

AD\_\_\_\_\_

Award Number: W81XWH-06-1-0417

TITLE: Immunology, Systems Biology, and Immunotherapy of Breast Cancer

PRINCIPAL INVESTIGATOR: Peter P. Lee, M.D.

CONTRACTING ORGANIZATION: Stanford University  
Stanford, CA 94305

REPORT DATE: March 2009

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command  
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT:

Approved for public release; distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

# REPORT DOCUMENTATION PAGE

*Form Approved*  
*OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

<b>1. REPORT DATE (DD-MM-YYYY)</b> 01-03-2009		<b>2. REPORT TYPE</b> Annual Progress Report		<b>3. DATES COVERED (From - To)</b> 01 Mar 08 - 28 Feb 09	
<b>4. TITLE AND SUBTITLE</b> Immunology, Systems Biology, and Immunotherapy of Breast Cancer				<b>5a. CONTRACT NUMBER</b>	
				<b>5b. GRANT NUMBER</b> W81XWH-06-1-0417	
				<b>5c. PROGRAM ELEMENT NUMBER</b>	
<b>6. AUTHOR(S)</b> Peter P. Lee, M.D.				<b>5d. PROJECT NUMBER</b>	
				<b>5e. TASK NUMBER</b>	
				<b>5f. WORK UNIT NUMBER</b>	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b>  Stanford University Stanford, CA 94305				<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>				<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b>	
				<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b>	
<b>12. DISTRIBUTION / AVAILABILITY STATEMENT</b> Approved for public release; distribution unlimited					
<b>13. SUPPLEMENTARY NOTES</b>					
<b>14. ABSTRACT</b>  In year 3, we built upon the foundation from the first two years of this award, and are making progress in multiple areas of this project. We now have an efficient system in place to recruit patients into this study and procure their samples. Limited numbers of subjects available and limited amounts of clinical materials available from each subject remain major challenges to the success of this project – we continually attempt to address and solve this issue by reducing the cell numbers that we need for each assay. We have developed a powerful set of immunological assays and molecular tools to study these samples in greater detail than previously possible. We continue to uncover dramatic changes in the immune cell populations within tumors, TDLNs, and peripheral blood from breast cancer patients. These findings are reported above, and have led to 3 manuscripts under review. We look forward in the coming year to build upon this momentum and specifically to elucidate insights into the immunobiology of breast cancer. In the coming year and beyond, we will begin to focus on translating our early findings into novel therapeutic strategies for the immunotherapy of breast cancer.					
<b>15. SUBJECT TERMS</b> Breast cancer, immunology, immunotherapy, systems biology					
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>  UU	<b>18. NUMBER OF PAGES</b>  49	<b>19a. NAME OF RESPONSIBLE PERSON</b> USAMRMC
<b>a. REPORT</b> U	<b>b. ABSTRACT</b> U	<b>c. THIS PAGE</b> U			<b>19b. TELEPHONE NUMBER (include area code)</b>

## Table of Contents

	<b><u>Page</u></b>
Introduction .....	4
BODY .....	4 - 47
Key Research Accomplishments .....	48
Reportable Outcomes .....	48
Conclusion .....	49
References .....	49
Appendices .....	49

**Annual Progress Report 3/1/08-2/28/09****DoD Era of Hope Scholar Award**Immunology, Systems Biology, and Immunotherapy of Breast Cancer

Peter P. Lee, M.D.

Stanford University

**INTRODUCTION**

Breast cancer patients with similar tumor characteristics may have vastly different clinical courses, response to therapy, and outcome. Several lines of evidence now suggest that the host immune response may play a significant role in modulating disease progression in cancer. A complex interplay exists between the host immune response and tumor cells as a critical determinant in clinical outcome. These factors remain poorly understood. By comprehensively studying the dynamics between breast cancer and the immune response using an integrative systems approach, we hope to uncover opportunities for vastly different immunotherapy approaches than what are available today. We seek to move beyond the current paradigm of eliciting immune responses against defined antigens via vaccination, as this strategy alone does not appear to be effective in a number of clinical trials for melanoma and other cancers. Rather, we seek strategies that specifically modulate tumor-immune cell interactions and block cancer-induced immune dysfunction on a systemic and local level (at tumor sites and draining lymph nodes). In this project, we use a number of novel immunological approaches to look for evidence of immune cell dysfunction within the tumor or tumor-draining lymph nodes (TDLNs) from breast cancer patients. This includes archived samples from patients with at least five year survival data, and fresh samples from newly diagnosed patients. We use DNA microarrays to analyze the gene expression patterns of purified tumor and immune cells, focusing on gene networks and cross-talk between tumor and immune cells. We generate high-resolution images of tumor and TDLN sections and develop image analysis algorithms to assess the spatial arrangement and grouping of tumor and immune cells with respect to each other that may have biological significance. Using statistics and mathematical tools, we will integrate the complex data generated from all of these studies and correlate them with clinical parameters. Lastly, our observations will be combined into a mathematical model that will enable us to perform *in silico* experiments to quickly test novel therapeutic strategies for breast cancer. This work may lead to novel diagnostic tools to help predict clinical outcome and guide therapy in breast cancer patients. We also hope to find new insights into the mechanisms of immune evasion by breast cancer cells and ultimately new treatment strategies for breast cancer directed specifically at altering the biology of TDLNs.

**BODY**

Our team currently consists of two excellent postdoctoral fellows, one research assistant, and several faculty collaborators (only one of which draws a modest amount of salary support from this award). We work closely with our surgery, medical oncology, and pathology colleagues to identify, recruit, and consent subjects, and to obtain samples from the operating room to pathology and eventually to my laboratory. In addition, we continue to refine our protocols to maximize recovery of immune cells from tumor and lymph node specimens, and to optimize methods for analysis of fresh and archive samples by flow cytometry, immunohistology, immunofluorescence, function assays, and DNA microarray analysis using the smallest numbers

of cells possible. Below is a summary of our progress in year 3 of our EHSA in relation to my proposed SOW.

## Experiment Strategy

To fully understand tumor-immune cell interactions in breast cancer, our strategy is to compare the immune cells and tumor cells within three distinct compartments: the tumor, TDLNs, and blood. We approach this at both the molecular and cellular levels. At the molecular level, gene expression profiling of immune cells and tumor cells within the tumor site and TDLNs are being carried out. At the cellular level, immunologic functions of immune cells are being studied and compared across these three compartments.

### A. Immunological Analyses

Originally proposed in the SOW:

1. Analysis of archived samples of tumor and TDLN from breast cancer patients with at least 5 years of clinical follow-up data. Tumor and immune cell markers will be identified via immunohistochemical (IHC) staining and in-situ hybridization (ISH). Images will be acquired in high resolution using an automated imaging system (BLISS), and data will be acquired using automated software. Over 50 immune and tumor markers will be assessed. To facilitate these complex studies, we will also explore the use of tissue microarrays (TMA). This would enable us to analyze sections from 100-400 samples on each slide. We will first perform a pilot study to ensure that the TMA method is compatible for our studies and would not be negatively impacted by the architectural heterogeneity within TDLN. (months 0-60)
2. Analysis of live cells from fresh tumor, TDLN, blood, and possibly bone marrow from newly diagnosed or relapsed breast cancer patients undergoing surgery or treatment. Assays include flow cytometry (up to 12 colors), peptide-MHC tetramer analysis, sorting, functional responses (e.g. cytotoxicity, cytokine release, anergy, apoptosis, proliferation), and others. (months 6-60)
3. Generation of T cell lines and tumor cell lines from fresh tumor and TDLN samples for further detailed analyses. (months 6-60)
4. If the above studies demonstrate immune cell dysfunction within tumor or TDLN, but by themselves do not reveal any definitive mechanisms, then we will undertake in vivo studies with mouse models of de novo breast cancer to address the early events in immune dysfunction. (months 24-60)

### **Sample Acquisition**

At the end of year 3, over a 170 breast cancer patients have been enrolled into this study. All participants were newly diagnosed, had recurrent or metastatic disease and had their surgical and/or oncological treatments at Stanford University Medical Center. Written informed consents were obtained from all participants according to Stanford IRB, DoD HSRRB, and HIPAA regulations. Patients' heparinized peripheral blood samples, breast tumor tissue, TDLNs (non-sentinel lymph node and/or sentinel lymph node), and tumor/TDLN aspirates have been collected for this study. The samples acquired thus far are shown in Table 1. Clinical data (stage,

grade, ER/PR/Her-2/neu status, treatment, and clinical outcome) for each participant has been recorded.

Tissue	Storage	Quantity
Blood	Trizol	52
	Cryopreserved	117
SLN	Trizol	98
	Cryopreserved	126
Non-SLN	Trizol	82
	Cryopreserved	101
Breast Tumor	Trizol	39
	Cryopreserved	74

### **Optimized Sample Processing Procedures**

#### Peripheral blood immune cell isolation

Peripheral blood mononuclear cells (PBMCs) from peripheral blood are separated by Ficoll-Hypaque density gradient centrifugation. For RNA isolation and microarray analysis, RBC lysis buffer treatment is used to remove residual red blood cells and to ensure an accurate counting using hemocytometer. A total of 1 million isolated immune cells are preserved in Trizol for RNA isolation and microarray analysis while the remaining cells are cryopreserved in liquid nitrogen until further use. Blood samples not intended for microarray analysis are spun down in a centrifuge first to collect plasma for Luminex multi-plex bead assays prior to PBMC collection.

#### Breast tissue dissociation and immune cell/tumor cell isolation

After surgery, breast tumor tissues are minced and dissociated enzymatically with type III Collagenase and DNase I for 1-2 hour to generate single cell suspensions. For microarray analysis, the cells are stained with pan leukocyte marker CD45, epithelial surface antigen ESA, fibroblast marker CD140- $\beta$ , and a dead cell exclusion marker ViViD and purified via FACS sorting. The two populations of interest in breasts tumor tissue are immune cells (CD45<sup>+</sup>ESA<sup>-</sup>CD140 $\beta$ <sup>-</sup>) and epithelial/tumor cells (ESA<sup>+</sup>CD45<sup>-</sup>CD140 $\beta$ <sup>-</sup>). Up to 1 million sorted immune cells or tumor cells are preserved in Trizol for RNA isolation and microarray analysis and the remaining cells are cryopreserved in liquid nitrogen until further use. Samples not intended for microarray analysis are digested as described above and the heterogeneous single cell suspensions are cryopreserved. To ensure the breast tissue specimen does indeed contain tumor cells, the tumor tissue we receive is bisected and portion is submitted for histological processing using hematoxylin and eosin staining, and examined by a pathologist specializing in cytology.

#### TDLNs

After lymph node dissection, fine-needle aspirates of sentinel lymph nodes (SLN) are collected and are stained as above for RNA isolation and microarray analysis or are cryopreserved for future use. For non-sentinel lymph nodes (NSLN), a very small portion of identified lymph

nodes are excised for research purposes, differentially inked and then submitted to pathology for diagnostic purposes. Once processed by pathology, we then utilize the paraffin-embedded lymph nodes, both sentinel and non-sentinel, for further Immunohistochemical (IHC) analysis. Fresh tissues received are minced to generate single cell suspensions. For grossly tumor involved lymph nodes, the minced specimen is subjected to enzymatic dissociation as described above. For RNA and microarray analysis, immune cells and/or tumor cells are purified through FACS sorting where up to 1 million isolated immune cells or tumor cells are preserved in Trizol for RNA isolation. The remaining cells are cryopreserved in liquid nitrogen until further use. Samples not intended for microarray analysis are minced and/or enzymatically digested (if grossly tumor positive) and the single cell suspensions are cryopreserved for later use.

### **Interferon Signaling Defect in Lymphocytes from Breast Cancer, Melanoma, and Gastrointestinal Patients**

Recently, we demonstrated Interferon (IFN) signaling defects in melanoma patient lymphocytes as measured by microarray, Q-PCR, and Phosflow analysis (Critchley-Thorne, R., *et al*, 2007). To determine whether this may be a common occurrence in cancer, we assessed the functional responses of lymphocytes from breast cancer, gastrointestinal (GI), and melanoma patients, compared to age-matched healthy controls, by measuring IFN-stimulated genes (ISGs), Phosflow analysis (detection of STAT1-pY701) upon stimulation with IFN- $\alpha$  and IFN- $\gamma$ , downstream functional responses in breast cancer T cells to IFN- $\alpha$ , and additional protein phosphorylation in melanoma lymphocytes via JAK/STAT and MAPK pathways.

#### IFN-stimulated genes

Basal expression levels of ISGs were assessed in breast cancer lymphocytes. Lymphocytes from 5-healthy and 7-breast cancer PBMC samples were enriched (~90% purity) using granulocyte and monocyte depletion cocktails. RNA was isolated, quantified, and cDNA was synthesized. Quantitative RT-PCR analysis was performed for five ISGs: *STAT1*, *IFI44*, *IFIT1*, *IFIT2*, and *MX1*. All gene expression data presented were normalized to GAPDH levels for each sample. We showed the expression of basal levels of these five ISGs were downregulated in peripheral blood lymphocytes from breast cancer patients versus healthy controls indicating a defect in IFN signaling in lymphocytes in vivo from breast cancer patients (Figure A1)

#### Phosflow analysis (detection of STAT1-pY701) upon stimulation with IFN- $\alpha$ and IFN- $\gamma$

PBMCs from 12 new melanoma, 27 breast cancer, 11 GI patients and 27 age-matched healthy controls were stimulated with either 1000U/ml IFN- $\alpha$  or 1000U/ml IFN- $\gamma$  or left unstimulated for 15 minutes. The samples were then fixed and permeabilized and acquired on the FACS Aria. Data was analyzed using two-sided Wilcoxon-Mann-Whitney test (95% CI) by statistical application, R, to calculate exact p-values for each of the comparisons. P-values of < 0.05 were considered significant.

