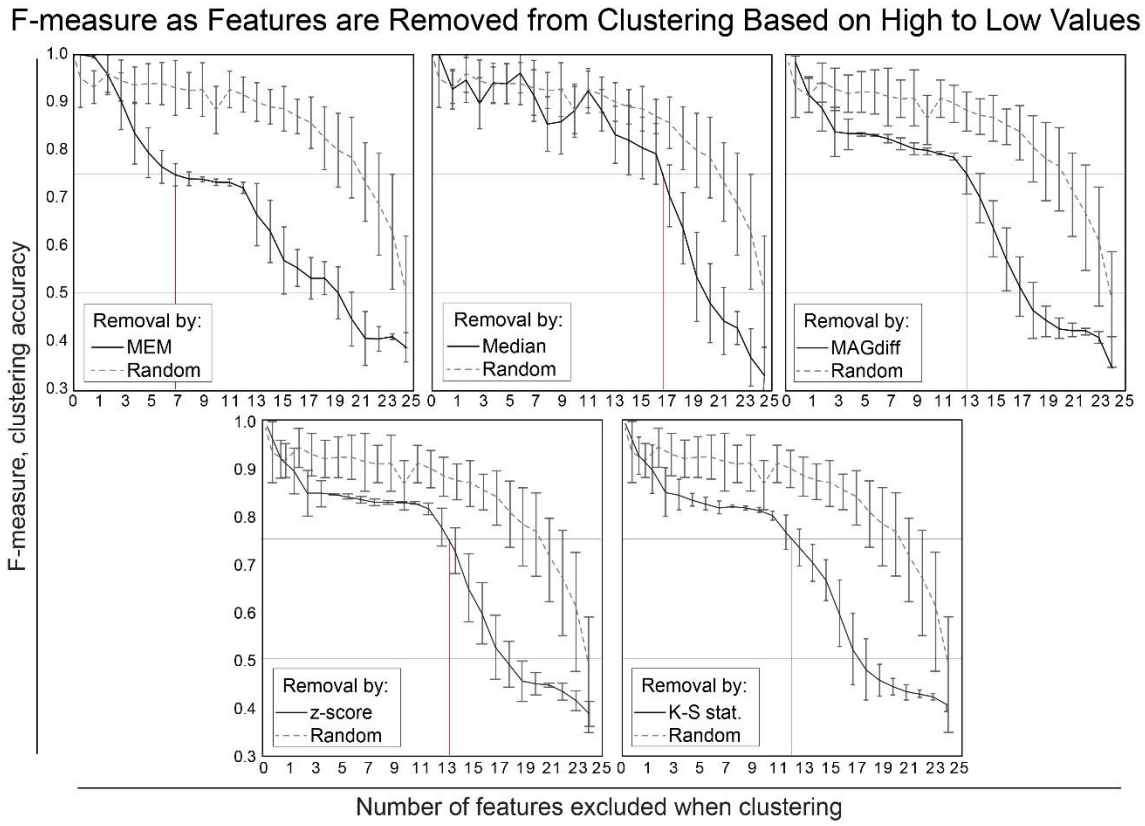
F-measure = 2*(sensitivity*specificity)/ (sensitivity + specificity)

An F-measure was calculated for each round of clustering, where truth was the cell cluster ID resulting from clustering on all 25 markers. The moving average of f-measure with an interval of 3 was calculated in Microsoft Excel. The F-measures for random marker exclusion are the average at each point of 15 different rounds of random marker exclusion from clustering.

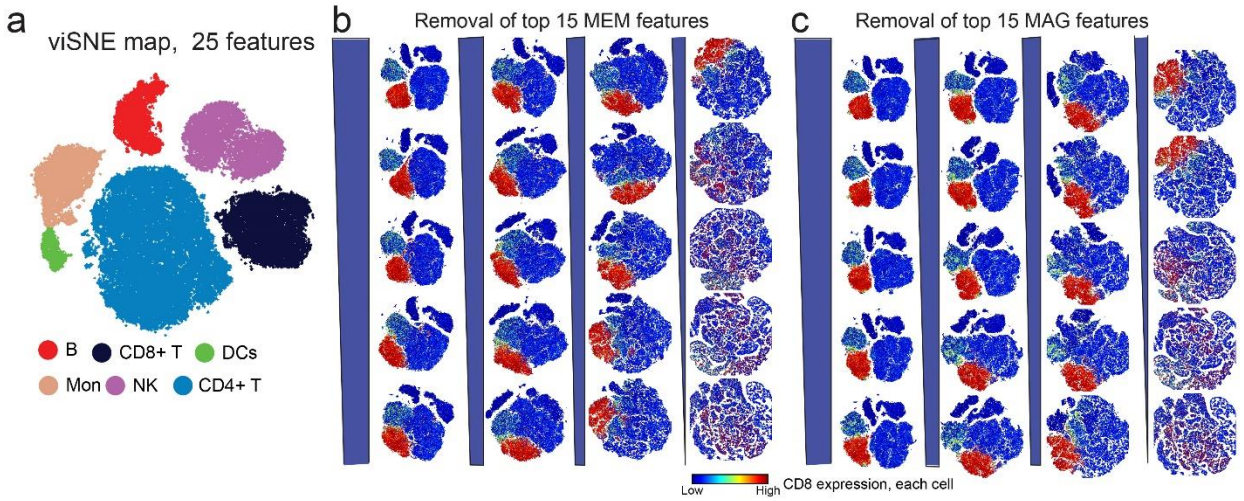
## Supplement



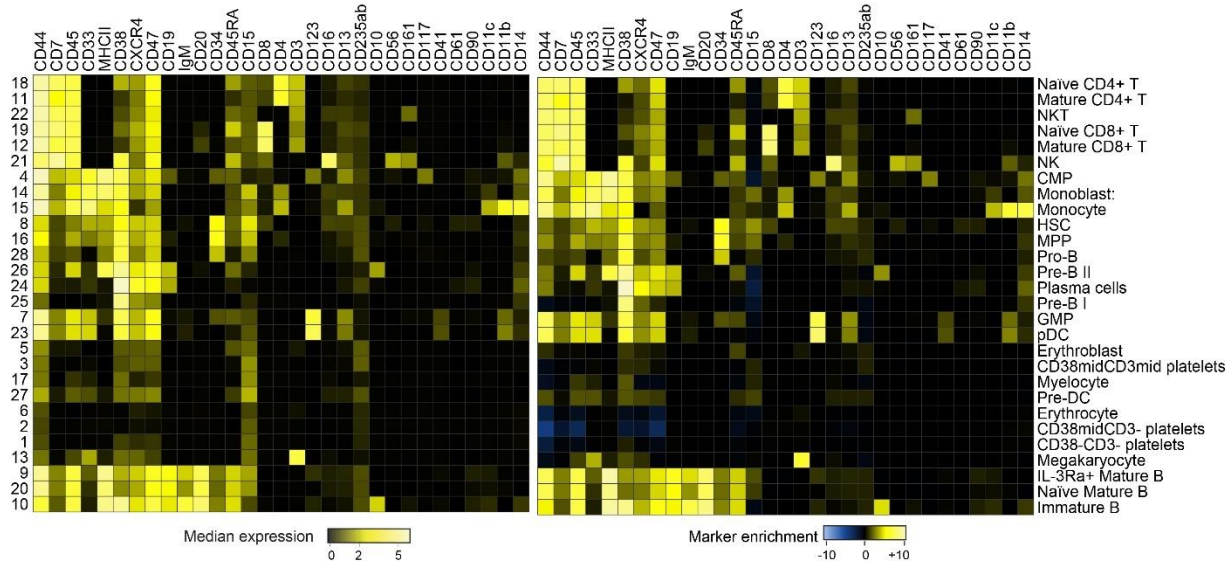

*5-6 Supplementary Figure 1 Examples of MEM reference population selection to capture different contexts. Alternative reference populations (REF) can be used to capture how features of the test population (POP) are enriched in different contexts. Reference comparisons include a) all non-population cells in the sample or experiment (default), b) a population from another sample in the same study, c) a population from the same sample, d) multiple subsets of non-population cells from the same sample, e) a standard control population, and f) pairwise comparison between all populations in a sample.*



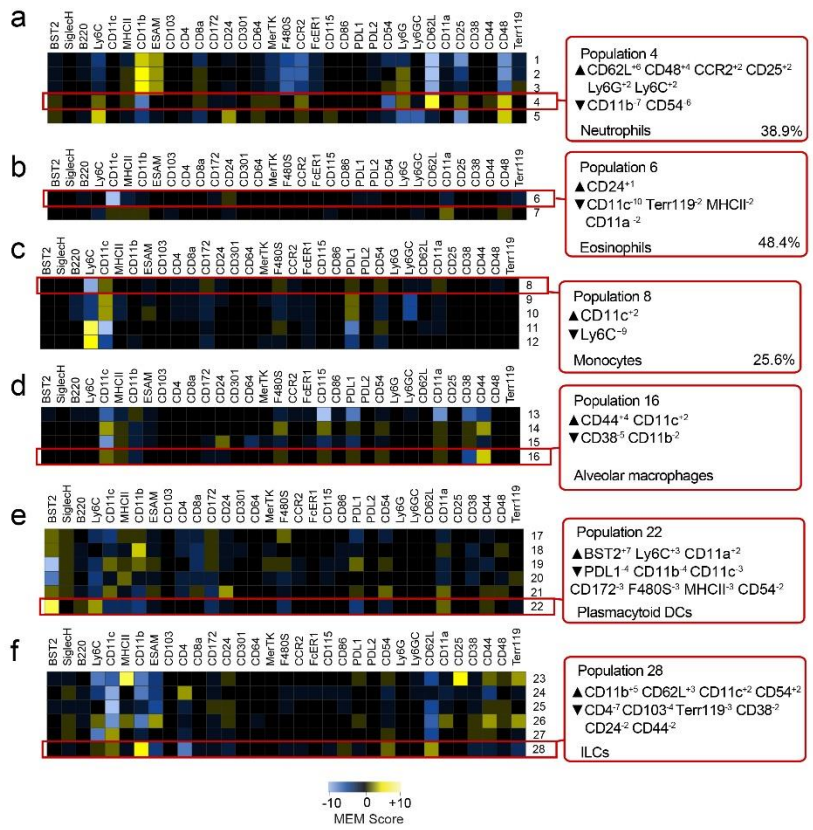

*5-7 Supplementary Figure 2 MEM highly scores markers that are important to clustering accuracy.*
*Markers were sequentially and cumulatively excluded from k-means cluster analysis of Dataset A, from high to low, sorted based on 5 different statistics or scores (marker order shown in Supplementary Table 3): MEM, median, median difference (MAGDIFF), z-score, and Kolmogorov-Smirnov (K-S) statistic. Clustering accuracy was quantified as the f-measure where true cluster identity was assumed to be the clusters formed by clustering on all 25 markers in the dataset. The moving average of the f-measure is shown. Error bars represent the standard error. The vertical red line indicates the number of excluded features at which the f-measure reached 0.75.*




*5-8 Supplementary Figure 3 MEM highly scores markers that are important to viSNE mapping.*
*a) viSNE map for healthy human blood, built using 25 surface protein markers. Populations were identified by expert analysis and color coded.*
*b). Top to bottom, left to right: viSNE maps generated as markers were iteratively, cumulatively excluded based on their MEM scores (high to low absolute value). Heat intensity for each cell indicates CD8 expression.*
*c) Top to bottom, left to right: viSNE maps generated as markers were iteratively, cumulatively excluded based on their median scores (high to low absolute value). Heat intensity for each cell indicates CD8 expression.*

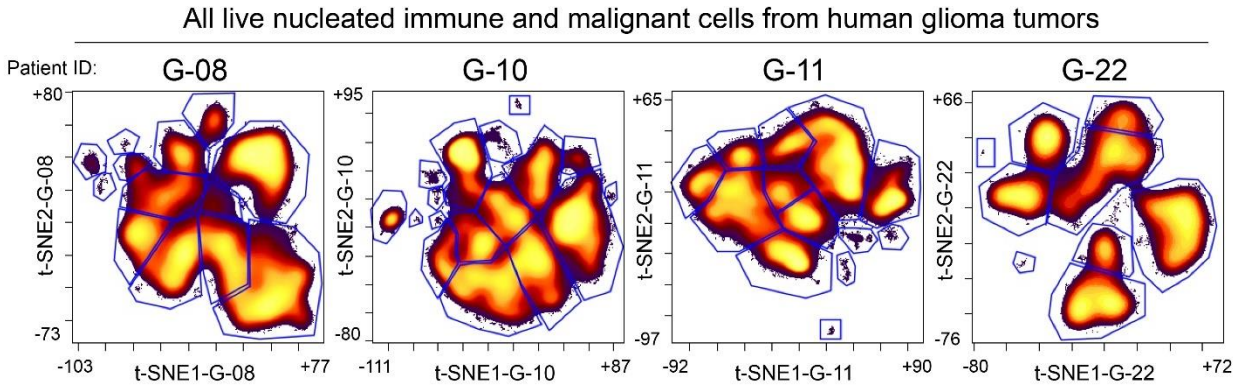



*5-9 Supplementary Figure 4 MEM scores largely reflect median expression values for relatively homogenous populations.*
*Heatmaps show median intensity of protein expression (left) and protein enrichment by MEM (right) for measured proteins in 28 populations characterized as relatively homogeneous for established cell types by expert analysis (rows). Each population was compared to the other 27 subsets for the MEM analysis. MEM scores approach median expression values in homogeneous populations because the contribution of variance approaches zero.*




