5-2015

Correlation Matrix Analysis Identifies Gene Signatures of Immune Cell Subsets and Their Interactions in Follicular Lymphoma

Jason R. Westin

CORRELATION MATRIX ANALYSIS IDENTIFIES GENE SIGNATURES OF IMMUNE CELL SUBSETS AND THEIR INTERACTIONS IN FOLLICULAR LYMPHOMA

A

THESIS

Presented to the Faculty of the University of Texas

Health Science Center at Houston

and

M. D. Anderson Cancer Center

Graduate School of Biomedical Sciences

in partial fulfillment of the requirements

for the Degree of

MASTER OF SCIENCE

by

Jason R. Westin, MD

Houston, Texas

May, 2015

Dedication

This thesis is dedicated to our patients, my mentors, and my family.

Acknowledgements
These studies were supported in part by Waun Ki Hong, MD, institutional startup funds to Dr. Davis, by National Cancer Institute R01 CA155143 (Davis and Neelapu), R21 CA143785 (Neelapu), and Paul Calabresi Career Development Award in Clinical Oncology K12 2K12CA088084-11 (Westin).

Correlation Matrix Analysis Identifies Gene Signatures of Immune Cell Subsets and their Interactions in Follicular Lymphoma

Jason R. Westin, MD

Advisory Professor: R. Eric Davis, MD

ABSTRACT TEXT

There are important but ill-defined interactions between benign immune cell subsets and neoplastic B cells within follicular lymphoma (FL). Using the novel technique of correlation matrix analysis (CMA) of publicly available FL whole-tumor gene expression profiling (GEP) data, we have identified signatures of immune cell subsets. Overall survival correlated most highly with a model using signatures of macrophages, T cells, and stroma, which was able to add significantly to existing clinical prognostic tools. From our own data of a cohort of 43 FL tumors sorted into B-cell and non-B cell (NB) fractions for GEP, CMA of the tumor infiltrating NB fraction revealed additional immune cell subset signatures, including T follicular helper (TFH) cells. Comparison of gene signatures between FL and tonsils (n=24) suggested that TFH cells and macrophages are qualitatively distinct in FL from normal tissue. “Cross-correlation”, between FL NB fraction signatures and individual B fraction genes, suggests that TFH cells promote proliferation, germinal center stage differentiation, B-cell receptor signaling, and induction of CCL17 and CCL22 by tumor B cells. This novel analytical approach may be broadly applicable to define gene signatures of rare immune cell subsets in the tumor microenvironment,

determine their prognostic impact, discover novel therapeutic targets, and identify patients likely to benefit from therapies targeting tumor-stroma interactions.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

COMMONLY USED ABBREVIATIONS

CR	Complete response
OS	Overall Survival
PFS	Progression Free Survival
PR	Partial response
SD	Stable disease
FL	Follicular Lymphoma
T_{FH}	T Follicular Helper
NB	Non-B Cell
GEP	Gene Expression Profiling
CMA	Correlation Matrix Analysis
NHL	Non-Hodkins Lymphoma
IPI	International Prognostic Index
LDH	Lactate Dehydrogenase
ESR	Erythrocyte Sedimentation Rate
FLIPI	Follicular Lymphoma International Prognostic Index
GELF	Group d'Etudades Lymphomes Folliculaires
ORR	Overall Response Rate
FDA	Food and Drug Administration
KLH	Keyhole-Limpet Hemocyanin
GM-CSF	Granulocyte-Macrophage Colony Stimulating Factor
DFS	Disease-Free Survival
FFS	Failure Free Survival
CHOP	Cyclophosphamide, Doxorubicin, Vincristine, Prednisone
IR	Immune Response
IHC	Immunohistochemistry
PACE	Prednisone, Doxorubicin, Cyclophosphamide, Etoposide
pDC	Plasmacytoid Dendritic Cells
Mac	Macrophage
NK	Natural Killer
mDC	Myeloid Dendritic Cells
GSEA	Gene Set Enrichment Analysis
GC	Germinal Center
BCR	B-cell Receptor

Chapter 1

INTRODUCTION

Follicular Lymphoma

Follicular lymphoma (FL) is the second most common subtype of non-Hodgkin lymphoma (NHL) with ~15,000 new cases annually in the United States.(1) For reasons that remain unclear, the incidence of FL has increased in the United States and Europe over the past 2 decades.(2, 3) FL arises from maturing B cells within the lymph node germinal center. The neoplastic cells are typically organized in follicles (hence the name of the disease), and express germinal center-associated markers including BCL6, CD10, SERPINA9 (GCET1), and LMO2.(4) In addition, FL cells have a gene expression profile similar to that of centrocytes, a common cell found in the germinal center.(5) FL cells have an identical immunoglobulin (Ig) gene rearrangement pattern, which indicates that the malignant transformation occurs after VDJ gene recombination.(6, 7) Nearly all cases of FL feature the genetic hallmark of an acquired t(14;18)(q32;q21) translocation, resulting in deregulation of BCL2, a critical gene which regulates apoptosis and cell death. The t(14;18) translocation is generally thought to represent the initial genomic alteration in FL, but alone is insufficient for lymphomagenesis.(7) Up to 50% of healthy individuals harbor circulating cells that contain t(14;18) translocation, and rarely progress to FL. (8-12) Peripheral blood cells which harbor the t(14;18) translocation in healthy individuals were thought to represent naïve B cells, but this assumption has been challenged by the discovery of many similarities to FL cells, including class-switch recombination and surface expression of IgM and IgD.(9) The significance of the

presence of the t(14;18) translocation in healthy individuals is unclear, and to date has not correlated with any prognostic or therapeutic relevance, but does suggests a potential common pre-malignant stage. It is noted that the incidence of FL, and the prevalence of t(14;18)-containing cells, both increase with age.(13) A subsequent discovery of MLL2 mutation, a histone methyltransferase gene, has been found in 89% of FL patients, representing a potential second founder mutation.(14) However, elegant studies utilizing exome sequencing of sorted FL patients samples has demonstrated a large amount of intratumoral diversity in FL, and that MLL2 mutations may be late common events, and not founder mutations.(6) In addition, these experiments identified mutations in another histone-modifying enzyme, CREBBP, that appear to be early events in the clonal evolution of FL.

Prognostic Systems for FL

Over 70% of newly diagnosed FL patients present with advanced disease, either due to early dissemination from the initial site or to a long asymptomatic phase prior to diagnosis. In the modern era, many patients are diagnosed with asymptomatic disease due to serendipitous imaging findings obtained for other indications. It is unclear if this potential lead time bias will ultimately have any impact on the natural history of FL. Follicular lymphoma is characterized by the Ann Arbor staging system, originally devised nearly 40 years ago to account for radiation field size. The system of scoring is based upon a stepwise anatomic progression of lymph node involvement, a pattern found commonly in Hodgkin's lymphoma but not NHL, and thus its applicability to FL has been questioned.(15) Generally, it is suspected that patients whose extent of

involvement by FL appears to be only stage I (localized to a single lymph node) or stage II (localized to more than one lymph node on one side of the diaphragm) are likely to have undetectable disease at distant sites. A clinical trial conducted at M.D. Anderson found that patients with localized FL who were treated with aggressive chemotherapy and radiation still had a 24% chance of disease relapse at 10 years.(16) Indeed, clinical outcomes are not uniform within FL stages, confirming the suspicion that factors other than the simple location of the disease may affect outcomes.

The original prognostic model for NHL was the international prognostic index (IPI), which accounted for clinically apparent factors.(17) The IPI incorporates age (>60), performance status (>2), lactate dehydrogenase (elevated), extranodal sites of disease (•2), and disease stage (•3) as adverse factors, and is able to separate patients into general categories. Unfortunately, these categories are heterogeneous in that patient the high risk group may do well, and low risk may do poorly, and thus their applicability is not predictive for a given patient. In 2000, the Intergruppo Italiano Linfomi retrospectively evaluated for these additional factors and created a prognostic model.(18) This initial model accounted for known risk associated factors including: pre-treatment age, gender, number of sites of extranodal disease, B symptoms (fevers, chills, night sweats, significant weight loss), and the level of serum lactate dehydrogenase (LDH) and erythrocyte sedimentation rate (ESR) in the blood. Building upon their work, an international cooperative group created the Follicular Lymphoma International Prognostic Index (FLIPI) based on the data from over 4000 FL patients.(19) This model included age (• 60 years vs < 60 years), Ann Arbor stage (III-

IV vs I-II), hemoglobin level (< 12 dg/L vs • 12dg/L), number of nodal areas (> 4 vs < 4), and serum LDH level (above normal vs normal or below). By counting the adverse features present, patients were classified into three risk groups with 10 year overall survival (OS) of 70.7%, 50.9%, and 35.5%. The most predictive factor for the greatest risk of death evaluated in the multivariate analyses was advanced age.

The FLIPI is widely utilized, but has several important weaknesses. First, the FLIPI was generated from a retrospective database analysis of patients that were treated with chemotherapies that did not include the currently ubiquitous immunotherapy rituximab. Second, the FLIPI dataset did not include significant data including the performance status and blood tests of ESR and 2M. Third, the FLIPI had a primary endpoint of OS, which is challenging to study in an indolent disease like FL. The long natural history of the disease, serial responses to multiple effective therapies, and duration of follow up required to evaluate OS make this endpoint problematic. In order to attempt to overcome these known weaknesses, the International Follicular Lymphoma Prognostic Factor Project created the F2 study in 2003.(20) The primary endpoint of the model now referred to as FLIPI2 was progression free survival (PFS), which is the preferred metric in lymphoma clinical trials.(21) In multivariate analyses, factors found to achieve significance for increased risk included elevated blood 2M, a single lymph node with a length greater than 6 cm, involvement of the bone marrow with FL, low hemoglobin level (< vs • 12dg/L), and advanced age (• vs < 60 years. Based on the number of factors present, patients were classified into three risk groups with 5 year PFS of 79.5%, 51.2%, and 18.8%.

Indications for therapy of FL

FL generally is highly sensitivity to multiple therapeutic classes, although the disease is considered incurable as relapse after initial response is nearly uniform.(22) Due to the fact that the majority of patients with FL have an indolent disease that is treatable but incurable, the option of deferred therapy has been found to have no negative impact on overall survival.(23, 24) In order to identify patients who do not require immediate therapy, the Group d'Etude des Lymphomes Folliculaires (GELF) criteria were developed from a prospective clinical trial.(25, 26) Patients with • 3 nodal sites with greatest length of • 3 cm, a single lymph node with a length of • 7 cm, symptomatic splenomegaly, B symptoms, and patients with cytopenias or effusions were found to have inferior outcomes with observation. As a result, the presence of any "GELF criteria" implies the need for therapy.

Watch and Wait in FL

Spontaneous regressions of biopsy proven sites of FL occur in 5 – 25% patients, which is likely related to the interaction of the intact immune-FL interaction.(27) In an immediate vs. delayed therapy clinical trial, patients with FL were randomized to chemotherapy or observation, and were found to have equivalent survival.(28) Of note, ~10% of the observation patients did not require systemic therapy during the 10 years of follow up. Many other clinical trials have demonstrated similar excellent outcomes with delayed therapy.(29, 30) To date, there are no conclusive studies that show that a strategy of "watch and wait" results in inferior long term survival outcomes, despite the

current availability of many novel effective therapies. Further arguing for observation, therapy with a single chemotherapy drug or combination of drugs was shown to result in responses, but not a significant change in the overall survival of FL patients.(31) Modern therapy which includes biologically relevant agents has resulted in moderate improvements in overall survival, although disease eradication is still only achieved in an extreme minority, if any FL patients. This “responding, remission, relapsing” pattern remains poorly understood.

Immunotherapy

The immune system has been successfully manipulated to combat cancer for over 100 years. In 1891, Dr. William Coley treated cancer patients with a mixture of bacterial toxins (Coley’s toxin), and reportedly achieved dramatic results in lymphoma and other malignancies. Other investigators could not achieve similar results, and thus early attempts at immune-based cancer therapy fell out of favor compared with radiation or chemotherapy.(32)

Dr. Ronald Levy from Stanford and others have proposed that the use of monoclonal antibodies, a passive immunization strategy, could have efficacy via targeting lymphoma cell surface markers. In 1994, the phase I trial of IDEC-C2B8, a monoclonal antibody now commonly known as rituximab which targets the B-cell marker CD20, resulted in tumor regression in 40% of heavily pre-treated B-cell lymphoma patients.(33) The mechanism of action of rituximab is a combination of antibody-dependent cellular cytotoxicity, complement-mediated cytotoxicity, and direct signaling. The subsequent phase II and III rituximab trials evaluated 4 doses, administered once a week, of

375mg/m2 rituximab in relapsed low grade NHL and demonstrated overall response rates (ORR) of 46 and 48%, with a median duration of response of 10.2 and 13.0 months, respectively.(34, 35) Of note, these large trials found rituximab to be very well tolerated with toxicities limited to allergic type reactions during the infusion.

The Food and Drug Administration (FDA) approved immunomodulatory drug lenalidomide has shown promising efficacy in both newly diagnosed and relapsed FL.(36-39) A phase II trial in newly diagnosed FL patients found that lenalidomide and rituximab had an ORR of 98% and CR of 87%.(39-41) An ongoing phase III trial randomizes newly diagnosed FL patients who meet criteria for needing therapy to receive rituximab with lenalidomide or standard chemotherapy, followed by maintenance rituximab.

Other immunomodulating agents have shown preliminary impressive activity against FL. We recently reported that pidilizumab, a monoclonal antibody directed against programmed death receptor 1 (PD1), a co-inhibitory receptor expressed by activated T cells, B cells, NK cells, and myeloid cells, had minimal toxicity and impressive efficacy when combined with rituximab in relapsed FL.(42) Pidilizumab is a humanized IgG1-kappa recombinant monoclonal antibody that blocks the interaction of PD-1 with its ligands. The antitumor activity of the antibody in preclinical models was associated with increased numbers of CD4+ and CD8+ T cells and NK cells, and induction of immunologic memory. Based on our preliminary studies, administration of pidilizumab was expected to augment the naturally induced antitumor T-cell immunity in FL and enhance ADCC mediated by natural killer (NK) cells in the presence of rituximab. It was hypothesized that activation of both innate (NK cells) and adaptive (T cells) arms of the

immune system would likely minimize the emergence of immune escape variants and lead to improved remission duration. Indeed, we found that the overall response rate (ORR) of pidilizumab and rituximab was 66% (19/29) and the complete response (CR) rate was 52% (15/29).(42) In comparison, previously reported data in a similar population found rituximab alone to result in an ORR of 40% and CR rate of 11%.(43)

Tumor immunoglobulin (idiotype, Id) is expressed in a clonal fashion on FL cells and has been demonstrated as safe and immunogenic in Phase I/II vaccination trials.(44-51) In a randomized, double-blind multicenter phase III clinical trial, patient-specific tumor-derived Id protein was conjugated with a carrier protein (keyhole-limpet hemocyanin, KLH) and administered together with an adjuvant (granulocyte-macrophage colony stimulating factor, GM-CSF) to patients with advanced stage, previously untreated FL, in CR following standard induction chemotherapy.(52) Patients receiving vaccination with Id-KLH+GM-CSF had a significantly prolonged disease-free survival (DFS) when compared with the control group that received a non-specific immune stimulant (KLH+GM-CSF). This trial demonstrated the first positive result for a Phase III vaccine trial against lymphoma. Although significant, the therapeutic benefit of Id vaccination was small, and no change in overall survival was detectable. Further improvements to cancer vaccine therapy are needed if it is to be a viable treatment strategy for patients with FL.

Chemotherapy

Follicular lymphoma is generally sensitive to chemotherapy, with numerous combinations demonstrating significant activity over the past 40 years. A retrospective review from M. D. Anderson Cancer Center found that over 25 years, each subsequent

generation of initial therapy, outcomes for patients with FL improved (both failure free survival (FFS) and OS).(53) Rituximab-based therapies had not yet reached their median FFS and OS, but appeared to result in a significant improvement in outcomes.

The addition of rituximab to chemotherapy is known to be superior to chemotherapy alone.(54) As a result, rituximab is nearly ubiquitous in FL therapies, alone and combined with chemotherapy. A trial by the German Low-Grade Lymphoma Study Group randomized FL patients to CHOP with or without rituximab and found improved outcomes with the rituximab treated patients.(55)

Interaction of the benign immune system and cancer

It is now established that that the innate and adaptive immune systems can prevent the development of or eradicate tumors (cancer immunosurveillance or cancer elimination).(56, 57) Multiple studies have demonstrated that deficiencies of various immune components (such as interferon gamma or FasL) increase the rate of development of malignancies, including lymphomas.(58, 59) Mice lacking recombinase activating gene (RAG2), an essential gene for somatic rearrangement of lymphocyte antigen receptors, have a total lack of peripheral effector T cells, B cells, and NK/T cells. In a RAG2 deficient mouse model, Shankaran et al. found that the development of sarcomas after carcinogen exposure was significantly increased as compared to RAG2 wild type mice, which along with other similar studies definitively established the principle of cancer immunosurveillance/elimination.

Beyond prevention or eradication, the immune system can also modulate the immune profile of the tumor cells. Tumors from immune deficient models can be rejected when transplanted into immune competent models, speaking to their being more

immunogenic than those which developed in the presence of a functional immune system (cancer immunoediting).(56, 59) In order for cancer to escape the immunosurveilance or elimination in an immune competent host, malignant cells may display fewer antigens, and thus be less immunogenic.(60) In addition, cancers have also been shown to express inhibitory T-cell receptors, such as CTLA-4 and PD-1, which result in an immune tolerance phenotype, and eventual progression of cancer.(61) Those cancer cells with the ability to impair an immune surveillant response would have a large evolutionary advantage in comparison to their immune regulated counterparts, and thus the immune system provides a selection pressure. Together, these data speak to the "cancer immunoediting hypothesis" with three phases in the tumor modulation process: the "three Es of cancer immunoediting": elimination, equilibrium, and escape.(62)

Follicular lymphoma evades endogenous antitumor immune responses

Follicular lymphoma has a relatively unique biology. Patients may have durable stable disease, spontaneous regressions, or even remissions, and yet the majority of patients eventually suffer disease progression. This unusual behavior for a malignancy is related to the complex interaction between the benign immune system and FL. This "waxing and waning" clinical course suggests that FL may go through an equilibrium phase where tumor progression is impeded or partly reversed by the immune system. Eventually, escape mechanisms develop resulting in progression of FL.