Healthy controls composed of both males and females, so we first determined whether samples from different genders were statistically different from each other before we compared the complete healthy population to breast cancer patients (all female). We found that fold induction of STAT1-pY701 (pSTAT1) was not statistically different between male and female healthy controls, or male and females from either the melanoma and GI groups for both IFN- $\alpha$  and IFN- $\gamma$

stimulation (data not shown), and therefore, we included all healthy patients in this analysis. The mean age for breast cancer, melanoma, GI patients and healthy controls were 52.4, 51.6, 66.0, and 54.0, respectively. Since the mean age for the GI group was higher than the healthy, breast cancer, and melanoma groups, we compared the GI group to an older subset of the healthy population (healthy: n=17, mean age 61.8; GI: n=11, mean age 66.0).

Fold induction of pSTAT1 was significantly reduced in breast cancer, melanoma, and GI patients compared to healthy controls in response to IFN- $\alpha$  stimulation in CD3+ T cells, CD19+ B+ cells, CD16+ NK cells, (Figure A2a-c) and CD4+CD45RO+CD25<sup>HI</sup> regulatory T cells (Treg) (Figure A2g), regardless of neo-adjuvant/adjuvant therapy (data not shown). In contrast, the fold induction of pSTAT1 in response to IFN- $\gamma$  stimulation in T cells, NK cells, and Treg cells was not significant. However, IFN- $\gamma$  stimulation of B cells showed a profound reduction of pSTAT1 (Figure A2e). Likewise, the reduced fold change in pSTAT1 in response to IFN- $\alpha$  stimulation was observed in breast cancer stages II, III and IV for T cells, B cells, NK cells (Figure A3a-c) and B cells upon IFN- $\gamma$  stimulation (Figure A3e) illustrating this immune defect occurs in early stage breast cancer.

#### Downstream Functional Responses in Breast Cancer T cells to IFN- $\alpha$

Downstream functional responses to IFN- $\alpha$  were assessed in breast cancer patients' lymphocytes. Lymphocytes were enriched using a monocyte and granulocyte depletion cocktail to a purity >95%. Lymphocytes were stimulated with beads coated with anti-CD3/anti-CD28 antibodies alone or in the presence of 1000U/ml IFN- $\alpha$  or IFN- $\gamma$ , or left unstimulated for 48 hours. Expression of CD25 as a general activation marker, HLA-DR, CD95 and CD54 as activation markers that are further induced by IFNs, and activation-induced cell death (AICD) were measured by flow cytometry. T cells from breast cancer patients showed reduced expression of CD25, HLA-DR, CD54 and CD95 in response to anti-CD3/CD28 stimulation alone and in combination with IFN- $\alpha$  or IFN- $\gamma$  (Figure A4a-d). AICD due to anti-CD3/CD28 stimulation was higher in T cells from breast cancer patients versus healthy controls, while stimulation with anti-CD3/CD28 plus IFN- $\alpha$  or IFN- $\gamma$  resulted in less apoptosis in breast cancer patient T cells (Figure A4e).

#### Additional Protein Phosphorylation via JAK/STAT and MAPK Pathways

Cytokine receptors signal through two main types of pathways: JAK-STAT and MAPK (Murray, PJ, *et al*, 2007). To determine whether similar players in JAK/STAT signaling cascades of other pathways are also perturbed, we utilized Phosflow and stimulated PBMCs from melanoma patients and healthy controls with IL-6 and IL-2 to measure the levels of pSTAT3 and pSTAT5, respectively. The target cells for IL-6 stimulation are T cells and B cells, whereas, the target cells for IL-2 stimulation include T cells, B cells, and NK cells. To investigate the MAPK pathway, which does not signal via JAK/STAT proteins, we stimulated PBMCs with phorbol 12-myristate 13-acetate (PMA) and measured the levels of pERK1/2.

Our results showed that IL-2 responsiveness decreased in CD4+ T cells of melanoma patients, but not in CD8+ T cells and NK cells (Figure A5). B cells were irresponsive to IL-2 stimulation, as demonstrated by no increase in STAT5 phosphorylation. In contrast, IL-6 responsiveness is higher in CD4+ and CD8+ T cells of melanoma patients than in healthy controls, while B and NK cells were irresponsive to IL-6 stimulation (Figure A6). All PBMC lymphocyte subsets

showed an increase in ERK1/2 phosphorylation upon PMA stimulation (Figure A7). We observed an increase trend in the fold change of pERK1/2 in lymphocyte subsets from melanoma patients compared to healthy controls, particularly in B and NK cells; however, the p-values were not significant based on the current sample size.

### **Analysis of STAT1-pY701 by Firefly 3000**

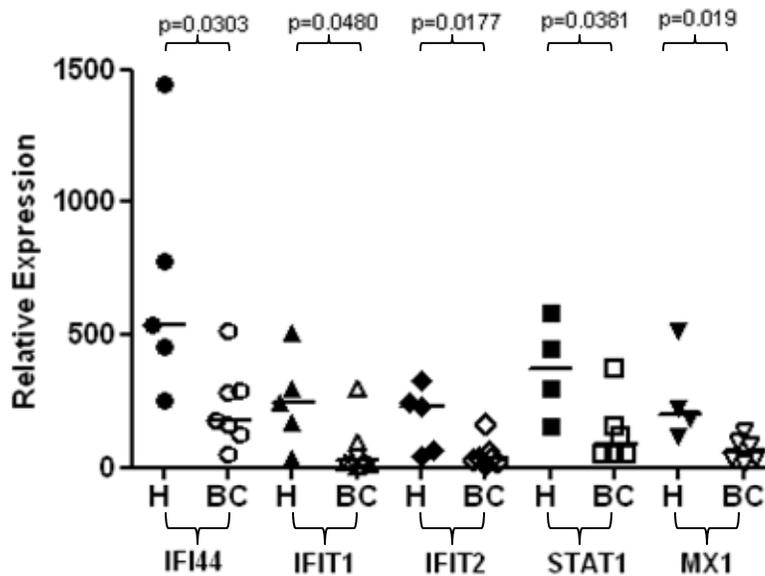
We are currently optimizing protocols to use with a new system, Firefly 3000 (Cell Biosciences). Firefly 3000 is a capillary-based nanofluidic platform with which samples are separated by isoelectric focusing, immobilized, and probed with antibodies. Thus, charge-based changes such as phosphorylation states can be resolved and percentages of protein in different states can be determined. Furthermore, because of the small scale sensitivity of this assay, we will be able to analyze samples with limited sizes such as immune and tumor cells isolated from breast cancer tumors, tumor-draining lymph nodes (TDLNs), and/or peripheral blood leukocytes. At this point, we have successfully tested the ability of the system to detect total STAT1 and STAT1-pY701. The isoelectric points of STAT1-alpha and STAT1-beta were estimated at 5.7 and 6.0, respectively, by the UCSC Proteome Browser (<http://genome.ucsc.edu>). Figure A8 shows the ability of this assay to detect total STAT1 and p-Y701-STAT1 in IFN $\alpha$ -stimulated peripheral blood lymphocytes. Two peaks were identified for each of the possible unphosphorylated and p-Y701-STAT1 states, and may represent the STAT1-alpha and STAT1-beta isoforms.

### **Summary of major findings and plans:**

- We have demonstrated a defect of IFN signaling in breast cancer patients' peripheral blood leukocytes, regardless of therapy. To determine the extent of the IFN defect, we will continue to optimize the use of Firefly 3000 and Phosflow to examine pSTAT1 induction in leukocytes isolated from breast cancer patients' non-sentinel lymph nodes (NSLN) and tumor infiltrating leukocytes (TILs).
- We will determine the molecular basis of IFN signaling perturbations in immune cells from breast cancer patients versus healthy controls.
  - To determine if IFN signaling is perturbed due to altered expression of signaling molecules in the IFN pathway, we will measure basal and IFN-induced levels of proteins in the IFNAR pathway, upstream activators of IFN expression, and negative regulators in peripheral blood B cells, T cells, and NK cells. Since IFN signaling has different effects on immune cell types of different lineages and the IFN response has not yet been analyzed in myeloid lineage cells, we will also assay peripheral blood macrophages, dendritic cells, neutrophils, and other myeloid populations to determine the extent of the IFN signaling defect. RNA collected from unstimulated or IFN-alpha-stimulated peripheral blood populations (FACS or magnetically separated lymphoid and myeloid cells) will be analyzed for gene expression using Q-PCR. Protein expression will be examined by flow-cytometry or Firefly 3000 on FACS sorted populations.
  - To determine if IFN signaling is perturbed due to altered signaling, we will examine IFN-induced phosphorylation of JAK/Tyk kinases and STAT1/2 substrates. To do this, we will continue to optimize the use of Firefly 3000 and Phosflow to examine basal and IFN-alpha induced phosphorylation of JAK/Tyk kinases and STAT1/2 substrates.

- We will determine whether suppression of the IFN pathway in breast cancer patients is a consequence of extracellular signals such as reduced IFN levels or altered cytokines that may positively or negatively cross-talk with the IFN pathway. Cytokine multiplex assays (Luminex) will be used to measure the concentrations of 37 cytokines and chemokines (including IFN-alpha), in serum from healthy vs. breast cancer patients. Cytokines showing altered expression levels that correlate with IFN pathway alterations will be selected for further examination for their potential effects on the IFN pathway.

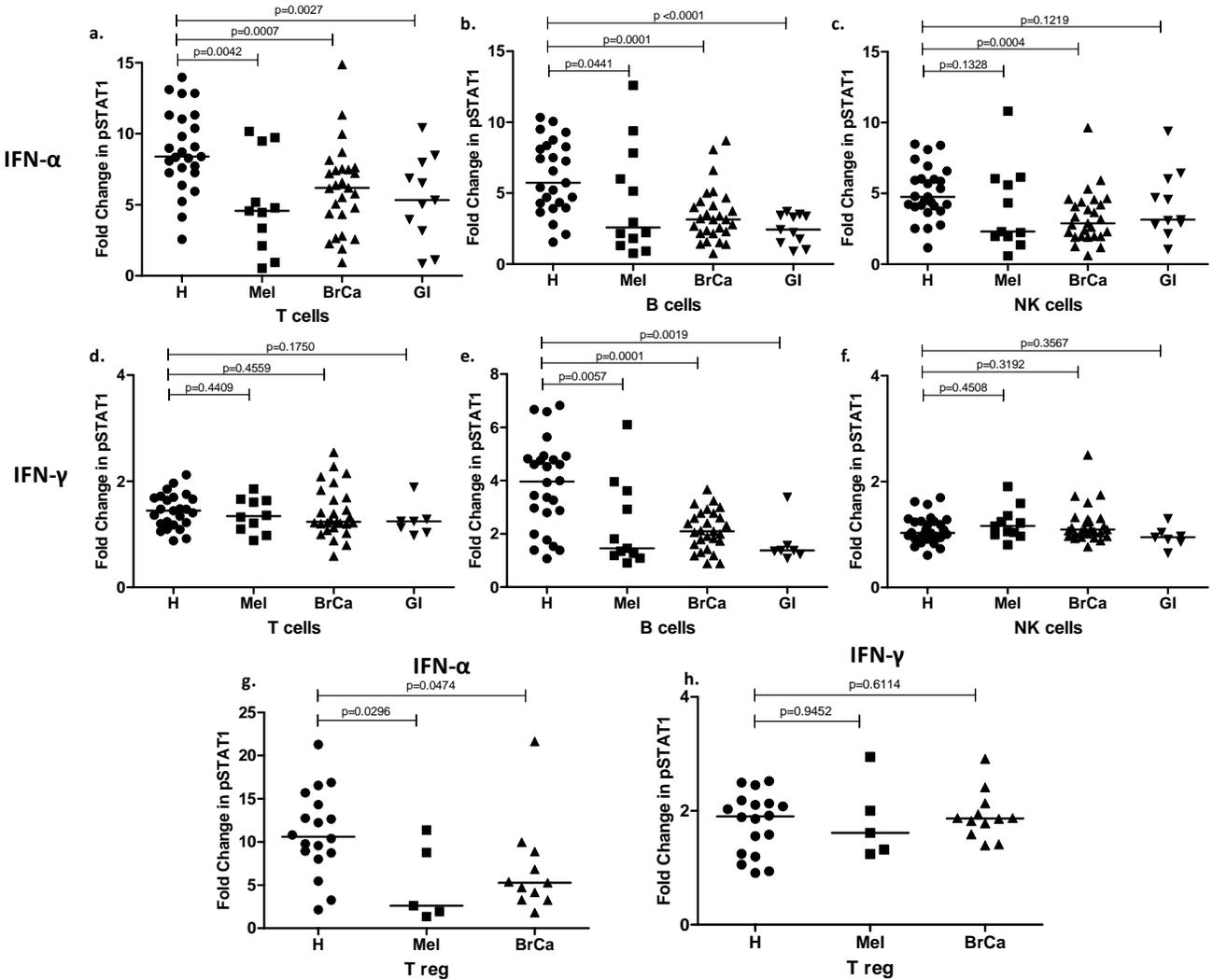
**Figure A1**



**Figure A1. Real Time Quantitative PCR Analysis of ISG Expression in Lymphocytes from Breast Cancer Patients and Healthy Controls.**

The expression levels of ISGs: STAT1, IFI44, IFIT1, IFIT2 and MX1 were measured in unstimulated lymphocytes from breast cancer patients (BC) and age-matched healthy controls (H) by real time quantitative PCR. Expression of each gene was normalized to GAPDH. Medians are indicated by the bar in each data set.