*5-10 Supplementary Figure 5 Focused MEM analysis quantifies feature enrichment within phenotypically similar groups of cells.*

*a-f) Focused MEM analysis on murine myeloid cell subsets. A MEM label for one population within each group is shown as an example. Groups were defined as the 6 major murine subgroups identified by t-SNE and DensVM by (Becher, Schlitzer et al. 2014).*




*5-11 Supplementary Figure 6 Unsupervised clustering and gating of 52 populations of malignant and immune cells in glioma.*

*Live nucleated immune and malignant cells were gated from glioma tumors as described in Leelatian and Doxie et al., Cytometry B 2016 (Leelatian, Doxie et al. 2016). Patient-specific t-SNE axes were created in separate viSNE analyses of each tumor (e.g. t-SNE1-G-08 for glioma tumor G-08). Shown here is density of cells on t-SNE1 vs. t-SNE2 from each tumor-specific viSNE analysis. Expert analysis of density was then used to identify 52 cell clusters from the 4 glioma tumors. These 52 populations were subsequently grouped by MEM in Fig. 5a using 9 proteins expressed on malignant cells.*

**Supplementary Table 1 – Healthy human CD4⁺ T cells from various mass cytometry studies were labeled consistently by MEM**

| Tissue | Study | Panel focus | # | MEM label for CD4⁺ T cells vs. other live cells |
|---|---|---|---|---|
| Bone marrow | Bendall et al. | Hematopoiesis | 4 | ▲CD4$^{+10}$ CD3$^{+6}$<br>▼MHCII$^{-3}$ CD8$^{-2}$ CD45RA$^{-2}$ CD11b$^{-2}$ |
| | Amir et al. | Canonical immune | 5 | ▲CD4$^{+10}$ CD3$^{+4}$<br>▼CD45RA$^{-2}$ |
| | Ferrell et al. | AML myeloid cells | &2 | ▲CD3$^{+10}$ CD7$^{+7}$ CD4$^{+6}$ CD62L$^{+4}$<br>▼CD11b$^{-6}$ CD11c$^{-6}$ MHCII$^{-6}$ CD64$^{-5}$ CD61$^{-4}$ CD13$^{-3}$ CD38$^{-3}$ CD123$^{-2}$ CD33$^{-2}$ CD14$^{-2}$ |
| Healthy human PBMCs | Leelatian et al. | Canonical immune | 7 | ▲CD4$^{+10}$ CD3$^{+7}$<br>▼CD8a$^{-3}$ CD16$^{-3}$ CD11b$^{-2}$ CD69$^{-2}$ MHCII$^{-2}$ |
| | Greenplate AR§ | T cells | 1.3 | ▲CD4$^{+10}$ CD3$^{+9}$ CCR7$^{+5}$ CD27$^{+5}$ CD28$^{+3}$ CD64$^{+2}$ |
| | | | 1.7 | ▲CD4$^{+10}$ CD3$^{+9}$ CD45RO$^{+3}$ CD27$^{+3}$ CD64$^{+2}$ CD28$^{+2}$<br>▼CD45RA$^{-4}$ |
| | | | 1.1 | ▲CD4$^{+10}$ CD3$^{+8}$ CD27$^{+4}$ CCR7$^{+3}$ CD64$^{+2}$<br>▼CD45RA$^{-2}$ |
| | | | 1.8 | ▲CD4$^{+10}$ CD3$^{+9}$ CD45RO$^{+4}$ CD27$^{+3}$ CD43$^{+2}$<br>▼CD45RA$^{-5}$ |
| | | | 1.6 | ▲CD4$^{+10}$ CD3$^{+7}$ CD27$^{+2}$<br>▼CD45RA$^{-4}$ |
| | | | 1.4 | ▲CD3$^{+10}$ CD4$^{+9}$ CD27$^{+4}$ CCR7$^{+3}$ CD28$^{+2}$<br>▼CD45RA$^{-3}$ |
| | | | 1.2 | ▲CD4$^{+10}$ CD3$^{+10}$ CD43$^{+4}$ CD27$^{+3}$ CCR7$^{+2}$ CD28$^{+2}$ CD45RO$^{+2}$<br>▼MHCII$^{-4}$ CD38$^{-2}$ CD45RA$^{-2}$ |
| Healthy human tonsil | Polikowsky et al. | B cells | 3.1 | ▲CD4$^{+10}$ CD3$^{+9}$ CD5$^{+5}$ CD27$^{+5}$<br>▼MHCII$^{-10}$ CD19$^{-9}$ CD20$^{-7}$ CD40$^{-7}$ CD22$^{-7}$ Igκ$^{-5}$ Igλ$^{-4}$ IgD$^{-2}$ SHP1$^{-2}$ CD16$^{-2}$ CD33$^{-2}$ IgM$^{-2}$ |
| | | | 3.4 | ▲CD4$^{+9}$ CD3$^{+9}$ CD5$^{+5}$ CD27$^{+5}$<br>▼MHCII$^{-10}$ CD19$^{-10}$ CD20$^{-9}$ CD22$^{-7}$ CD40$^{-7}$ Igλ$^{-3}$ Igκ$^{-3}$ IgG$^{-2}$ CD33$^{-2}$ CD16$^{-2}$ SHP1$^{-2}$ |
| | | | 3.5 | ▲CD4$^{+10}$ CD3$^{+10}$ CD27$^{+6}$ CD5$^{+4}$<br>▼MHCII$^{-7}$ Igλ$^{-5}$ CD20$^{-5}$ CD19$^{-5}$ CD22$^{-5}$ CD40$^{-4}$ IgD$^{-3}$ IgM$^{-3}$ CD16$^{-2}$ Igκ$^{-2}$ CD79B$^{-2}$ SHP1$^{-2}$ |

# Sample numbers match study numbers in Figure 4.

§New data from AR Greenplate, prepared as in Leelatian et al.

*5-1 Supplementary Table 1 – Healthy human CD4+ T cells from various mass cytometry studies were labeled consistently by MEM*