Numerous recent studies support the hypothesis that the immune system plays a significant role in the control of FL. Among types of NHL, FL has a relatively high

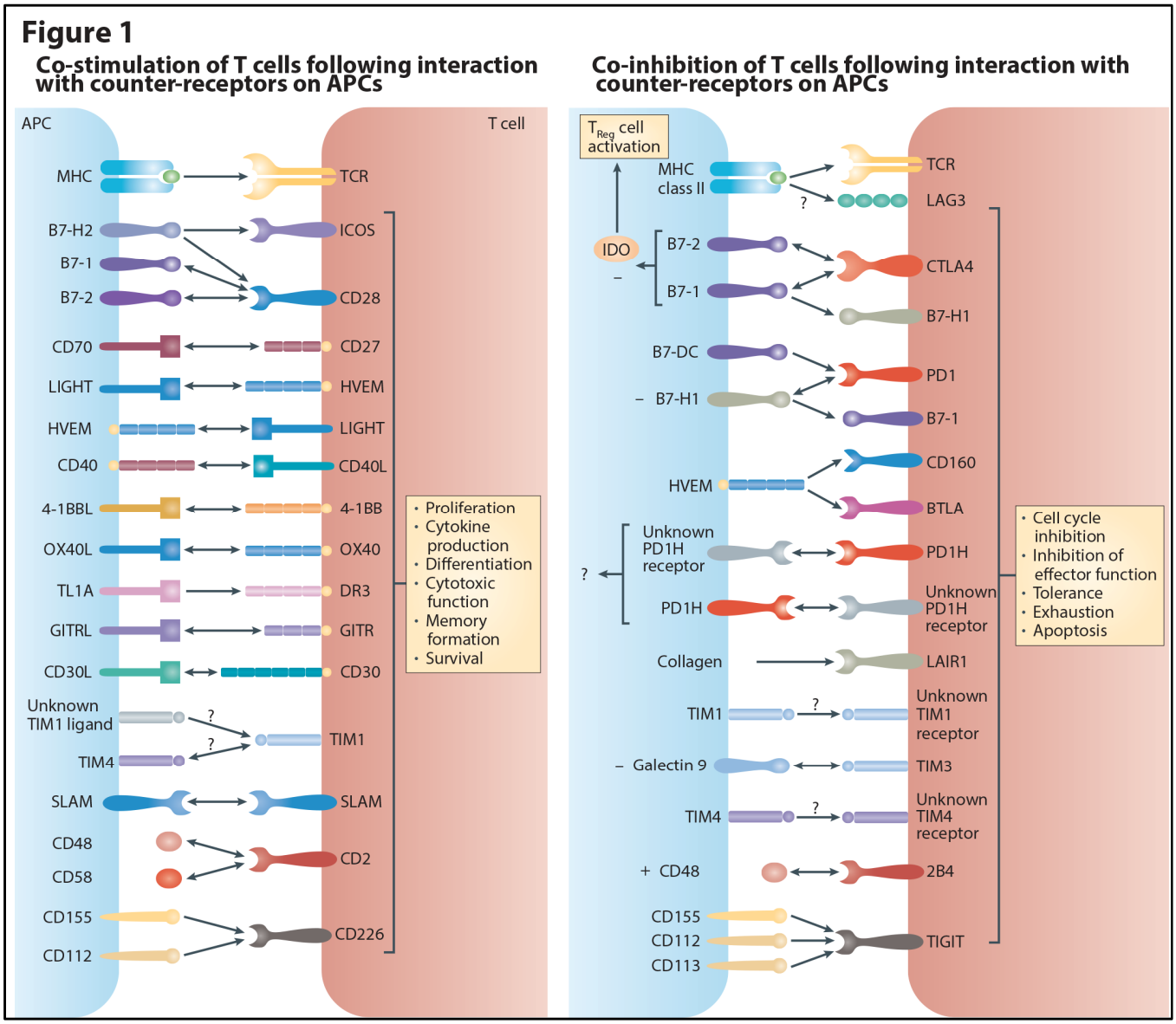
proportion of non-neoplastic immune cells infiltrating involved lymph nodes (T, NK, and monocytes/macrophages). The extensive differences between neoplastic B cells and non-neoplastic cells may allow the overall gene expression profiles, determined by microarrays on the whole tumor, to be de-convoluted into “signatures” that are attributable to different elements within the tumor. Indeed, the landmark GEP study of FL by Dave et al. identified multiple signatures, two of which were attributed to infiltrating T cells and macrophages, with respectively positive and negative effects on outcome, adding to the effect of clinical prognostic tools.(63) Of note, this study did not find a predictive signature which could be assigned to the neoplastic FL cells, in contrast to other types of cancers. The prognostic significance of immune-response signatures has largely been validated by subsequent studies using microarrays and comparable techniques to measure mRNA (qRT-PCR),(64) including in patients treated with Rituximab,(65) and is among the evidence that the immune environment is an important determinant of outcome in FL. Although Dave et al. observed that “the immune-response 1 signature is not merely a surrogate for the number of T cells in the tumor-biopsy specimen, since many other standard T-cell genes…were not associated with survival,” to date GEP has not revealed specific mechanisms affecting the immune response in FL.

Further speaking to the immune-FL relationship, an increased level of CD8+ T-cells in FL involved lymph nodes correlated with an improved prognosis.(66) In addition, the prognosis of patients with FL is positively correlated with an immunosurveillance pattern (CD8+ T cells) and negatively with an immune-escape pattern (CD57+ T cells).(67) Finally, tumor-specific T cells can be isolated from the tumor microenvironment and

peripheral blood in patients with FL.(68, 69) Together, these data suggest that an endogenous antitumor immune response is present in patients who harbor FL, but eventually becomes ineffective at controlling the tumor. This escape by the tumor is likely due to the development of various immunosuppressive mechanisms in the tumor microenvironment.

Over the past decade, there have been significant advances in our understanding of the exceedingly complex and tightly controlled interaction of immune stimulating and inhibiting receptors.(70) Importantly, tumor-specific effector T cells have been found to be impaired by immunoregulatory mechanisms in the tumor microenvironment in various cancer models.(71) Cancers, including FL, are now known to modulate tumor-beneficial immunosuppressive mechanisms which have the potential to result in progression from immune “equilibrium” to the “escape” phase.(61, 72, 73) Important inhibitory pathways that decrease tumor-specific T-cell effector function include extrinsic suppression by regulatory T cells (Tregs), direct inhibition through inhibitory ligands (Figure 1), and soluble factors such as transforming growth factor and interleukin (IL)-10.(71)

Figure 1. Co-stimulatory and co-inhibitory receptors expressed on T cells and their corresponding ligands on antigen presenting cells. (Adapted from (70))



Summary and Application of The Above Information To Thesis Project

Gene expression profiling (GEP) provides a wealth of high-content molecular data regarding the genomic state of tumors, which are complex mixtures of cancer and non-cancer cells. In the landmark study by Dave et al., whole-tumor GEP data from 191

pretreatment samples of FL patients were utilized to seek features (predictors) which significantly correlated with OS.(63) A model based on two multi-gene signatures (one favorable, the other unfavorable) was found to add to the predictive power of the IPI. Remarkably, neither of these signatures was determined to originate from tumor B cells, but instead from benign tumor-infiltrating immune cells. As the signatures contained numerous genes related to the immune system, they were named immune response (IR)-1, attributed to T cells (OS-favorable), and IR-2, attributed to macrophages (IR-2, OS-unfavorable). The IR-1 signature "included several T-cell restricted genes but was not merely a measure of the number of tumor-infiltrating T cells, since a signature of pan-T-cell genes was not associated with survival."(63) The conclusion of this statement is that the outcome in FL could be influenced by differences in infiltrating T cells, not only quantitative (i.e., their frequency) but also qualitative, such as proportions of the multiple T-cell subsets and physiologic states now known. Subsequent studies, most employing immunohistochemistry (IHC), have validated prior data regarding a role of the host immune microenvironment as a major determinant of outcome in FL, but consistent evidence is lacking for the cell types and mechanisms involved.

In the 10 years since the study of Dave et al. was published, there have been significant advances in the understanding of both the immune system and its interaction with cancer. Knowledge of basic and tumor immunology has rapidly evolved, including the identification of multiple T-cell subsets and physiologic states of immune cells (e.g., T-cell exhaustion, and states of macrophage polarization). In particular, the T follicular helper (T_{FH}) cell, first reported in 2000, has been identified as a critical promoter of normal germinal center B cells. However, the ability of GEP to discover interactions

within and between the tumor cell and host immune cell compartments of FL, and their impact on survival, greatly depends on the technical and analytical methods employed. When GEP data derives from whole (unseparated) tumor samples, as used by Dave et al., the cell type(s) from which individual mRNA transcripts originate is not readily apparent. However, newer methods applied to GEP data from whole tumor biopsies, including correlation matrix analysis, are able to identify the relative proportions of T-cell subsets and other non-neoplastic cells infiltrating the tumor, and their correlation with outcome.(74) We report here our results from applying these methods to the data of Dave et al., as well as results from an independent series of FL tumors in which GEP was performed after separation into B-cell and non-B-cell fractions.

Chapter 2

METHODS

Cell Sorting

Follicular lymphoma tumor biopsy samples were obtained from patients treated on the BiovaxID phase III idiotype vaccine clinical trial.(75) All tissue samples were obtained after written informed consent was obtained from patients through this institutional review board-approved protocol. Samples derived from a pre-treatment biopsy, obtained in part for idiotype vaccine generation, were processed into single cell suspension and viably preserved in -80°C refrigeration until removal for this project. All patients were subsequently treated with PACE (prednisone, doxorubicin, cyclophosphamide, etoposide)(47) chemotherapy for 6 cycles. If a patient achieved a clinical response which proved durable for at least 6 months, he/she was subsequently randomized to receive one of two adjuvant therapies: idiotype vaccination with immune stimulation vs. immune stimulation alone.

A single sample vial for a given patient was identified by a patient-specific barcode and removed from -80°C refrigeration and immediately placed in a 37°C water bath. After 2 minutes, 2ml of “Thawing Media” was added. “Thawing Media is defined as:

RPMI liquid medium with 1x Glutamax = 500 mL bottle, added:

1. HEPES buffer 1M – 10 mL or 20mM,
2. sodium pyruvate 100mM – 5mL or 1mM,
3. Penicillin 10,000 U/mL + Streptomycin 10,000 ug/ml – 5mL
4. Gentamycin 10mg/ml – 500 uL
5. 2-Mercapto Ethanol stock – 1000X – 500uL

6. 125mL of heat inactivated fetal bovine serum

Cells were allowed to stand at room temperature for an additional 2 minutes prior to being centrifuged at 4°C with 1500 RPM for 10 minutes. The supernatant was removed, and the cell pellet was re-suspended in 2ml of MACS buffer (Miltenyi), and again centrifuged at 4°C with 1500 RPM for 10 minutes. The supernatant was removed, and the cell pellet was re-suspended in 10ml of MACS buffer. After gentle vortexing to ensure relatively homogeneity of suspended cells, a 2ml aliquot was removed and placed in a separate vial. Both the 8ml and 2ml samples were centrifuged at 4°C with 1500 RPM for 10 minutes, and supernatant was discarded. The 8ml sample (referred to as “Non-B”) had 5ul of both Miltenyi CD19 and CD20 MicroBeads added per 1e7 cells, along with 90ul of MACS buffer and was gently vortexed for 30 seconds. The 2ml sample (referred to as “B”) had 10ul of Miltenyi CD3 MicroBeads per 1e7 cells added, along with 90ml of MACS buffer and was gently vortexed for 30 seconds. Both vials were then placed in an 8°C refrigerator for 30 minutes. While the cells were incubating, the LD columns (Miltenyi) were prepared by placing them in the magnetic field of a MACS separator, and rinsing the LD column with 2 mL of MACS buffer. The cells were removed from the refrigerator and washed by adding 1mL of MACS buffer, and centrifuged at 4°C with 1500 RPM for 10 minutes. The supernatant was removed, and the cell pellet was re-suspended in 500ul of MACS buffer. The labeled cells and MACS buffer wash were then added to the LD column in MACS separator, and the elution was collected (negatively selected cells). The column was washed with an additional 1 mL of MACS buffer twice to ensure elution of the maximal amount of unbound cells. The column was then removed from the magnetic separator, placed in a new collection tube,

and washed with gentle manual pressure to elute 1.5 mL of MACS buffer to allow collection of the positively selected cells.

The negatively selected cells were counted and assessed for viability based with trypan blue dye uptake. The cells were then aliquoted in 1e6 portions and placed in the centrifuge at 4°C with 1500 RPM for 7 minutes, and the supernatant was discarded. The cell pellet was snap frozen in -80°C and stored until all desired samples had undergone the above process.

Prior to freezing, a portion of the cells were set aside for flow cytometry evaluation (~1e5 cells) from unsorted, negative and positive selection of both the CD19 and CD20 and the CD3 populations. These cells were then assessed for purity with labeling with 5uL CD3 PE and CD20 PerCP, along with 1uL of Kappa FITC via standard flow cytometry methods, and analysis with FloJo software.

Nucleic acid extraction:

Samples with an adequate number of negatively selected cells (>= 1e6 cells) were thawed for nucleic acid extraction. Vials were removed from the -80°C freezer and placed in ice for transportation to the bench. Immediately after arrival, 350uL of Buffer RLT Plus (Qiagen - AllPrep DNA/RNA Mini) was added to disrupt the cell membrane and cellular enzymes which may destroy nucleic acids. The vial was gently flicked to ensure the pellet was completely dissolved in buffer, and the homogenized lysate was then transferred to an AllPrep DNA spin column. The column was placed in a 2 mL collection tube, and then centrifuged at 10,000RPM for 30 seconds.

The AllPrep DNA spin column was then transferred to a new collection tube and stored at room temperature. The flow through from the initial column wash had 350 uL of 70%

ethanol added, and was mixed well by pipetting up and down gently. The now ~700uL mixture, including any precipitate, was transferred to an RNeasy spin column placed in a 2 mL collection tube, and centrifuged at 10,000 RPM for 15 seconds. The flow through was processed for microRNA collection, and stored for future potential studies. The RNeasy spin column then had 700uL of Buffer RW1 (Qiagen) added, and was centrifuged at 10,000 RPM for 15 seconds, with the flow through being discarded. The RNeasy spin column then had 500 uL of Buffer RPE (Qiagen) added and was centrifuged at 10,000 RPM for 15 seconds, with the flow through discarded. An additional 500 uL of Buffer RPE (Qiagen) was then added to the RNeasy spin column, centrifuged at 10,000 RPM for 2 minutes, and the flow through was discarded. To dry the column completely, the RNeasy spin column was centrifuged at full speed for 1 minute, and then placed in a new 1.5mL collection tube. The now dry RNeasy spin column had 30uL of RNase-free water added, and was centrifuged at 10,000 RPM for 1 minute to elute the RNA. An additional 20uL of RNase-free water was added, and again centrifuged at 10,000 RPM for 1 minute to elute any remaining RNA. The tube containing the eluted RNA was then placed in an 8C refrigerator.

The DNA was then eluted from the DNA spin column used earlier per Qiagen's AllPrep DNA protocol.

Quantification of DNA and RNA

To ensure an adequate quantity of nucleic acid was present, a Nanodrop Spectrophotometer 2000 was utilized. After cleaning the device, 1uL of blank (RNase free water for RNA, Buffer EB for DNA) was pipetted to the pedestal to calibrate machine. Per standard protocol, 1uL of the nucleic solution from the previous extraction

steps was then pipetted onto the pedestal and measured, including the 260/280 ratio (absorbance of light a 260nm and 280nm), with a goal ratio of 1.8 – 2.2.

Determination of quality of RNA

To ensure an adequate quality of nucleic acid was present, an Agilent 2100 Bioanalyzer was utilized. After cleaning the device, the standard RNA ladder (Agilent) was gently thawed in an ice bath, which serves as a reference RNA to allow standardization between assays and ensure an external quality measurement. The electrode cleaner chip had 350 uL of RNase free water added and was placed in the Agilent 2100 Bioanalyzer for 5 minutes. To prepare the gel matrix for the chip, 550 uL of RNA 6000 Pico (Agilent) gel matrix was added to a supplied spin filter, and centrifuged at 4000RPM for 10 minutes. Then 65uL of the gel matrix was combined with 1uL of RNA 6000 Pico dye, and the mixture was centrifuged at 14,000 RPM for 10 minutes. An RNA Pico chip was placed in the priming station, and 9 uL of dye-gel mixture was pipetted into the appropriate well. The chip was then primed by depressing 1mL of air from the syringe into the appropriate well to force the dye-gel mixture into the microtubes. After priming, 9 uL of the dye-gel mixture was then pipetted into each well, and 9 uL of RNA 6000 Pico conditioning solution and 5 uL of RNA 6000 Pico marker were then pipetted into their respective wells. In order to minimize secondary structure formation of the RNA, the sample was then heated to 70C for 2 minutes. After heating, 1 uL of the RNA ladder and 1 uL of sample RNA were then pipetted into the correct wells, and vortexed in the Agilent vortex at 2400 RPM for 60 seconds. The RNA 6000 Pico chip was then placed into the Bioanalyzer, and the software package was then instructed to “Run” the assay. The Bionanalyzer then determined the

“electropherogram” for the control markers, and the pattern of the 18S and 28S peaks. The software calculates an “RNA integrity number” (RIN), which correlates with the degree of RNA degradation and overall quality. To ensure high quality data, only samples with both a B and non B fractions having a RIN number of at least 6.6 were taken to the next steps.

RNA amplification and cDNA generation

For preparation of gene expression profiling, the RNA from previous steps requires amplification. Into a sterile RNase free 0.5 mL microcentrifuge tube, ~1000ng (100ng minimum, 10ug maximum) of RNA was aliquoted, and enough nuclease free water to bring to 11 uL total volume was added.