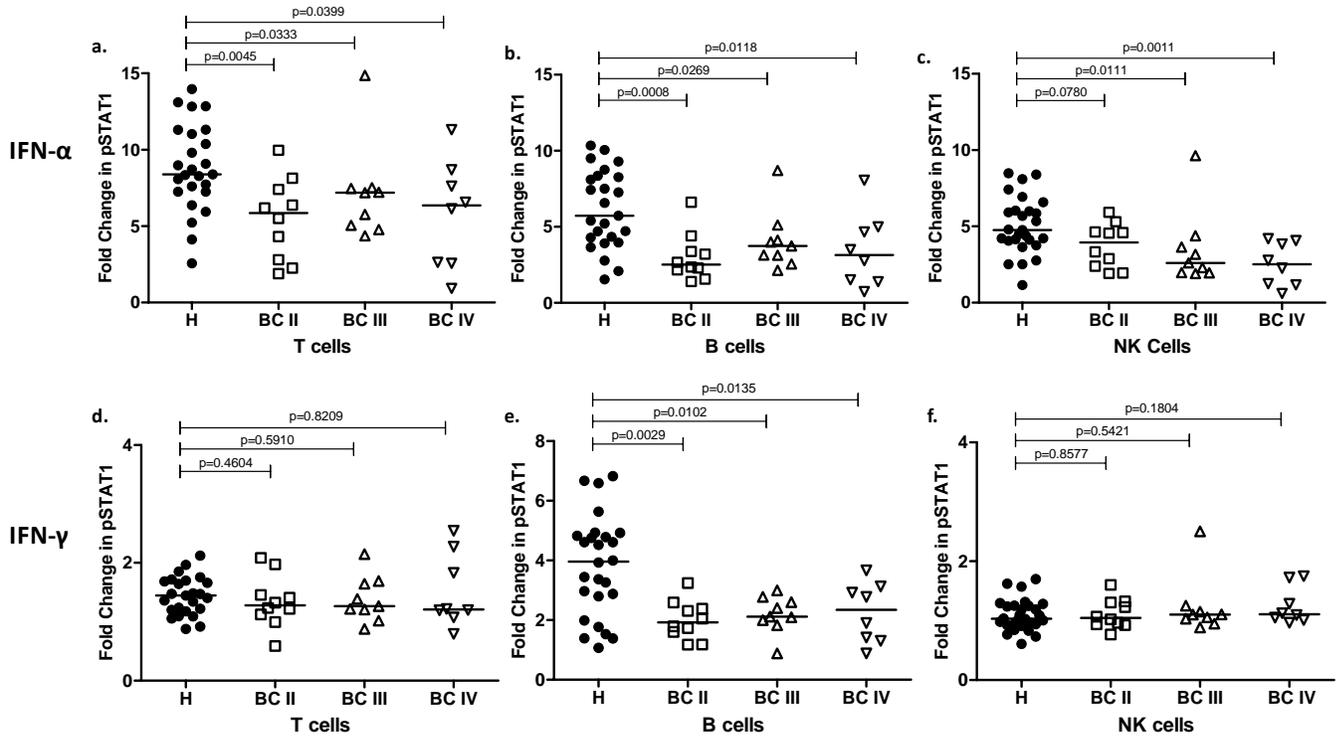
**Figure A2**



**Figure A2. IFN- $\alpha$ - and IFN- $\gamma$ -Stimulated Fold Change in pSTAT1 (Y701) in PBMC Subsets from Breast Cancer Patients, Melanoma Patients and Gastrointestinal Cancer Patients Versus Healthy Controls.**

PBMCs were stimulated with IFN- $\alpha$ , IFN- $\gamma$ , or unstimulated and pSTAT1-Y701 was measured by Phosflow. The fold change in pSTAT1 in T cells (CD3+), B cells (CD19+) and NK cells (CD16+) was calculated by dividing the mean fluorescence intensity (MFI) of pSTAT1 staining in IFN-stimulated cells by the MFI of pSTAT1 staining in the corresponding unstimulated cell subset from healthy controls (H●), patients with breast cancer (BC ■) patients with melanoma (Mel ▲) and patients with gastrointestinal cancer (GI ▼). The median is indicated by the bar in each data set.

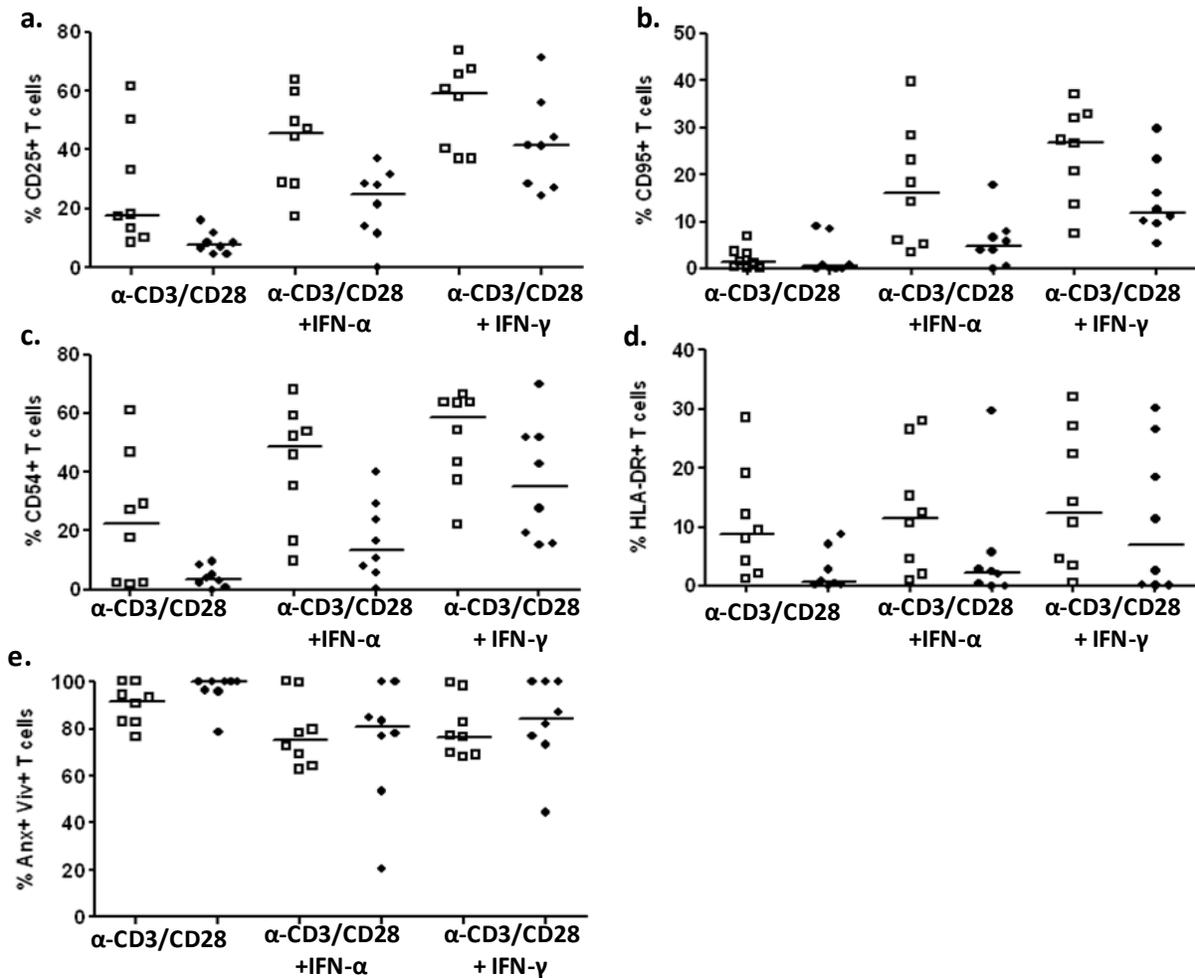
**Figure A3**



**Figure A3. Effect of breast cancer stage on IFN- $\alpha$ - and IFN- $\gamma$ -stimulated fold changes in pSTAT1 in cancer patients and healthy controls.**

PBMCs were stimulated with IFN- $\alpha$ -, IFN- $\gamma$ -, or unstimulated and pSTAT1 was measured by Phosflow in T, B and NK cells from healthy controls (H  $\blacklozenge$ ), patients with stage II breast cancer (BC II  $\square$ ), patients with stage III breast cancer (BC III  $\triangle$ ), patients with stage IV breast cancer (BC IV  $\nabla$ ). The median is indicated by the bar in each data set.

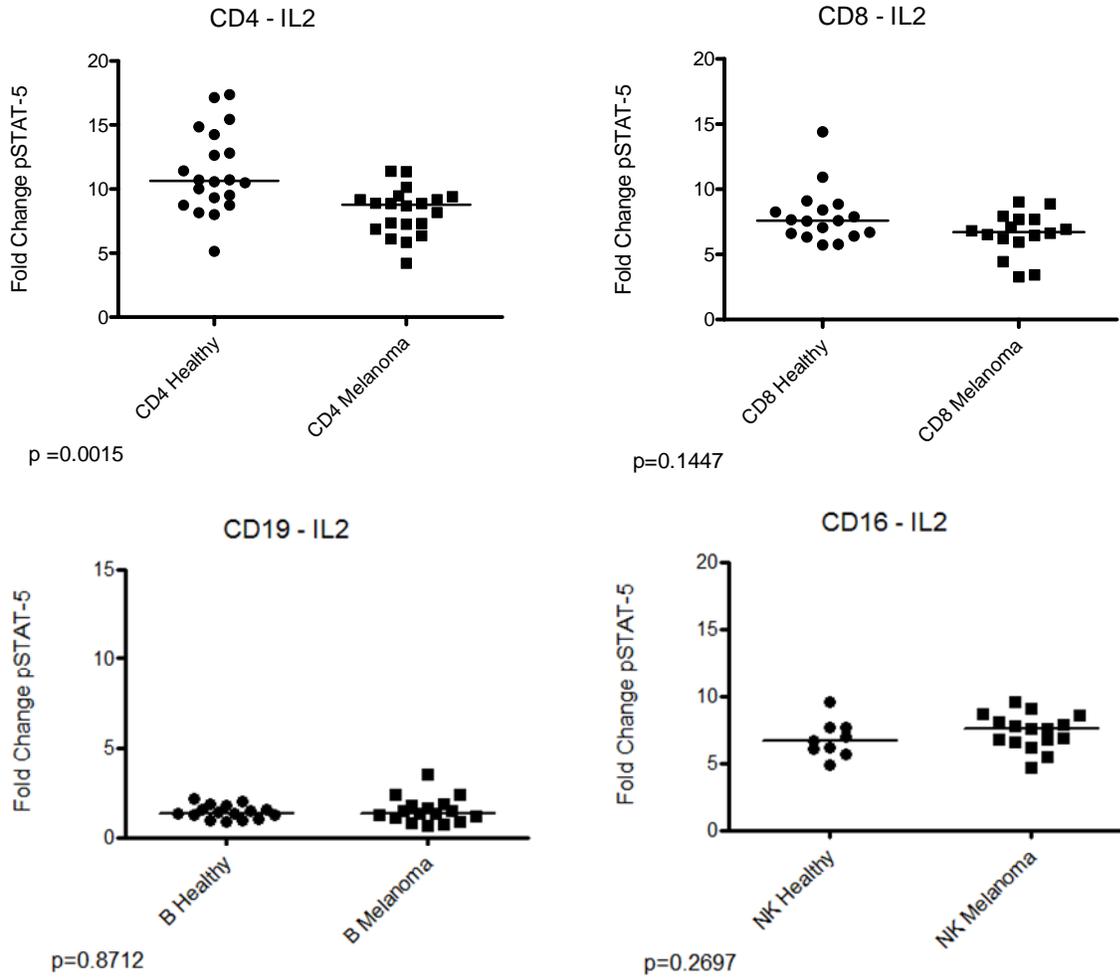
Figure A4



**Figure A4. Expression of Activation Markers and Apoptosis of T cells Stimulated with anti-CD3/CD28 antibodies and IFN- $\alpha$  or IFN- $\gamma$  in breast cancer patients and healthy controls.**

Lymphocytes from breast cancer patients ( $\blacklozenge$ ) and healthy controls ( $\square$ ) were stimulated with beads coated with anti-CD3 and anti-CD28 antibodies alone or in combination with IFN- $\alpha$  or IFN- $\gamma$  or left unstimulated. The percentages of cells that stained positive for CD25, HLA-DR, CD54, CD95, Annexin V (Anx) and ViViD (Viv) were measured by flow cytometry at 48 hours following stimulation. The median is indicated by the bar in each data set.

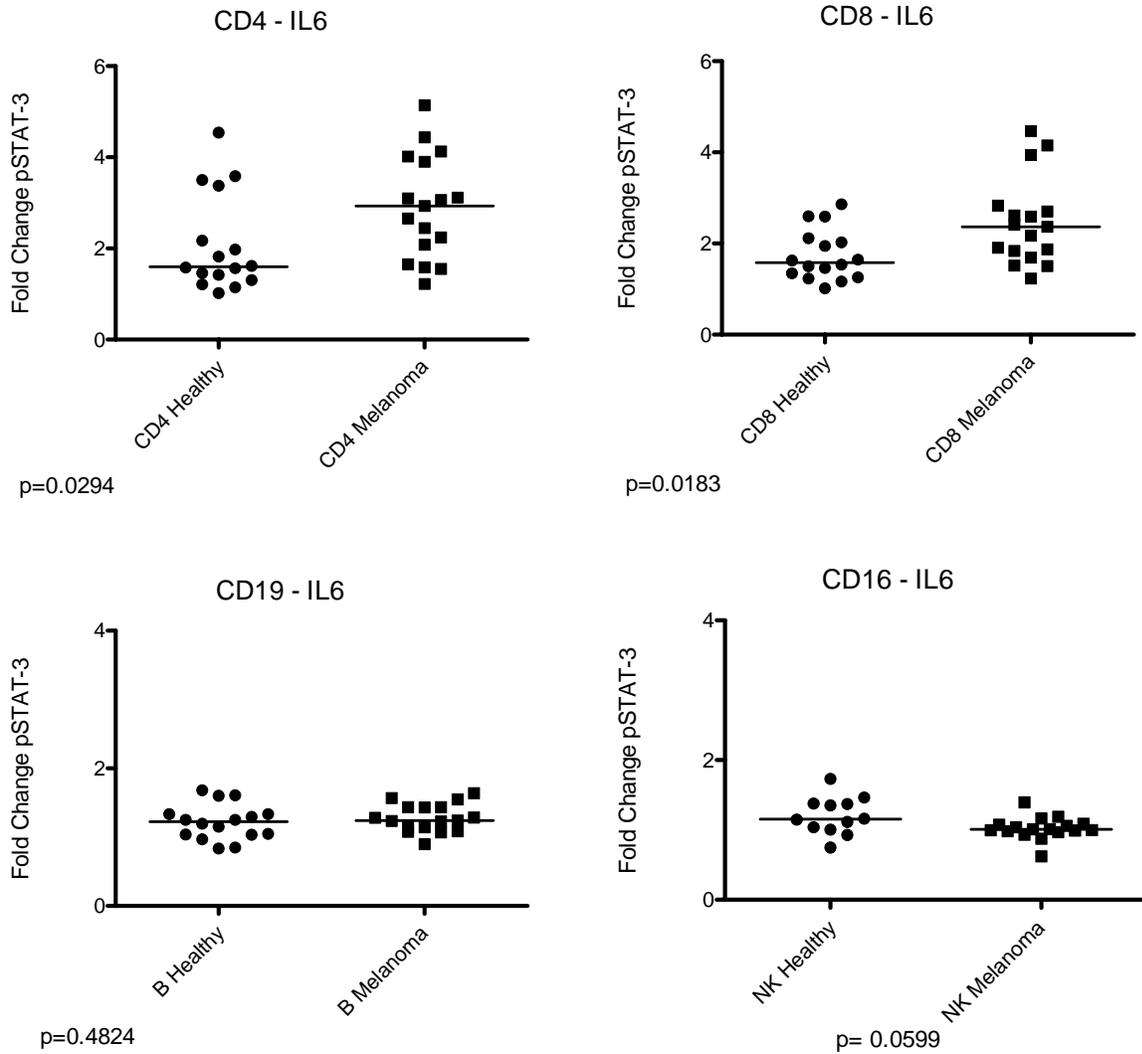
**Figure A5**



**Figure A5. Comparison of IL2-stimulated fold changes of pSTAT5 in PBMC lymphocyte subsets from melanoma patients and healthy subjects.**

PBMCs were stimulated with IL2 or left unstimulated. pSTAT5 levels were measured by phosflow analysis. The fold change in pSTAT5 was calculated by dividing the mean fluorescence intensity (MFI) of pSTAT5 in IL2-stimulated cells with the MFI of pSTAT5 in unstimulated cells.