**Supplementary Table 2. MEM equation components for PBMC subsets in Fig. 1***

| | | CD45 | CD45RA | CD44 | HLADR | IgM | CD19 | CD8 | CD11c | CD33 | CD11b | CD14 | CD61 | CD16 | CD3 | CD4 | CD38 | CD69 | CD123 | CD56 | CD34 | CD25 | CD10 | CD15 | CD117 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CD4 T cells | $MAG_{POP}$ | 4.3 | 2.3 | 3.2 | 0.0 | 0.0 | 0.0 | 0.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 2.4 | 3.1 | 0.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 |
| | $MAG_{REF}$ | 4.0 | 2.6 | 2.6 | 0.1 | 0.0 | 0.0 | 0.8 | 0.1 | 0.0 | 0.4 | 0.0 | 0.0 | 0.2 | 0.1 | 0.0 | 0.8 | 0.3 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $IQR_{POP}$ | 0.7 | 1.7 | 0.9 | 0.5 | 0.5 | 0.5 | 0.8 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.9 | 0.6 | 0.8 | 0.5 | 0.5 | 0.5 | 0.5 | 0.6 | 0.5 | 0.5 | 0.5 |
| | $IQR_{REF}$ | 0.8 | 1.1 | 1.6 | 2.3 | 0.5 | 0.5 | 4.8 | 2.2 | 0.7 | 1.2 | 0.5 | 0.8 | 3.7 | 1.8 | 0.6 | 1.2 | 1.4 | 0.5 | 0.7 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| | MEM | 0.6 | 0.1 | 1.9 | -5.0 | 0.0 | 0.0 | -7.4 | -4.7 | 0.7 | -2.5 | 0.0 | 0.7 | -8.8 | 4.2 | 4.1 | -1.0 | -2.7 | 0.0 | -0.8 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $MAG_{DIFF-S}$ | 0.5 | -0.6 | 1.3 | -0.2 | 0 | 0 | -0.5 | -0.2 | 0 | -0.7 | 0 | 0 | -0.4 | 4.8 | 6.4 | -0.4 | -0.6 | 0 | -0.2 | 0 | 0.2 | 0 | 0 | 0 |
| | $MAG_{DIFF}$ | 0.3 | -0.3 | 0.6 | -0.1 | 0.0 | 0.0 | -0.3 | -0.1 | 0.0 | -0.4 | 0.0 | 0.0 | -0.2 | 2.3 | 3.0 | -0.2 | -0.3 | 0.0 | -0.1 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 |
| | $IQR_{DIFF}$ | -0.1 | 0.7 | -0.7 | -1.8 | 0.0 | 0.0 | -4.1 | -1.7 | -0.2 | -0.7 | 0.0 | -0.3 | -3.2 | -0.8 | 0.0 | -0.4 | -0.9 | 0.0 | -0.2 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 |
| CD8 T cells | $MAG_{POP}$ | 4.1 | 2.7 | 2.6 | 0.0 | 0.0 | 0.0 | 5.1 | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 0.0 | 2.2 | 0.1 | 0.2 | 1.7 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $MAG_{REF}$ | 4.2 | 2.4 | 3.1 | 0.0 | 0.0 | 0.0 | 0.4 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 1.7 | 2.6 | 0.7 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $IQR_{POP}$ | 0.7 | 1.0 | 0.8 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.6 | 0.5 | 0.5 | 0.5 | 0.8 | 0.5 | 0.6 | 0.8 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| | $IQR_{REF}$ | 0.7 | 1.4 | 1.3 | 0.9 | 0.5 | 0.5 | 0.9 | 0.7 | 0.5 | 0.6 | 0.5 | 0.5 | 0.6 | 2.5 | 3.0 | 1.0 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| | MEM | -0.2 | 0.9 | -1.3 | 1.0 | 0.0 | 0.0 | 7.3 | 0.5 | 0.0 | 0.3 | 0.0 | 0.0 | 0.3 | 3.5 | -10.0 | -1.7 | 1.7 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $MAG_{DIFF-S}$ | -0.2 | 0.5 | -0.9 | 0 | 0 | 0 | 10 | 0 | 0 | 0.2 | 0 | 0 | 0 | 1.1 | -5.3 | -1 | 3.5 | 0 | 0 | 0.1 | 0 | 0 | 0 | 0 |
| | $MAG_{DIFF}$ | -0.1 | 0.3 | -0.5 | 0.0 | 0.0 | 0.0 | 4.7 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.5 | -2.5 | -0.5 | 1.7 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $IQR_{DIFF}$ | -0.1 | -0.4 | -0.5 | -0.4 | 0.0 | 0.0 | -0.4 | -0.2 | 0.0 | -0.1 | 0.0 | 0.0 | -0.1 | -1.7 | -2.5 | -0.4 | 0.3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| DCs | $MAG_{POP}$ | 3.9 | 3.0 | 3.4 | 2.7 | 0.0 | 0.0 | 0.1 | 3.1 | 0.9 | 0.0 | 0.1 | 0.3 | 1.3 | 0.0 | 0.7 | 0.7 | 0.0 | 1.3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $MAG_{REF}$ | 4.1 | 2.5 | 3.0 | 0.0 | 0.0 | 0.0 | 0.6 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 1.9 | 2.0 | 0.6 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $IQR_{POP}$ | 1.1 | 1.6 | 1.0 | 2.0 | 0.5 | 0.5 | 0.5 | 3.9 | 1.7 | 0.5 | 0.5 | 1.5 | 3.9 | 0.5 | 1.4 | 1.2 | 0.5 | 2.3 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| | $IQR_{REF}$ | 0.7 | 1.3 | 1.3 | 0.5 | 0.5 | 0.5 | 1.5 | 0.5 | 0.5 | 0.7 | 0.5 | 0.5 | 0.5 | 2.4 | 3.1 | 1.0 | 0.6 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| | MEM | 0.2 | 0.5 | 0.8 | 2.6 | 0.0 | 0.0 | -3.2 | 3.0 | 0.2 | -0.5 | 0.1 | -0.5 | 0.5 | -7.6 | -3.4 | -0.1 | -0.4 | 0.7 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $MAG_{DIFF-S}$ | -0.4 | 1.1 | 0.7 | 5.6 | 0 | 0 | -1 | 6.6 | 1.8 | -0.1 | 0.1 | 0.6 | 2.6 | -3.9 | -2.8 | 0.2 | 0 | 2.6 | 0 | 0 | 0 | 0 | 0 | 0 |
| | $MAG_{DIFF}$ | -0.2 | 0.5 | 0.4 | 2.7 | 0.0 | 0.0 | -0.5 | 3.1 | 0.9 | -0.1 | 0.1 | 0.3 | 1.3 | -1.9 | -1.3 | 0.1 | 0.0 | 1.3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $IQR_{DIFF}$ | 0.4 | 0.3 | -0.2 | 1.5 | 0.0 | 0.0 | -1.0 | 3.4 | 1.2 | -0.2 | 0.0 | 1.0 | 3.4 | -1.9 | -1.7 | 0.2 | -0.1 | 1.8 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| IgM- B cells | $MAG_{POP}$ | 4.4 | 3.5 | 3.1 | 3.1 | 0.0 | 2.5 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.7 | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $MAG_{REF}$ | 4.1 | 2.5 | 3.0 | 0.0 | 0.0 | 0.0 | 0.6 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 1.9 | 2.0 | 0.6 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $IQR_{POP}$ | 0.8 | 1.4 | 1.0 | 1.3 | 0.5 | 0.7 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 1.0 | 0.5 | 0.5 | 0.5 | 0.6 | 0.5 | 0.5 | 0.5 | 0.5 |
| | $IQR_{REF}$ | 0.7 | 1.3 | 1.3 | 0.5 | 0.5 | 0.5 | 1.5 | 0.5 | 0.5 | 0.6 | 0.5 | 0.5 | 0.5 | 2.4 | 3.0 | 1.0 | 0.6 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| | MEM | 0.2 | 1.3 | 0.5 | 3.3 | 0.0 | 3.0 | -3.3 | 0.0 | 0.1 | -0.5 | 0.0 | 0.0 | -0.1 | -7.6 | -9.4 | 0.1 | -0.4 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $MAG_{DIFF-S}$ | 0.4 | 2.1 | 0.2 | 6.5 | 0 | 5.2 | -1.1 | 0 | 0.1 | -0.2 | 0 | 0 | 0 | -3.9 | -4.2 | 0.2 | 0 | 0 | 0 | 0.3 | 0 | 0 | 0 | 0 |
| | $MAG_{DIFF}$ | 0.2 | 1.0 | 0.1 | 3.1 | 0.0 | 2.5 | -0.5 | 0.0 | 0.1 | -0.1 | 0.0 | 0.0 | 0.0 | -1.9 | -2.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $IQR_{DIFF}$ | 0.0 | 0.0 | -0.3 | 0.8 | 0.0 | 0.2 | -1.0 | 0.0 | 0.0 | -0.1 | 0.0 | 0.0 | 0.0 | -1.9 | -2.5 | 0.0 | -0.1 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 |
| IgM+ B cells | $MAG_{POP}$ | 4.3 | 2.6 | 2.7 | 3.7 | 3.0 | 2.5 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.9 | 0.0 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $MAG_{REF}$ | 4.1 | 2.5 | 3.0 | 0.0 | 0.0 | 0.0 | 0.6 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 2.0 | 2.3 | 0.6 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $IQR_{POP}$ | 0.7 | 1.2 | 0.9 | 1.4 | 2.6 | 0.7 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 1.1 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| | $IQR_{REF}$ | 0.7 | 1.4 | 1.3 | 0.5 | 0.5 | 0.5 | 1.6 | 0.5 | 0.5 | 0.7 | 0.5 | 0.5 | 0.6 | 2.4 | 3.0 | 0.9 | 0.7 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| | MEM | 0.4 | 0.5 | -1.0 | 4.0 | 2.9 | 3.0 | -3.6 | 0.0 | 0.1 | -0.6 | 0.0 | 0.0 | -0.3 | -7.5 | -9.6 | 0.2 | -0.6 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $MAG_{DIFF-S}$ | 0.4 | 0.3 | -0.7 | 7.7 | 6.3 | 5.2 | -1.2 | 0 | 0.2 | -0.2 | 0 | 0 | 0 | -4 | -4.8 | 0.5 | 0 | 0 | 0 | 0.2 | 0 | 0 | 0 | 0 |
| | $MAG_{DIFF}$ | 0.2 | 0.2 | -0.3 | 3.7 | 3.0 | 2.5 | -0.6 | 0.0 | 0.1 | -0.1 | 0.0 | 0.0 | 0.0 | -1.9 | -2.3 | 0.3 | 0.0 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $IQR_{DIFF}$ | -0.1 | -0.2 | -0.4 | 0.9 | 2.1 | 0.2 | -1.1 | 0.0 | 0.0 | -0.2 | 0.0 | 0.0 | -0.1 | -1.9 | -2.5 | 0.1 | -0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Monocytes | $MAG_{POP}$ | 4.1 | 2.3 | 4.2 | 2.1 | 0.0 | 0.0 | 0.2 | 3.6 | 3.5 | 2.7 | 2.9 | 2.9 | 0.1 | 0.0 | 1.1 | 1.3 | 0.2 | 0.4 | 0.1 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 |
| | $MAG_{REF}$ | 4.2 | 2.5 | 2.9 | 0.0 | 0.0 | 0.0 | 0.6 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 2.0 | 2.4 | 0.6 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $IQR_{POP}$ | 0.7 | 1.1 | 0.7 | 1.5 | 0.5 | 0.5 | 0.5 | 0.8 | 0.6 | 0.7 | 0.7 | 1.2 | 0.5 | 0.5 | 0.9 | 0.8 | 0.6 | 0.8 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| | $IQR_{REF}$ | 0.7 | 1.4 | 1.2 | 0.5 | 0.5 | 0.5 | 1.7 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.6 | 2.2 | 3.1 | 0.9 | 0.6 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| | MEM | -0.2 | -0.5 | 2.6 | 1.9 | 0.0 | 0.0 | -3.5 | 4.3 | 4.5 | 3.1 | 3.5 | 3.1 | 0.2 | -7.2 | -5.0 | 1.1 | 0.3 | 0.0 | 0.1 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 |
| | $MAG_{DIFF-S}$ | -0.2 | -0.3 | 2.6 | 4.4 | 0 | 0 | -0.9 | 7.5 | 7.4 | 5.5 | 6 | 6.1 | 0 | -4.2 | -2.7 | 1.5 | 0.4 | 0.7 | 0.1 | 0 | 0 | 0.1 | 0 | 0 |
| | $MAG_{DIFF}$ | -0.1 | -0.2 | 1.3 | 2.1 | 0.0 | 0.0 | -0.5 | 3.6 | 3.5 | 2.6 | 2.9 | 2.9 | 0.0 | -2.0 | -1.3 | 0.8 | 0.2 | 0.4 | 0.1 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 |
| | $IQR_{DIFF}$ | 0.0 | -0.2 | -0.5 | 1.0 | 0.0 | 0.0 | -1.1 | 0.3 | 0.1 | 0.2 | 0.2 | 0.7 | -0.1 | -1.7 | -2.2 | -0.1 | 0.0 | 0.3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| NK cells | $MAG_{POP}$ | 3.8 | 2.6 | 1.6 | 0.0 | 0.0 | 0.0 | 0.5 | 0.8 | 0.0 | 0.6 | 0.0 | 0.0 | 4.3 | 0.0 | 0.0 | 1.1 | 0.1 | 0.0 | 0.8 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $MAG_{REF}$ | 4.2 | 2.4 | 3.2 | 0.0 | 0.0 | 0.0 | 0.5 | 0.0 | 0.0 | 0.1 | 0.0 | 0.0 | 0.0 | 2.1 | 2.6 | 0.6 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $IQR_{POP}$ | 0.7 | 0.9 | 0.9 | 0.5 | 0.5 | 0.5 | 2.4 | 1.6 | 0.5 | 0.9 | 0.5 | 0.5 | 0.8 | 0.5 | 0.5 | 1.3 | 0.5 | 0.5 | 1.0 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| | $IQR_{REF}$ | 0.7 | 1.5 | 1.0 | 1.0 | 0.5 | 0.5 | 1.3 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 1.6 | 2.8 | 0.9 | 0.7 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| | MEM | -0.6 | 1.0 | -2.3 | 1.2 | 0.0 | 0.0 | 0.6 | 0.1 | 0.0 | 0.1 | 0.0 | 0.0 | 5.2 | -5.6 | -9.6 | 0.3 | 0.5 | 0.0 | 0.4 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $MAG_{DIFF-S}$ | -0.9 | 0.2 | -3.3 | 0 | 0 | 0 | 0 | 1.6 | 0 | 1 | 0 | 0 | 9.1 | -4.3 | -5.3 | 1.1 | 0 | 0 | 1.7 | 0 | 0 | 0 | 0 | 0 |
| | $MAG_{DIFF}$ | -0.4 | 0.1 | -1.6 | 0.0 | 0.0 | 0.0 | 0.0 | 0.8 | 0.0 | 0.5 | 0.0 | 0.0 | 4.3 | -2.1 | -2.5 | 0.5 | 0.0 | 0.0 | 0.8 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | $IQR_{DIFF}$ | 0.0 | -0.6 | -0.1 | -0.5 | 0.0 | 0.0 | 1.1 | 1.1 | 0.0 | 0.4 | 0.0 | 0.0 | 0.3 | -1.1 | -2.3 | 0.4 | -0.2 | 0.0 | 0.5 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

* Rounded values are shown from raw calculations

*5-2 Supplementary Table 2. MEM equation components for PBMC subsets in Fig. 1*

**Supplementary Table 3. Order of marker exclusion for clustering and f-measure (high to low)**

| MEM | MAG | $MAG_{DIFF}$ | z-score | K-S |
|---|---|---|---|---|
| CD4 | CD8a | CD8a | CD8 | CD14 |
| CD16 | CD45 | CD16 | CD33 | IgM |
| CD3 | CD16 | CD11c | CD16 | CD33 |
| CD8a | CD44 | HLA-DR | CD11b | CD19 |
| CD11c | CD11c | CD33 | CD19 | CD16 |
| HLA-DR | HLA-DR | CD4 | CD14 | CD61 |
| CD33 | CD33 | CD61 | CD4 | CD8 |
| CD61 | CD4 | CD14 | CD11c | CD11b |
| CD14 | CD61 | CD19 | IgM | CD11c |
| CD19 | CD45RA | CD11b | CD61 | CD4 |
| CD11b | CD14 | IgM | HLADR | CD123 |
| CD69 | CD11b | CD3 | CD20 | CD69 |
| CD44 | CD19 | CD20 | CD69 | CD20 |
| CD20 | IgM | CD69 | CD3 | HLADR |
| IgM | CD3 | CD44 | CD44 | CD44 |
| CD38 | CD20 | CD56 | CD56 | CD56 |
| CD45RA | CD69 | CD38 | CD123 | CD3 |
| CD56 | CD38 | CD123 | CD45RA | CD25 |
| CD45 | CD56 | CD45 | CD38 | CD117 |
| CD34 | CD123 | CD45RA | CD25 | CD34 |
| CD10 | CD34 | CD34 | CD117 | CD45RA |
| CD117 | CD25 | CD25 | CD34 | CD38 |
| CD123 | CD10 | CD10 | CD45 | CD45 |
| CD25 | CD117 | CD15 | CD15 | CD15 |
| CD15 | CD15 | CD117 | CD10 | CD10 |

*5-3 Supplementary Table 3. Order of marker exclusion for clustering and f-measure (high to low)*