Following the protocol for Illumina TotalPrepTM RNA Amplification Kit instructions, the Reverse Transcriptase Master Mix is prepared as follows (multiply below by number of assays to be run):

1. 1 uL of T7 Oligo(dT) primer
2. 2 uL of 10X First Strand Buffer
3. 4 uL of dNTP mix
4. 1 uL RNase Inhibitor
5. 1 uL of ArrayScript

Into each RNA sample, 9 uL of Reverse Transcriptase Master Mix is transferred to bring to 20 uL total volume, and mixed well. The tube is then placed into the thermal cycler and incubated for 2 hours at 42C, allowing the first reaction to occur. After 2 hours elapse, the tube is placed on ice to stop the reaction.

The Second Strand Master Mix (Illumina TotalPrepTM RNA Amplification Kit, multiply below by number of assays to be run) was then prepared as follows:

1. 63 uL of nuclease free Water
2. 10 uL of 10x second strand buffer
3. 4 uL of dNTP mix
4. 2 uL of DNA Polymerase
5. 1 uL of RNase H

Into each sample from the Reverse Transcription first step, 80 uL of Second Strand Master Mix is transferred (total volume of 100 uL) and mixed well. The tube is placed into the thermal cycler and incubated for 2 hours at 16C. After 2 hours is completed, the sample tube is placed on ice to stop the reaction. To prepare for eventual sample elution, sufficient nuclease-free water to allow for at least 20uL for sample is preheated to 55C. The tube is removed from the ice bath and 250 uL of cDNA Binding Buffer (Illumina) was added to each sample, and mixed well. A new cDNA filter cartridge was placed in a wash tube supplied with the kit. The cDNA and Binding Buffer mixture were then pipetted on to the cDNA filter cartridge membrane, and centrifuged at 10,000 RPM for 1 minute, with flow through being discarded. To each cDNA filter cartridge, 500 uL of Wash Buffer (Illumina) was added and centrifuged at 10,000 RPM for 1 minute, with flow through being discarded. The cDNA filter cartridge was then transferred to a cDNA elution tube, and 20 uL of preheated nuclease free water was pipetted to the cDNA filter membrane. The membrane was allowed to stand at room temperature for 2 minutes, and then centrifuged at 10,000 RPM for 1.5 minutes. The flow through which contained the cDNA sample was then transferred into a PCR tube and placed on ice.

The IVT Master Mix (Illumina TotalPrepTM RNA Amplification Kit, multiply below by number of assays to be run) was then prepared as follows:

1. 2.5 uL of T7 10x Reaction Buffer
2. 2.5 uL of T7 Enzyme Mix
3. 2.5 uL of Biotin-NTP Mix

Into each cDNA sample, 7.5 uL of IVT Master Mix is transferred and mixed well. The mixture is then placed in the thermal cycler and incubated for 14 hours at 37C. After 14 hours was completed, 75 uL nuclease free water was added to stop the reaction, and the sample was placed on ice. Prior to stopping the reaction, sufficient nuclease free water for at least 100uL per sample was preheated to 55C.

To each sample, 350 uL of cRNA Binding Buffer (Illumina TotalPrepTM RNA Amplification Kit) was added, followed by 250 uL of 100% ethanol, which is mixed by pipetting. The resulting mixture was pipetted onto the filter of a cRNA Filter Cartridge and was centrifuged at 10,000 RPM for 1 minute, and the flow through was discarded. Each cRNA filter cartridge then had 650 uL of Wash Buffer (Illumina TotalPrepTM RNA Amplification Kit) added, and was centrifuged at 10,000 RPM for 1 minute, discard flow through with the flow through discarded. The filter was further dried by centrifuging at 10,000 RPM for an additional minute, and then transferred to a cRNA collection tube. The cRNA filter then had 100 uL of preheated nuclease free water added and was incubated for 10 minutes in a 55C heat block. The heated tube was then removed, and centrifuged at 10,000 RPM for 1.5 minutes. The flow through containing the cRNA was then transferred to an eppendorf tube for storage at -80C.

cRNA Hybridization

The Illumina Hybrization Oven was pre-heated to 58^{O}C for at least 30 minutes, and the HYB and HCB buffer tubes were placed in the oven for 10 minutes to ensure any precipitation was dissolved. The cRNA was quantified with the nanospectrometer, and 750ng was removed from the storage eppendorf tube for hybridization, with RNase-free water added to bring the total volume to 5uL. To the cRNA and RNase-free water mixture, 10uL of HYB buffer (Illumina) was added. The Hyb Chamber was then prepared following the protocol from Illumina, with Hyb Chamber inserts and Gaskets installed and 200uL of HCB buffer added into the 8 humidifying buffer reservoirs. A single HT12 Illumina BeadChip was then removed from its package for each 12 samples to be hybridized, and placed in the Hyb Chamber Insert. A total of 15 uL of solution containing the cRNA was then loaded into the inlet port of the BeadChip and observed for evidence of air bubbles, which were recorded to allow for adequate interpretation of results. The Hyb Chamber inserts with BeadChips were then placed in the Hyb Chamber, and the Chamber was locked prior to being placed in the Hybridization Oven for between 14 – 20 hours.

The Illumina High-Temp wash buffer was then prepared per protocol, and allowed to heat to 55^{O}C overnight. The following day, the Wash E1BC buffer was prepared by adding 6mL of E1BC buffer to 2L of RNase-Free water. The Hyb Chamber was removed from the Hybridization Oven and disassembled. All BeadChips were removed and placed face-up in a beaker with Wash E1BC buffer. While submerged, the cover seal of the BeadChip was removed, and then the BeadChip was transferred to a submerged slide rack in a staining dish filled with Wash E1BC buffer. The slide rack

was then transferred to the Hybex Waterbath containing the High-Temp wash buffer and allowed to incubate for 10 minutes. After incubation, the slide rack was then immediately transferred to a staining dish with fresh Wash E1BC buffer and plunged in and out of the solution 5-10 times. The staining dish was then placed on an orbital shaker and agitated at room temperature for 5 minutes. The slide rack was then transferred to a new staining dish containing 100% ethanol, plunged in and out of the solution 5-10 times, and returned to the orbital shaker for an additional 10 minutes. The slide rack was then transferred to the staining dish with Wash E1BC buffer and plunged in and out of the solution 5-10 times, and returned to the orbital shaker for 2 minutes. The BeadChip wash tray was then prepared on the Rocker Mixer, with 4 mL of Block E1 buffer (Illumina) added. Using tweezers, the BeadChip was removed from the slide rack and transferred into the BeadChip wash tray face up, ensuring the entire chip is covered by the buffer. The tray was then placed on the rocker and agitated at medium speed for 10 minutes. The BeadChip was then grasped with tweezers and transferred to a fresh BeadChip wash tray 2 mL of Block E1 buffer containing a 1:1000 dilution of Cy3-Streptavidin. The tray was then covered completely by placing an empty opaque ice bath container over the wash trays, and agitated gently for 10 minutes. The BeadChip was then removed from the wash tray and transferred with tweezers to a slide rack submerged in Wash E1BC buffer, and the slide rack was then plunged in and out of the solution 5-10 times. The staining dish containing the Wash E1BC buffer was then agitated gently for 5 minutes.

To fully dry the BeadChips, the centrifuge was prepared with paper towels and allowed to rise to room temperature. The slide rack containing BeadChips was then removed

from the Wash E1BC buffer and centrifuged at 1400 RPM for 4 minutes, and transferred to a dry opaque storage container. The BeadChips were then taken to the University of Texas Medical Center Microarray facility and processed per manufacturer recommendations using an iScan device. The resulting data files were then transferred back to our secure hard drive for further processing.

Samples from normal tonsils from 24 children undergoing elective tonsillectomy were similarly prepared as single cell populations and viably frozen, and were depleted of B cells to generate NB fractions using the same approach as described above for FL samples.

Flow cytometry

Immunophenotyping of single-cell suspensions from FL lymph node biopsies and tonsils was done by 10-color, 12-parameter flow cytometry (LSR Fortessa, Becton Dickinson), using a panel of antibodies against CD3, CD4, CD8, CD20, CD16, CD56, CD68, CD45RA, CCR7, CXCR5, PD-1, and Foxp3 (all from Becton Dickinson) to determine various subsets of immune cellular elements. For each subset, the relative percentage of total live cells was determined with the software package FlowJo version 9.3.3.

Data Analysis

GEP data and clinical information (prognostic factors and censored survival time data points) were downloaded from http://llmpp.nih.gov/follicularlymphoma from the Dave et al dataset.(63)

Affymetrix gene expression data were handled essentially as originally detailed, with a few minor modifications. According to the most recent version of the annotation file, the

Affymetrix probe sets were matched to the corresponding gene symbols (Affymetrix U133 set, 2011/6/9), and the probe sets which did not correspond to a unique gene symbol annotation were removed. Finally, the median expression value of all remaining probe sets was selected as the expression value for further analysis of the corresponding individual gene. These data processing steps, principally the probe set with gene symbols matching, resulted in minor differences from the original analysis of Dave et al. From the initial IR-1 and IR-2 signatures, not all genes were retained in our analysis. The Cox model was used to identify either signature in multivariate models associated with survival outcomes up to 10 years after diagnosis. The log-rank test and Kaplan-Meier curves were employed to evaluate for differences in survival outcomes, using the Bioconductor “survival” software package.

Data from the B and Non-B fractions from the sorting procedural steps described in this thesis were processed and normalized both together and independently, using methods previously described.(76) Raw bead-level fluorescence intensities for a given probe set was corrected by the model-based background correction method,(77) subtracting the mean value of negative control probes. Next, the intensities were trimmed by the 3-MAD method offered with Illumina software, quantile-normalized to equalize the distribution, and subsequently averaged resulting in a single value for a probe. A lower level of detection threshold value was also generated for each array from the negative control value distribution, which was subsequently used as the “floor” for experimental probe values. Next, probes were converted to gene symbols via the Bioconductor "illuminaHumanv4.db" software package. Standard methods of hierarchical clustering and heat mapping using Cluster and Treeview software packages were employed,(78)

and were used to screen for batching effects. When batching effects were suspected, they were addressed by discarding unsatisfactory samples, re-arraying samples, or using batch-effect correction software methods.(79) The samples from normal tonsils that were similarly prepared as our FL samples and underwent the GEP process to generate control NB fractions, and subsequently underwent a combined normalization along with the FL NB fractions to allow for comparisons.

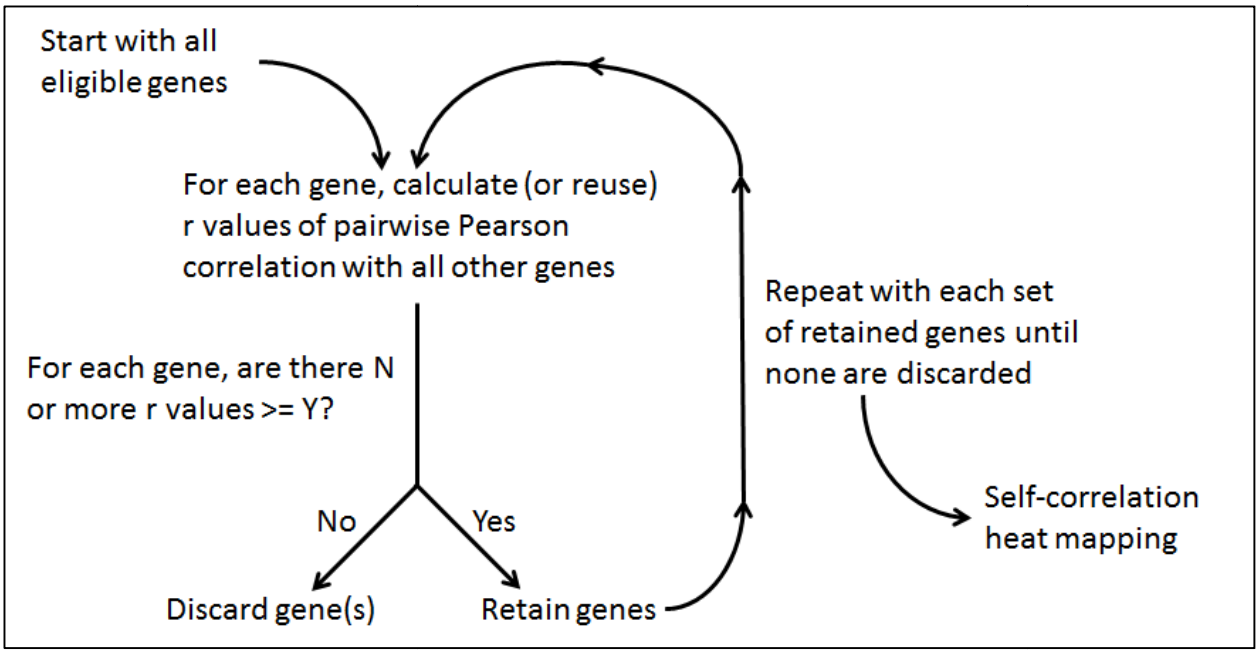
For the jointly-normalized dataset of combined B and non-B fraction array results, 11,959 genes were expressed above background in 25% or more of samples. For each gene in this dataset, an NB-B fold-difference in expression between NB and B fractions was calculated as the average difference in paired log2-transformed values. For each of the B and NB fraction datasets normalized separately, genes were also eliminated that were not expressed above background in 25% or more of samples, and which were judged to represent contamination from the other fraction. After this, totals of 12,087 and 10,652 genes remained in the NB and B gene expression profiles respectively.

From the unsorted tumor GEP from the Dave et al. dataset, a master matrix of pairwise r values was determined from the Pearson correlation of all possible pairs from the 14,383 genes with an average log2 expression value $\geq$ 6. The master matrix then was iteratively filtered, depicted in Figure 2, to arrive at a final set of genes with a target number (N) or more correlations with other genes in the set exceeding a target correlation level of a positive r value threshold (Y). This process results in a list of genes with at least N correlations with the other genes in the set exceeding a positive r value of Y, of which N and Y can be modified for variable gene list creation to identify an

optimal fit. Visual inspection of the resulting hierarchically-clustered heat maps of the correlation matrix of r values of the remaining genes, as illustrated in the Results, was used to select the optimal values of N and Y, and to identify highly associated gene clusters of genes, which were then selected as gene signatures. These heat maps have a 45 degree axis of full correlation of a gene with itself (r=1), and result in a small number of large highly correlating on-diagonal clusters (red color). The clustering of genes, and thus their order in the list, is not determined by similarities in pairwise r values, but takes into account all respective vectors of their pairwise r values with all others genes including in the list. If two genes have similar vectors, and thus a similar pairwise r pattern with non-self genes, they would be ranked closely together in the list and thus their intersection pixels would be near the diagonal. If their vectors are not similar, and thus a dissimilar pairwise r pattern with non-self genes, they would not be ranked closely together and the red pixels which demonstrate their tight correlation with each other would be off the diagonal.

These large clusters generally have significantly more than the minimal N number of correlations, and thus can be designated as a signature.

Figure 2. Schematic of iterative filtering approach



For the B and NB fractions, the initial filtering step using N = 1 and Y = 0.8 resulted in correlation matrix heat maps which highlighted the presence of genes that could represent contamination from cells from the "other type" (B genes in the NB gene set, or NB genes in the B gene set). In the B fraction matrix, a total of 886 genes likely originating from non-B cell types were identified and removed. From the NB fraction, 64 genes which were likely resulting from contaminating B cells (including CD19, CD22, TNFRSF13C (BAFFR), MS4A1 (CD20), CD40, and CD79B) were identified and removed. Even after removing likely contaminating genes from the correlation matrix, heat mapping of single-filtered genes were consistently weakly organized. The red clusters, indicating highly-correlated genes, were uniformly small and many were off the diagonal axis, representing genes with high pairwise r values but dissimilar vectors. This incongruence is likely due to the effects of other genes in the matrix, which raised suspicion of the strength of the initial correlation. The resulting heat maps were difficult to interpret, and essentially did not allow for the identification of signatures.

To deal with this significant roadblock, we employed the iterative filtering described above for its initial use in our analysis of the Dave dataset. We believed the weak correlations found in the separated tumor datasets were likely due to undesirable genes with low numbers of correlations which were obscuring those genes with higher numbers of correlations. By setting our iterative filtering criteria, as outlined in Figure 2, we were able account for this issue.

To assign an attribution for a given signature from CMA, we tested the genes in the cluster for their correlation with previously published gene sets, or enrichment for containing genes from within these sets, via the hypergeometric distribution test.(80) The full methods of this test are published, but in brief can be described by Figure 3.

Figure 3. Hypergeometric Distribution Test

$$p = 1 - \sum_{i=0}^{k-1} \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}}.$$

Legend: In this equation, p is the p value of enrichment. n is the number of genes included in a given cluster identified by our CMA. M is the number of genes in a previously published gene set that has a validated correlation with a biologic term, such as a cell type or cellular process. k is the number genes from n which are included in M. N is the total number of genes evaluated in the microarray studies and are members of one or more gene sets in a signature collection.

For the “self-correlation” matrix analysis of the separated FL tumor cell suspensions, the B and NB fraction datasets were normalized separately. Genes that were not expressed above background in at least 25% of the dataset samples were removed. This filtering resulted in totals of 12,087 and 10,652 genes in the NB and B gene expression profiles, respectively. For each fraction, the initial master matrix of pairwise r values was filtered to retain only those genes with at least one r value $\geq$ 0.8 with another gene. For the remaining genes, hierarchically-clustered heat maps of pairwise r values were then examined to identify clusters of genes that were likely to represent contaminating cells from the other fraction.

In the correlation matrix of the NB fraction, 64 genes were identified to have an apparent origin from B cells (including CD19, CD22, TNFRSF13C (BAFFR), MS4A1 (CD20), CD40, and CD79B), and were subsequently removed. In the correlation matrix of the B fraction, 886 genes were identified to have an apparent origin from non-B cell were similarly identified and removed. For both fractions, the remaining genes were then iteratively filtering as described above, with a visual inspection of clustered heat maps to select N and Y values, as illustrated in the Results, to identify highly associated gene clusters of genes, which were then selected as gene signatures.

For “cross-correlation”, additional genes were eliminated from the B fraction that were more prevalent in the NB fraction, as defined as an NB-B fold-difference $\geq$ 0, as determined in the jointly-normalized B and NB fraction data. The value for each NB signature, defined as the average value of its constituent genes, was then calculated

from the NB fraction of each sample. These values allowed the creation of correlation metric to quantify the correlation with the value of each remaining gene in the B fraction of corresponding samples:

$$correlation\ metric = sign(r) \times \frac{var(B\ gene) \times var(nonB\ signature)}{abs(log2(abs(r)))}$$

The absolute values of this correlation metric would increase with increasing variation in expression of the B gene and/or the non-B signature, and with correlation (either positive or negative) between these two values.