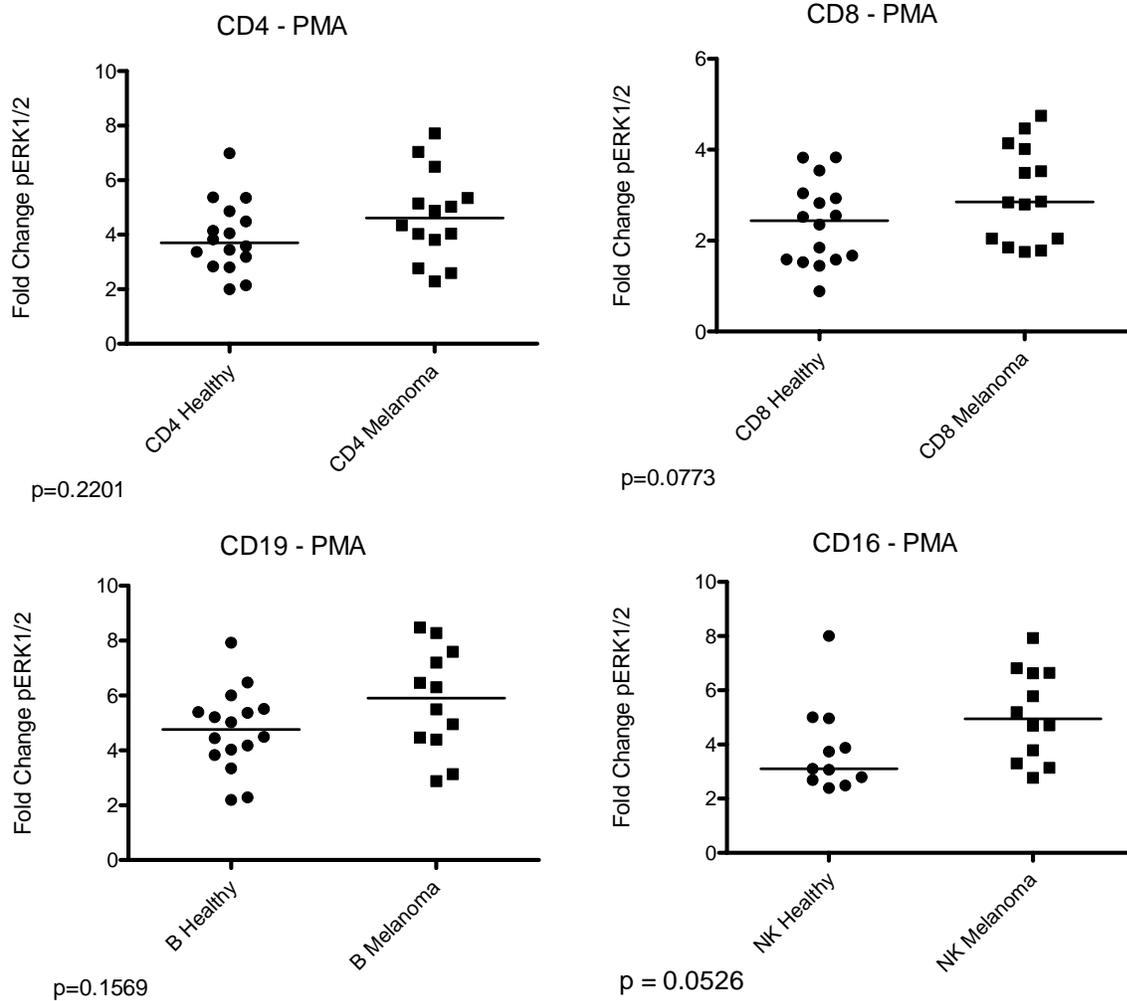
**Figure A6**



**Figure A6. IL6-stimulated fold change of pSTAT3 in PBMC lymphocyte subsets from melanoma patients and healthy subjects.**

PBMCs were stimulated with IL6 or left unstimulated. pSTAT3 levels were measured by phosflow analysis. The fold change in pSTAT3 was calculated by dividing the MFI of pSTAT3 in IL6-stimulated cells with the MFI of pSTAT3 in unstimulated cells.

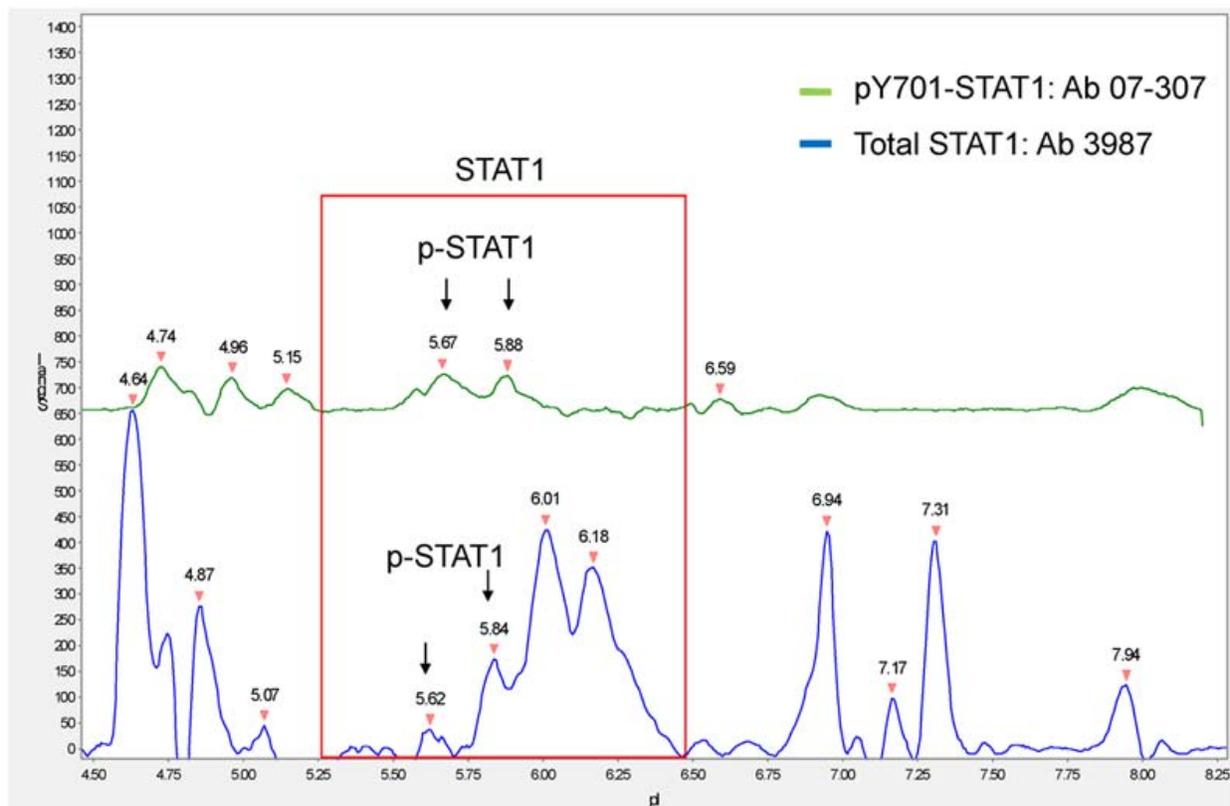
**Figure A7**



**Figure A7. PMA-stimulated fold change of pERK1/2 in PBMC lymphocyte subsets from melanoma patients and healthy subjects.**

PBMCs were stimulated with PMA or left unstimulated. pERK1/2 levels were measured by phosflow analysis. The fold change in pERK1/2 was calculated by dividing the MFI of pERK1/2 in PMA-stimulated cells with the MFI of pERK1/2 in unstimulated cells.

Figure A8



**Figure A8. Detection of total STAT1 and STAT1-pY701 by Firefly 3000.**

Peripheral blood lymphocytes were isolated from healthy controls and stimulated with IFN-alpha for 15 minutes. Cells were lysed and run on the Firefly 3000 capillary-based platform with which samples were separated by isoelectric focusing, immobilized, and probed with antibodies against total STAT1 (blue, Ab 3987) or p-Y701-STAT1 (green, Ab 07-307). Images were taken of capillaries and peaks were determined by Compass software (Cell Biosciences). The isoelectric points of STAT1-alpha and STAT1-beta were estimated at 5.7 and 6.0 respectively by the UCSC Proteome Browser and peaks in that region were identified as likely to be STAT1 (red box). Two peaks were identified for each of the possible unphosphorylated and p-Y701-STAT1 (arrows) states.

Personnel: Lee, Johnson, Dirbas, Schwartz, Yu, Simons.

## B. Microarray analysis of immune and tumor cells independently

Originally proposed in SOW:

1. Microarray analysis of gene expression of purified tumor and immune cells, isolated from fresh tumor or TDLN samples, and peripheral blood mononuclear cells (PBMC) from breast cancer patients. (months 6-60)
2. Detailed analyses of gene expression data focusing on gene networks and cross-talk between tumor and immune cells. (months 12-60)

This project utilizes a systematic approach to study the dynamics between breast cancer and the immune responses by directly comparing the gene expression patterns from TDLNs with the tumor site and peripheral blood. An increasing number of studies have used microarray to profile breast tumor specimens, which in fact represent heterogeneous cell populations consisting of tumor cells and tumor infiltrating immune cells. Our strategy is to profile purified tumor and immune cells independently, isolated from tumors and/or TDLNs.

### Microarray analysis of immune and tumor cells independently

We took an integrative systems approach to study the dynamics between breast cancer and the immune responses by directly comparing the gene expression patterns between tumor, TDLNs, and PBMCs. Our strategy was to profile purified tumor and immune cells, isolated from tumors and/or TDLNs independently. Key accomplishments for year 3 include the completion of the first batch microarray experiment and microarray data pre-processing. Currently, we are in the process of identifying immune profile changes across three anatomical compartments and validating these findings through histology. Our data set allows independent analysis of expression profiles in tumor and immune cells as well as correlation of genes expressed in tumor cells with the transcriptional alterations in the paired immune cells. We are using this gene network cross talk analysis to generate hypotheses regarding how and where tumor cells are modulating the host immune system.

#### 1. Summary of sample composition for microarray analysis

We have completed gene expression profiling of the initial set of patient specimens which comprises 156 samples collected from 24 newly diagnosed breast cancer patients. These samples include immune cells and/or tumor cells from peripheral blood, primary breast tumor tissues, and TDLNs (tumor free or tumor involved). Of these, a total of 10 patient complete sets including tumor cells and their paired immune cells from tumor tissues, TDLNs, and blood are undergoing analysis to directly compare the gene expression patterns across three anatomical compartments.

#### 2. Experimental approach

Samples were acquired and processed as shown in Figure B1. Live immune cells and tumor cells from each compartment were sorted from heterogeneous cell populations to high purity by flow cytometry. Total RNA was isolated through Trizol method and amplified in two consecutive rounds using TrueLabeling-PicoAMP™ kit (SuperArray), followed by the Cy3/Cy5 labeling (Amersham Biosciences Corp.). High sensitivity quality control of amplified/labeled RNA samples was carried out using the RNA 6000 Nano LabChip kit and 2100 Bioanalyzer

(Agilent). In the current study we used a two-color platform with dye swap design to correct dye bias and, in addition, serves as replicates for each sample. Stratagene's Universal Human Reference (UHR) RNA was amplified and labeled using the same protocol and used for all the arrays as a common reference.