## Supplementary Table 4. K-S and z-score values for immune cell populations in Fig. 1

| | | CD19 | CD117 | CD11b | CD4 | CD8 | CD20 | CD34 | CD61 | CD123 | CD45RA | CD45 | CD10 | CD33 | CD11c | CD14 | CD69 | CD15 | CD16 | CD44 | CD38 | CD25 | CD3 | IgM | HLADR | CD56 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| K-S | CD4+ T | 1.8 | 1.2 | 3.3 | 9.7 | 4.1 | 3.8 | 2.7 | 2.2 | 1.8 | 2.0 | 2.0 | 0.1 | 2.3 | 3.8 | 2.5 | 3.7 | 1.3 | 3.4 | 3.2 | 1.7 | 3.6 | 6.7 | 1.2 | 3.7 | 3.6 |
| | CD8+ T | 1.0 | 1.2 | 1.3 | 6.2 | 10.0 | 6.5 | 3.5 | 1.1 | 1.0 | 1.4 | 1.0 | 0.7 | 1.5 | 2.3 | 1.2 | 8.3 | 0.1 | 1.9 | 2.8 | 3.3 | 2.1 | 4.1 | 0.8 | 2.0 | 0.3 |
| | DC | 0.8 | 0.7 | 4.5 | 4.0 | 3.1 | 1.5 | 0.3 | 3.1 | 5.8 | 2.6 | 1.8 | 0.2 | 4.7 | 4.8 | 2.8 | 1.8 | 0.6 | 3.4 | 1.8 | 0.9 | 2.1 | 6.7 | 0.9 | 7.8 | 0.8 |
| | IgM- B | 9.3 | 3.6 | 1.7 | 5.9 | 4.8 | 7.6 | 3.4 | 0.7 | 0.4 | 4.3 | 1.6 | 0.4 | 2.2 | 1.0 | 0.9 | 3.0 | 0.6 | 1.7 | 0.9 | 0.7 | 0.5 | 6.6 | 1.5 | 7.9 | 1.1 |
| | IgM+ B | 9.8 | 3.6 | 2.9 | 6.4 | 5.4 | 8.4 | 3.5 | 0.8 | 0.9 | 1.3 | 1.5 | 0.5 | 2.6 | 1.7 | 1.2 | 3.1 | 0.4 | 1.9 | 2.1 | 1.4 | 0.7 | 6.8 | 9.4 | 8.5 | 0.9 |
| | Mono | 0.9 | 0.5 | 9.8 | 4.9 | 3.0 | 2.5 | 1.2 | 9.2 | 5.6 | 1.0 | 0.8 | 1.5 | 9.9 | 9.4 | 9.7 | 2.7 | 1.4 | 1.7 | 6.6 | 4.2 | 2.5 | 7.3 | 0.6 | 8.2 | 1.8 |
| | NK | 1.8 | 1.4 | 3.6 | 6.6 | 1.9 | 2.0 | 1.6 | 1.1 | 0.9 | 1.5 | 3.2 | 0.6 | 1.3 | 4.7 | 1.1 | 1.0 | 1.8 | 9.8 | 7.3 | 2.7 | 2.8 | 7.7 | 2.0 | 3.2 | 6.1 |
| z-score | CD4+ T | -0.6 | -0.3 | -0.9 | 4.6 | -0.7 | -1.0 | -0.6 | -0.6 | -0.5 | -0.8 | 0.6 | 0.0 | -0.7 | -1.0 | -0.7 | -1.0 | -0.3 | -0.9 | 0.7 | -0.4 | 1.6 | 1.9 | -0.4 | -0.9 | -0.9 |
| | CD8+ T | -0.3 | 0.3 | -0.1 | -1.5 | 5.6 | 1.4 | 0.7 | -0.4 | -0.3 | 0.4 | -0.2 | -0.2 | -0.5 | -0.7 | -0.5 | 4.2 | 0.0 | -0.4 | -0.4 | -1.0 | -0.5 | 0.9 | -0.3 | -0.5 | 0.1 |
| | DC | -0.1 | -0.2 | -0.9 | -0.6 | -0.8 | -0.2 | 0.1 | 0.7 | 4.4 | 0.6 | -0.4 | 0.0 | 1.1 | 1.9 | 0.2 | -0.5 | 0.1 | 1.0 | 0.4 | 0.3 | -0.5 | -1.7 | 0.2 | 2.3 | 0.0 |
| | IgM- B | 4.4 | 1.2 | -0.5 | -1.6 | -1.0 | 3.1 | 1.2 | -0.2 | 0.0 | 1.2 | 0.4 | 0.1 | 0.0 | -0.3 | -0.3 | -0.7 | -0.1 | -0.6 | 0.2 | 0.3 | 0.0 | -1.6 | -0.3 | 2.6 | -0.3 |
| | IgM+ B | 7.6 | 1.4 | -0.7 | -1.7 | -1.1 | 3.9 | 1.2 | -0.3 | 0.1 | 0.5 | 0.5 | 0.2 | 0.0 | -0.5 | -0.4 | -0.8 | -0.1 | -0.6 | -0.3 | 0.4 | -0.2 | -1.7 | 9.9 | 3.8 | -0.2 |
| | Mono | -0.3 | 0.1 | 5.1 | -0.3 | -0.8 | -0.7 | 0.3 | 6.2 | 1.9 | 0.0 | -0.2 | 0.5 | 8.7 | 5.1 | 10.0 | 0.1 | 0.4 | -0.4 | 1.8 | 1.1 | -0.6 | -1.9 | -0.3 | 2.3 | 0.0 |
| | NK | -0.4 | -0.1 | 0.5 | -1.9 | -0.2 | -0.6 | -0.4 | -0.4 | -0.3 | 0.3 | -1.0 | -0.1 | -0.5 | 0.8 | -0.2 | -0.2 | 0.5 | 7.4 | -2.8 | 0.9 | -0.7 | -2.2 | -0.5 | -0.8 | 2.5 |

Each population (POP) was compared to all non-population cells (REF). K-S statistic was calculated as a comparison between POP and REF distributions for each protein. Z-score was calculated as the number of standard deviations the POP mean was from the REF mean. K-S and z-score were signed and scaled to -10 to +10 as for MEM values.

*5-4 Supplementary Table 4. K-S and z-score values for immune cell populations in Fig. 1*

## Supplementary Table 5. Full antibody panels for immune cell datasets in Fig. 2

| Dataset | Panel |
|---|---|
| Nicholas KJ et al.(Nicholas, Greenplate et al. 2016) | HLADR, CD8, CD25, CD38, CD4, CD3, CD62L, CD69 |
| Nicholas KJ et al.(Nicholas, Greenplate et al. 2016) | HLADR, CD19, CD27, CD38, CD86, CD20 |
| Greenplate AR[§] | ICOS, CD19, TIM3, CCR5, CD4, CD64, CD20, CD38, CCR4, CD43, CD14, TCRγδ, CD45RA, CD45, CXCR3, CD33, CCR7, CD28, CD32, CD69, HLADR, CD45RO, CD16, CD44, CD27, CD8, CD2, CD3, CXCR5, CD57, PD1, PDL1, CD56 |
| Ferrell et al.(Ferrell, Diggins et al. 2016) | CD235a, CD19, CD117, CD11b, CD4, CD64, CD7 , CD34, CD61, CD123, CD13, CD62L, CD45, CD183, CD33, CD11c, CD14, CD15, CD16, CD24, CD38, CD25, CD3, CD185, HLA-DR, CD184, CD56 |
| Leelatian et al.(Leelatian, Diggins et al. 2015) | CD19, CD117, CD11b, CD4, CD8a, CD20, CD34, CD61, CD123, CD45RA, CD45, CD10, CD33, CD11c, CD14, CD69, CD15, CD16, CD44, CD38, CD25, CD3, IgM, HLADR, CD56 |
| Bendall et al.(Bendall, Simonds et al. 2011) | CD45, CD45RA, CD235ab, CD19, CD11b, CD4, CD8, CD34, CD161, CD20, CD41, CD11c, CD123, IgM, CD10, CD33, CD14, CD38, CD15, CD16, CD44, CD7, CD13, CD56, CD61, CD117, CD47, HLADR, CD90, CXCR4, CD3 |
| Polikowsky et al.(Polikowsky, Wogsland et al. 2015) | CD19, CD5, IgG, CD4, IgD, CD20, CD16, Igλ, CD45, CD27, CD86, CD33, CD22, Igκ, CD79B, CD40, CD44, CD38, CD8, CD3, IgM, HLADR, SHP1, CD56 |
| Amir et al.(Amir el, Davis et al. 2013) | CD45, CD3, CD45RA, CD19, CD11b, CD4, CD8, CD34, CD20, Ki67, CD33, CD123, IκBα, CD38, CD90 |
| HSC REF**:** Amir et al.(Amir el, Davis et al. 2013) & Ferrell et al.(Ferrell, Diggins et al. 2016) | CD45, CD45RA, CD235ab, CD19, CD11b, CD4, CD8, CD34, CD161, CD20, CD41, CD11c, CD123, IgM, CD10, CD33, CD14, CD38, CD15, CD16, CD44, CD7, CD13, CD56, CD61, CD117, CD47, HLADR, CD90, CXCR4 |

Note: MEM comparisons were made using all markers in common between each dataset and the HSCs (combined from Amir et al. and Ferrell et al. datasets). Pairwise RMSD comparisons of MEM scores were made using all markers in common between the pairs of datasets. [§]Unpublished data from AR Greenplate, prepared as in Leelatian et al.(Leelatian, Diggins et al. 2015)

*5-5 Supplementary Table 5. Full antibody panels for immune cell datasets in Fig. 2*

# APPENDIX F : ALLERGY RESEARCH IN COLLABORATION WITH UVA

Full Title: High-Dimensional Single-Cell Monitoring of Circulating Cells in Allergic Asthmatics Infected with Rhinovirus Reveals Dynamic Flux in Diverse Immune Cells (Muehling, Wogsland et al. 2015)

Lyndsey Muehling*, Cara E. Wogsland*, Rachana Agrawal, Peter W. Heyman, Jonathan M. Irish, Judith A. Woodfolk

* Equal contributions

## My Contributions

I hosted Lyndsey Muehling at Vanderbilt University for two weeks during which time the mass cytometry data for the following abstract was collected. Prior to her visit, we collaborated to design a T cell allergy focused panel. During her visit I instructed and assisted with the conjugation of metal tagged antibodies, titration of the conjugated antibodies, use of the CyTOF1, and use of Cytobank software to analyze mass cytometry data and how to interpret the data. I introduced her to data analysis with viSNE and how useful a tool it is when comparing different sample types.

We used viSNE to compare lymphocyte changes in the peripheral blood of patients with allergic asthma over the course of infection. The Woodfolk group at UVA has a rare setup where human subjects are infected with the common cold and monitored during the course of infection. Using mass cytometry was an ideal way to look at many cell types simultaneously. The data collected shed light on T cell populations that are lost from the periphery over the course of infection and thought to be sequestered and those that expanded after infection.

**Abstract**

*Rationale:* Infection with rhinovirus (RV) exacerbates allergic asthma. The immune mechanisms remain poorly understood, in part owing to practical and technical limitations. Using mass cytometry, we applied a systems biology approach to identify novel transitions in circulating cells during an experimental RV infection in allergic asthmatics.

*Methods:* Allergic asthmatics were challenged intranasally with RV-16. Peripheral blood mononuclear cells (PBMCs) were collected immediately before inoculation, and then at days 4 and 21 post-inoculation. Mass cytometry was performed using a 35-parameter phenotyping panel, and data analysis was performed using viSNE.

*Results:* At each time point, viSNE created visual maps of phenotypically similar cell types, including CD4+, CD8+, and γδ T-cell populations, as well as B-cells. During the acute phase, loss of cells expressing the Th1 transcription factor T-bet, was observed, suggesting egress from the periphery. These T-bet+ populations included memory B-cells and memory CD4+ T-cells with lung-homing potential (CCR5+). Further inspection of the data revealed complete loss of memory CCR5+CD4+T-cells that co-expressed the Th2 marker CCR4. These cells re-emerged and were expanded at day 21, along with both CCR4+CD8+ and γδ T-cells.

*Conclusions:* For the first time, we have used mass cytometry to identify dynamic fluctuations in complex cell populations in asthmatics infected with RV. Sequestration of diverse immune cell types during the acute phase, including both Th1- and Th2-like cells, coupled with expansion of both CD4+ and non-CD4+ T-cell types several weeks after

infection, were notable findings. These novel observations provide new insight into the immunopathogenesis of RV-induced asthma.

# REFERENCES


(1984). "Nomenclature for clusters of differentiation (CD) of antigens defined on human leukocyte populations. IUIS-WHO Nomenclature Subcommittee." Bull World Health Organ **62**(5): 809-815.

(2007). "What is systems biology?" THE SEVEN STONES http://blogs.nature.com/sevenstones/2007/07/what_is_systems_biology_3.html 2017.

(2016). Obinutuzumab. fda.gov, FDA.

(2017). "Department of Systems Biology | Harvard Medical School." Retrieved 02/24/2017, 2017, from http://sysbio.med.harvard.edu/.

Alduaij, W., A. Ivanov, J. Honeychurch, E. J. Cheadle, S. Potluri, S. H. Lim, K. Shimada, C. H. Chan, A. Tutt, S. A. Beers, M. J. Glennie, M. S. Cragg and T. M. Illidge (2011). "Novel type II anti-CD20 monoclonal antibody (GA101) evokes homotypic adhesion and actin-dependent, lysosome-mediated cell death in B-cell malignancies." Blood **117**(17): 4519-4529.

Alizadeh, A. A., M. B. Eisen, R. E. Davis, C. Ma, I. S. Lossos, A. Rosenwald, J. C. Boldrick, H. Sabet, T. Tran, X. Yu, J. I. Powell, L. Yang, G. E. Marti, T. Moore, J. Hudson, L. Lu, D. B. Lewis, R. Tibshirani, G. Sherlock, W. C. Chan, T. C. Greiner, D. D. Weisenburger, J. O. Armitage, R. Warnke, R. Levy, W. Wilson, M. R. Grever, J. C. Byrd, D. Botstein, P. O. Brown and L. M. Staudt (2000). "Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling." Nature **403**(6769): 503-511.

Allen, C. D., K. M. Ansel, C. Low, R. Lesley, H. Tamamura, N. Fujii and J. G. Cyster (2004). "Germinal center dark and light zone organization is mediated by CXCR4 and CXCR5." Nat Immunol **5**(9): 943-952.

Allen, C. D., T. Okada and J. G. Cyster (2007). "Germinal-center organization and cellular dynamics." Immunity **27**(2): 190-202.

Amir el, A. D., K. L. Davis, M. D. Tadmor, E. F. Simonds, J. H. Levine, S. C. Bendall, D. K. Shenfeld, S. Krishnaswamy, G. P. Nolan and D. Pe'er (2013). "viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia." Nat Biotechnol **31**(6): 545-552.

Bandura, D. R., V. I. Baranov, O. I. Ornatsky, A. Antonov, R. Kinach, X. Lou, S. Pavlov, S. Vorobiev, J. E. Dick and S. D. Tanner (2009). "Mass Cytometry: Technique for Real Time Single Cell Multitarget Immunoassay Based on Inductively Coupled Plasma Time-of-Flight Mass Spectrometry." Analytical Chemistry **81**(16): 6813-6822.

Barmania, F. and M. S. Pepper (2013). "C-C chemokine receptor type five (CCR5): An emerging target for the control of HIV infection." Applied & Translational Genomics **2**: 3-16.

Basit, A., J. Reutershan, M. A. Morris, M. Solga, C. E. Rose, Jr. and K. Ley (2006). "ICAM-1 and LFA-1 play critical roles in LPS-induced neutrophil recruitment into the alveolar space." Am J Physiol Lung Cell Mol Physiol **291**(2): L200-207.

Bax, M., T. W. J. Huizinga and R. E. M. Toes (2014). "The pathogenic potential of autoreactive antibodies in rheumatoid arthritis." Seminars in Immunopathology **36**(3): 313-325.