For those genes with multiple probes, the correlation metric values were averaged. Genes were then ranked from highest to lowest by their correlation metric value, which was subsequently utilized for gene set enrichment analysis (GSEA),(81) with standard methods and using default settings and the “Run Gsea on a Pre-Ranked gene list” option.. Gene sets originated from the Molecular Signatures Database v4.0 C2, C3, C5, C6, and C7 categories (http://www.broadinstitute.org/gsea/msigdb/index.jsp), and from custom gene sets created from relevant literature.

Chapter 3

RESULTS

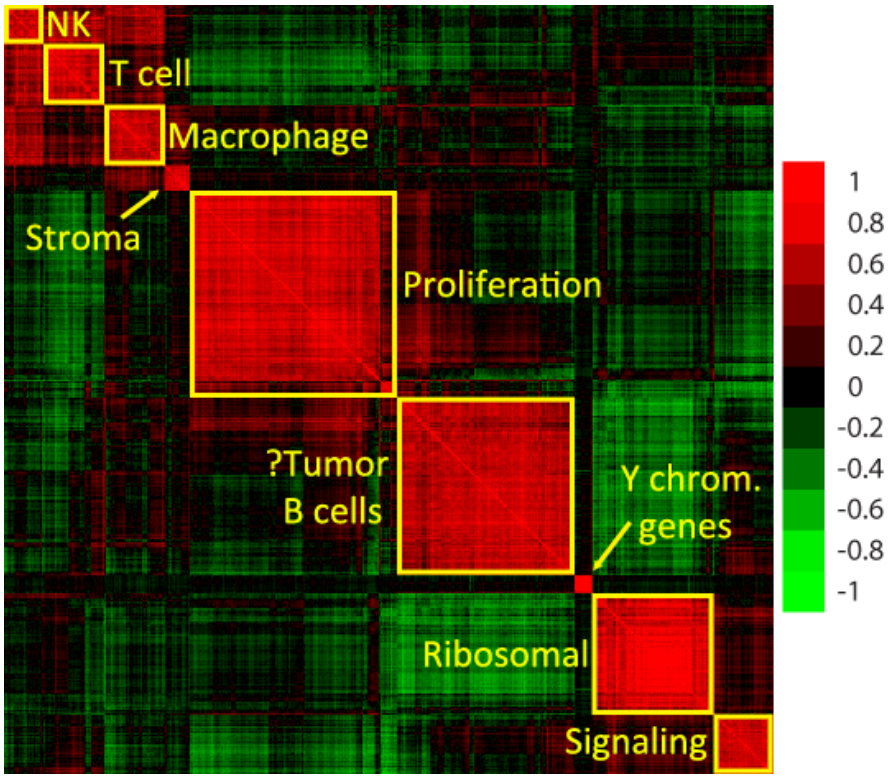
Correlation matrix analysis of whole-tumor GEP data reveals signatures of immune cell types and biological features in FL tumors

We utilized our correlation matrix analysis (CMA) technique to analyze GEP data from whole-tumor pretreatment biopsies of 191 FL patients, which were treated with a variety of therapies, from the Dave et al. dataset.(63) In contrast, the previous CMA work by Galon et al. utilized a small pre-selected list of T-cell relevant genes in their study of colon carcinoma.(82) Additionally distinct from our approach, Dave only assessed for correlation in genes which displayed a univariate correlation with overall survival (OS). In contrast, we took an unbiased approach, as detailed in the Methods section, to analyze the Dave dataset. We generated a master “correlation matrix” of pairwise Pearson correlation r values between all genes. This matrix was then iteratively filtered to arrive at a final set of genes with N equaling 5 or more r values (correlation with other genes in the set) exceeding a certain positive threshold of Y equaling r =0.7 for $p<0.0001$ in the 191 samples, resulting in a list of 468 genes. Data exploration involved selection of trial N and r value thresholds, followed by assessment of hierarchically-clustered heat maps of the r values of final genes. This approach differs from typical heat maps of expression values as it is symmetrical, with the same genes in the same order in both rows and columns. Thus, the diagonal represents a self-correlation. The large red squares along the diagonal of Figure 3 identify groups of highly-correlated genes, containing >N number of genes, which are suitable candidates for designation as signatures. Goodness of heat maps was based on the degree of clustering of high r

values into discrete squares along the diagonal, and the degree to which genes in these squares plausibly represented immune cell types or biological features (e.g., proliferation).

Analysis of the CMA heat map resulted in generation of eight signatures, which were subsequently analyzed for their constituent genes (figure 4). Based upon analysis of the genes which were included, the signatures were attributed to immune cell types or processes as described in the methods using the hypergeometric distribution test. Most of the genes in the attributed signatures were related to immune cell types in the tumor microenvironment, but some (Proliferation and Ribosomal) are features whose restriction to cell types, if any, are unclear.

Figure 4. Whole-FL tumor correlation matrix reveals outcome-correlated signatures

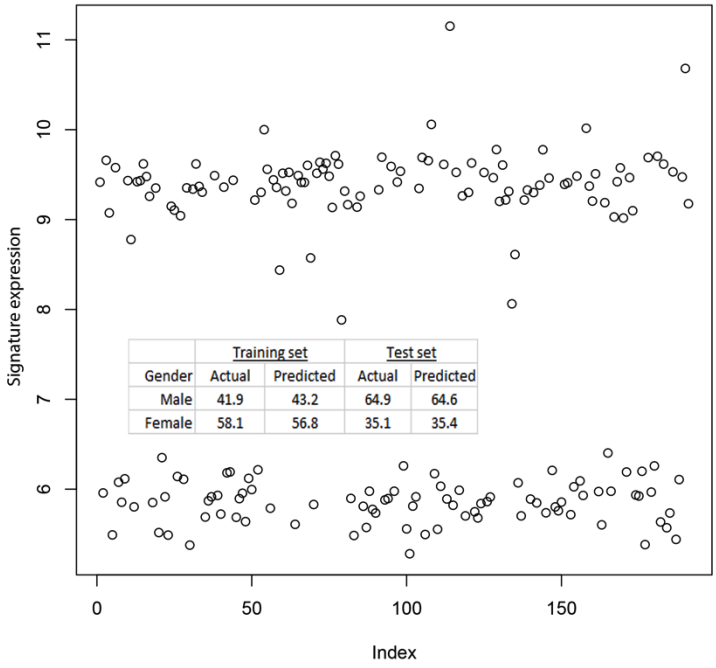


Legend: The self-correlation heat map shows pairwise r values of correlation for

468 genes, selected by iterative filtering with N of 5 and Y of 0.7 for 191 samples in the GEP data of Dave et al. The heat map is symmetrical, with the same order of genes in rows and columns, and hierarchical clustering of any two genes is based on the similarity of their pairwise r values with other genes in the matrix. Pairwise r values are colored according to the scale of the color bar. Large red squares along the diagonal identify groups of highly-correlated genes, which become signatures with the attributions shown in yellow (NK = natural killer).

A ninth group of 11 Y-chromosome associated genes, highly correlated due to the presence of male and female patients in the dataset, served as quality control. The levels of the Y-chromosome signature were distributed in a biphasic pattern which corresponded with the gender proportions in the Dave dataset (Figure 5).

Figure 5. The Y-chromosome gene signature, defined by correlation matrix analysis of whole-tumor GEP data from Dave et al, identifies patient gender.

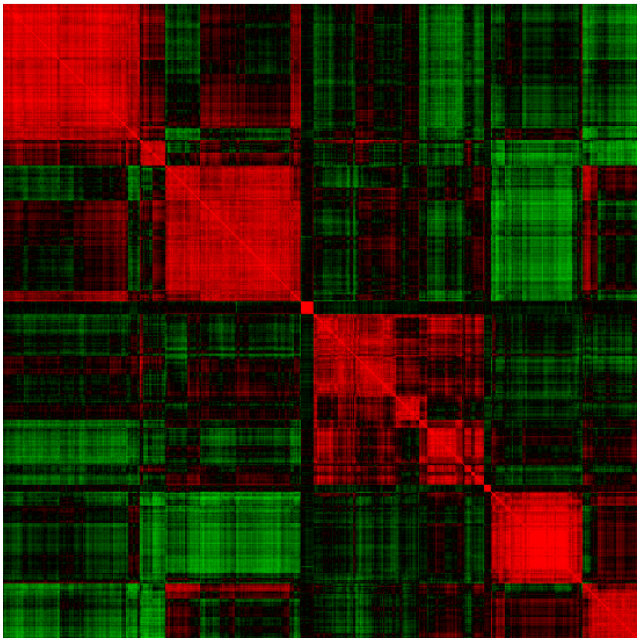


Legend: Y-chromosome signature levels (average log2 levels of signature

genes) are shown for each of the 191 patients, ranked by their patient identifier index number. The inset shows the degree to which signature levels, dichotomized at the log2 = 7 threshold, correctly identifies the gender of patients in the training and test sets of patients

To explore the validity of the iterative filtering, we analyzed the CMA without iterative filtering to evaluate for the presence of organized signatures. A correlation matrix heat map of the 547 genes resulting from the same N and Y criteria but without iterative filtering (Figure 6) is not as well-organized, showing that exclusion of genes by iterative filtering helps in defining signatures. The lack of iterative filtering did not prevent identification of clustering genes, however there were more genes with a strong correlation (red) that were off the diagonal, with much larger and less tightly defined clusters.

Figure 6. Correlation Matrix Heat Map Without Iterative Filtering



Signatures of immune cell types in FL tumors are predictive of outcome

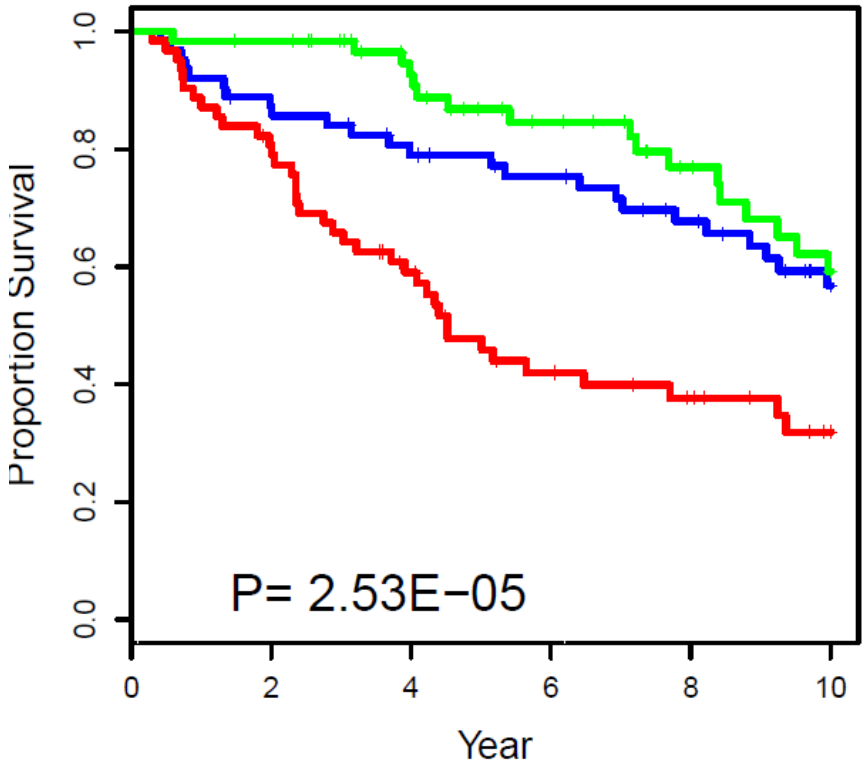
The signatures created with the CMA method were then evaluated in a multivariate Cox model to determine their correlation with survival outcomes (10-year OS) in the 164 patients for which clinical parameters allowing for calculation of the IPI score data. We required that each signature was able to contribute significantly ($p \leq 0.05$) to the creation of the model, both with and without inclusion of the IPI. Based upon this requirement, a 3-signature predictive model was found (Table 1) that stratified patients into groups with significantly-different OS (Figure 7).

Table 1. Multivariate Cox Model to determine the correlation with survival outcomes

Overall model	IPI	T cell	Macrophage	Stroma
7.602e-05		0.000585 **(-0.764)**	0.000106 **(1.0389)**	0.000919 **(-0.6556)**
1.976 E-06	0.000691 **(0.6996)**	0.015134 **(-0.5967)**	0.001203 **(0.9396)**	0.014737 **(-0.512)**

The p values for the individual parameters, and for the models overall (by the Wald test) are without parenthesis. The model coefficients, with **positive** values indicating parameters with an unfavorable effect on survival and **negative** values indicating a favorable effect on survival, in parenthesis.

Figure 7. Overall survival in the three signature model without the IPI

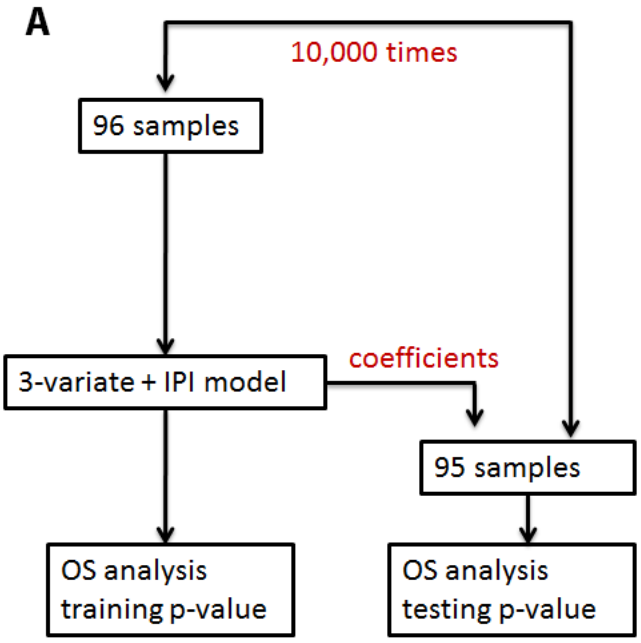


As a confirmatory finding, our model included signatures for T cells and macrophages that were respectively OS-favorable and OS-unfavorable in the Dave dataset, and no signature of apparent B-cell origin similar to the Dave dataset. Distinct from the Dave initial findings included a favorable effect of stroma, as previously reported in diffuse large B-cell lymphoma,(83) and that our signature genes were more clearly linked to T cells (e.g., CDs 2, 3, and 5; ZAP70; ITK; ICOS) and MP (e.g., CDs 14 and 63).

Highly filtered data such as our signatures is often evaluated for the possibility of "overfitting", which is ideally addressed by utilizing the model in an independent dataset. As there is no such available comparable dataset, we simulated testing our 3-signature model 10,000 times. Each test allowed determination of the coefficients in a randomly-selected training set (50% of patients), with subsequent evaluation with the coefficients in the test set (50% of the patients). Utilizing this method of confirmation with inclusion

of the IPI to correct for patient factor imbalances between sets, the model was predictive ($p \leq 0.05$) in both the randomly-selected training and test sets in 90.5% of the 10,000 tests, far greater than by chance (figure 8). This is not equivalent to an external dataset, but does lessen concerns regarding overfitting the data.

Figure 8. Permutation test for robustness of the 3-signature+IPI model



B

	test.p>0.05	test.p<=0.05
train.p>0.05	2	490
train.p<=0.05	454	9054

	test.p>0.01	test.p<=0.01
train.p>0.01	32	2038
train.p<=0.01	1630	6300

Legend: A) Schematic of the iterative procedure used. For the 164 patients for whom IPI data were available, each iteration involved random division into training and test sets of 82 patients each. For the training set patients, coefficients and the significance of correlation with overall survival (OS) were determined in a multivariate Cox model using levels of the IPI and the 3 most predictive CMA-defined signatures (Macrophage, T cell, and Stroma). The significance of correlation with OS was then determined in the test set, using the same coefficients. The process was repeated for a total of 10,000 times.

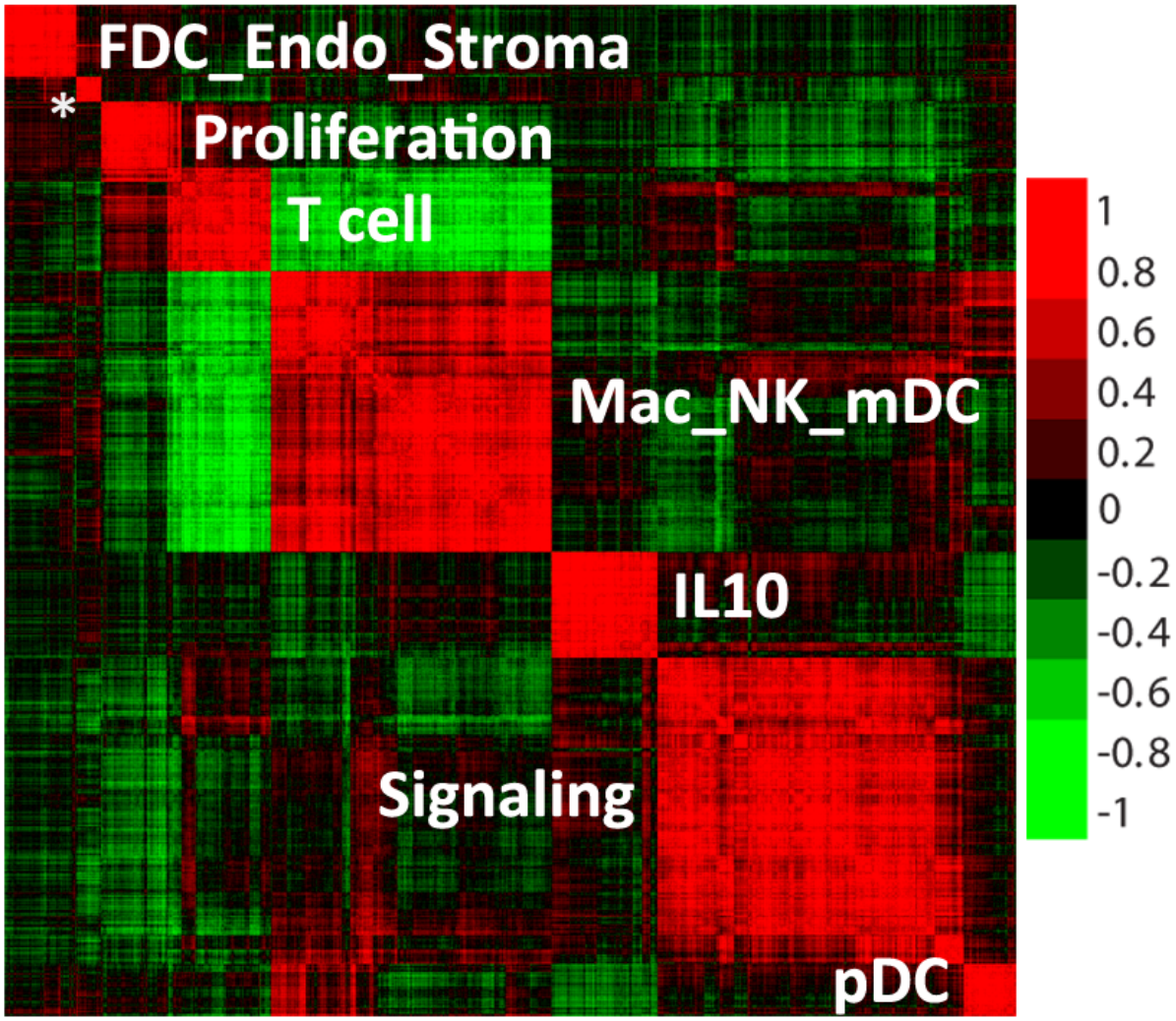
B) Totals for the 4 possible outcomes, according to threshold of significance.