### 3. Microarray Data pre-processing

We used Agilent's Whole Human Genome Microarray 4x44K G4112F to generate our microarray data. This array has Platform ID GPL6480 in NCBI's GEO database. Our data processing procedure involved the following steps:

- Image files were generated from microarray slides using Agilent Microarray Scanner G2505B and raw data was extracted from the image files using Agilent's Feature Extraction Software (version 9.5). The raw data was transformed to log<sub>2</sub> scale where we used our LIS algorithm to identify a group of approximately 1500 probes with highly consistent ranking patterns across all arrays, and with intensity levels spanning the entire data range. We subsequently used these probes as "surrogate house-keeping probes" for the purpose of normalization.
  - This is a data-driven approach which takes advantage of the internal pattern within each microarray dataset to identify a likely group of non-changing probes specifically for that data. Our data-driven approach works better than the standard use of a "house-keeping gene list" compiled from prior biological knowledge, because very few genes can be considered truly stable under all experimental conditions.
- We blended two standard normalization methods: quantile normalization and LOESS normalization to transform the raw data for each array into normalized data.
  - Our approach based on a relatively small group of "surrogate house-keeping probes" requires that the distribution of expression levels for the genes corresponding to these probes remain the same across samples, which is a modest requirement, and therefore, our approach can adapt better to many different experimental conditions and sample types.
- All arrays were calibrated to the same scale and we applied a "variance stabilizing transformation" to the data.
  - In general not subtracting the background or subtracting too little effectively compresses the fold change values at low intensity range, while an over-estimation of background would inflate the fold change. The influence of intensity level on fold change may also be observed when comparing middle to higher intensity ranges, typically a compression occurs at the highest intensity range, probably due to the influence of saturation. In statistical terms, in microarray data, the spread (fold change) is not independent of location (intensity level). Such lack of independence between spread and location will cause problems when we rank probes by fold change: a probe is more likely to be ranked high if its intensity level happens to be in an inflation range and more likely to be ranked low if it falls in an compression range. Moreover, many standard statistical methods implicitly assume equal variance (independence of spread and location). Therefore, it is helpful to apply a "variance stabilizing transformation" to the data

prior to analysis. Intuitively such a transformation would stretch and compress the data at different ranges as needed to equalize the spread for all locations.

- The normalized data values from replicated probes were aggregated by taking the median across replicates of each unique probe sequence. Because these arrays come in dye swap pairs, we took the average of each dye swap pair to obtain one data column for each sample. In addition, we went through a parallel procedure following all the steps above but this time including the Universal Human Reference (UHR) columns and produced log ratio values (sample vs. UHR) by subtracting the UHR columns from the corresponding sample columns.

### The LIS algorithm

We developed the LIS (Long Increasing Sequence) algorithm as a general data-driven approach to identify a special kind of rank-based substructure within a larger numerical dataset. The simplest example of such a substructure is the well-studied case of a longest increasing subsequence within a finite sequence of real numbers. In the context of microarray data analysis, the substructure is an optimal subset of probes showing consistent ranking patterns across all the arrays from an experimental study. Given the typical large size of microarray data, carefully designed non-parametric approaches based on reasonably mild assumptions can be expected to work well for the purpose of data normalization. Our approach is based on the rather modest and intuitive non-parametric assumption that the array effect is an order-preserving (i.e. monotonically increasing) transformation.

### Data Quality

The first level of quality control is to view the QC reports produced by the feature extraction Software and then examine the dye swap pair plots utilizing the raw data. Figure B2 shows representative plots from 8 dye swap pairs. Most of these plots show a characteristic curve due to dye effect, which will be corrected in our calibration step before an average profile is calculated for each pair. Sporadic data points away from the curve often occurs due to minor local artifacts during either hybridization or scanning; occasionally we see clouds of data points away from the curve which may be due to a slightly larger scale artifact. Again these effects can often be corrected by careful comparison of the dye swap pair. IF29269 ANANI illustrates a much higher degree of inconsistency between the two arrays in that pair. In such cases, we flag the data for this sample as problematic and we may do another dye swap pair for that sample if more material is available.

Normalized data from two representative NSLN immune cell samples, BC#5 ANBLI and BC#5 ANGRI, respectively, is presented in Figure B3. Each of these plots represents a leukocyte population isolated from separate NSLNs from the same patient. Figure B3a compares normalized single-channel intensity values. The clean diagonal shape demonstrates high similarity between the gene transcription profiles of the two samples, and more importantly, the fact that we can clearly see this similarity highlights the reliability of our data generation protocol as well as the success of our data normalization procedure. This plot also demonstrates that the manufacturing quality of current generation arrays is good enough to allow confident use of single channel data (MAQC Consortium, 2006). The more traditional two-channel log-ratio of sample vs. reference is shown in Figure B3b showing the expected spindle shape with a bulk in the middle formed by data points with log-ratio values close to zero in both samples. Figure

B3c and Figure B3d are rank-transformed version of Figure B3a and Figure B3b, respectively. These rank-transformed plots are very useful for highlighting noisy regions in the data. We see the single channel data is noisier at the lower end of the intensity range, where the signal is close to background level. By contrast, the log-ratio data is noisier in the middle range. These are well understood characteristic features of microarray data.

#### The “Cell Number Effect”

When the tissue samples were processed to isolate tumor or immune cell populations, the resulting number of cells varied widely from sample to sample. In some cases very small cell populations provided sub-optimal amounts of total RNA input for the amplification step, resulting in a skewed expression profile after hybridization onto the array. This produced a “cell number effect”, which we demonstrate in Figure B4. We compared immune cells from NSLNs in patient BC#5 ANBLI, where 1.3 million immune cells were isolated, to another NSLN sample from the same individual, BC#5 ANOGI, which had a cell count of 0.2 million. The plots are similar to Figure B3 but they are much less centered around the red diagonal line, showing a large scale difference in the measurements of expression profile. It becomes clear that the difference is due to the “cell number effect” when we compare many more profiles and correlate the differences to cell counts. In particular, Figure B5 compares BC#5 ANBLI to BC#5 ANYLI, the latter of which has an even smaller immune cell count consisting of 0.05 million cells. As expected, the cell number effect produced even larger differences as shown by the scattered dots away from the red line. We have been very careful in using data from samples with small cell counts and our standard protocol calls for a cell count of 500,000 per NSLN sample for hybridization and analysis.

#### 4. Preliminary Results

##### Gene set enrichment (GSEA) analysis: Upregulation of cell cycle checkpoint pathways and TGF $\beta$ pathway in lymph node compared to PBMCs.

Gene set enrichment analysis (GSEA) was used to compare the gene expression pattern between PBMCs and TDLNs from breast cancer patients. GSEA is a computational method that determines whether an a priori defined set of genes shows statistically significant, concordant differences between two biological states. A total of fourteen gene sets are significantly upregulated in TDLNs compared to PBMCs at nominal p-value < 5%. Four out of fourteen gene sets are involved in cell cycle checkpoint pathways and the expression pattern of these genes is presented in the expression heatmap in Figure B6. Hierarchical clustering separated PBMCs from TDLNs on the basis of the expression patterns of these genes. Several genes involved in these cell cycle checkpoints (ATM, CHEK1, WEE1, MYT1, PTC1) are required for checkpoint mediated cell cycle arrest in response to DNA damage or the presence of unreplicated DNA. These genes may also negatively regulate cell cycle progression during unperturbed cell cycles. It is plausible that these genes are part of the negative regulation loop which is upregulated upon clonal expansions of activated lymphocytes in breast cancer TDLNs. To resolve this issue, we are currently collecting non-cancer lymph nodes from gastrointestinal patients. Gene expression profiling for up to 50 non-cancer lymph nodes will be performed using the same approach as the breast cancer patients. This will enable us to compare the proliferation profile between breast cancer TDLNs and non-cancer reactive lymph nodes.

In addition, the TGF $\beta$  pathway is also upregulated in TDLNs compared to PBMCs (Figure B6). TGF $\beta$  has a well-established role in controlling T cell activation, proliferation, differentiation and survival. It inhibits T cell proliferation by blocking IL-2 and cyclins, hindering Th1 and Th2 differentiation and downregulates cytotoxic T cell development. TGF $\beta$  is also required for the sustained expression of FoxP3, which controls the function of regulatory T cells (Tregs) and impedes non-regulatory T cell activation, proliferation and differentiation. Therefore, we hypothesize that the clonal expansion of lymphocytes in TDLNs is blunted and the cell cycle checkpoint regulators are upregulated by tumor cells or other suppressor cells, such as Tregs, to modulate anti-tumor immunity.

#### CFSE-based proliferation assay: Blunted proliferative capacity of lymphocytes from TDLNs upon *in vitro* challenge

We have developed a 9 color, 11 parameter flow cytometry panel to assess the proliferation capacity of immune cells. Thus far, we have assayed a total of 30 patient samples, which include immune cells from 20 tumor free lymph nodes, 7 tumor involving lymph nodes and 3 peripheral blood samples. We also assayed 7 peripheral blood samples from healthy donors as controls. Data show that lymphocytes from breast cancer lymph nodes (regardless of tumor involvement) have less capability to proliferate compared with lymphocytes from peripheral blood (Figure B7a), which is consistent with our microarray results. To determine whether the difference between proliferation capacities of lymphocytes from peripheral blood and lymph nodes is due to the different anatomical locations, we performed the CFSE-based proliferation assay on five C57BL/6J wild type mice and compared lymphocyte proliferation in mouse peripheral blood (PBL) and lymph nodes (LN). In C57BL/6J wild type mice, lymphocytes from PBL and LN showed similar proliferation capacity (Figure B7b). This suggests that lymphocytes from different anatomical compartment respond in similar magnitude upon *in vitro* stimulation. The proliferation data of breast cancer TDLN were stratified according to patients' ER/PR/Her2 status (Figure B7c), sentinel lymph node (SLN)/non-sentinel lymph node (NSLN) (Figure B7d) which showed no significant difference. Paired peripheral blood lymphocytes for these lymph node samples will be analyzed in a batch to increase the sample size and allow paired analysis of proliferation capability of lymphocytes in the same individuals.

#### 4-color IHC Staining: Decreased Ki67 expression in immune cells in breast cancer TDLNs compared to non-cancer patients

The CFSE-based proliferation assay is to measure the proliferation capacity of cells upon *in vitro* challenge. In addition, we optimized a quadruple immunohistochemical (IHC) staining panel to assess the lymphocyte proliferation status *in vivo*. Ki67 is a proliferation marker expressed by proliferating cells in all phases of the active cell cycle (G1, S, G2 and M phase) and absent in resting (G0) cells. In addition to Ki67, we also selected a panel of antibodies directed against cell of interests: CD3 T cells, CD20 B cells, AE1/AE3 for tumor cells. Colors were developed using DAB for B cells, Ferangie blue for T cells, Vulcan red for Ki67, Bajoran purple for tumor cells and sections were counterstained with hematoxylin. Images are acquired using the automated imaging system (Olympus and Ludl) with Nuance<sup>TM</sup>. Acquired images are then analyzed with our custom image analysis software Gemident to identify and enumerate the following

phenotypes: T cells, B cells, tumor cells, Ki67+ T cells, Ki67+ B cells, Ki67+ tumor cells. The Ki67 positivity for each cell type is being compared in the following groups:

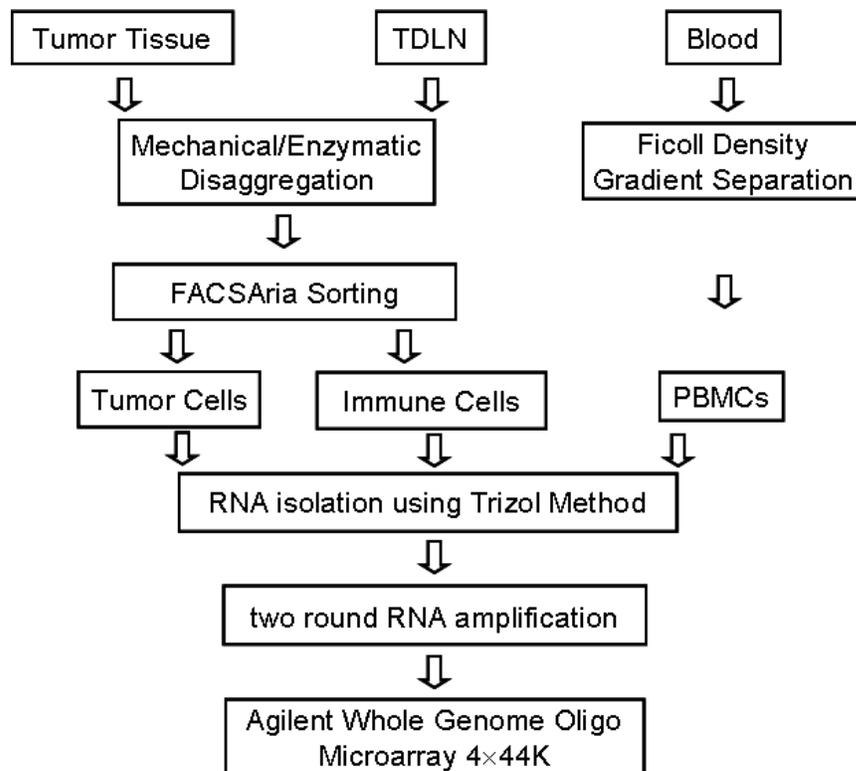
1. breast cancer TDLNs and non-cancer reactive lymph nodes
2. breast cancer TDLNs with different clinical outcome
2. tumor free and tumor positive lymph nodes
3. SLNs and NSLNs

This will help to confirm the proliferation defect in breast cancer TDLNs and gain insight into the mechanisms involved.

A total of 15 TDLN sections have been stained and data analysis is in progress. Preliminary data has shown a decreased Ki67 positivity in breast cancer TDLNs compared to reactive lymph nodes from non-cancer patients, indicating a perturbed proliferation immune response in breast cancer TDLNs. Remarkably, TDLNs from five patients showed less than ten Ki67 positive cells for the whole section regardless of tumor involvement. The clinical history of these patients showed recurrence, metastasis, and development of second primary tumor after their initial diagnosis. On the other hand, TDLNs from five disease free patients showed abundant Ki67 positivity. Representative data is shown in Figure B8. Furthermore, we identified interesting Ki67 staining patterns of germinal center (GC) B cells in these Ki67 low patients. In reactive lymph nodes, most GC B cells are positive for Ki67, often clustering in the dark zone and surrounded by Ki67 negative B cells in the light zone. In four out of the five breast cancer TDLNs mentioned above, germinal center architecture is identifiable, however, Ki67 positivity is lost from these GC B cells indicating a quiescent or suppressed GC (Figure B9). Local Ki67 positive cells are present in each of those sections to ensure this is not due to technical issues. None of the four disease free patients showed Ki67 negative germinal centers.