Becher, B., A. Schlitzer, J. Chen, F. Mair, H. R. Sumatoh, K. W. Teng, D. Low, C. Ruedl, P. Riccardi-Castagnoli, M. Poidinger, M. Greter, F. Ginhoux and E. W. Newell (2014). "High-dimensional analysis of the murine myeloid cell system." Nat Immunol **15**(12): 1181-1189.

Behbehani, G. K., S. C. Bendall, M. R. Clutter, W. J. Fantl and G. P. Nolan (2012). "Single Cell Mass Cytometry Adapted to Measurements of the Cell Cycle()." Cytometry A **81**(7): 552-566.

Bendall, S. C., K. L. Davis, A. D. Amir el, M. D. Tadmor, E. F. Simonds, T. J. Chen, D. K. Shenfeld, G. P. Nolan and D. Pe'er (2014). "Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development." Cell **157**(3): 714-725.

Bendall, S. C., E. F. Simonds, P. Qiu, A. D. Amir el, P. O. Krutzik, R. Finck, R. V. Bruggner, R. Melamed, A. Trejo, O. I. Ornatsky, R. S. Balderas, S. K. Plevritis, K. Sachs, D. Pe'er, S. D. Tanner and G. P. Nolan (2011). "Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum." Science **332**(6030): 687-696.

Berkowska, M. A., G. J. A. Driessen, V. Bikos, C. Grosserichter-Wagener, K. Stamatopoulos, A. Cerutti, B. He, K. Biermann, J. F. Lange, M. van der Burg, J. J. M. van Dongen and M. C. van Zelm (2011). "Human memory B cells originate from three distinct germinal center-dependent and -independent maturation pathways." Blood **118**(8): 2150-2158.

Brachtel, E. F., M. Washiyama, G. D. Johnson, K. Tenner-Racz, P. Racz and I. C. MacLennan (1996). "Differences in the germinal centres of palatine tonsils and lymph nodes." Scand J Immunol **43**(3): 239-247.

Brack, C., M. Hirama, R. Lenhard-Schuller and S. Tonegawa (1978). "A complete immunoglobulin gene is created by somatic recombination." Cell **15**(1): 1-14.

Braza, F., J. Chesne, S. Castagnet, A. Magnan and S. Brouard (2014). "Regulatory functions of B cells in allergic diseases." Allergy **69**(11): 1454-1463.

Breitling, R. (2010). "What is Systems Biology?" Frontiers in Physiology **1**: 9.

Brocker, E. B., L. Suter and C. Sorg (1984). "HLA-DR antigen expression in primary melanomas of the skin." J Invest Dermatol **82**(3): 244-247.

Brodin, P. and M. M. Davis (2017). "Human immune system variation." Nat Rev Immunol **17**(1): 21-29.

Browning, J. L. (2006). "B cells move to centre stage: novel opportunities for autoimmune disease treatment." Nat Rev Drug Discov **5**(7): 564-576.

Bruggner, R. V., B. Bodenmiller, D. L. Dill, R. J. Tibshirani and G. P. Nolan (2014). "Automated identification of stratifying signatures in cellular subpopulations." Proc Natl Acad Sci U S A **111**(26): E2770-2777.

Buchner, M. and M. Müschen (2014). "Targeting the B cell receptor signaling pathway in B lymphoid." Curr Opin Hematol **21**(4): 341-349.

Cerutti, A. (2008). "The regulation of IgA class switching." Nat Rev Immunol **8**(6): 421-434.

Cha, S.-C., H. Qin, S. Kannan, S. Rawal, L. S. Watkins, F. E. Baio, W. Wu, J. Ong, J. Wei, B. Kwak, S. Kim, M. S. Popescu, D. S. Paick, K. Kim, A. Luong, R. E. Davis, H. W. Schroeder, L. W. Kwak and S. S. Neelapu (2013). "Nonstereotyped Lymphoma B Cell Receptors Recognize Vimentin as a Shared Autoantigen." The Journal of Immunology **190**(9): 4887-4898.

Chattopadhyay, P. K., T. M. Gierahn, M. Roederer and J. C. Love (2014). "Single-cell technologies for monitoring immune systems." Nat Immunol **15**(2): 128-135.

Chen, X. and P. E. Jensen (2008). "The role of B lymphocytes as antigen-presenting cells." Arch Immunol Ther Exp (Warsz) **56**(2): 77-83.

Chester, C. and H. T. Maecker (2015). "Algorithmic Tools for Mining High-Dimensional Cytometry Data." The Journal of Immunology **195**(3): 773-779.

Cho, S. H., A. L. Raybuck, K. Stengel, M. Wei, T. C. Beck, E. Volanakis, J. W. Thomas, S. Hiebert, V. H. Haase and M. R. Boothby (2016). "Germinal centre hypoxia and regulation of antibody qualities by a hypoxia response system." Nature **537**(7619): 234-238.

Civin, C. I., L. C. Strauss, C. Brovall, M. J. Fackler, J. F. Schwartz and J. H. Shaper (1984). "Antigenic analysis of hematopoiesis. III. A hematopoietic progenitor cell surface antigen defined by a monoclonal antibody raised against KG-1a cells." J Immunol **133**(1): 157-165.

Clem, A. S. (2011). "Fundamentals of Vaccine Immunology." J Glob Infect Dis **3**(1): 73-78.

Coelho, V., S. Krysov, A. M. Ghaemmaghami, M. Emara, K. N. Potter, P. Johnson, G. Packham, L. Martinez-Pomares and F. K. Stevenson (2010). "Glycosylation of surface Ig creates a functional bridge between human follicular lymphoma and microenvironmental lectins." Proceedings of the National Academy of Sciences **107**(43): 18587-18592.

Conley, M. E., A. K. Dobbs, D. M. Farmer, S. Kilic, K. Paris, S. Grigoriadou, E. Coustan-Smith, V. Howard and D. Campana (2009). "Primary B cell immunodeficiencies: comparisons and contrasts." Annu Rev Immunol **27**: 199-227.

Cox, C., J. E. Reeder, R. D. Robinson, S. B. Suppes and L. L. Wheeless (1988). "Comparison of frequency distributions in flow cytometry." Cytometry **9**(4): 291-298.

Craig, F. E. and K. A. Foon (2008). "Flow cytometric immunophenotyping for hematologic neoplasms." Blood **111**(8): 3941-3967.

Crotty, S. (2015). "A brief history of T cell help to B cells." Nat Rev Immunol **15**(3): 185-189.

D'Amato, G., A. Salzillo, A. Piccolo, M. D'Amato and G. Liccardi (2007). "A review of anti-IgE monoclonal antibody (omalizumab) as add on therapy for severe allergic (IgE-mediated) asthma." Ther Clin Risk Manag **3**(4): 613-619.

Dave, S. S. (2006). "Gene expression signatures and outcome prediction in mature B-cell malignancies." Current Treatment Options in Oncology **7**(4): 261-269.

Dave, S. S., G. Wright, B. Tan, A. Rosenwald, R. D. Gascoyne, W. C. Chan, R. I. Fisher, R. M. Braziel, L. M. Rimsza, T. M. Grogan, T. P. Miller, M. LeBlanc, T. C. Greiner, D. D. Weisenburger, J. C. Lynch, J. Vose, J. O. Armitage, E. B. Smeland, S. Kvaloy, H. Holte, J. Delabie, J. M. Connors, P. M. Lansdorp, Q. Ouyang, T. A. Lister, A. J. Davies, A. J. Norton, H. K. Muller-Hermelink, G. Ott, E. Campo, E. Montserrat, W. H. Wilson, E. S. Jaffe, R. Simon, L. Yang, J. Powell, H. Zhao, N. Goldschmidt, M. Chiorazzi and L. M. Staudt (2004). "Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells." N Engl J Med **351**(21): 2159-2169.

Davidson, B., B. Risberg, A. Berner, E. B. Smeland and E. Torlakovic (1999). "Evaluation of lymphoid cell populations in cytology specimens using flow cytometry and polymerase chain reaction." Diagn Mol Pathol **8**(4): 183-188.

Detry, G., B. Drenou, A. Ferrant, I. Theate, L. Michaux, J. M. Scheiff, D. Latinne, P. Leveugle, A. M. Mazzon and V. Deneys (2004). "Tracking the follicular lymphoma cells in flow cytometry: characterisation of a new useful antibody combination." Eur J Haematol **73**(5): 325-331.

Diggins, K. E., P. B. Ferrell, Jr. and J. M. Irish (2015). "Methods for discovery and characterization of cell subsets in high dimensional mass cytometry data." Methods **82**: 55-63.

Diggins, K. E., A. R. Greenplate, N. Leelatian, C. E. Wogsland and J. M. Irish (2017). "Characterizing cell subsets using marker enrichment modeling." Nat Methods **14**(3): 275-278.

Dotan, E., C. Aggarwal and M. R. Smith (2010). "Impact of Rituximab (Rituxan) on the Treatment of B-Cell Non-Hodgkin's Lymphoma." P t **35**(3): 148-157.

Doulatov, S., F. Notta, E. Laurenti and J. E. Dick (2012). "Hematopoiesis: a human perspective." Cell Stem Cell **10**(2): 120-136.

Eberth, S., B. Schneider, A. Rosenwald, E. M. Hartmann, J. Romani, M. Zaborski, R. Siebert, H. G. Drexler and H. Quentmeier (2010). "Epigenetic regulation of CD44 in Hodgkin and non-Hodgkin lymphoma." BMC Cancer **10**: 517.

Eibel, H., H. Kraus, H. Sic, A. K. Kienzler and M. Rizzi (2014). "B cell biology: an overview." Curr Allergy Asthma Rep **14**(5): 434.

Eide, M. B., K. Liestøl, O. C. Lingjærde, M. E. Hystad, S. H. Kresse, L. Meza-Zepeda, O. Myklebost, G. Trøen, H. V. Aamot, H. Holte, E. B. Smeland and J. Delabie (2010). "Genomic alterations reveal potential for higher grade transformation in follicular lymphoma and confirm parallel evolution of tumor cell clones." Blood **116**(9): 1489-1497.

Elenitoba-Johnson, K. S. J., S. D. Jenson, R. T. Abbott, R. A. Palais, S. D. Bohling, Z. Lin, S. Tripp, P. J. Shami, L. Y. Wang, R. W. Coupland, R. Buckstein, B. Perez-Ordonez, S. L. Perkins, I. D. Dube and M. S. Lim (2003). "Involvement of multiple signaling pathways in follicular lymphoma transformation: p38-mitogen-activated protein kinase as a target for therapy." Proceedings of the National Academy of Sciences **100**(12): 7259-7264.

Evans, S. S. and A. B. Clemmons (2015). "Obinutuzumab: A Novel Anti-CD20 Monoclonal Antibody for Chronic Lymphocytic Leukemia." J Adv Pract Oncol **6**(4): 370-374.

Ezawa, K., M. Yamamura, H. Matsui, Z. Ota and H. Makino (1997). "Comparative analysis of CD45RA- and CD45RO-positive CD4+T cells in peripheral blood, synovial fluid, and synovial tissue in patients with rheumatoid arthritis and osteoarthritis." Acta Med Okayama **51**(1): 25-31.

Ferrell, P. B., Jr., K. E. Diggins, H. G. Polikowsky, S. R. Mohan, A. C. Seegmiller and J. M. Irish (2016). "High-Dimensional Analysis of Acute Myeloid Leukemia Reveals Phenotypic Changes in Persistent Cells during Induction Therapy." PLoS One **11**(4): e0153207.

Feuillard, J., D. Taylor, M. Casamayor-Palleja, G. D. Johnson and I. C. MacLennan (1995). "Isolation and characteristics of tonsil centroblasts with reference to Ig class switching." Int Immunol **7**(1): 121-130.

Freedman, A. (2015). "Follicular lymphoma: 2015 update on diagnosis and management." Am J Hematol **90**(12): 1171-1178.

Frei, A. P., F.-A. Bava, E. R. Zunder, E. W. Y. Hsieh, S.-Y. Chen, G. P. Nolan and P. F. Gherardini (2016). "Highly multiplexed simultaneous detection of RNAs and proteins in single cells." Nat Meth **13**(3): 269-275.

Friend, C., W. Marovitz, G. Henie, W. Henie, D. Tsuei, K. Hirschhorn, J. G. Holland and J. Cuttner (1978). "Observations on cell lines derived from a patient with Hodgkin's disease." Cancer Res **38**(8): 2581-2591.

Furze, R. C. and S. M. Rankin (2008). "Neutrophil mobilization and clearance in the bone marrow." Immunology **125**(3): 281-288.

Gadol, N., M. A. Peacock and K. A. Ault (1988). "Antigenic phenotype and functional characterization of human tonsil B cells." Blood **71**(4): 1048-1055.

Gascan, H., J. F. Gauchat, G. Aversa, P. Van Vlasselaer and J. E. de Vries (1991). "Anti-CD40 monoclonal antibodies or CD4+ T cell clones and IL-4 induce IgG4 and IgE switching in purified human B cells via different signaling pathways." J Immunol **147**(1): 8-13.

Gascan, H., J. F. Gauchat, M. G. Roncarolo, H. Yssel, H. Spits and J. E. de Vries (1991). "Human B cell clones can be induced to proliferate and to switch to IgE and IgG4 synthesis by interleukin 4 and a signal provided by activated CD4+ T cell clones." J Exp Med **173**(3): 747-750.