Features of immune cell types in the FL microenvironment revealed by CMA of separated tumor cell suspensions

Although our correlation matrix analysis of the data from Dave et al. confirmed the previously reported importance of the microenvironment in FL, featuring opposing effects of macrophages and T cells, it did not provide further insights into important interactions between various cell populations. As the detection of GEP features of a particular cell type are likely obscured by the presence of other cell types in a heterogeneous sample,(84) the presence of B cells in whole-tumor biopsies may have limited or influenced our ability to examine the microenvironment, as well as its effects on B cells. To account for this limitation, and attempt to correct for it, we analyzed viably-frozen cell suspensions from untreated FL tumors, as described in the Methods section. Technically adequate results for both B and NB fractions from a given patient were achieved for 43 patient samples.

CMA of NB fraction data ("self-correlation"), using N of 5 and Y of 0.8 ($p<0.0001$), resulted a highly organized heat map of 1097 genes (Figure 9). Without the potentially confounding presence of B cells, we could detect signatures of immune cell types aside from T cells and macrophages, including five highly self-correlating signatures of immune cell types and three highly self-correlating signatures of mechanistic processes (Signaling, IL10, and Proliferation).

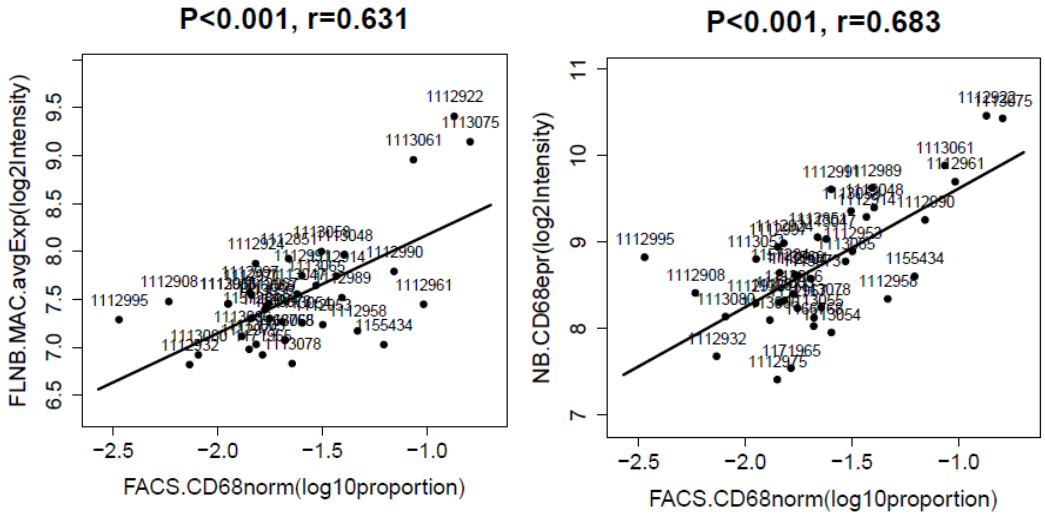
Figure 9. Self Correlation heat map from the NB fraction



Legend: The self-correlation heat map shows pairwise r values of correlation for 1097 genes from GEP data of NB fractions from 43 FL tumors, selected by iterative filtering with N of 5 and Y of 0.8. Signature attributions are in white: Mac = macrophage, NK = natural killer, mDC = myeloid dendritic cell, FDC = follicular dendritic cell, Endo = endothelial, pDC = plasmacytoid dendritic cells. The signature marked with an asterisk is an FDC_Endo_Stroma-associated signature of 27 genes.

Of interest, plasmacytoid dendritic cells (pDC) manifested a small self-correlating gene signature, despite their typically only representing ~0.1% of all cellular events by flow cytometry (FACS) in FL. Some of our identified immune cell signatures were not attributable to a single cell type including macrophage (Mac) genes which were mixed with genes of natural killer (NK) and myeloid dendritic cells (mDC) in a Mac_NK_mDC signature. However, the Mac_NK_mDC signature did correspond to the presence of macrophages in the tumor samples. For confirmation, 39 samples from this dataset were analyzed with FACS for CD68+ cells. The proportion of CD68+ cells among non-B cells correlated highly with NB fraction levels of CD68 gene expression alone and with the Mac_NK_mDC signature (Figure 10)

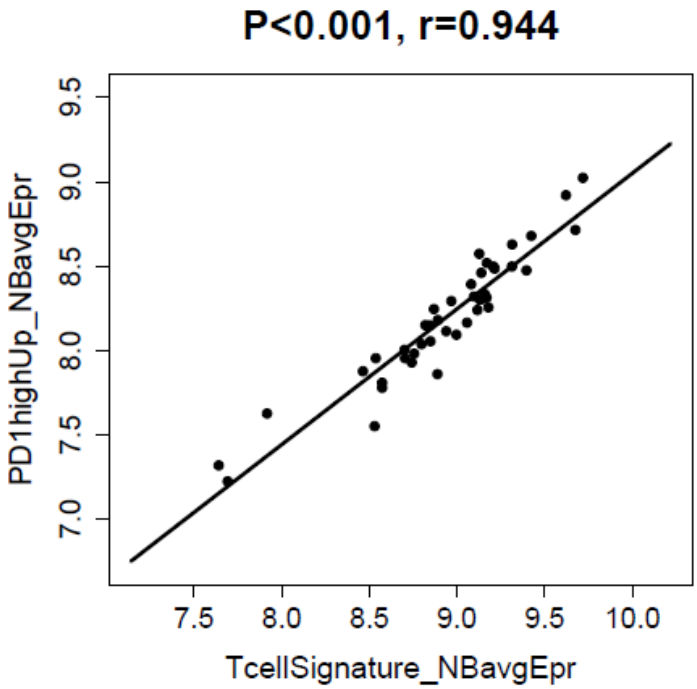
Figure 10. Flow cytometry of CD68 expression compared with Mac_NK_mDC signature and CD68 gene expression



Legend: The expression of CD68 was determined as outlined in the Methods. In each tumor sample, the median expression value of the Mac_NK_mDC signature (left panel) and gene expression of CD68 (right panel) were correlated with the normalized CD68 flow cytometry data, showing good correlation.

The self-correlating T cell signature found in our NB fraction (T_{NB}) was compared to signatures of effector T cells (T_{eff}) and T_{FH} cells previously published by us,(42) derived from GEP comparing flow cytometry defined T_{eff} (PD1(PDCD1)^{int}CXCR5^{int} or PD1^{lo}CXCR5^{lo}) vs. T_{FH} cells (PD1^{hi}CXCR5^{hi}) in 3 independent FL samples. Although only 18 genes were identified in both the 113-gene T_{NB} signature created by our CMA and the 55-gene T_{FH} signature from our prior work, we found a highly-significant positive correlation (r=0.944, p<0.001) in their levels in the NB fraction (Figure 11).

Figure 11. The CMA-defined T_{NB} signature compared with a previous T_{FH} signature

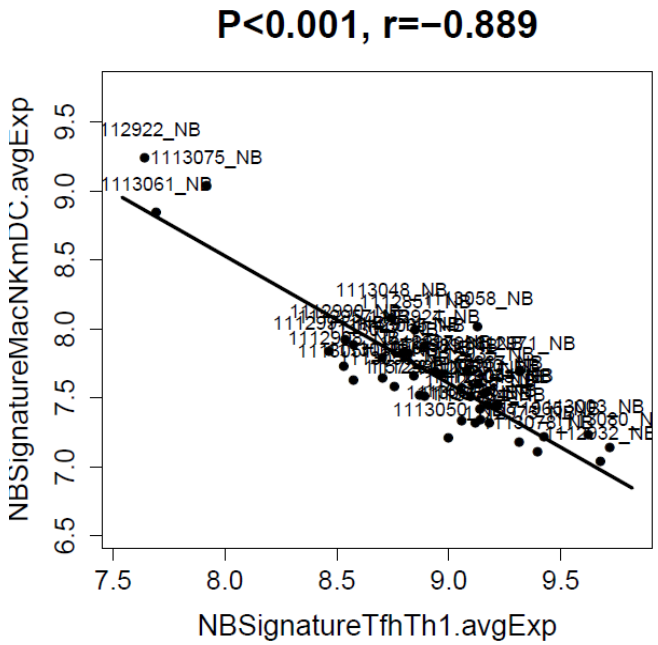


Legend: The median gene expression value for the T_{NB} signature was determined from each patient and compared with the median gene expression value for the previously published $T_{FH,}$ from the same patient, showing excellent positive concordance.

Based upon this strikingly high correlation, we conclude that T_{NB} signature from the separated FL samples essentially represents a T_{FH} signature.

In Figure 9, the increased negative correlation manifested by green off-diagonal clusters of negative r values at the intersections between T_{NB} and Mac_NK_mDC clusters suggests an inverse relationship between these cell types. This finding is confirmed by Pearson correlation between average levels of signature genes ($r=-0.889$, $p<0.00001$; Figure 12). Additional negative correlations were noted between the pDC and IL-10 signatures, and between the Proliferation and both Signaling and Macrophage signatures. A positive correlation was observed between the Proliferation and T cells signatures in the FL microenvironment.

Figure 12. The T_{NB} signature compared with the Mac_NK_mDC signature



Legend: The median gene expression value for the T_{NB} signature was determined from each patient and compared with the median gene expression value for the Mac_NK_mDC signature from the same patient, showing strong negative concordance.

In theory, the attribution of a signature to a particular cell type could allow the conclusion that other genes in the signature, not previously reported to be associated with a specific cell type, are expressed by the putative cell of origin. This feature of self correlation matrixes applies to many genes, including chemokines, cytokines, and surface or intracellular molecules such as CCL19 from FDC/endothelial cells, and CCL3, CXCL16, FGL2, S100A8, S100A9, IL8, and TNFSF13B from macrophages.

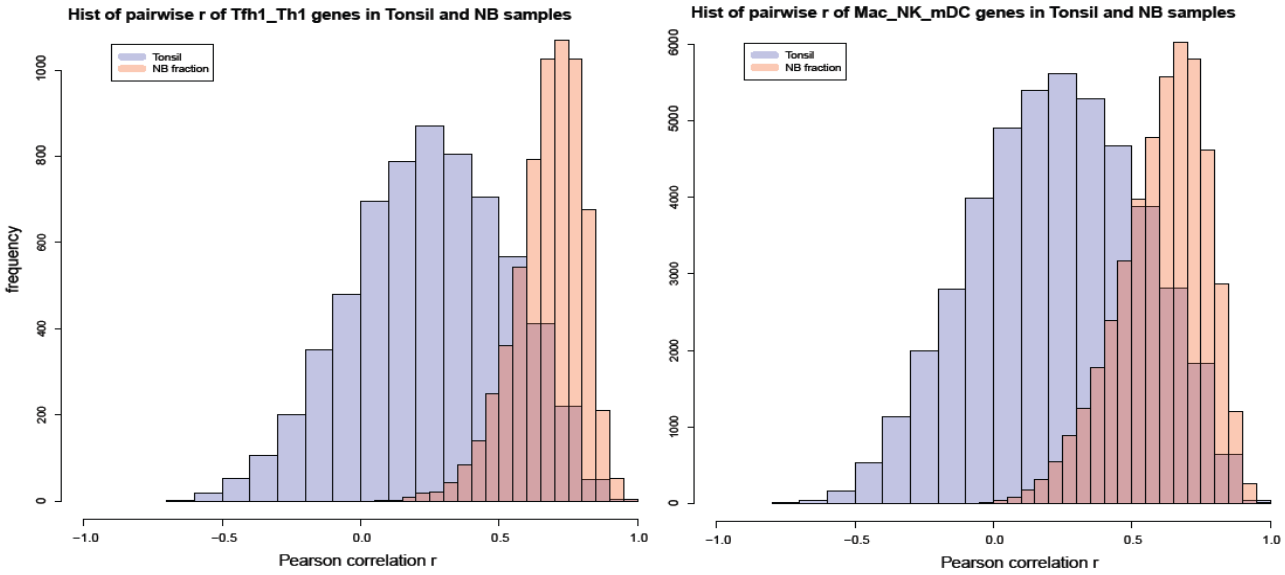
A single signature for T cells was identified via our self correlation matrix, containing the genes CXCR5, PDCD1, ICOS, SH2D1A, CD40LG, CD200, BTLA, TIGIT, CD200 that are known to have varying levels of expression levels in distinct T cells subsets, such as Tfh, Th1, Tc1, and Treg.

Comparison of follicular lymphoma microenvironment with the normal lymphoid microenvironment

The interactions suggested by our correlation matrix analyses of FL patient tumor biopsies may not be unique to the FL microenvironment. To address the potential that immune cell types found in the FL microenvironment are qualitatively different from immune cells in a benign lymphoid organ, we performed GEP on the NB fraction of tonsils from 24 healthy children. In the NB fraction of the tonsils, we evaluated the T_{NB} and Mac_NK_mDC signatures that we had previously generated in the FL NB fraction CMA. The individual genes from each signature were evaluated for their pairwise correlation in the benign tonsil NB fractions. In addition, we compared the expression of individual genes from these signatures in the NB tonsil population to the gene expression in the FL NB fractions. We found that the pairwise r values of the expression of the T_{NB} cell and macrophage signature genes demonstrated a significantly weaker

correlation with each other in benign tonsils than in FL samples, as graphed in histograms from the two sample types (Figure 13). The Pearson correlation r values are much lower in the NB fraction of tonsils than FL, which suggests that the T cells and macrophages (or NK or mDC cells) in the FL cells are distinct from their benign microenvironment counterparts.

Figure 13. Histogram of T_{NB} and Mac_NK_mDC signature correlation in FL and Tonsils



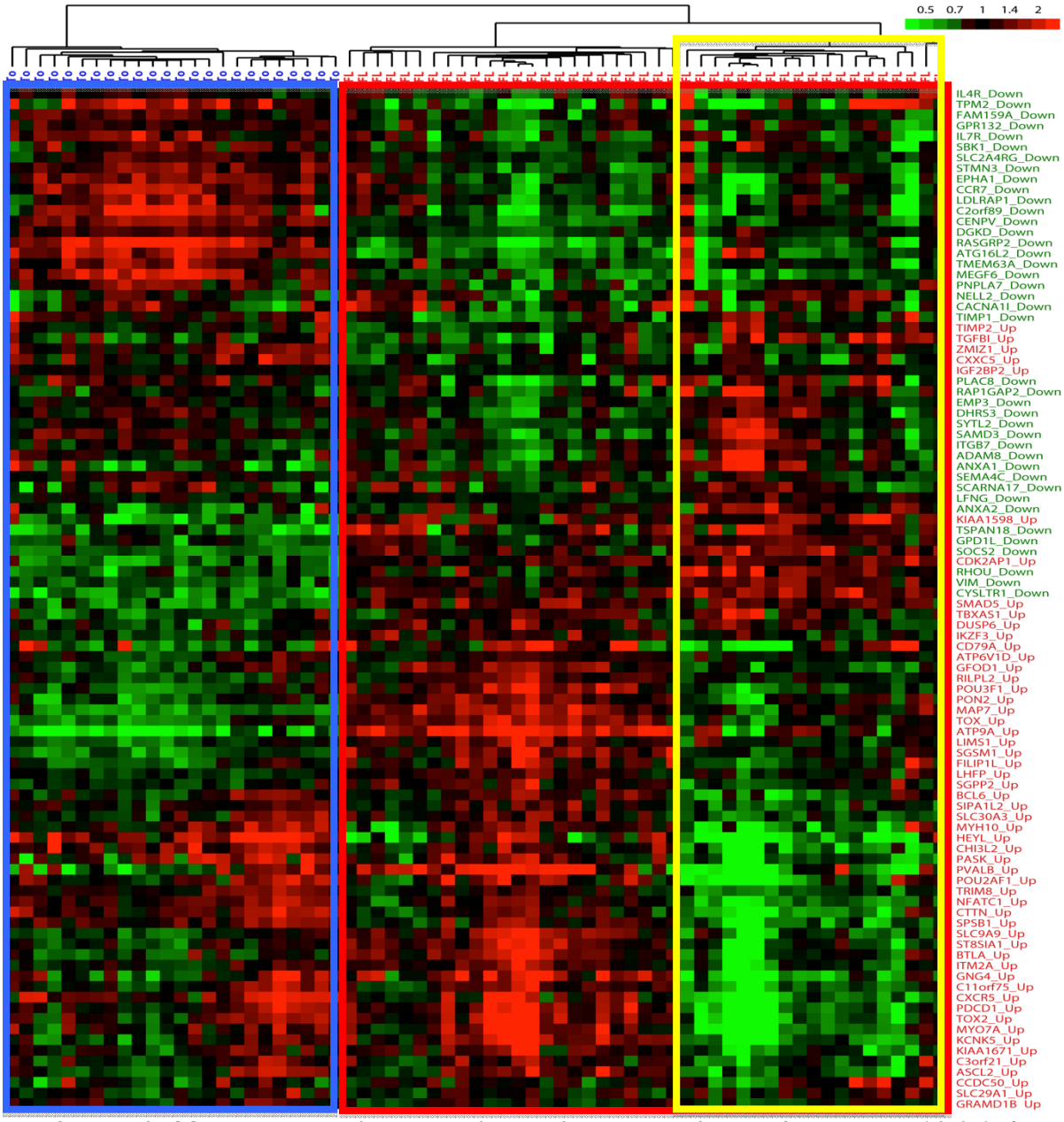
Legend: The genes comprising the T_{NB} signature and the Mac_NK_mDC were evaluated for correlation with other genes of in the signature. The histogram shows the tonsil (blue) samples had a much weak correlation than did the FL (orange) samples, demonstrating the differences between the microenvironment in the benign and malignant context.