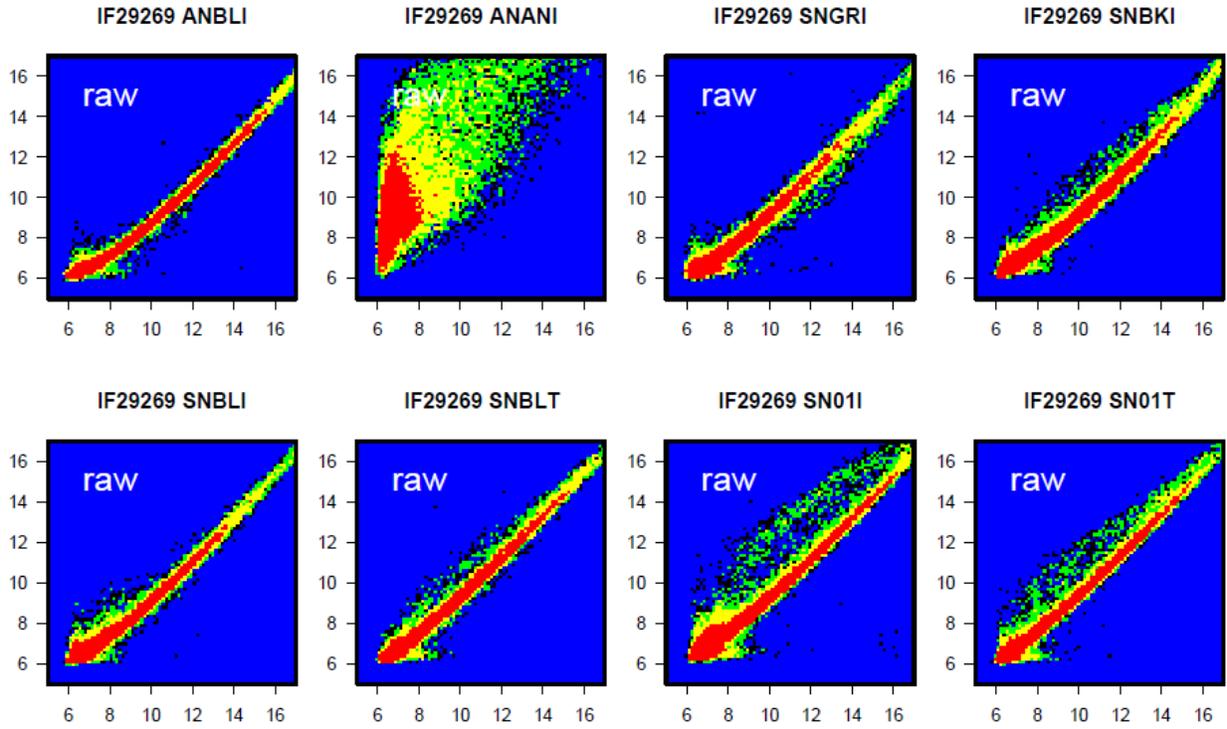
We plan to further validate the current findings by performing quadruple Ki67 staining on a total of 20 disease free patients (5 years or 10 years) and 20 breast cancer patients with recurrence/metastasis/second primary tumor development. Ki67 positivity for these TDLN sections will be quantitated using GemIdent. The number of Ki67+ T cell, Ki67+ B cells, Ki67+ tumor cells will be normalized to the total number of each cell type to generate the percentage of proliferating T cells/B cells/tumor cells for each patient. Statistical comparison will be performed to correlate Ki67 positivity with clinical outcome. Quadruple Ki67 staining on 15 lymph nodes from non-cancer patients will also be performed to serve as control samples.

**Figure B1.**

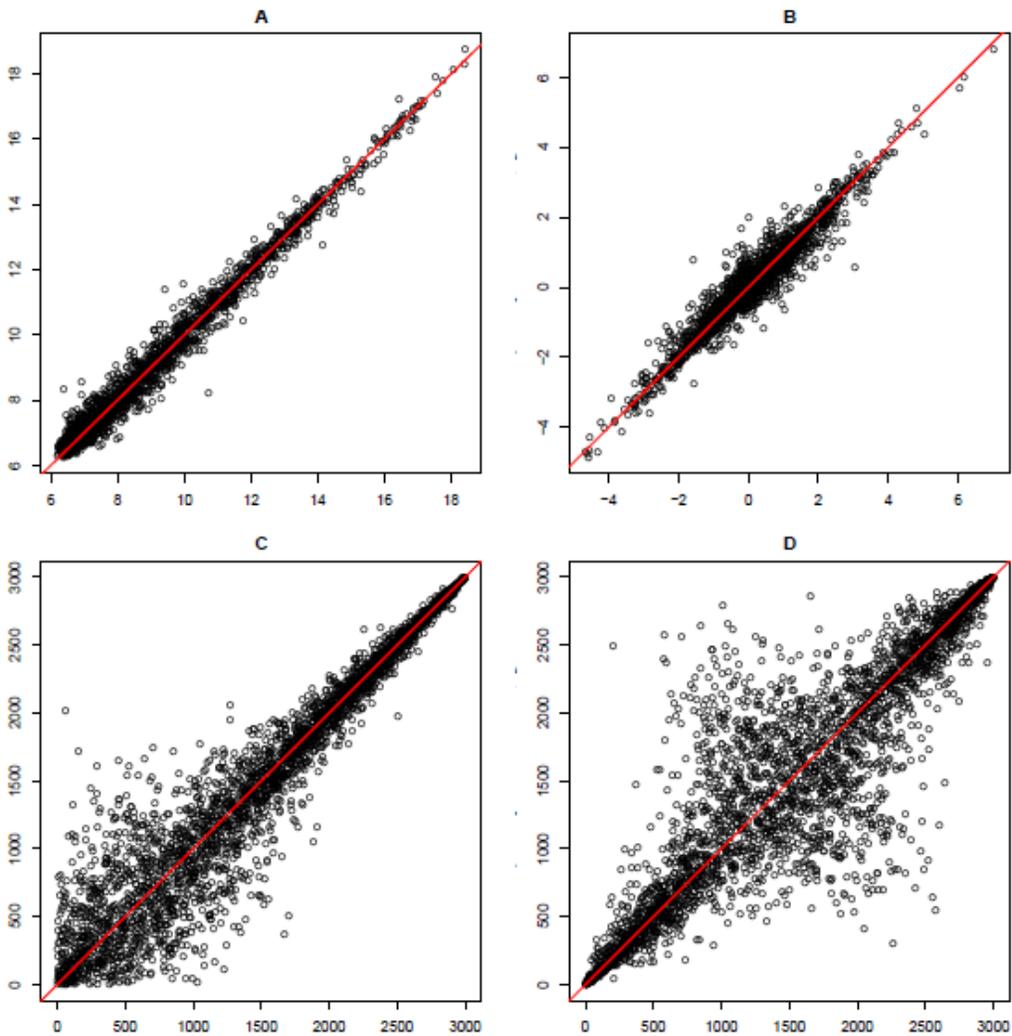


**Figure B1. Sample Processing**

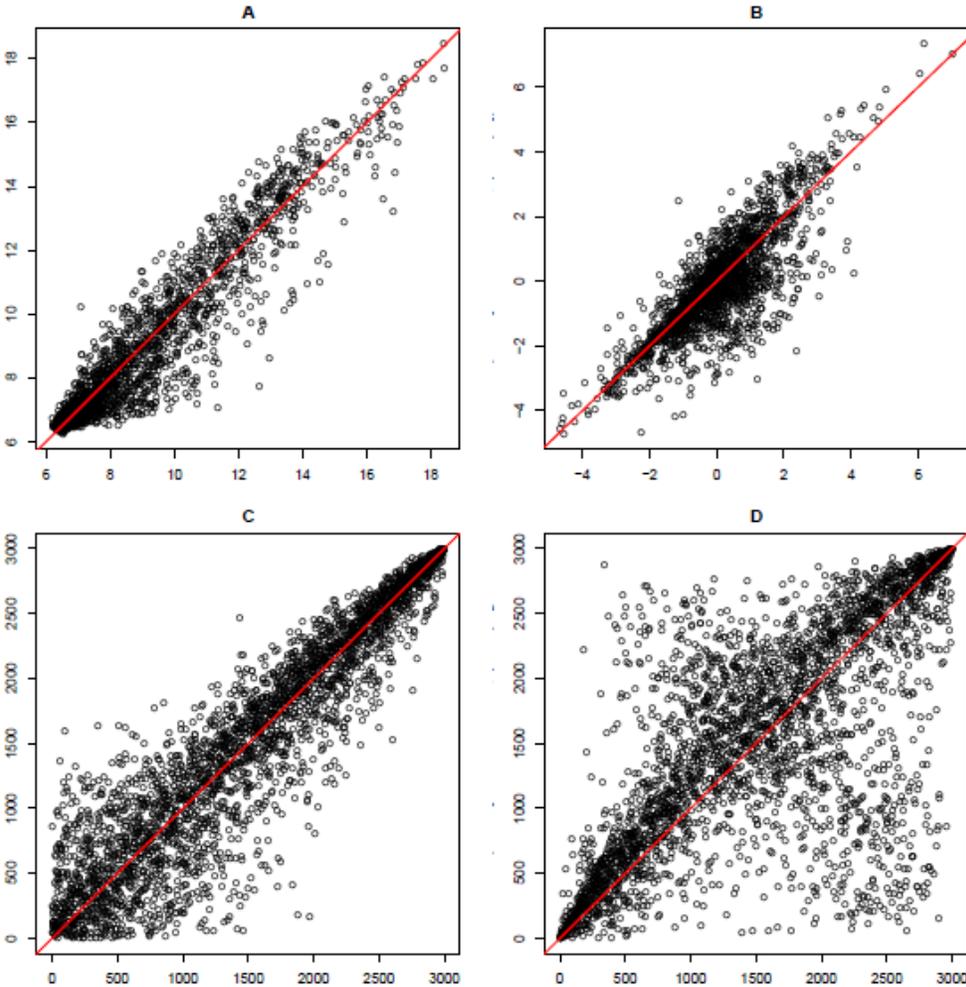
**Figure B2.**



**Figure B2. Dye swap pair plots for 8 representative samples from the same patient.** Most of these plots show a characteristic curve due to dye effect, which will be corrected in our calibration step before an average profile is calculated for each pair.

**Figure B3.****Figure B3. Normalization data from two representative samples.**

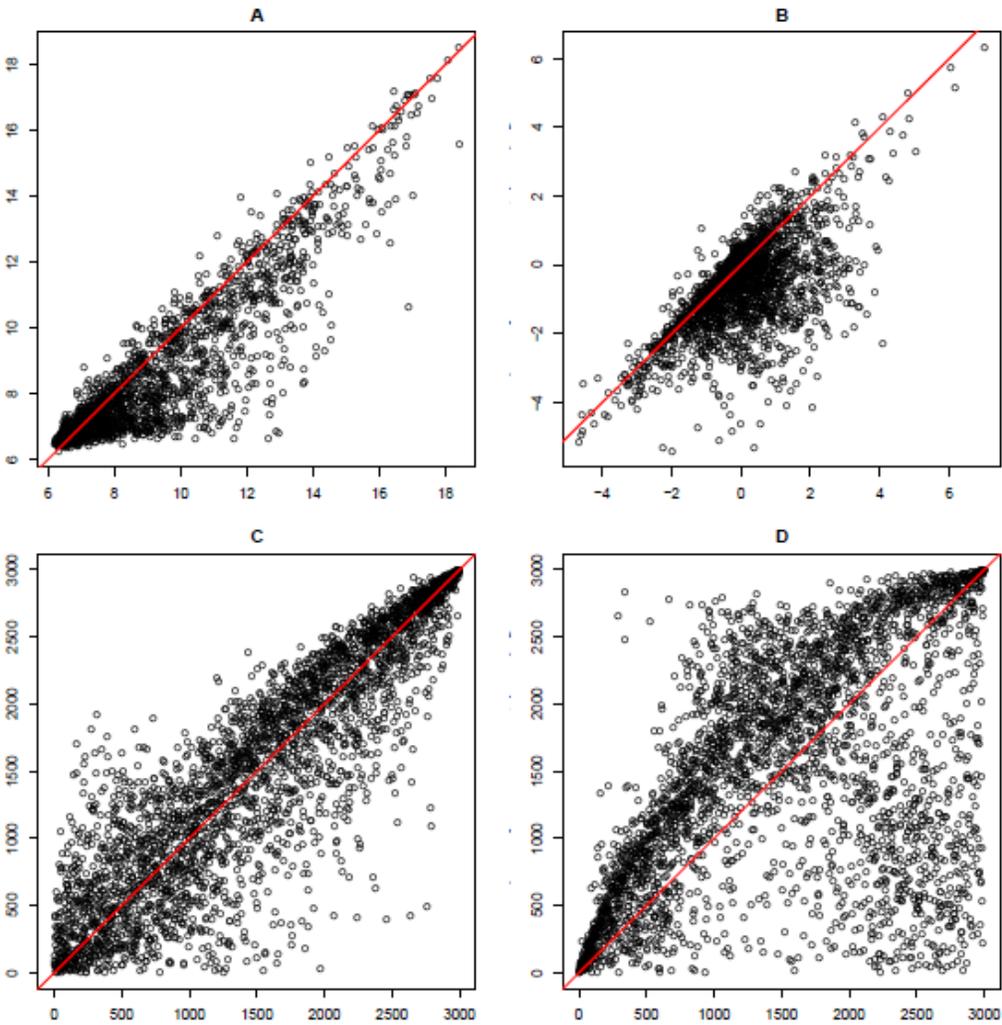
The horizontal and vertical axes each corresponds to one sample. (A) Comparison of normalized single-channel intensity values. (B) Comparison of normalized two-channel log-ratio values. (C) and (D) are rank transformed version of (A) and (B), respectively. Each plot is clearly centered around the red diagonal line. All plots are based on the same random sample of 3000 probes.

**Figure B4.**

**Figure B4. The cell number effect comparing 1.3 million immune cells to 0.2 million immune cells in NSLNs.**

The horizontal axis corresponds to a NSLN sample containing 1.3 million immune cells and the vertical axis corresponds to another NSLN sample from the same patient containing 0.2 million immune cells. (A) Comparison of normalized single-channel intensity values. (B) Comparison of normalized two-channel log-ratio values. (C) and (D) are rank transformed version of (A) and (B), respectively. All plots are based on the same random sample of 3000 probes.