Gaudilliere, B., G. K. Fragiadakis, R. V. Bruggner, M. Nicolau, R. Finck, M. Tingle, J. Silva, E. A. Ganio, C. G. Yeh, W. J. Maloney, J. I. Huddleston, S. B. Goodman, M. M. Davis, S. C. Bendall, W. J. Fantl, M. S. Angst and G. P. Nolan (2014). "Clinical recovery from surgery correlates with single-cell immune signatures." Sci Transl Med **6**(255): 255ra131.

Gerstein, R., Z. Zhou, H. Zhang, A. Evens, N. Walsh, L. Shultz, R. Bortell, M. Brehm, J. Cerny, G. Whalen, J. Leonard, B. Woda, D. Greiner and A. G. Rosmarin (2015). "Patient-Derived Xenografts (PDX) of B Cell Lymphoma in NSG Mice: A Mouse Avatar for Developing Personalized Medicine." Blood **126**(23): 5408-5408.

Giesen, C., H. A. Wang, D. Schapiro, N. Zivanovic, A. Jacobs, B. Hattendorf, P. J. Schuffler, D. Grolimund, J. M. Buhmann, S. Brandt, Z. Varga, P. J. Wild, D. Gunther and B. Bodenmiller (2014). "Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry." Nat Methods **11**(4): 417-422.

Gitlin, A. D., Z. Shulman and M. C. Nussenzweig (2014). "Clonal selection in the germinal centre by regulated proliferation and hypermutation." Nature **509**(7502): 637-640.

Givan, A. L. (2001). Flow cytometry : first principles. New York, NY, Wiley-Liss.

Gould, H. J., B. J. Sutton, A. J. Beavil, R. L. Beavil, N. McCloskey, H. A. Coker, D. Fear and L. Smurthwaite (2003). "The biology of IGE and the basis of allergic disease." Annu Rev Immunol **21**: 579-628.

Green, M. R., A. J. Gentles, R. V. Nair, J. M. Irish, S. Kihira, C. L. Liu, I. Kela, E. S. Hopmans, J. H. Myklebust, H. Ji, S. K. Plevritis, R. Levy and A. A. Alizadeh (2013). "Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma." Blood **121**(9): 1604-1611.

Green, M. R., S. Kihira, C. L. Liu, R. V. Nair, R. Salari, A. J. Gentles, J. Irish, H. Stehr, C. Vicente-Dueñas, I. Romero-Camarero, I. Sanchez-Garcia, S. K. Plevritis, D. A. Arber, S. Batzoglou, R. Levy and A. A. Alizadeh (2015). "Mutations in early follicular lymphoma progenitors are associated with suppressed antigen presentation." Proc Natl Acad Sci U S A **112**(10): E1116-1125.

Greenplate, A. R., D. B. Johnson, P. B. Ferrell and J. M. Irish (2016). "Systems immune monitoring in cancer therapy." Eur J Cancer **61**: 77-84.

Greenplate, A. R., D. B. Johnson, M. Roussel, M. R. Savona, J. A. Sosman, I. Puzanov, P. B. Ferrell, Jr. and J. M. Irish (2016). "Myelodysplastic Syndrome Revealed by Systems Immunology in a Melanoma Patient Undergoing Anti-PD-1 Therapy." Cancer Immunol Res.

Gregory R. Warnes, B. B., Lodewijk Bonebakker, Robert Gentleman, Wolfgang Huber Andy Liaw, Thomas Lumley, Martin Maechler, Arni Magnusson, Steffen Moeller, Marc Schwartz, Bill Venables (2015). "gplots: Various R Programming Tools for Plotting Data."

Grier, D. D., S. Z. Al-Quran, D. M. Cardona, Y. Li and R. C. Braylan (2012). "Flow cytometric analysis of immunoglobulin heavy chain expression in B-cell lymphoma and reactive lymphoid hyperplasia." Int J Clin Exp Pathol **5**(2): 110-118.

Grosmaire, L. S., M. S. Hayden-Ledbetter, J. A. Ledbetter, P. A. Thompson, S. A. Simon and W. Brady (2014). B-cell reduction using cd37-specific and cd20-specific binding molecules, Google Patents.

Guy, K., A. S. Krajewski and A. E. Dewar (1986). "Expression of MHC class II antigens in human B-cell leukaemia and non-Hodgkin's lymphoma." Br J Cancer **53**(2): 161-173.

Hagan, T., H. I. Nakaya, S. Subramaniam and B. Pulendran (2015). "Systems vaccinology: Enabling rational vaccine design with systems biological approaches." Vaccine **33**(40): 5294-5301.

Hagman, J. and K. Lukin (2006). "Transcription factors drive B cell development." Current Opinion in Immunology **18**(2): 127-134.

Hagman, J., J. Ramirez and K. Lukin (2012). "B lymphocyte lineage specification, commitment and epigenetic control of transcription by early B cell factor 1." Curr Top Microbiol Immunol **356**: 17-38.

Hahne, F., N. LeMeur, R. R. Brinkman, B. Ellis, P. Haaland, D. Sarkar, J. Spidlen, E. Strain and R. Gentleman (2009). "flowCore: a Bioconductor package for high throughput flow cytometry." BMC Bioinformatics **10**: 106.

Hara-Chikuma, M., S. Chikuma, Y. Sugiyama, K. Kabashima, A. S. Verkman, S. Inoue and Y. Miyachi (2012). "Chemokine-dependent T cell migration requires aquaporin-3-mediated hydrogen peroxide uptake." J Exp Med **209**(10): 1743-1752.

Hardwick, J. M. and L. Soane (2013). "Multiple functions of BCL-2 family proteins." Cold Spring Harb Perspect Biol **5**(2).

Herzenberg, L. A., D. Parks, B. Sahaf, O. Perez, M. Roederer and L. A. Herzenberg (2002). "The history and future of the fluorescence activated cell sorter and flow cytometry: a view from Stanford." Clin Chem **48**(10): 1819-1827.

Hoffman, W., F. G. Lakkis and G. Chalasani (2016). "B Cells, Antibodies, and More." Clin J Am Soc Nephrol **11**(1): 137-154.

Horna, P., H. Olteanu, S. H. Kroft and A. M. Harrington (2011). "Flow cytometric analysis of surface light chain expression patterns in B-cell lymphomas using monoclonal and polyclonal antibodies." Am J Clin Pathol **136**(6): 954-959.

Huse, K., M. Bakkebo, M. P. Oksvold, L. Forfang, V. I. Hilden, T. Stokke, E. B. Smeland and J. H. Myklebust (2011). "Bone morphogenetic proteins inhibit CD40L/IL-21-induced Ig production in human B cells: differential effects of BMP-6 and BMP-7." Eur J Immunol **41**(11): 3135-3145.

Irish, J. M. (2014). "Beyond the age of cellular discovery." Nat Immunol **15**(12): 1095-1097.

Irish, J. M., D. K. Czerwinski, G. P. Nolan and R. Levy (2006). "Altered B-cell receptor signaling kinetics distinguish human follicular lymphoma B cells from tumor-infiltrating nonmalignant B cells." Blood **108**(9): 3135-3142.

Irish, J. M., D. K. Czerwinski, G. P. Nolan and R. Levy (2006). "Altered B-cell receptor signaling kinetics distinguish human follicular lymphoma B cells from tumor-infiltrating nonmalignant B cells." Blood **108**(9): 3135-3142.

Irish, J. M., D. K. Czerwinski, G. P. Nolan and R. Levy (2006). "Kinetics of B Cell Receptor Signaling in Human B Cell Subsets Mapped by Phosphospecific Flow Cytometry." The Journal of Immunology **177**(3): 1581-1589.

Irish, J. M., D. K. Czerwinski, G. P. Nolan and R. Levy (2006). "Kinetics of B cell receptor signaling in human B cell subsets mapped by phosphospecific flow cytometry." J Immunol **177**(3): 1581-1589.

Irish, J. M. and D. B. Doxie (2014). "High-dimensional single-cell cancer biology." Curr Top Microbiol Immunol **377**: 1-21.

Irish, J. M., R. Hovland, P. O. Krutzik, O. D. Perez, O. Bruserud, B. T. Gjertsen and G. P. Nolan (2004). "Single cell profiling of potentiated phospho-protein networks in cancer cells." Cell **118**(2): 217-228.

Irish, J. M., J. H. Myklebust, A. A. Alizadeh, R. Houot, J. P. Sharman, D. K. Czerwinski, G. P. Nolan and R. Levy (2010). "B-cell signaling networks reveal a negative prognostic human lymphoma cell subset that emerges during tumor progression." Proceedings of the National Academy of Sciences.

Irish, J. M., J. H. Myklebust, A. A. Alizadeh, R. Houot, J. P. Sharman, D. K. Czerwinski, G. P. Nolan and R. Levy (2010). "B-cell signaling networks reveal a negative prognostic human lymphoma cell subset that emerges during tumor progression." Proceedings of the National Academy of Sciences **107**(29): 12747-12754.

Jackson, S. M., P. C. Wilson, J. A. James and J. D. Capra (2008). "Human B cell subsets." Adv Immunol **98**: 151-224.

Jacques, B. and R. Françoise (1991). "Growing human B lymphocytes in the CD40 system." Nature **353**(6345): 678-679.

Janeway, C. J., P. Travers and M. Walport, et al. (2001). The major histocompatibility complex and its functions. Immunobiology: The Immune System in Health and Disease. New York, Garland Science.

Jeannin, P., S. Lecoanet, Y. Delneste, J. F. Gauchat and J. Y. Bonnefoy (1998). "IgE versus IgG4 production can be differentially regulated by IL-10." Journal of Immunology **160**(7): 3555-3561.

Johnson, D. B., M. V. Estrada, R. Salgado, V. Sanchez, D. B. Doxie, S. R. Opalenik, A. E. Vilgelm, E. Feld, A. S. Johnson, A. R. Greenplate, M. E. Sanders, C. M. Lovly, D. T. Frederick, M. C. Kelley, A. Richmond, J. M. Irish, Y. Shyr, R. J. Sullivan, I. Puzanov, J. A. Sosman and J. M. Balko (2016). "Melanoma-specific MHC-II expression represents a tumour-autonomous phenotype and predicts response to anti-PD-1/PD-L1 therapy." Nature Communications **7**: 10582.

Johnson, D. B., M. V. Estrada, R. Salgado, V. Sanchez, D. B. Doxie, S. R. Opalenik, A. E. Vilgelm, E. Feld, A. S. Johnson, A. R. Greenplate, M. E. Sanders, C. M. Lovly, D. T. Frederick, M. C. Kelley, A. Richmond, J. M. Irish, Y. Shyr, R. J. Sullivan, I. Puzanov, J. A. Sosman and J. M. Balko (2016). "Melanoma-specific MHC-II expression represents a tumour-autonomous phenotype and predicts response to anti-PD-1/PD-L1 therapy." Nat Commun **7**: 10582.

Jourdan, M., A. Caraux, J. De Vos, G. Fiol, M. Larroque, C. Cognot, C. Bret, C. Duperray, D. Hose and B. Klein (2009). "An in vitro model of differentiation of memory B cells into plasmablasts and plasma cells including detailed phenotypic and molecular characterization." Blood **114**(25): 5173-5181.

Jung, S. N., W. K. Yang, J. Kim, H. S. Kim, E. J. Kim, H. Yun, H. Park, S. S. Kim, W. Choe, I. Kang and J. Ha (2008). "Reactive oxygen species stabilize hypoxia-inducible factor-1 alpha protein and stimulate transcriptional activity via AMP-activated protein kinase in DU145 human prostate cancer cells." Carcinogenesis **29**(4): 713-721.

Karnell, J. L., N. Dimasi, F. G. Karnell, R. Fleming, E. Kuta, M. Wilson, H. Wu, C. Gao, R. Herbst and R. Ettinger (2014). "CD19 and CD32b Differentially Regulate Human B Cell Responsiveness." J Immunol **192**(4): 1480-1490.

Karpova, M. B., J. Schoumans, I. Ernberg, J. I. Henter, M. Nordenskjold and B. Fadeel (2005). "Raji revisited: cytogenetics of the original Burkitt's lymphoma cell line." Leukemia **19**(1): 159-161.

Kaur, G. and J. M. Dufour (2012). "Cell lines: Valuable tools or useless artifacts." Spermatogenesis **2**(1): 1-5.

Khalil, A. M., J. C. Cambier and M. J. Shlomchik (2012). "B Cell Receptor Signal Transduction in the GC Is Short-Circuited by High Phosphatase Activity." Science **336**(6085): 1178-1181.

Kim, D., V. S. Donnenberg, J. W. Wilson and A. D. Donnenberg (2016). "The use of simultaneous confidence bands for comparison of single parameter fluorescent intensity data." Cytometry A **89**(1): 89-97.

Kinet, J. P. (1999). "The high-affinity IgE receptor (Fc epsilon RI): from physiology to pathology." Annu Rev Immunol **17**: 931-972.

Kirschner, M. W. (2005). "The Meaning of Systems Biology." Cell **121**(4): 503-504.

Klein, U. and R. Dalla-Favera (2008). "Germinal centres: role in B-cell physiology and malignancy." Nat Rev Immunol **8**(1): 22-33.

Klein, U., Y. Tu, G. A. Stolovitzky, J. L. Keller, J. Haddad, V. Miljkovic, G. Cattoretti, A. Califano and R. Dalla-Favera (2003). "Transcriptional analysis of the B cell germinal center reaction." Proceedings of the National Academy of Sciences **100**(5): 2639-2644.