We concluded that this weaker correlation in the NB fraction of the benign tonsils is due to the possibility that our signatures correspond to with a cell type that is not normally found in benign lymphatic tissue. Based upon the above analyses, this indicates that

TFH and at least one of the Mac_NK_mDC cell types are qualitatively distinct in FL vs. the normal lymph node microenvironment.

Similar findings were identified when the expression of previously determined signatures of effector T cells (T_{eff})(42) and T_{FH} cells in FL and tonsil samples were compared. When the genes from these signatures were combined in a single heat map (Figure 14), we identified clusters which corresponded largely according to signature, and perfectly distinguished the FL samples from the tonsil samples. Furthermore, the samples from the NB FL population clustered into two distinct groups, predominantly on the basis of genes more associated with T_{FH} cells. These data suggest that T_{FH} cell infiltration of FL is not uniform, and may speak to the clinical heterogeneity seen in this disease.

Figure 14. Heat map comparing FL and tonsil samples for the expression of previously-determined signatures of effector T cells (Teff) and TFH cells



Legend: Mean-centered expression values are shown for genes (right) from the combined signatures distinguishing sorted Teff (green) and TFH (red) cells in 3 previously-studied FL samples. The color bar indicates the fold value of expression relative to the mean of each gene. Data come from GEP of NB fractions of 43 FL

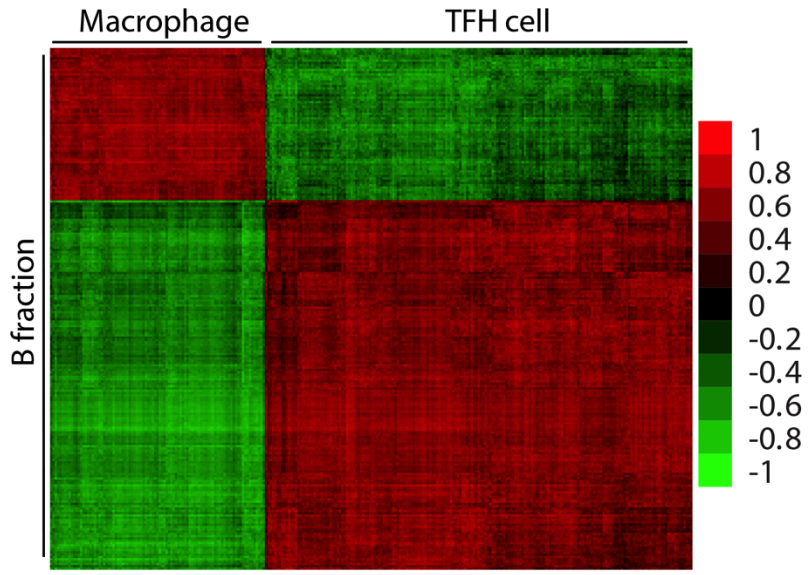
samples (red) and 24 normal tonsils (blue). Hierarchical clustering for similar variation in expression shows that genes cluster largely according to signature, and samples cluster perfectly according to sample group. The FL samples (red box) have 2 relatively distinct populations, one of which is highlighted (yellow box).

"Cross-correlation" analysis identifies interactions between tumor cells and immune cells in the immune microenvironment

A fundamental hypothesis underlying this thesis is that the biology of FL involves significant cross talk, either symbiotic and/or antagonistic, between FL cells and multiple benign immune cell types in the tumor microenvironment. To examine this hypothesis, we explored whether these interactions might be revealed by "cross-correlation" analysis, similar to the "self-correlation" analysis of the NB fraction described above, but involving correlation between levels of genes in one cell fraction (B or NB) and those of genes in the opposite fraction. Individual genes in each fraction for cross correlation were limited to those whose levels were at least as high as in the other fraction, to minimize the risk for identifying highly cross-correlated genes originating from cell types present in both fractions. Iterative filtering, using N of 8 and Y of 0.65 applied to genes of both fractions, produced a cross-correlation matrix of 226 B fraction genes and 278 NB genes. The heat map of this matrix (Figure 15) clustered into non-overlapping paired groups of positively-correlated genes from each fraction, which were negatively-correlated with each other. Evaluation of the gene lists via the hypergeometric distribution test showed that one group of NB fraction genes was characteristic of T_{FH} cells. The positively-correlated group of B fraction genes which corresponds to the T_{FH} genes contained many genes associated with proliferation, including CDC20, TOP2A,

CCNB2, AURKA, AURKB, suggesting that tumor *T_{FH} cells support proliferation of FL B cells*. This group of B genes also included SERPINA9, previously reported to be highly characteristic of germinal center B cells and FL.(85, 86) The other group of NB fraction genes was characteristic of macrophages. Upon evaluation with the hypergeometric distribution test, the corresponding positively-correlated group of B fraction genes were not clear as to their implications, but overlapped significantly with several stem cell gene sets.

Figure 15. Cross-correlation heat map of B vs. NB fraction genes in FL tumors



In another approach to cross-correlation, levels of individual B fraction genes were compared to those of T cell signatures in the NB fraction; other signatures were not considered, since the methods used to prepare the fractions did not exclude the potential presence of non-B, non-T cells (such as macrophages) in both fractions. This approach also differed in that the metric for comparison was not based solely on the Pearson r value, but also considered the variances of expression of each gene and signature as described in the Methods section. This approach created a cross-

correlation matrix which represents a set of independent correlations between the T_{NB} signature and all B fraction genes. These values allow ranking of individual B fraction genes by their correlation with each NB signature. Table 2 shows results for the top individual 20 genes from the B fraction with the highest positive correlation with the T cell signature from the NB self correlation matrix. The B fraction gene with the strongest correlation with the T cell signature was SERPINA9, and several other genes associated with proliferation. We also identified genes previously described to be associated with proliferation (e.g., CDC20, TOP2A, AURKA, AURKB) were highly correlated, which imply that the NB T cell signature and/or tumor infiltrating T_{FH} cells play a positive role in supporting B-cell proliferation, which here would mean malignant B-cell proliferation. As additional findings, B fraction levels of CCL17 and CCL22 were highly ranked, consistent with our previous observation that T_{FH} induce these chemokines in FL B cells.(87)

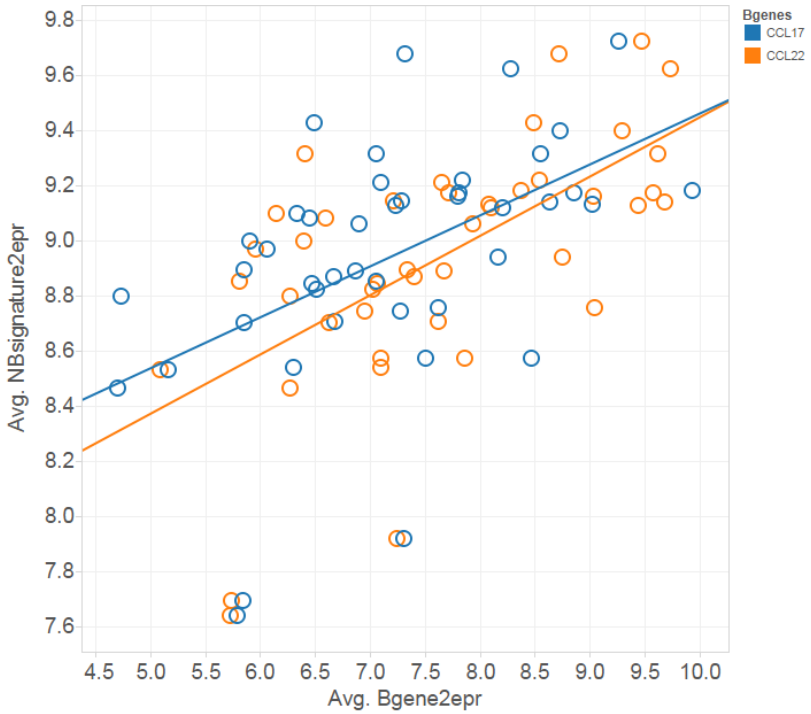
Table 2. B genes having expressions with top 20 highest correlation with the T-cell NB signature

	TFH,ECMA-defined	
symbol	**Metric**	**Rank**
SERPINA9	0.79959191	1
CDC20	0.75779624	2
KIAA0101	0.51064588	3
TOP2A	0.50500598	4
TYMS	0.47261181	5
RPS4Y1	0.46898551	6
CCL22	0.45483511	7
IGJ	0.41632578	8
CCNB2	0.37535751	9
CAMP	0.33334485	10
UBE2C	0.32712714	11
KCNK12	0.30605593	12
ELL3	0.30466897	13
CCL17	0.29107038	14
GTSF1	0.27116205	15
AURKA	0.27057119	16
CDCA5	0.26226925	17
AURKB	0.25810637	18
KIFC1	0.25697333	19
SLC2A5	0.2460317	20

In addition, we also examined the previously determined T_{eff} and T_{FH} signatures(42) in cross-correlation analysis between the B and NB fractions. B fraction genes that ranked most positively by the correlation metric with the T_{FH} signature in the NB fraction, and most negatively with the T_{eff} signature, were similar to those ranked positively with the T_{NB} signature. This finding is consistent with our previous conclusion that the T_{NB} signature is effectively a surrogate for a T_{FH} signature. Further supporting our prior conclusion that T_{FH} cells support FL proliferation, we also identified that the prior published T_{FH} signature in the NB fraction correlated highly with B fraction levels of a proliferation signature identified by CMA of the B fraction.

Cross-correlation analysis was able to identify other previously reported biological interactions between FL B cells and immune cells. We previously identified that T_{FH} cells induce expression of the chemokines CCL17 and CCL22 by FL B cells.(87) In our cross correlation analysis, we identified that B fraction levels of these chemokines were highly ranked as positively correlated with NB fraction levels of both the T_{NB} and T_{FH} signatures. This finding may imply that the expression of these chemokines by malignant FL B-cells likely plays an important role in shaping the FL immune microenvironment, including recruitment of T_{FH} cells. We also found a positive correlation between individual values of the T_{NB} signature in NB fractions and values of CCL22 and CCL17 in their corresponding B fractions, is shown in Figure 16.

Figure 16. Correlation of CCL17 and CCL22 in the B tumor fraction with the TNB signature in the NB fraction



Legend: The individual genes CCL17 and CCL22 were determined in the B tumor fraction from each patient and compared with the median expression value of the T_{NB} signature from the NB fraction in the same patient, showing good concordance.

To allow for further interpretation of the cross correlation matrix, we employed Gene Set Enrichment Analysis (GSEA)(88) which we conducted by ranking B fraction genes based upon their correlation with the T_{NB} signature. We identified previously known findings, including that ranked B fraction genes based upon correlation with the T cell signature had a strong positive enrichment for gene sets related to proliferation. However, our GSEA also positively enriched two gene sets which are based the comparison of germinal center (GC) B cells to cells at neighboring stages of differentiation, including before (naïve B cells; GSE12366_GC_VS_NAIVE_BCELL_UP)

or after (memory B cells; GSE12366_GC_VS_MEMORY_BCELL_UP) the GC stage.(89) This finding was not anticipated as all samples were from FL, which is known to correspond to the GC stage. This finding suggests that the FL samples with B fractions that have strong similarities to normal GC B cells are those with increased T_{FH} cells. Another possible mechanism for this finding may be that T_{FH} cells are associated with FL variants with a greater reliance on B-cell receptor (BCR) signaling, as a signature comparing splenic B cells from wild type mice against BTK-mutant mice (GSE2826_WT_VS_XID_BCELL_UP)(90) was also enriched by correlation with the T_{NB} cell signature.

Chapter 4

DISCUSSION

The creation and use of signatures to interpret GEP data is a common practice as it allows for conceptualization of the complex gene interactions that occur in normal and diseased cells. When a signature is determined in a given context, it is generally applicable to other datasets, at least for hypothesis generation. This includes signatures found in the public repository Molecular Signatures Database. The novel technique of correlation matrix analysis (CMA), assisted by iterative filtering, is distinctive in that it is able to identify groups of coordinately expressed genes from the data themselves, which thus fit the data well. This allows for a "dimension reduction" of the GEP data, which can greatly assist analysis.

The work in my thesis is an unbiased use of all potential genes for CMA, as opposed to previous applications of CMA, which have generally utilized a preselected set of genes, for which the proper attribution to a given cell type is known. Our novel approach poses new challenges, the first being whether signatures can be identified that are discrete (i.e., composed of genes that correlate more highly with each other than with genes of other signatures) and biologically plausible rather than random collections of covariant genes. The use of iterative filtering, which we developed for this study, contributed greatly to our ability to identify discrete signatures. Although results depended on the N and Y criteria used for the iterative filtering, we found that the number of signatures was relative small, and that their constituent genes generally provided a clear basis for attribution. In other words, at least in FL, there is substantial "structure" to the data, reflecting the presence of well-defined immune cell types and highly-orchestrated

biological processes (such as proliferation) in the tumor. In other words, just as there are not an unlimited number of different cell types in an FL tumor, so there are not an unlimited number of signatures.

In the dataset from the landmark study of intact whole-tumor of FL GEP by Dave, our method of CMA was able to identify multiple groups of coordinately expressed genes which we could attribute to biologically relevant immune cell subsets and other essential cellular function programs. Our signatures included the OS-favorable T cells and OS-unfavorable macrophages, similar to the initial non-CMA findings. We did not consider these similarities to be surprising, as the initial group used hierarchical correlation, similar to the pairwise correlation we used, to group genes into a total of 10 signatures, 2 of which were T cells and macrophages. However, their data analysis did not incorporate all potential genes and thus had an introduction of bias. The genes that Dave et al selected for clustering had to first meet a minimum level of significance ($p \leq 0.1$) with OS in a univariate correlation analysis. In contrast, our analysis included all genes, and performed correlation with OS only on CMA-defined groups of genes, or signatures. This distinction likely resulted in our CMA-defined T cell and macrophage signatures including a greater number of total genes, and a greater number of genes that are characteristic of their putative cell of origin. In our T cell signature, we identified genes associated with CDs 2, 3, and 5, ZAP70, ITK, and ICOS. In our macrophage signature, we identified genes associated with CDs 14 and 63, suggesting biologic plausibility. Our analysis also identified a stroma-related signature which was OS-favorable, consistent with the findings of others in another subtype of NHL, diffuse large B-cell lymphoma.(83) Together, these findings lead us to conclude that our method of

signature creation is valid and can create biologically relevant signatures that are attributable to immune cell subtypes and other cellular processes.

Our CMA-defined T cell signature from the Dave et al data was not specific to a particular T cell subset or physiologic state, comparable to original analysis. As a result of this caveat, our signature is limited in mechanistic implications. However, our CMA of the separated untreated FL tumor samples, specifically the NB fractions resulted in our T_{NB} signature, larger than that of the T cell signature of Dave. Based on further correlative analyses, we concluded that our T_{NB} signature is effectively representative of T_{FH} cells. Not all tumor infiltrating T cells in FL are T_{FH} cells, as has been shown by immunostaining and other analyses.(87) However, our CMA-defined T cell signature implies that T_{FH} cells are the most consistent T cell subset across our FL biopsy samples, although varying in their frequency. Based upon review of the literature, we did not find this to be unexpected as FL is often described as originating from the normal germinal center of benign lymph nodes in which T_{FH} cells play an essential role. Our findings regarding T_{FH} cells were also unsurprising as we found that T_{FH} cells were increased in the NB fraction of FL tumors when compared to reactive lymph tissue.(91) However, our T_{NB} signature did not have significant applicability in the NB fractions isolated from the tonsil samples, which we concluded to imply that the T_{FH} in FL and the normal lymph node microenvironment have distinctive qualitative differences.

Based upon our conclusion that T cells, specifically T_{FH} cells, have significant self correlation in FL samples, it is logical to assume that this may occur due to interactions occurring between FL tumor B cells and immune cells in the tumor microenvironment. To further support this hypothesis, we identified “cross-correlations” between genes

expressed in isolated malignant B cells from FL and the level of expression of NB fraction signatures. The nature of these relationships, antagonistic or tumor promoting, is less clear, but their existence suggest a bi-directional interaction on and by both B and non-B cells. Pangault et al. also performed GEP on sorted FL tumors or reactive lymph nodes and tonsils, utilizing different methods than ours including positive selection of cells during isolation. Their analysis generated “interfaces”, defined as sets of genes which could distinguish B and NB fractions.(92) Their use of this analytic technique concluded that T_{FH} cells were responsible for the majority of the differences between the FL and non-FL interfaces. In addition, they also concluded that the T_{FH} cell population affected B cells. Further supporting their implication of the role of T_{FH}, they also found that IL-4 expression was distinctive for FL T_{FH} when compared with normal T cells. The FL B cells demonstrated effects of this T_{FH} IL-4 signaling, including high expression of IL4I1, an IL-4 target gene whose B fraction levels were highly ranked by metric correlation with T_{NB} signature levels in our data. Ame-Thomas et al also identified IL-4 as the most highly expressed gene by FL T_{FH} cells as compared to tonsil T_{FH} cells. In addition, they also found that antibodies which block CD40L or IL-4 were able to reduce the ability of autologous FL T_{FH} cells to promote survival via protection of tumor B cells from spontaneous apoptosis.(93) Myklebust et al. found that FL T_{FH} cells can have IL-4 production induced by stimulation, but are insensitive to STAT6 phosphorylation via IL-4, and thus can avoid a negative feedback loop.(73) Lastly, Kiaii et al. identified several genes, including PMCH, ETV1, and NAMPT, that were more highly expressed in $CD4^+$ and/or $CD8^+$ T cells from FL samples than from tonsils. In normal T cells co-

cultured with FL B cells, but not with tonsillar B cells, these genes were up-regulated implying a distinctive FL-T cell interaction.(94)

Recent innovative studies using multiparameter confocal immunofluorescent microscopy have evaluated the FL microenvironment. This technique allows for evaluation of multiple parameters, similar to flow cytometry, without disruption of the tissue architecture, similar to immunohistochemistry. These analyses support our findings that T_{FH} cells have effects on FL B cells, including promotion of their proliferation, as they demonstrate enrichment of FL tumor follicles for T_{FH} cells, defined as $CD4^+/PD1^+/ICOS^+$ cells. They confirmed these cells were T_{FH} cells by protein expression of BCL6, the master T_{FH} transcription factor.(95) The multiparameter microscopy also found that 42% of proliferating tumor B cells, defined as $Ki67^+$, were in direct contact with T_{FH} cells including synapse formation.