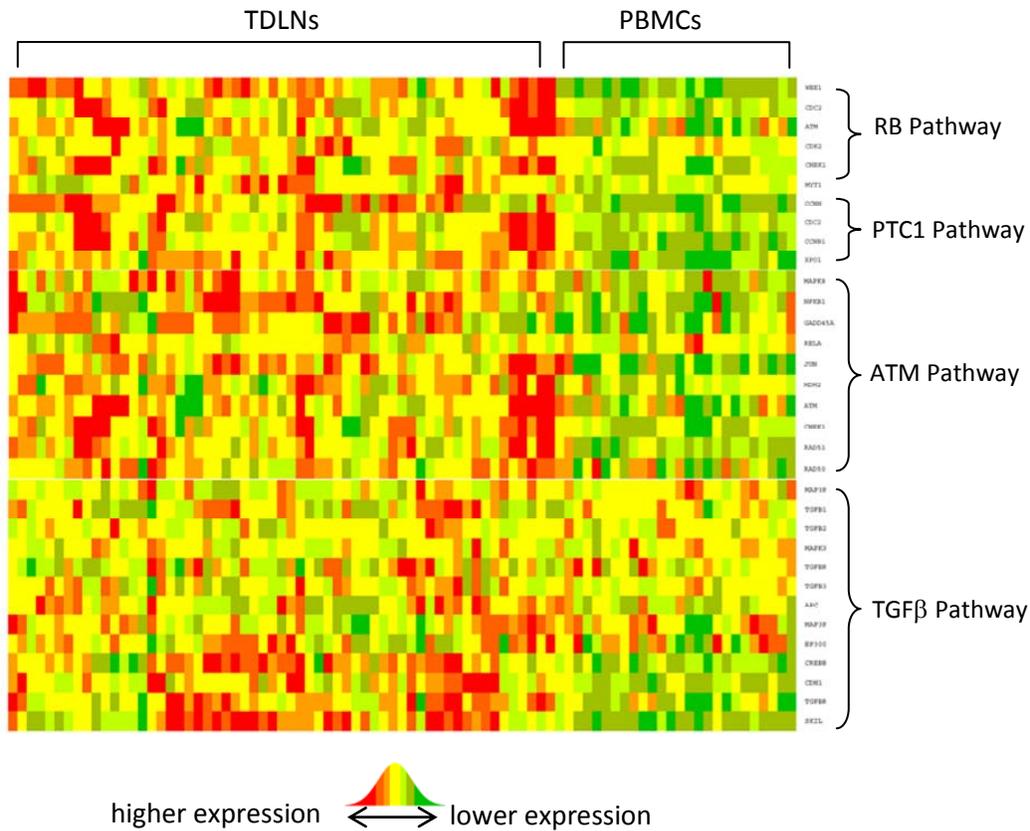
Figure B5.



**Figure B5. The cell number effect comparing fewer than 0.1 million immune cells to 1.3 million immune cells in NSLNs.**

The horizontal axis corresponds to a NSLN sample containing 1.3 million immune cells and the vertical axis corresponds to another NSLN sample from the same patient containing 0.05 million immune cells. (A) Comparison of normalized single-channel intensity values. (B) Comparison of normalized two-channel log-ratio values. (C) and (D) are rank transformed version of (A) and (B), respectively. All plots are based on the same random sample of 3000 probes.

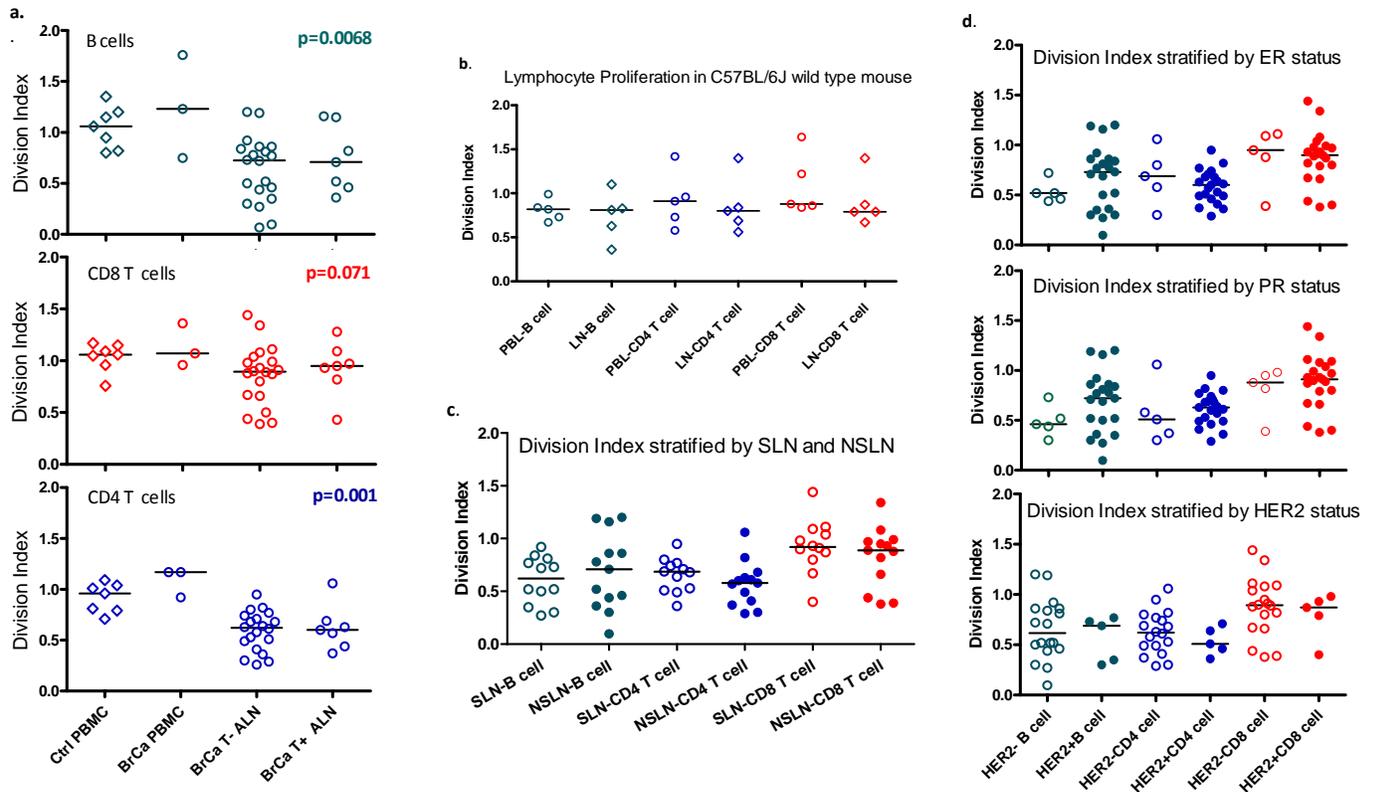
**Figure B6.**



RBPATHWAY	The ATM protein kinase recognizes DNA damage and blocks cell cycle progression by phosphorylating chk1 and p53, which normally inhibits Rb to allow G1/S transitions.
CDC25PATHWAY	The protein phosphatase Cdc25 is phosphorylated by Chk1 and activates Cdc2 to stimulate eukaryotic cells into M phase.
PTC1PATHWAY	The binding of extracellular signaling protein Sonic hedgehog to the Patched receptor (Ptc1) allows progression through G1 and may inhibit the G2/M transition.
ATMPATHWAY	The tumor-suppressing protein kinase ATM responds to radiation-induced DNA damage by blocking cell-cycle progression and activating DNA repair.
TGFBPATHWAY	The TGF-beta receptor responds to ligand binding by activating the SMAD family of transcriptional regulations, commonly blocking cell growth.

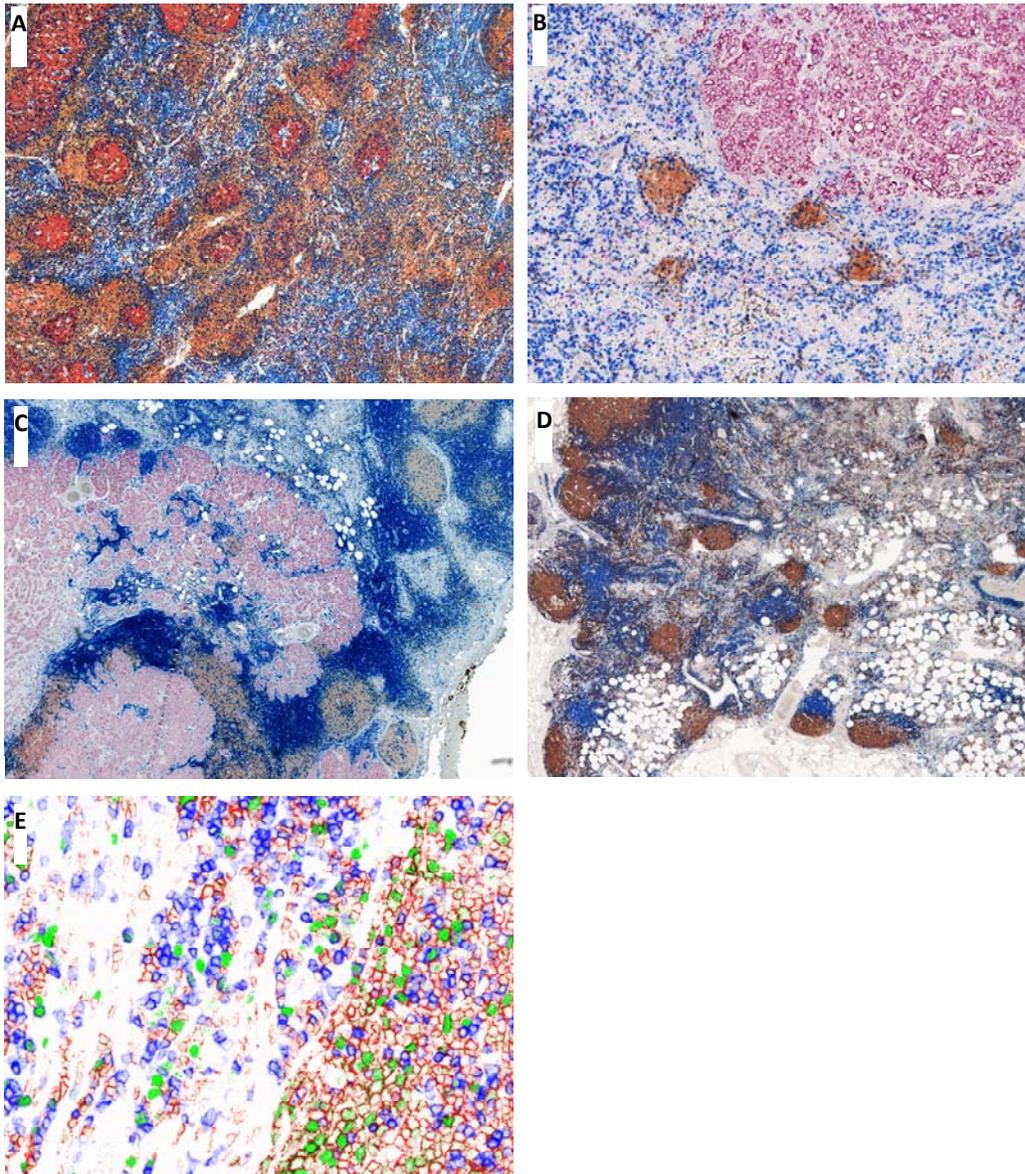
**Figure B6. Heatmap of pathways upregulated in TDLNs compared to PBMCs.** Hierarchical clustering separated PBMCs from TDLNs on the basis of the expression patterns of these genes. Red indicates higher gene expression and green indicates lower gene expression as illustrated in the legend. The biological significance of these pathways is listed below the heatmap.