Koh, H. J., H. H. Park and C. E. Lee (2000). "Regulation of IgE and type II IgE receptor expression by insulin-like growth factor-1: Role of STAT6 and NF-kappa B." Journal of Biochemistry and Molecular Biology **33**(6): 454-462.

Kondo, T., H. Takata and M. Takiguchi (2007). "Functional expression of chemokine receptor CCR6 on human effector memory CD8+ T cells." Eur J Immunol **37**(1): 54-65.

Kossev, P. M., P. N. Raghunath, A. Bagg, S. Schuster, J. E. Tomaszewski and M. A. Wasik (2001). "SHP-1 expression by malignant small B-cell lymphomas reflects the maturation stage of their normal B-cell counterparts." Am J Surg Pathol **25**(7): 949-955.

Kotecha, N., P. O. Krutzik and J. M. Irish (2010). "Web-based analysis and publication of flow cytometry experiments." Curr Protoc Cytom **Chapter 10**: Unit10 17.

Kraus, M., M. B. Alimzhanov, N. Rajewsky and K. Rajewsky (2004). "Survival of resting mature B lymphocytes depends on BCR signaling via the Igalpha/beta heterodimer." Cell **117**(6): 787-800.

Kremmidiotis, G. and H. Zola (1995). "Changes in CD44 expression during B cell differentiation in the human tonsil." Cell Immunol **161**(2): 147-157.

Kumar, S., A. K. Verma, M. Das and P. D. Dwivedi (2012). "Molecular mechanisms of IgE mediated food allergy." Int Immunopharmacol **13**(4): 432-439.

Kuppers, R. (2005). "Mechanisms of B-cell lymphoma pathogenesis." Nat Rev Cancer **5**(4): 251-262.

Lam, K. P., R. Kuhn and K. Rajewsky (1997). "In vivo ablation of surface immunoglobulin on mature B cells by inducible gene targeting results in rapid cell death." Cell **90**(6): 1073-1083.

Larbi, A. and T. Fulop (2014). "From “truly naïve” to “exhausted senescent” T cells: When markers predict functionality." Cytometry Part A **85**(1): 25-35.

LeBien, T. W. (2000). "Fates of human B-cell precursors." Blood **96**(1): 9-23.

LeBien, T. W. and T. F. Tedder (2008). "B lymphocytes: how they develop and function." Blood **112**(5): 1570-1580.

Leelatian, N., K. E. Diggins and J. M. Irish (2015). "Characterizing Phenotypes and Signaling Networks of Single Human Cells by Mass Cytometry." Methods Mol Biol **1346**: 99-113.

Leelatian, N., D. B. Doxie, A. R. Greenplate, B. C. Mobley, J. M. Lehman, J. Sinnaeve, R. M. Kauffmann, J. A. Werkhaven, A. M. Mistry, K. D. Weaver, R. C. Thompson, P. P. Massion, M. A. Hooks, M. C. Kelley, L. B. Chambless, R. A. Ihrie and J. M. Irish (2016). "Single cell analysis of human tissues and solid tumors with mass cytometry." Cytometry B Clin Cytom.

Levine, J. H., E. F. Simonds, S. C. Bendall, K. L. Davis, A. D. Amir el, M. D. Tadmor, O. Litvin, H. G. Fienberg, A. Jager, E. R. Zunder, R. Finck, A. L. Gedman, I. Radtke, J. R. Downing, D. Pe'er and G. P. Nolan (2015). "Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis." Cell **162**(1): 184-197.

Li, X. and R. P. Kimberly (2014). "Targeting the Fc receptor in autoimmune disease." Expert Opin Ther Targets **18**(3): 335-350.

Lo K, H. F., Brinkman R and Gottardo R (2009). "FlowClust: a Bioconductor package for automated gating of flow cytometry data." BMC Bioinformatics **10**(R package version 3.1.0.).

Maecker, H. T., J. P. McCoy and R. Nussenblatt (2012). "Standardizing immunophenotyping for the Human Immunology Project." Nat Rev Immunol **12**(3): 191-200.

Majonis, D., I. Herrera, O. Ornatsky, M. Schulze, X. Lou, M. Soleimani, M. Nitz and M. A. Winnik (2010). "Synthesis of a Functional Metal-Chelating Polymer and Steps towards Quantitative Mass Cytometry Bioassays." Anal Chem **82**(21).

Mangan, N. E., R. E. Fallon, P. Smith, N. van Rooijen, A. N. McKenzie and P. G. Fallon (2004). "Helminth infection protects mice from anaphylaxis via IL-10-producing B cells." J Immunol **173**(10): 6346-6356.

Mantei, K. W., Brent L. (2009). "Flow cytometric evaluation of CD38 expression assists in distinguishing follicular hyperplasia from follicular lymphoma." Cytometry Part B: Clinical Cytometry **76B**(5).

McLaughlin, P. (2002). "Progress and Promise in the Treatment of Indolent Lymphomas." The Oncologist **7**(3): 217-225.

Medina, F., C. Segundo, A. Campos-Caro, I. González-García and J. A. Brieva (2002). "The heterogeneity shown by human plasma cells from tonsil, blood, and bone marrow reveals graded stages of increasing maturity, but local profiles of adhesion molecule expression." Blood **99**(6): 2154-2161.

Medvedovic, J., A. Ebert, H. Tagoh and M. Busslinger (2011). "Pax5: a master regulator of B cell development and leukemogenesis." Adv Immunol **111**: 179-206.

Mietzner, B., M. Tsuiji, J. Scheid, K. Velinzon, T. Tiller, K. Abraham, J. B. Gonzalez, V. Pascual, D. Stichweh, H. Wardemann and M. C. Nussenzweig (2008). "Autoreactive IgG memory antibodies in patients with systemic lupus erythematosus arise from nonreactive and polyreactive precursors." Proceedings of the National Academy of Sciences **105**(28): 9727-9732.

Montecino-Rodriguez, E. and K. Dorshkind (2012). "B-1 B Cell Development in the Fetus and Adult." Immunity **36**(1): 13-21.

Mosmann, T. R., I. Naim, J. Rebhahn, S. Datta, J. S. Cavenaugh, J. M. Weaver and G. Sharma (2014). "SWIFT-scalable clustering for automated identification of rare cell populations in large, high-dimensional flow cytometry datasets, Part 2: Biological evaluation." Cytometry A.

Muehling, L., C. Wogsland, R. Agrawal, P. W. Heymann, J. M. Irish and J. A. Woodfolk (2015). "High-Dimensional Single-Cell Monitoring of Circulating Cells in Allergic Asthmatics Infected with Rhinovirus Reveals Dynamic Flux in Diverse Immune Cells." Journal of Allergy and Clinical Immunology **135**(2, Supplement): AB382.

Mueller, J., M. Matloubian and J. Zikherman (2015). "Cutting Edge: An In Vivo Reporter Reveals Active B Cell Receptor Signaling in the Germinal Center." The Journal of Immunology **194**(7): 2993-2997.

Muramatsu, M., K. Kinoshita, S. Fagarasan, S. Yamada, Y. Shinkai and T. Honjo (2000). "Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme." Cell **102**(5): 553-563.

Murphy, K., P. Travers and M. Walport (2008). Janeway's Immunobiology 7th Edition. New York, NY, Garland Science.

Myklebust, J. H., J. M. Irish, J. Brody, D. K. Czerwinski, R. Houot, H. E. Kohrt, J. Timmerman, J. Said, M. R. Green, J. Delabie, A. Kolstad, A. A. Alizadeh and R. Levy (2013). "High PD-1 expression and suppressed cytokine signaling distinguish T cells infiltrating follicular lymphoma tumors from peripheral T cells." Blood **121**(8): 1367-1376.

Nathan, C. and A. Cunningham-Bussel (2013). "Beyond oxidative stress: an immunologist's guide to reactive oxygen species." Nat Rev Immunol **13**(5): 349-361.

Nemazee, D. (2006). "Receptor editing in lymphocyte development and central tolerance." Nat Rev Immunol **6**(10): 728-740.

Newell, E. W. and Y. Cheng (2016). "Mass cytometry: blessed with the curse of dimensionality." Nat Immunol **17**(8): 890-895.

Newell, E. W. and Y. Cheng (2016). "Mass cytometry: blessed with the curse of dimensionality." Nat Immunol **17**(8): 890-895.

Newell, E. W., N. Sigal, S. C. Bendall, G. P. Nolan and M. M. Davis (2012). "Cytometry by time-of-flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of CD8+ T cell phenotypes." Immunity **36**(1): 142-152.

Ngan , B.-Y., Z. Chen-Levy , L. M. Weiss , R. A. Warnke and M. L. Cleary (1988). "Expression in Non-Hodgkin's Lymphoma of the bcl-2 Protein Associated with the t(14;18) Chromosomal Translocation." New England Journal of Medicine **318**(25): 1638-1644.

Nicholas, K. J., A. R. Greenplate, D. K. Flaherty, B. K. Matlock, J. S. Juan, R. M. Smith, J. M. Irish and S. A. Kalams (2016). "Multiparameter analysis of stimulated human peripheral blood mononuclear cells: A comparison of mass and fluorescence cytometry." Cytometry A **89**(3): 271-280.

Nimmerjahn, F. and J. V. Ravetch (2008). "Fc[gamma] receptors as regulators of immune responses." Nat Rev Immunol **8**(1): 34-47.

Okada, T., M. J. Miller, I. Parker, M. F. Krummel, M. Neighbors, S. B. Hartley, A. O'Garra, M. D. Cahalan and J. G. Cyster (2005). "Antigen-engaged B cells undergo chemotaxis toward the T zone and form motile conjugates with helper T cells." PLoS Biol **3**(6): e150.

Orlova, D. Y., N. Zimmerman, S. Meehan, C. Meehan, J. Waters, E. E. Ghosn, A. Filatenkov, G. A. Kolyagin, Y. Gernez, S. Tsuda, W. Moore, R. B. Moss, L. A. Herzenberg and G. Walther (2016). "Earth Mover's Distance (EMD): A True Metric for Comparing Biomarker Expression Levels in Cell Populations." PLoS One **11**(3): e0151859.

Ornatsky, O., D. Bandura, V. Baranov, M. Nitz, M. A. Winnik and S. Tanner (2010). "Highly multiparametric analysis by mass cytometry." J Immunol Methods **361**(1-2): 1-20.

Patel, A. P., I. Tirosh, J. J. Trombetta, A. K. Shalek, S. M. Gillespie, H. Wakimoto, D. P. Cahill, B. V. Nahed, W. T. Curry, R. L. Martuza, D. N. Louis, O. Rozenblatt-Rosen, M. L. Suva, A. Regev and B. E. Bernstein (2014). "Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma." Science **344**(6190): 1396-1401.

Pene, J., F. Rousset, F. Briere, I. Chretien, J. Y. Bonnefoy, H. Spits, T. Yokota, N. Arai, K. Arai, J. Banchereau and et al. (1988). "IgE production by normal human lymphocytes is induced by interleukin 4 and suppressed by interferons gamma and alpha and prostaglandin E2." Proc Natl Acad Sci U S A **85**(18): 6880-6884.

Perez, M. E., L. A. Billordo, P. Baz, L. Fainboim and E. Arana (2014). "Human memory B cells isolated from blood and tonsils are functionally distinctive." Immunol Cell Biol.

Perosa, F., M. Prete, V. Racanelli and F. Dammacco (2010). "CD20-depleting therapy in autoimmune diseases: from basic research to the clinic." Journal of Internal Medicine **267**(3): 260-277.

Pillai, S., H. Mattoo and A. Cariappa (2011). "B cells and Autoimmunity." Curr Opin Immunol **23**(6): 721-731.

Polikowsky, H. G., C. E. Wogsland, K. E. Diggins, K. Huse and J. M. Irish (2015). "Cutting Edge: Redox Signaling Hypersensitivity Distinguishes Human Germinal Center B Cells." J Immunol **195**(4): 1364-1367.

Preffer, F. and D. Dombkowski (2009). "Advances in Complex Multiparameter Flow Cytometry Technology: Applications in Stem Cell Research." Cytometry B Clin Cytom **76**(5): 295-314.

Punnonen, J., G. Aversa, B. G. Cocks, A. N. McKenzie, S. Menon, G. Zurawski, R. de Waal Malefyt and J. E. de Vries (1993). "Interleukin 13 induces interleukin 4-independent IgG4 and IgE synthesis and CD23 expression by human B cells." Proc Natl Acad Sci U S A **90**(8): 3730-3734.

Punnonen, J., H. Yssel and J. E. de Vries (1997). "The relative contribution of IL-4 and IL-13 to human IgE synthesis induced by activated CD4+ or CD8+ T cells." J Allergy Clin Immunol **100**(6 Pt 1): 792-801.

Qiu, P., E. F. Simonds, S. C. Bendall, K. D. Gibbs, Jr., R. V. Bruggner, M. D. Linderman, K. Sachs, G. P. Nolan and S. K. Plevritis (2011). "Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE." Nat Biotechnol **29**(10): 886-891.

R Core Team (2013) (URL https://www.R-project.org). R: A Language and Environment for Statistical Computing. Vienna, Austria, R Foundation for Statistical Computing.