An additional implication of this FL B cell – T_{FH} cell interaction was shown with cross-correlation with NB fraction levels of the T_{NB} signature, and those from a previously-defined T_{FH} signature, suggesting induction of CCL22 and CCL17 in FL B cells by infiltrating T_{FH} cells. This supports a conclusion of our previous work(87) that identified CCL17 and CCL22 levels to be significantly higher in both serum and tumor B cells from FL patients when compared with healthy donors. We also previously found that T_{FH} cells can strongly induce CCL17 and CCL22 production by FL tumor B cells through IL-4 and CD40L. Lastly, we also previously demonstrated that cultured FL supernatant could induce migration of both IL-4-producing T cells and regulatory T cells (T_{reg}), more so than IFN-gamma-producing T cells.(87) Other investigators have previously found

similarities between T_{FH} cells and subsets of T_{reg} cells (Follicular T_{reg}, T_{FR}), both which express $CD4^+/CXCR5^+/ICOS^+$ cells in FL.(91, 93)

We found substantial differences in our CMA results of the whole tumor Dave dataset and of the separated 43 FL samples from our dataset. These differences included the gene lists selected by our iterative filtering, with only 87 genes overlapping between the Dave list (468 genes) and from our NB fraction (1097 genes). Both datasets yielded T cell and macrophage-related signatures generated from CMA, of which there were few genes in common, although more than expected by chance. These differences in CMA do not imply that the datasets are non-compatible, as some of the differences could result from whole vs. sorted tumors. Another potential explanation for these differences could be the use of the Affymetrix microarray platform for Dave, and Illumina for our dataset. Another potential limitation of our approach is illustrated by our identification of signatures of rare cell types including FDC and pDC in the NB fraction, but inability to identify unique signatures of other immune cell subsets (e.g., macrophages, NK cells, and mDC), or identify an attributable signature of others which would be expected to be present at low levels (e.g., T_{regs}, and T_{eff} subsets like T_{H1}, T_{H2}, and T_{H17}). An explanation of this limitation may be the potential similarities or overlaps in the expression of genes between these cell types, or because they may display significant heterogeneity across patient samples.

The work in this thesis confirms the findings of others and contributes additional data to support the hypothesis of a symbiotic feed-forward relationship between the microenvironment and FL cells. In this interaction, T_{FH} cells induce both proliferation and production of CCL17 and CCL22 by FL tumor cells, which subsequently leads to

attracting T_{reg} and IL-4-producing T cells. These tumor-infiltrating benign immune cells further stimulate chemokine production, and perpetuate additional pro-tumor cell recruitment.

In conclusion, the results of my thesis project suggest that the novel technique of CMA may have broad applicability to characterize interactions found in cancer biopsies, and shed light on complex but potentially targetable abnormal interactions. Our CMA method can define signatures of both cancer cells, as well as infiltrating benign immune and stromal cell subsets in the tumor microenvironment. These novel, or modified from prior, signatures may subsequently be utilized to characterize and potentially understand these tumor-microenvironment interactions, determine the prognostic impact of the presence or absences of microenvironmental cell types, discover novel therapeutic targets for future study with clinical trials, and to identify subsets of patients with a potential increased likelihood of response to CMA signature informed targeted therapies.

Chapter 5

STRENGTHS AND WEAKNESSES

This thesis is the first attempt at using CMA paired with separated cell population analyses in follicular lymphoma. The significant strengths of my project include its unique ability to determine the complex interactions between various cell types in the tumor microenvironment, and thus determine not just which cell types are present, but are also exerting an influence on other types.

The work in this thesis demonstrates that the novel approach of CMA can discern signatures consisting of genes attributable to various immune cell types in FL. We believe that this approach can be utilized to implicate mechanisms which are essential to FL biology, and could be applied to other cancer types for similar analyses. We believe that our T_{NB} signature was validated by its high correlation in our analyses with a previously determined T_{FH} signature from cells sorted with flow cytometry.

However, it is important to note there are several limitations to the work detailed in this thesis, including the technique of CMA and the confidence of the results. By definition, CMA requires multiple samples in order to evaluate for correlations, and becomes increasingly powerful with an increasing sample size. The implication of this dependence on sample size is that findings from CMA, including the signatures that it can generate, may be specific to the sample set and thus not applicable globally, especially with smaller or highly selected datasets. Furthermore, the determination of the variables of N and Y for iterative filtering is subjective, and may have implications on the ability of the conclusions to be applied to other datasets. The evaluation of goodness of correlation matrix heat maps is also admittedly operator-dependent.

However, in our analysis, we found that for a particular dataset, the CMA generated signatures did not change significantly as the variables were modified.

Chapter 6

FUTURE DIRECTIONS

The work in this thesis suggests previously unexplored potential targets for therapeutic intervention. Agents which target the pro-tumor B-T_{FH} interaction which we have demonstrated to be present and robust include antibodies which block the interaction between CD40L and CD40, which have been tested in clinical trials. To date, these agents have had only modest efficacy in patients with relapsed FL,(96) however this was as a single agent. In isolation, interfering with the CD40L and CD40 interaction, or other pro-tumor B-T_{FH} interactions, may not be sufficient for an impressive clinical response as they affect one of the many important interactions between FL B and T_{FH} cells, leaving others (e.g., IL-4) unaffected. It is doubtful that this complex interaction would collapse if a single portion is interrupted, as has been seen in other complex network interactions. Removal of a single interaction of T_{FH} and B cells, even if it appears to be essential, is likely insufficient to cause significant damage to a complex network such as the cancer-microenvironment interaction.(97) Therefore, we hypothesize that agents which target and potentially kill T_{FH} cells directly may need to be developed in order to eliminate their pro-tumor effects, and combined with other therapeutic agents for maximal effect. Our T_{NB} signature from this thesis may be utilized to identify a subset of FL patients most likely to benefit from eradication of T_{FH} cells due to their symbiotic interaction in that subset. Similarly, other signatures derived from our CMA methods may help to identify additional patients likely to benefit from microenvironment-directed targeted therapies, including anti-CSF-1R therapies in patients with high levels of macrophage signatures and anti-PD-1 therapies in patients with high levels of the T_{eff}

signature.(42, 98) This hypothesis would need to be formally evaluated in clinical trials, as animal models which accurately depict this complex interaction are currently lacking.

Vita

Jason R. Westin was born in Phoenix, Arizona and grew up in Gainesville, Florida. He graduated from the University of Florida with honors, majoring in Microbiology, minoring in Chemistry, English, and History. He attended the University of Florida College of Medical, where he was elected to Alpha Omega Alpha honor society. He completed his Internal Medicine Residency at the University of North Carolina Chapel Hill, subsequently joining the faculty as a founding member of the Hospitalist Department. He then transitioned to MD Anderson Cancer Center in Houston, Texas as a Hospitalist in the Stem Cell Transplant and General Internal Medicine Departments. He completed his Hematology Oncology Fellowship at MD Anderson Cancer Center in Houston, Texas and served as Chief Fellow, followed by a year as an Advanced Scholar. He joined the faculty of the Department of Lymphoma & Myeloma in 2012 as an Assistant Professor, tenure track. Since joining the faculty, he has received numerous awards including being named to the inaugural class of the R. Lee Clark Fellows and an American Society of Clinical Oncology Career Development Award.

REFERENCES

1. Siegel R, Ward E, Brawley O, Jemal A. Cancer statistics, 2011. CA: A Cancer Journal for Clinicians. 2011;61(4):212-36.
2. Groves FD, Linet MS, Travis LB, Devesa SS. Cancer Surveillance Series: Non-Hodgkin's Lymphoma Incidence by Histologic Subtype in the United States From 1978 Through 1995. Journal of the National Cancer Institute. 2000 August 2, 2000;92(15):1240-51.
3. Cartwright R, Brincker H, Carli PM, Clayden D, Coebergh JW, Jack A, et al. The rise in incidence of lymphomas in Europe 1985-1992. European Journal of Cancer. 1999;35(4):627-33.
4. Swerdlow S, Campo E, Harris N. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. 2008 Lyon. France IARC.358.
5. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature. 2000 Feb 3;403(6769):503-11. PubMed PMID: 10676951.
6. Green MR, Gentles AJ, Nair RV, Irish JM, Kihira S, Liu CL, et al. Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma. Blood. 2013 Feb 28;121(9):1604-11. PubMed PMID: 23297126. Pubmed Central PMCID: 3587323.
7. Kridel R, Sehn LH, Gascoyne RD. Pathogenesis of follicular lymphoma. The Journal of clinical investigation. 2012 Oct 1;122(10):3424-31. PubMed PMID: 23023713. Pubmed Central PMCID: 3461914.
8. Cong P, Raffeld M, Teruya-Feldstein J, Sorbara L, Pittaluga S, Jaffe ES. In situ localization of follicular lymphoma: description and analysis by laser capture microdissection. Blood. 2002 May 1;99(9):3376-82. PubMed PMID: 11964306. Epub 2002/04/20. eng.
9. Roulland S, Navarro J-M, Grenot P, Milili M, Agopian J, Montpellier B, et al. Follicular lymphoma-like B cells in healthy individuals: a novel intermediate step in early lymphomagenesis. The Journal of Experimental Medicine. 2006 October 30, 2006;203(11):2425-31.
10. Limpens J, de Jong D, van Krieken JH, Price CG, Young BD, van Ommen GJ, et al. Bcl-2/JH rearrangements in benign lymphoid tissues with follicular hyperplasia. Oncogene. 1991 Dec;6(12):2271-6. PubMed PMID: 1766674. Epub 1991/12/01. eng.
11. Jaffe ES. The 2008 WHO classification of lymphomas: implications for clinical practice and translational research. Hematology. 2009 January 1, 2009;2009(1):523-31.
12. Staudt LM. A Closer Look at Follicular Lymphoma. New England Journal of Medicine. 2007;356(7):741-2.
13. Schuler F, Dolken L, Hirt C, Kiefer T, Berg T, Fusch G, et al. Prevalence and frequency of circulating t(14;18)-MBR translocation carrying cells in healthy individuals. International journal of cancer Journal international du cancer. 2009 Feb 15;124(4):958-63. PubMed PMID: 19030176.
14. Morin RD, Mendez-Lago M, Mungall AJ, Goya R, Mungall KL, Corbett RD, et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. Nature. 2011 Aug 18;476(7360):298-303. PubMed PMID: 21796119. Pubmed Central PMCID: 3210554.

15. Staudt LM, Dave S. The Biology of Human Lymphoid Malignancies Revealed by Gene Expression Profiling. In: Frederick W. Alt KFATKFMJWU, Emil RU, editors. Advances in Immunology. Volume 87: Academic Press; 2005. p. 163-208.
16. Seymour JF, Pro B, Fuller LM, Manning JT, Hagemeister FB, Romaguera J, et al. Long-term follow-up of a prospective study of combined modality therapy for stage I-II indolent non-Hodgkin's lymphoma. J Clin Oncol. 2003 Jun 1;21(11):2115-22. PubMed PMID: 12775737.
17. A predictive model for aggressive non-Hodgkin's lymphoma. The International Non-Hodgkin's Lymphoma Prognostic Factors Project. The New England journal of medicine. 1993 Sep 30;329(14):987-94. PubMed PMID: 8141877.
18. Federico M, Vitolo U, Zinzani PL, Chisesi T, Clo V, Bellesi G, et al. Prognosis of follicular lymphoma: a predictive model based on a retrospective analysis of 987 cases. Intergruppo Italiano Linfomi. Blood. 2000 Feb 1;95(3):783-9. PubMed PMID: 10648386.
19. Solal-Céligny P, Roy P, Colombat P, White J, Armitage JO, Arranz-Saez R, et al. Follicular Lymphoma International Prognostic Index. Blood. 2004 September 1, 2004;104(5):1258-65.
20. Federico M, Bellei M, Marcheselli L, Luminari S, Lopez-Guillermo A, Vitolo U, et al. Follicular Lymphoma International Prognostic Index 2: A New Prognostic Index for Follicular Lymphoma Developed by the International Follicular Lymphoma Prognostic Factor Project. Journal of Clinical Oncology. 2009 September 20, 2009;27(27):4555-62.
21. Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht L, Horning SJ, et al. Revised Response Criteria for Malignant Lymphoma. Journal of Clinical Oncology. 2007 February 10, 2007;25(5):579-86.
22. Cheson B. Targeted treatment and new agents in follicular lymphoma. International Journal of Hematology. 2010;92(1):5-11.
23. O'BRIEN M, EASTERBROOK P, POWELL J, BLACKLEDGE G, JONES L, MACLENNAN I, et al. The Natural History of Low Grade Non-Hodgkin's Lymphoma and the Impact of a No Initial Treatment Policy on Survival. QJM. 1991 August 1, 1991;80(2):651-60.
24. Horning SJ, Rosenberg SA. The Natural History of Initially Untreated Low-Grade Non-Hodgkin's Lymphomas. New England Journal of Medicine. 1984;311(23):1471-5.
25. Wang CY, Cusack Jr JC, Liu R, Baldwin Jr AS. Control of inducible chemoresistance: Enhanced anti-tumor therapy through increased apoptosis by inhibition of NF- B. Nature Medicine. 1999;5(4):412-7.
26. West KA, Sianna Castillo S, Dennis PA. Activation of the PI3K/Akt pathway and chemotherapeutic resistance. Drug Resistance Updates. 2002;5(6):234-48.
27. Kannan S, Neelapu S. Vaccination strategies in follicular lymphoma. Current Hematologic Malignancy Reports. 2009;4(4):189-95.
28. Ardeshna KM, Smith P, Norton A, Hancock BW, Hoskin PJ, MacLennan KA, et al. Long-term effect of a watch and wait policy versus immediate systemic treatment for asymptomatic advanced-stage non-Hodgkin lymphoma: a randomised controlled trial. The Lancet. 2003;362(9383):516-22.
29. Kloo B, Nagel D, Pfeifer M, Grau M, Düwel M, Vincendeau M, et al. Critical role of PI3K signaling for NF- B–dependent survival in a subset of activated B-cell–like diffuse large B-cell lymphoma cells. PNAS. 2011 January 4, 2011;108(1):272-7.

30. Abramson JS, Chen W, Juszczynski P, Takahashi H, Neuberg D, Kutok JL, et al. The heat shock protein 90 inhibitor IPI-504 induces apoptosis of AKT-dependent diffuse large B-cell lymphomas. BJH. 2009;144(3):358-66.
31. Horning SJ. Natural history of and therapy for the indolent non-Hodgkin's lymphomas. Semin Oncol. 1993 Oct;20(5 Suppl 5):75-88. PubMed PMID: 8211209. Epub 1993/10/01. eng.
32. Thotathil Z, Jameson MB. Early experience with novel immunomodulators for cancer treatment. Expert Opinion on Investigational Drugs. 2007;16(9):1391-403.
33. Maloney DG, Liles TM, Czerwinski DK, Waldichuk C, Rosenberg J, Grillo-Lopez A, et al. Phase I clinical trial using escalating single-dose infusion of chimeric anti-CD20 monoclonal antibody (IDEC-C2B8) in patients with recurrent B-cell lymphoma. Blood. 1994;84(8):2457-66.
34. Maloney DG, Grillo-López AJ, White CA, Bodkin D, Schilder RJ, Neidhart JA, et al. IDEC-C2B8 (Rituximab) Anti-CD20 Monoclonal Antibody Therapy in Patients With Relapsed Low-Grade Non-Hodgkin's Lymphoma. Blood. 1997 September 15, 1997;90(6):2188-95.
35. McLaughlin P, Grillo-Lopez A, Link B, Levy R, Czuczman M, Williams M, et al. Rituximab chimeric anti-CD20 monoclonal antibody therapy for relapsed indolent lymphoma: half of patients respond to a four-dose treatment program. Journal of Clinical Oncology. 1998 August 1, 1998;16(8):2825-33.
36. Zhu YX, Braggio E, Shi C-X, Bruins LA, Schmidt JE, Van Wier S, et al. Cereblon expression is required for the antimyeloma activity of lenalidomide and pomalidomide. Blood. 2011 November 3, 2011;118(18):4771-9.
37. Witzig TE, Wiernik PH, Moore T, Reeder C, Cole C, Justice G, et al. Lenalidomide Oral Monotherapy Produces Durable Responses in Relapsed or Refractory Indolent Non-Hodgkin's Lymphoma. Journal of Clinical Oncology. 2009 November 10, 2009;27(32):5404-9.
38. Leonard J, Jung S, Johnson J, Bartlett N, Blum K, Cheson B. CALGB 50401: A randomized trial of lenalidomide alone versus lenalidomide plus rituximab in patients with recurrent follicular lymphoma. J Clin Oncol 30, 2012 (suppl; abstr 8000).
39. N. H. Fowler PM, F. B. Hagemeister, L. W. Kwak, M. A. Fanale, S. S. Neelapu, L. Fayad, R. Z. Orlowski, M. Wang, F. Samaniego;. Complete response rates with lenalidomide plus rituximab for untreated indolent B-cell non-Hodgkin's lymphoma. J Clin Oncol 28:15s, 2010 (suppl; abstr 8036).
40. Fowler N, Hagemeister FB, Mclaughlin P, Kwak L, Romaguera J, Fanale M, et al. Lenalidomide plus rituximab is a highly effective and well-tolerated biologic therapy in untreated indolent B-cell non hodgkin lymphoma. Ann Oncol (2011) 22(suppl 4): iv128-iv129
41. Fowler NH, Davis RE, Rawal S, Nastoupil L, Hagemeister FB, McLaughlin P, et al. Safety, activity, and immune effects of lenalidomide and rituximab in untreated indolent lymphoma. Lancet Oncol. 2014;In Press.
42. Westin JR, Chu F, Zhang M, Fayad LE, Kwak LW, Fowler N, et al. Safety and activity of PD1 blockade by pidilizumab in combination with rituximab in patients with relapsed follicular lymphoma: a single group, open-label, phase 2 trial. Lancet Oncol. 2014 Jan;15(1):69-77. PubMed PMID: 24332512. Pubmed Central PMCID: 3922714.