Rao, D. A., M. F. Gurish, J. L. Marshall, K. Slowikowski, C. Y. Fonseka, Y. Liu, L. T. Donlin, L. A. Henderson, K. Wei, F. Mizoguchi, N. C. Teslovich, M. E. Weinblatt, E. M. Massarotti, J. S. Coblyn, S. M. Helfgott, Y. C. Lee, D. J. Todd, V. P. Bykerk, S. M. Goodman, A. B. Pernis, L. B. Ivashkiv, E. W. Karlson, P. A. Nigrovic, A. Filer, C. D. Buckley, J. A. Lederer, S. Raychaudhuri and M. B. Brenner (2017). "Pathologically expanded peripheral T helper cell subset drives B cells in rheumatoid arthritis." Nature **542**(7639): 110-114.

Reth, M. (2002). "Hydrogen peroxide as second messenger in lymphocyte activation." Nat Immunol **3**(12): 1129-1134.

Rodríguez-Pinto, D. (2005). "B cells as antigen presenting cells." Cellular Immunology **238**(2): 67-75.

Rojkind, M., J.-A. Domínguez-Rosales, N. Nieto and P. Greenwel (2002). "Role of hydrogen peroxide and oxidative stress in healing responses." Cellular and Molecular Life Sciences CMLS **59**(11): 1872-1891.

RStudio Team (2015) ( URL http://www.rstudio.com/). RStudio: Integrated Development for R. Boston, MA, RStudio, Inc.

Sachen, K. L., M. J. Strohman, J. Singletary, A. A. Alizadeh, N. H. Kattah, C. Lossos, E. D. Mellins, S. Levy and R. Levy (2012). "Self-antigen recognition by follicular lymphoma B-cell receptors." Blood **120**(20): 4182-4190.

Saeys, Y., S. V. Gassen and B. N. Lambrecht (2016). "Computational flow cytometry: helping to make sense of high-dimensional immunology data." Nat Rev Immunol **16**(7): 449-462.

Salzer, E., E. Santos-Valente, S. Klaver, S. A. Ban, W. Emminger, N. K. Prengemann, W. Garncarz, L. Müllauer, R. Kain, H. Boztug, A. Heitger, K. Arbeiter, F. Eitelberger, M. G. Seidel, W. Holter, A. Pollak, W. F. Pickl, E. Förster-Waldl and K. Boztug (2013). "B-cell deficiency and severe autoimmunity caused by deficiency of protein kinase C δ." Blood **121**(16): 3112-3116.

Schroeder, H. W., Jr. and L. Cavacini (2010). "Structure and function of immunoglobulins." J Allergy Clin Immunol **125**(2 Suppl 2): S41-52.

Schwartz, M., Y. Zhang and J. D. Rosenblatt (2016). "B cell regulation of the anti-tumor response and role in carcinogenesis." Journal for ImmunoTherapy of Cancer **4**(1): 40.

Scott, D. W. and R. D. Gascoyne (2014). "The tumour microenvironment in B cell lymphomas." Nat Rev Cancer **14**(8): 517-534.

Sen, N., G. Mukherjee, A. Sen, S. C. Bendall, P. Sung, G. P. Nolan and A. M. Arvin (2014). "Single-cell mass cytometry analysis of human tonsil T cell remodeling by varicella zoster virus." Cell Rep **8**(2): 633-645.

Shaffer, A. L., A. Rosenwald and L. M. Staudt (2002). "Lymphoid Malignancies: the dark side of B-cell differentiation." Nat Rev Immunol **2**(12): 920-933.

Shekhar, K., P. Brodin, M. M. Davis and A. K. Chakraborty (2014). "Automatic Classification of Cellular Expression by Nonlinear Stochastic Embedding (ACCENSE)." Proceedings of the National Academy of Sciences **111**(1): 202-207.

Shekhar, K., P. Brodin, M. M. Davis and A. K. Chakraborty (2014). "Automatic Classification of Cellular Expression by Nonlinear Stochastic Embedding (ACCENSE)." Proc Natl Acad Sci U S A **111**(1): 202-207.

Shen, P. and S. Fillatreau (2015). "Antibody-independent functions of B cells: a focus on cytokines." Nat Rev Immunol **15**(7): 441-451.

Shlomchik, M. J. and F. Weisel (2012). "Germinal center selection and the development of memory B and plasma cells." Immunol Rev **247**(1): 52-63.

Shulman, Z., A. D. Gitlin, J. S. Weinstein, B. Lainez, E. Esplugues, R. A. Flavell, J. E. Craft and M. C. Nussenzweig (2014). "Dynamic signaling by T follicular helper cells during germinal center B cell selection." Science **345**(6200): 1058-1062.

Siegrist, C.-A. (2008). Vaccine immunology. Vaccines. W. A. O. Stanley A. Plotkin, Paul A. Offit**:** 17-36.

Silva, G. B. R. F. d., T. G. A. Silva, R. A. Duarte, N. L. Neto, H. Carrara, #xe9, l. H. Angotti, E. A. Donadi, #xf4, nio, Gon, #xe7, M. A. G. alves, #xe3, es, E. G. Soares and C. P. Soares (2013).

"Expression of the Classical and Nonclassical HLA Molecules in Breast Cancer." International Journal of Breast Cancer **2013**: 9.

Singh, D. K., D. Kumar, Z. Siddiqui, S. K. Basu, V. Kumar and K. V. Rao (2005). "The strength of receptor signaling is centrally controlled through a cooperative loop between Ca2+ and an oxidant signal." Cell **121**(2): 281-293.

Solal-Celigny, P., P. Roy, P. Colombat, J. White, J. O. Armitage, R. Arranz-Saez, W. Y. Au, M. Bellei, P. Brice, D. Caballero, B. Coiffier, E. Conde-Garcia, C. Doyen, M. Federico, R. I. Fisher, J. F. Garcia-Conde, C. Guglielmi, A. Hagenbeek, C. Haioun, M. LeBlanc, A. T. Lister, A. Lopez-Guillermo, P. McLaughlin, N. Milpied, P. Morel, N. Mounier, S. J. Proctor, A. Rohatiner, P. Smith, P. Soubeyran, H. Tilly, U. Vitolo, P. L. Zinzani, E. Zucca and E. Montserrat (2004). "Follicular lymphoma international prognostic index." Blood **104**(5): 1258-1265.

Somasundaram, R., M. A. J. Prasad, J. Ungerbäck and M. Sigvardsson (2015). "Transcription factor networks in B-cell differentiation link development to acute lymphoid leukemia." Blood **126**(2): 144-152.

Spitzer, M. H., P. F. Gherardini, G. K. Fragiadakis, N. Bhattacharya, R. T. Yuan, A. N. Hotson, R. Finck, Y. Carmi, E. R. Zunder, W. J. Fantl, S. C. Bendall, E. G. Engleman and G. P. Nolan (2015). "IMMUNOLOGY. An interactive reference framework for modeling a dynamic immune system." Science **349**(6244): 1259425.

Srinivasan, L., Y. Sasaki, D. P. Calado, B. Zhang, J. H. Paik, R. A. DePinho, J. L. Kutok, J. F. Kearney, K. L. Otipoby and K. Rajewsky (2009). "PI3 kinase signals BCR-dependent mature B cell survival." Cell **139**(3): 573-586.

Stein, J. V. and C. Nombela-Arrieta (2005). "Chemokine control of lymphocyte trafficking: a general overview." Immunology **116**(1): 1-12.

Stewart, R., S. A. Hammond, M. Oberst and R. W. Wilkinson (2014). "The role of Fc gamma receptors in the activity of immunomodulatory antibodies for cancer." Journal for ImmunoTherapy of Cancer **2**(1): 29.

Suurmond, J. and B. Diamond (2015). "Autoantibodies in systemic autoimmune diseases: specificity and pathogenicity." J Clin Invest **125**(6): 2194-2202.

Tangye, S. G., A. Ferguson, D. T. Avery, C. S. Ma and P. D. Hodgkin (2002). "Isotype Switching by Human B Cells Is Division-Associated and Regulated by Cytokines." The Journal of Immunology **169**(8): 4298-4306.

Tsou, P., H. Katayama, E. J. Ostrin and S. M. Hanash (2016). "The Emerging Role of B Cells in Tumor Immunity." Cancer Res **76**(19): 5597-5601.

Victora, G. D. and M. C. Nussenzweig (2012). "Germinal Centers." Annual Review of Immunology **30**(1): 429-457.

Victora, G. D., T. A. Schwickert, D. R. Fooksman, A. O. Kamphorst, M. Meyer-Hermann, M. L. Dustin and M. C. Nussenzweig (2010). "Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter." Cell **143**(4): 592-605.

Vidal-Rubio, B., M. Sanchez-Carril, J. Oliver-Morales, A. Gonzalez-Femandez and F. Gambon-Deza (2001). "Changes in human lymphocyte subpopulations in tonsils and regional lymph nodes of human head and neck squamous carcinoma compared to control lymph nodes." BMC Immunol **2**: 2.

Wagner, E. F., N. Hanna, L. D. Fast, N. Kouttab, P. R. Shank, A. Vazquez and S. Sharma (2000). "Novel Diversity in IL-4-Mediated Responses in Resting Human Naive B Cells Versus Germinal Center/Memory B Cells." The Journal of Immunology **165**(10): 5573-5579.

Wang, J.-Z., Y.-H. Zhang, X.-H. Guo, H.-Y. Zhang and Y. Zhang (2016). "The double-edge role of B cells in mediating antitumor T-cell immunity: Pharmacological strategies for cancer immunotherapy." International Immunopharmacology **36**: 73-85.

Warnatz, K., A. Denz, R. Drager, M. Braun, C. Groth, G. Wolff-Vorbeck, H. Eibel, M. Schlesier and H. H. Peter (2002). "Severe deficiency of switched memory B cells (CD27(+)IgM(-)IgD(-)) in subgroups of patients with common variable immunodeficiency: a new approach to classify a heterogeneous disease." Blood **99**(5): 1544-1551.

Wei, C., J. Jung and I. Sanz (2011). "OMIP-003: Phenotypic analysis of human memory B cells." Cytometry Part A **79A**(11): 894-896.

Wirth, T. C., V. P. Badovinac, L. Zhao, M. O. Dailey and J. T. Harty (2009). "Differentiation of central memory CD8 T cells is independent of CD62L-mediated trafficking to lymph nodes." J Immunol **182**(10): 6195-6206.

Wogsland, C. E., A. R. Greenplate, A. Kolstad, J. H. Myklebust, J. M. Irish and K. Huse (2017). "Mass Cytometry of Follicular Lymphoma Tumors Reveals Intrinsic Heterogeneity in Proteins Including HLA-DR and a Deficit in Nonmalignant Plasmablast and Germinal Center B-Cell Populations." Cytometry B Clin Cytom **92**(1): 79-87.

Wong, M. T., J. Chen, S. Narayanan, W. Lin, R. Anicete, H. T. Kiaang, M. A. De Lafaille, M. Poidinger and E. W. Newell (2015). "Mapping the Diversity of Follicular Helper T Cells in Human Blood and Tonsils Using High-Dimensional Mass Cytometry Analysis." Cell Rep **11**(11): 1822-1833.

Woo, H. A., S. H. Yim, D. H. Shin, D. Kang, D. Y. Yu and S. G. Rhee (2010). "Inactivation of peroxiredoxin I by phosphorylation allows localized H(2)O(2) accumulation for cell signaling." Cell **140**(4): 517-528.

Woof, J. M. and M. A. Kerr (2006). "The function of immunoglobulin A in immunity." J Pathol **208**(2): 270-282.

Yanaba, K., J. D. Bouaziz, T. Matsushita, C. M. Magro, E. W. St Clair and T. F. Tedder (2008). "B-lymphocyte contributions to human autoimmune disease." Immunol Rev **223**: 284-299.

Young, I. T. (1977). "Proof without prejudice: use of the Kolmogorov-Smirnov test for the analysis of histograms from flow systems and other sources." J Histochem Cytochem **25**(7): 935-941.

Yuen, G. J., E. Demissie and S. Pillai (2016). "B Lymphocytes and Cancer: A Love–Hate Relationship." Trends in Cancer **2**(12): 747-757.

Yurchenko, M. Y., L. M. Kovalevska, L. M. Shlapatska, G. G. Berdova, E. A. Clark and S. P. Sidorenko (2010). "CD150 regulates JNK1/2 activation in normal and Hodgkin's lymphoma B cells." Immunol Cell Biol **88**(5): 565-574.

Zhu, D., H. McCarthy, C. H. Ottensmeier, P. Johnson, T. J. Hamblin and F. K. Stevenson (2002). "Acquisition of potential N-glycosylation sites in the immunoglobulin variable region by somatic mutation is a distinctive feature of follicular lymphoma." Blood **99**(7): 2562-2568.

Zinzani, P. L. (2005). "Lymphoma: diagnosis, staging, natural history, and treatment strategies." Semin Oncol **32**(1 Suppl 1): S4-10.

Zou, Y. R. and B. Diamond (2013). "Fate determination of mature autoreactive B cells." Adv Immunol **118**: 1-36.

Zuckerman, N. S., K. J. McCann, C. H. Ottensmeier, M. Barak, G. Shahaf, H. Edelman, D. Dunn-Walters, R. S. Abraham, F. K. Stevenson and R. Mehr (2010). "Ig gene diversification and selection in follicular lymphoma, diffuse large B cell lymphoma and primary central nervous system lymphoma revealed by lineage tree and mutation analyses." Int Immunol **22**(11): 875-887.