43. Davis TA, Grillo-Lopez AJ, White CA, McLaughlin P, Czuczman MS, Link BK, et al. Rituximab anti-CD20 monoclonal antibody therapy in non-Hodgkin's lymphoma: safety and efficacy of re-treatment. J Clin Oncol. 2000 Sep;18(17):3135-43. PubMed PMID: 10963642.
44. Kwak LW, Campbell MJ, Czerwinski DK, Hart S, Miller RA, Levy R. Induction of immune responses in patients with B-cell lymphoma against the surface-immunoglobulin idiotype expressed by their tumors. The New England journal of medicine. 1992 Oct 22;327(17):1209-15. PubMed PMID: 1406793. Epub 1992/10/22. eng.
45. Hsu FJ, Caspar CB, Czerwinski D, Kwak LW, Liles TM, Syrengelas A, et al. Tumor-specific idiotype vaccines in the treatment of patients with B-cell lymphoma--long-term results of a clinical trial. Blood. 1997 May 1;89(9):3129-35. PubMed PMID: 9129015. Epub 1997/05/01. eng.
46. Hsu FJ, Benike C, Fagnoni F, Liles TM, Czerwinski D, Taidi B, et al. Vaccination of patients with B-cell lymphoma using autologous antigen-pulsed dendritic cells. Nat Med. 1996 Jan;2(1):52-8. PubMed PMID: 8564842. Epub 1996/01/01. eng.
47. Bendandi M, Gocke CD, Kobrin CB, Benko FA, Sternas LA, Pennington R, et al. Complete molecular remissions induced by patient-specific vaccination plus granulocyte-monocyte colony-stimulating factor against lymphoma. Nat Med. 1999 Oct;5(10):1171-7. PubMed PMID: 10502821. Epub 1999/09/30. eng.
48. Timmerman JM, Czerwinski DK, Davis TA, Hsu FJ, Benike C, Hao ZM, et al. Idiotype-pulsed dendritic cell vaccination for B-cell lymphoma: clinical and immune responses in 35 patients. Blood. 2002 Mar 1;99(5):1517-26. PubMed PMID: 11861263. Epub 2002/02/28. eng.
49. Neelapu SS, Kwak LW, Kobrin CB, Reynolds CW, Janik JE, Dunleavy K, et al. Vaccine-induced tumor-specific immunity despite severe B-cell depletion in mantle cell lymphoma. Nat Med. 2005 Sep;11(9):986-91. PubMed PMID: 16116429. Epub 2005/08/24. eng.
50. Neelapu SS, Baskar S, Gause BL, Kobrin CB, Watson TM, Frye AR, et al. Human autologous tumor-specific T-cell responses induced by liposomal delivery of a lymphoma antigen. Clin Cancer Res. 2004 Dec 15;10(24):8309-17. PubMed PMID: 15623607. Epub 2004/12/30. eng.
51. Redfern CH, Guthrie TH, Bessudo A, Densmore JJ, Holman PR, Janakiraman N, et al. Phase II trial of idiotype vaccination in previously treated patients with indolent non-Hodgkin's lymphoma resulting in durable clinical responses. J Clin Oncol. 2006 Jul 1;24(19):3107-12. PubMed PMID: 16754937. Epub 2006/06/07. eng.
52. Schuster SJ, Neelapu SS, Santos CF, Popa-McKiver MA, McCord AM, Kwak LW. Idiotype vaccination as consolidation therapy: time for integration into standard of care for follicular lymphoma? J Clin Oncol. 2011 Dec 20;29(36):4845-6. PubMed PMID: 22042953. Pubmed Central PMCID: 3255991.
53. Liu Q, Fayad L, Cabanillas F, Hagemeister FB, Ayers GD, Hess M, et al. Improvement of Overall and Failure-Free Survival in Stage IV Follicular Lymphoma: 25 Years of Treatment Experience at The University of Texas M.D. Anderson Cancer Center. Journal of Clinical Oncology. 2006 April 1, 2006;24(10):1582-9.
54. Czuczman MS, Weaver R, Alkuzweny B, Berlfein J, Grillo-López AJ. Prolonged clinical and molecular remission in patients with low-grade or follicular non-Hodgkin's

lymphoma treated with rituximab plus CHOP chemotherapy: 9-Year follow-up. Journal of Clinical Oncology. 2004;22(23):4659-64.
55. Hiddemann W, Kneba M, Dreyling M, Schmitz N, Lengfelder E, Schmits R, et al. Frontline therapy with rituximab added to the combination of cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP) significantly improves the outcome for patients with advanced-stage follicular lymphoma compared with therapy with CHOP alone: Results of a prospective randomized study of the German Low-Grade Lymphoma Study Group. Blood. 2005;106(12):3725-32.
56. Dunn GP, Old LJ, Schreiber RD. The immunobiology of cancer immunosurveillance and immunoediting. Immunity. 2004 Aug;21(2):137-48. PubMed PMID: 15308095.
57. Afshar-Sterle S, Zotos D, Bernard NJ, Scherger AK, Rodling L, Alsop AE, et al. Fas ligand-mediated immune surveillance by T cells is essential for the control of spontaneous B cell lymphomas. Nat Med. 2014 Feb 2. PubMed PMID: 24487434.
58. Street SE, Trapani JA, MacGregor D, Smyth MJ. Suppression of lymphoma and epithelial malignancies effected by interferon gamma. J Exp Med. 2002 Jul 1;196(1):129-34. PubMed PMID: 12093877. Pubmed Central PMCID: 2194011.
59. Shankaran V, Ikeda H, Bruce AT, White JM, Swanson PE, Old LJ, et al. IFNgamma and lymphocytes prevent primary tumour development and shape tumour immunogenicity. Nature. 2001 Apr 26;410(6832):1107-11. PubMed PMID: 11323675.
60. Dunn GP, Bruce AT, Sheehan KC, Shankaran V, Uppaluri R, Bui JD, et al. A critical function for type I interferons in cancer immunoediting. Nature immunology. 2005 Jul;6(7):722-9. PubMed PMID: 15951814.
61. Mellman I, Coukos G, Dranoff G. Cancer immunotherapy comes of age. Nature. 2011 Dec 22;480(7378):480-9. PubMed PMID: 22193102.
62. Kim R, Emi M, Tanabe K. Cancer immunoediting from immune surveillance to immune escape. Immunology. 2007 May;121(1):1-14. PubMed PMID: 17386080. Pubmed Central PMCID: 2265921.
63. Dave SS, Wright G, Tan B, Rosenwald A, Gascoyne RD, Chan WC, et al. Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells. The New England journal of medicine. 2004 Nov 18;351(21):2159-69. PubMed PMID: 15548776.
64. Byers RJ, Sakhinia E, Joseph P, Glennie C, Hoyland JA, Menasce LP, et al. Clinical quantitation of immune signature in follicular lymphoma by RT-PCR-based gene expression profiling. Blood. 2008 May 1;111(9):4764-70. PubMed PMID: 18174380. Epub 2008/01/05. eng.
65. Bohen SP, Troyanskaya OG, Alter O, Warnke R, Botstein D, Brown PO, et al. Variation in gene expression patterns in follicular lymphoma and the response to rituximab. Proc Natl Acad Sci U S A. 2003 Feb 18;100(4):1926-30. PubMed PMID: 12571354. Epub 2003/02/07. eng.
66. Wahlin BE, Sander B, Christensson B, Kimby E. CD8+ T-cell content in diagnostic lymph nodes measured by flow cytometry is a predictor of survival in follicular lymphoma. Clin Cancer Res. 2007 Jan 15;13(2 Pt 1):388-97. PubMed PMID: 17255259.
67. Alvaro T, Lejeune M, Salvado MT, Lopez C, Jaen J, Bosch R, et al. Immunohistochemical patterns of reactive microenvironment are associated with

clinicobiologic behavior in follicular lymphoma patients. J Clin Oncol. 2006 Dec 1;24(34):5350-7. PubMed PMID: 17135637.
68. Lee ST, Liu S, Radvanyi L, Sukhumalchandra P, Molldrem JJ, Wieder ED, et al. A novel strategy for rapid and efficient isolation of human tumor-specific CD4(+) and CD8(+) T-cell clones. J Immunol Methods. 2008 Feb 29;331(1-2):13-26. PubMed PMID: 17959194. Pubmed Central PMCID: 2265521.
69. Schultze JL, Seamon MJ, Michalak S, Gribben JG, Nadler LM. Autologous tumor infiltrating T cells cytotoxic for follicular lymphoma cells can be expanded in vitro. Blood. 1997 May 15;89(10):3806-16. PubMed PMID: 9160688.
70. Chen L, Flies DB. Molecular mechanisms of T cell co-stimulation and co-inhibition. Nature reviews Immunology. 2013 Apr;13(4):227-42. PubMed PMID: 23470321. Pubmed Central PMCID: 3786574.
71. Zou W. Immunosuppressive networks in the tumour environment and their therapeutic relevance. Nature reviews. 2005 Apr;5(4):263-74. PubMed PMID: 15776005.
72. Yang ZZ, Novak AJ, Stenson MJ, Witzig TE, Ansell SM. Intratumoral CD4+CD25+ regulatory T-cell-mediated suppression of infiltrating CD4+ T cells in B-cell non-Hodgkin lymphoma. Blood. 2006 May 1;107(9):3639-46. PubMed PMID: 16403912. Pubmed Central PMCID: 1895773.
73. Myklebust JH, Irish JM, Brody J, Czerwinski DK, Houot R, Kohrt HE, et al. High PD-1 expression and suppressed cytokine signaling distinguish T cells infiltrating follicular lymphoma tumors from peripheral T cells. Blood. 2013 Feb 21;121(8):1367-76. PubMed PMID: 23297127. Pubmed Central PMCID: 3578953.
74. Mlecnik B, Tosolini M, Charoentong P, Kirilovsky A, Bindea G, Berger A, et al. Biomolecular network reconstruction identifies T-cell homing factors associated with survival in colorectal cancer. Gastroenterology. 2010 Apr;138(4):1429-40. PubMed PMID: 19909745.
75. Schuster SJ, Neelapu SS, Gause BL, Janik JE, Muggia FM, Gockerman JP, et al. Vaccination with patient-specific tumor-derived antigen in first remission improves disease-free survival in follicular lymphoma. J Clin Oncol. 2011 Jul 10;29(20):2787-94. PubMed PMID: 21632504. Pubmed Central PMCID: 3139394.
76. Zent CS, LaPlant BR, Johnston PB, Call TG, Habermann TM, Micallef IN, et al. The treatment of recurrent/refractory chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL) with everolimus results in clinical responses and mobilization of CLL cells into the circulation. Cancer. 2010 May 1;116(9):2201-7. PubMed PMID: 20166206. Pubmed Central PMCID: 2861142. Epub 2010/02/19. eng.
77. Ding LH, Xie Y, Park S, Xiao G, Story MD. Enhanced identification and biological validation of differential gene expression via Illumina whole-genome expression arrays through the use of the model-based background correction methodology. Nucleic Acids Res. 2008 Jun;36(10):e58. PubMed PMID: 18450815. Epub 2008/05/03. eng.
78. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A. 1998 Dec 8;95(25):14863-8. PubMed PMID: 9843981. Epub 1998/12/09. eng.
79. Smyth GK. Limma: linear models for microarray data. In: Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W, editors. Bioinformatics and Computational Biology Solutions using R and Bioconductor. New York: Springer; 2005. p. 397-420.

80. Ma W, Yang D, Gu Y, Guo X, Zhao W, Guo Z. Finding disease-specific coordinated functions by multi-function genes: insight into the coordination mechanisms in diseases. Genomics. 2009 Aug;94(2):94-100. PubMed PMID: 19427897.
81. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005 Oct 25;102(43):15545-50. PubMed PMID: 16199517. Pubmed Central PMCID: 1239896. Epub 2005/10/04. eng.
82. Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pages C, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. Science. 2006 Sep 29;313(5795):1960-4. PubMed PMID: 17008531. Epub 2006/09/30. eng.
83. Lenz G, Wright G, Dave SS, Xiao W, Powell J, Zhao H, et al. Stromal gene signatures in large-B-cell lymphomas. The New England journal of medicine. 2008 Nov 27;359(22):2313-23. PubMed PMID: 19038878. Epub 2008/11/29. eng.
84. Andreasson U, Dictor M, Jerkeman M, Berglund M, Sundstrom C, Linderoth J, et al. Identification of molecular targets associated with transformed diffuse large B cell lymphoma using highly purified tumor cells. Am J Hematol. 2009 Dec;84(12):803-8. PubMed PMID: 19844990. Epub 2009/10/22. eng.
85. Montes-Moreno S, Roncador G, Maestre L, Martinez N, Sanchez-Verde L, Camacho FI, et al. Gcet1 (centerin), a highly restricted marker for a subset of germinal center-derived lymphomas. Blood. 2008 Jan 1;111(1):351-8. PubMed PMID: 17898315.
86. Paterson MA, Hosking PS, Coughlin PB. Expression of the serpin centerin defines a germinal center phenotype in B-cell lymphomas. American journal of clinical pathology. 2008 Jul;130(1):117-26. PubMed PMID: 18550480. Epub 2008/06/14. eng.
87. Rawal S, Chu F, Zhang M, Park HJ, Nattamai D, Kannan S, et al. Cross talk between follicular Th cells and tumor cells in human follicular lymphoma promotes immune evasion in the tumor microenvironment. J Immunol. 2013 Jun 15;190(12):6681-93. PubMed PMID: 23686488. Pubmed Central PMCID: 3680117. Epub 2013/05/21. eng.
88. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. PNAS. 2005 Oct 25;102(43):15545-50. PubMed PMID: 16199517. Pubmed Central PMCID: 1239896. Epub 2005/10/04. eng.
89. Longo NS, Lugar PL, Yavuz S, Zhang W, Krijger PH, Russ DE, et al. Analysis of somatic hypermutation in X-linked hyper-IgM syndrome shows specific deficiencies in mutational targeting. Blood. 2009 Apr 16;113(16):3706-15. PubMed PMID: 19023113. Pubmed Central PMCID: 2670789. Epub 2008/11/22. eng.
90. Lindvall JM, Blomberg KE, Berglof A, Yang Q, Smith CI, Islam TC. Gene expression profile of B cells from Xid mice and Btk knockout mice. European journal of immunology. 2004 Jul;34(7):1981-91. PubMed PMID: 15214046.
91. Hilchey SP, Rosenberg AF, Hyrien O, Secor-Socha S, Cochran MR, Brady MT, et al. Follicular lymphoma tumor-infiltrating T-helper (T(H)) cells have the same polyfunctional potential as normal nodal T(H) cells despite skewed differentiation. Blood. 2011 Sep 29;118(13):3591-602. PubMed PMID: 21821704. Pubmed Central PMCID: 3186335. Epub 2011/08/09. eng.

92. Pangault C, Ame-Thomas P, Ruminy P, Rossille D, Caron G, Baia M, et al. Follicular lymphoma cell niche: identification of a preeminent IL-4-dependent T(FH)-B cell axis. Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, UK. 2010 Dec;24(12):2080-9. PubMed PMID: 20944673. Pubmed Central PMCID: 3317889. Epub 2010/10/15. eng.
93. Ame-Thomas P, Le Priol J, Yssel H, Caron G, Pangault C, Jean R, et al. Characterization of intratumoral follicular helper T cells in follicular lymphoma: role in the survival of malignant B cells. Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, UK. 2012 May;26(5):1053-63. PubMed PMID: 22015774. Pubmed Central PMCID: 3428269.
94. Kiaii S, Clear AJ, Ramsay AG, Davies D, Sangaralingam A, Lee A, et al. Follicular lymphoma cells induce changes in T-cell gene expression and function: potential impact on survival and risk of transformation. J Clin Oncol. 2013 Jul 20;31(21):2654-61. PubMed PMID: 23775959. Pubmed Central PMCID: 3709054. Epub 2013/06/19. eng.
95. Townsend WM, Marcus R, Salisbury J, Yallop D, Patten PE, Buggins AG, et al. Multiparameter Microscopy Analysis of the Follicular Lymphoma Microenvironment and Normal Germinal Centers: In Vivo evidence That Follicular Helper T Cells Form Synapses with Neoplastic B Cells and Are Associated with Proliferation and Expression of Acti…. Blood. 2014 2014-12-06 00:00:00;124(21):144-.
96. Fanale M, Assouline S, Kuruvilla J, Solal-Celigny P, Heo DS, Verhoef G, et al. Phase IA/II, multicentre, open-label study of the CD40 antagonistic monoclonal antibody lucatumumab in adult patients with advanced non-Hodgkin or Hodgkin lymphoma. British journal of haematology. 2014 Jan;164(2):258-65. PubMed PMID: 24219359.
97. Westin JR. Busting robustness: using cancer's greatest strength to our advantage. Future oncology. 2015 Jan;11(1):73-7. PubMed PMID: 24673642.
98. Gopisetty A, Foglietta M, Zhang M, Wang Z, Fowler N, Davis RE, et al. Prosurvival and Chemoprotective Effects of Tumor-Associated Macrophages Reversed By Anti-CSF-1R Monoclonal Antibody in B-Cell Lymphomas. Blood. 2014;124(21):498-.