**Figure B7.**

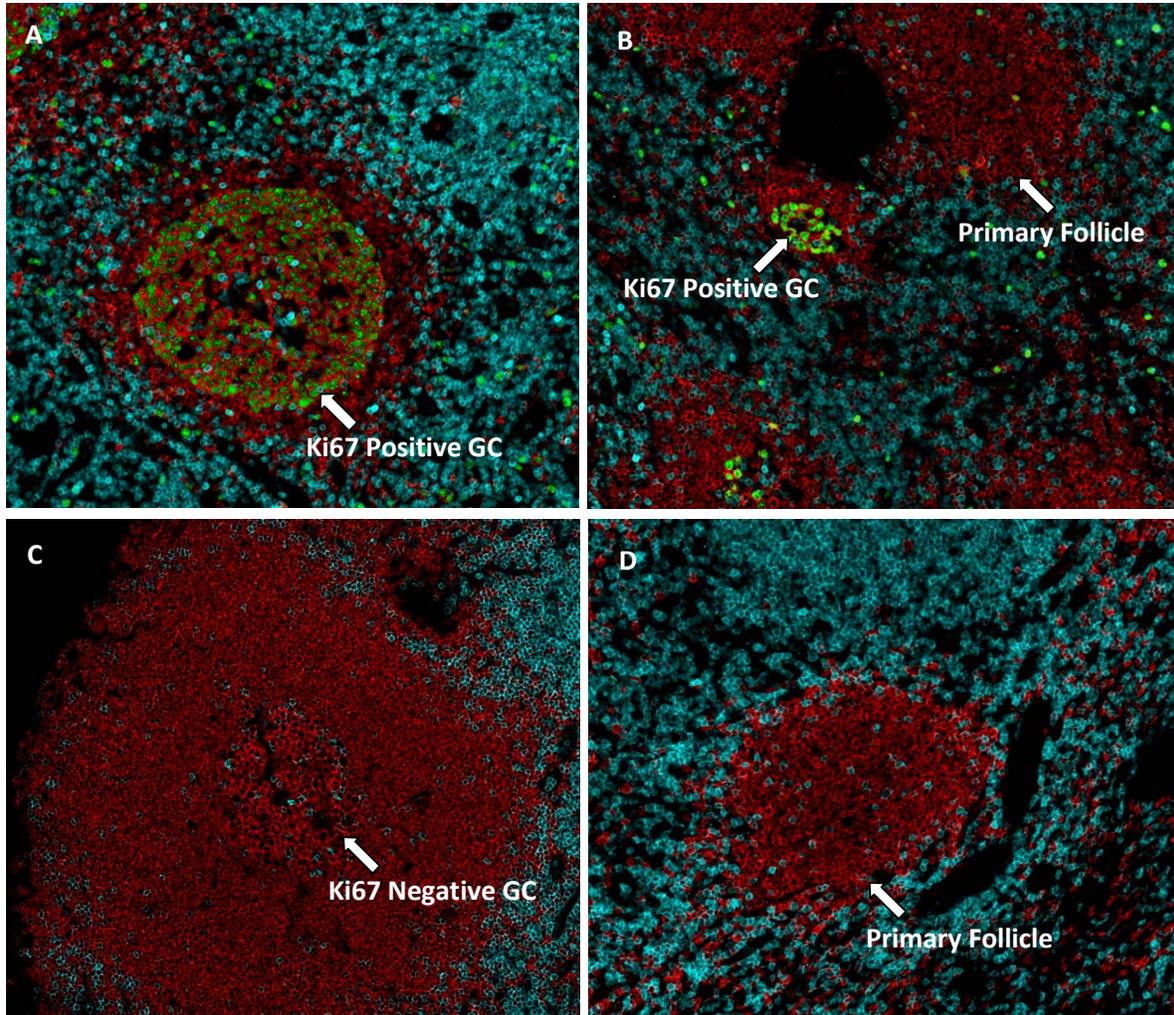


**Figure B7. CFSE-based proliferation assay.**

(a) Isolated immune cells were labeled with CFSE and stimulated with phorbol esters/Ionomycin for 115hour. Cell divisions were analyzed by FACSaria and cell proliferation analysis was performed by FlowJo. Proliferation capability of B cells, CD8 T cells, CD4 T cells for 7 healthy control blood samples (Ctrl PBMC), 3 patient blood samples (BrCa PBMC), 20 tumor free lymph nodes (BrCa T- ALN), 7 tumor involving lymph nodes (BrCa T+ ALN) are presented in separate scatter columns. Lymphocytes from breast cancer lymph nodes (regardless of tumor involvement) have less capability to proliferate compared with lymphocytes from peripheral blood. The p values were generated by comparing PBMCs (control and patients) with ALNs (tumor free and tumor involving) and are indicated in corresponding colors for B cells, CD8+ T cells and CD4+ T cells. Y axis indicates division index, the average number of divisions a cell has undergone. Medians are indicated by the bar in each scatter column. (b) Lymphocyte Proliferation in C57BL/6J wild type mouse. PBL: peripheral blood. LN: lymph node. (c) Proliferation data stratified by sentinel lymph node (SLN) versus non-sentinel lymph node (NSLN). (d) Proliferation data stratified by ER/PR/HER2 Status.

**Figure B8.****Figure B8. Immunohistochemical staining of Ki67 in lymph node sections.**

T cells: blue; B cells: brown; tumor cells: purple; Ki67 staining: red. The reactive lymph node (A) showed high Ki67 positivity, especially in germinal center (Ki67+ T cells: 5.52%; Ki67+ B cells: 31.5%). A tumor-invaded lymph node section from the disease free breast cancer patient (B) showing Ki67 positivity (Ki67+ T cells: 1.98%; Ki67+ B cells: 0.05%). Two TDLN sections from a recurrent patient are shown in C (tumor invaded lymph node) and D (tumor free lymph node). For both sections, there are less than five Ki67+ cells. Representative image for GemIdent to identify and enumerate cells is presented in pseudocolor in E. T cells are indicated by blue membrane staining. B cells are indicated by red membrane staining. Ki67 positivity is marked by green nuclear staining.

**Figure B9.****Figure B9. Immunohistochemical patterns of Ki67 expression in B cells.**

The images were presented in pseudo-colors: Ki67-green nuclear, B cells-red membrane, T cells-blue membrane. In reactive lymph nodes (A and B), most germinal center (GC) B cells are positive for Ki67, often clustered in the dark zone and surrounded by Ki67 negative B cells in the light zone. In breast cancer TDLNs, Ki67 positivity is lost from GC B cells, indicating a quiescent or suppressed GC. Representative image is shown in C. An example of Ki67 negative primary follicle (D) from breast cancer TDLNs was also shown to illustrate the architecture difference between a primary follicle and secondary follicle which contains a germinal center.

Personnel: Lee, Holmes, Johnson, Dirbas, Yu, Simons. **A third PhD postdoctoral fellow with expertise in bioinformatics, data integration and analysis would greatly enhance the success of this project.**

C. Analyzing the geometric relationships and interactions between cancer and immune cells in tumors and TDLN

Originally proposed in SOW:

1. Generate high-resolution images of tumor and TDLN sections. (months 0-60)
- B. Develop algorithms to identify cells/cell types and assign coordinates. (months 0-60)
2. Develop algorithms to assess the spatial arrangement and grouping of tumor and immune cells with respect to each other that may have biological significance. This will be done in collaboration with a Stanford mathematics professor, Dr. Doron Levy, using advanced image analysis and computational geometry techniques. (months 0-60)

Archived samples of tumor and TDLN from breast cancer patients with 1-13 years of clinical follow-up data are being analyzed. Tumor and immune cell markers are identified via immunohistochemical (IHC), immunofluorescence (IF) staining, and in-situ hybridization (ISH). Images are being acquired using a high-resolution, automated imaging system (Olympus and Ludl) with a special spectral imaging system (Nuance™). Acquired images are then analyzed with our custom image analysis software GemIdent. This software uses spatial statistics and machine learning algorithms to identify cells, cell types, and assign coordinates. We are also developing algorithms to assess the spatial arrangement and grouping of tumor and immune cells. By performing *in situ* analysis of tissue, our goal is to understand the mechanisms of cancer development by characterizing the spatial interactions between cell types. This is done in collaboration with Stanford statistics professor, Dr. Susan Holmes, who has expertise in novel image analysis and computational geometry techniques. Over 50 immune and tumor markers will, eventually, be assessed within tumor and TDLN sections.

Our key accomplishment in year 3 include optimization of 3- and 4-color IHC staining combinations to concurrently visualize breast cancer cells (via cytokeratin AE1/AE3) and various immune cells, such as CD4 and CD8 T cells, T-regulatory cells, mature and immature dendritic cells (DCs), and B cells within TDLN sections. Stained lymph node sections were scanned using a high-resolution, automated whole-section imaging system that is capable of un-mixing different spectra resulting from various chromogens and reconstructing the image in pseudo colors chosen by the researcher (Figure C1).

Another key accomplishment was the application of our new imaging system and custom image analysis software to quantify cell population size and analyze spatial relationships between various cell types in TDLN sections. All sub-images (200x magnification) that composed the whole lymph node section were analysed by our custom image analysis software, GemIdent, yielding a total number of cells for each stained phenotype and the Cartesian coordinates of each

cell identified. These data, collected via the multi-spectral approach and semi-automated features of the image acquisition and analysis tools, made various spatial statistical analyses possible.

We first tested the usefulness of our new imaging technology and analysis approach to analyze the relationship between CD1a+ DCs, CD3+ T cells and tumor cells via AE1/AE3 cytokeratin in triple-stained TDLN sections from 40 breast cancer patients aged 29-76 years, treated at Stanford University Medical Center between September 1995 and June 2003. One to thirteen-year clinical follow-up data were available for 34 of these patients. Fifty-two TDLN samples were randomly chosen, in order to avoid a biased sample set. Twenty-five of these nodes were tumor-involved, and the remaining 27 were tumor-free. Control nodes were obtained from 17 non-cancer patients. By using our novel image acquisition and analysis strategy in combination with R statistical program (Bivand, R., *et al*, 2001), we could generate heatmaps of the density of different cell types on lymph node sections (Figure C2). Using spatial regression statistical analysis (Ullah, A., and Giles, D.E.A., 1998), we found that the number of DCs within a region can be predicted from the number of T cells or tumor cells in the same lymph node. Furthermore, Moran's I spatial autocorrelation analysis (Moran, P.A., *et al*, 1948) showed that the degree of DC clustering in tumor-bordering regions correlated with disease-free survival (Figure C3). Representative images of low-versus-high DC clustering, imaged at 100x magnification from tumor-bordering regions of TDLNs from relapsed and disease-free patients are shown in Figure C4.

At present, we have also analyzed 26 axillary non-sentinel lymph node (NSLN) sections (15 tumor-positive and 11 tumor-negative) from 23 Stanford breast cancer patients, triple-stained using markers for T cells (CD3), B cells (CD20) and tumor cells (AE1/AE3 cytokeratin). One lymph node section per patient was analyzed from 23 of the patients, and two sections, a pair of tumor-positive and tumor-negative lymph nodes, were analyzed from each of the remaining 3 patients. Control nodes from 15 non-cancer patients have also been analyzed. We aim to increase our population size to at least 50 breast cancer patients from Stanford for each IHC staining combination.

Since all breast cancer patients included in our studies were adult females aged 29-76 year olds, and the non-cancer patients were males and females aged 1-70 year olds, we tested whether the age and gender differences affected the distribution of T and B cell proportions and ratios in the non-cancer samples. For these analyses, we divided the non-cancer patients into males and females, and into 2 age groups: 1-15 and 24-70 year olds, respectively. The results showed no significant difference in T and B cell proportions, as well as T:B ratio between the two age and gender groups (Figure C5a and C5b, respectively). Therefore, we included lymph node samples from all of the non-cancer patients in our subsequent analyses.

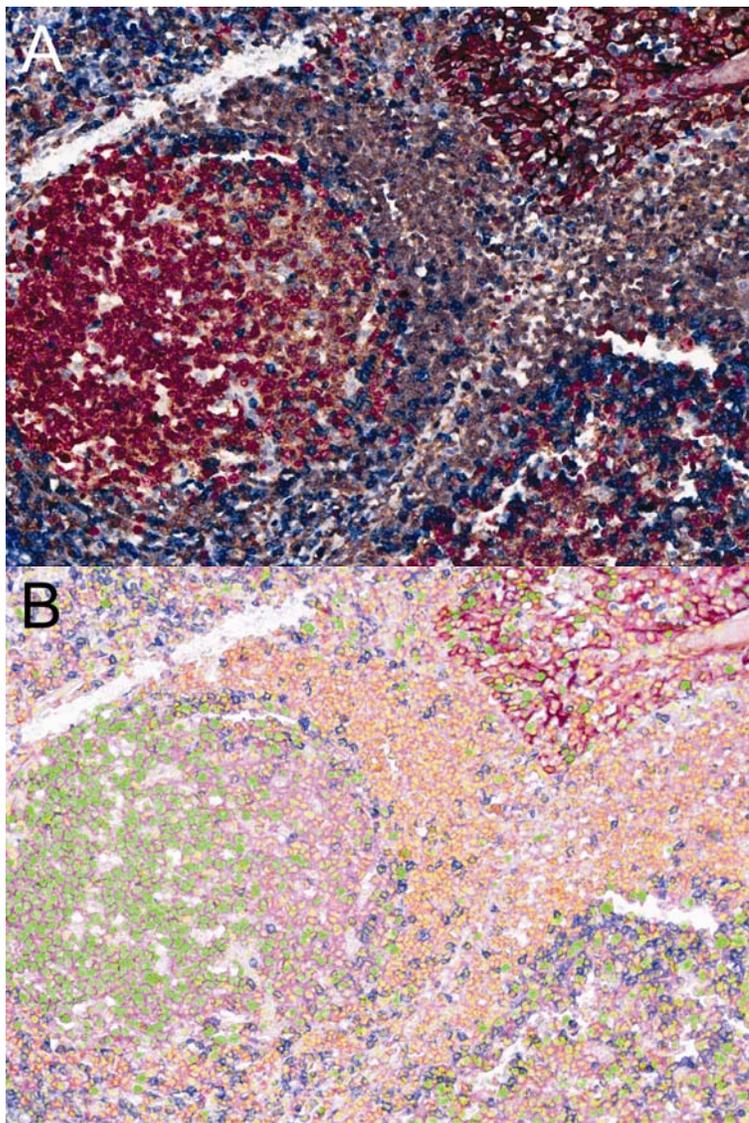
Our current result showed that B cell proportion was reduced in NSLNs from cancer patients, particularly in tumor-invaded lymph nodes (Figure C6a). The number of tumor cells was not included in the calculation of the total number of cells in tumor-positive NSLN sections to avoid distortion of numbers due to the large size of tumor cells. No significant difference was observed in the proportion of T cells from cancer patients compared to that from non-cancer patients (Figure C6b). We also observed a trend of increasing proportion of cell populations other than tumor, T or B cells in tumor-invaded NSLNs, although the difference is not significant ( $p > 0.05$ )

between the three groups of lymph nodes based on the current sample size (Figure C7a). The trend observed is partially due to an increased formation of fibrous-like cells that formed scar tissues in-between clusters of tumor cells, known as a desmoplastic reaction in tumor-invaded LNs (Figure C7b). Furthermore, we observed a strong trend in the ratio of T:B cells which was higher in NSLNs from breast cancer patients compared to the ratio in non-cancer patients, particularly in tumor-positive NSLNs (Figure C8).

To investigate whether the number of T cells and B cells could predict clinical outcomes, we divided our samples into those from disease-free (at least 3 years) and relapsed groups of patients. We found that the proportion of B cells in tumor-positive NSLNs was significantly lower than that in non-cancer LNs (Figure C9a). Although there is a general trend in the proportions of B cells were lower in all groups of NSLNs from breast cancer patients, the differences of values from groups other than the tumor-positive NSLNs were not significant compared to the non-cancers LNs. T cells' proportion was found to be higher in tumor-negative NSLNs from relapsed patients compared to that from disease-free and non-cancer patients (Figure C9b). Although the trend in the T:B ratio was higher in NSLNs from relapsed patients than the ratio in LNs from disease-free and non-cancer patients, a significant difference was observed only between T:B ratio in tumor-negative NSLNs and in LNs from non-cancer patients (Figure C10).

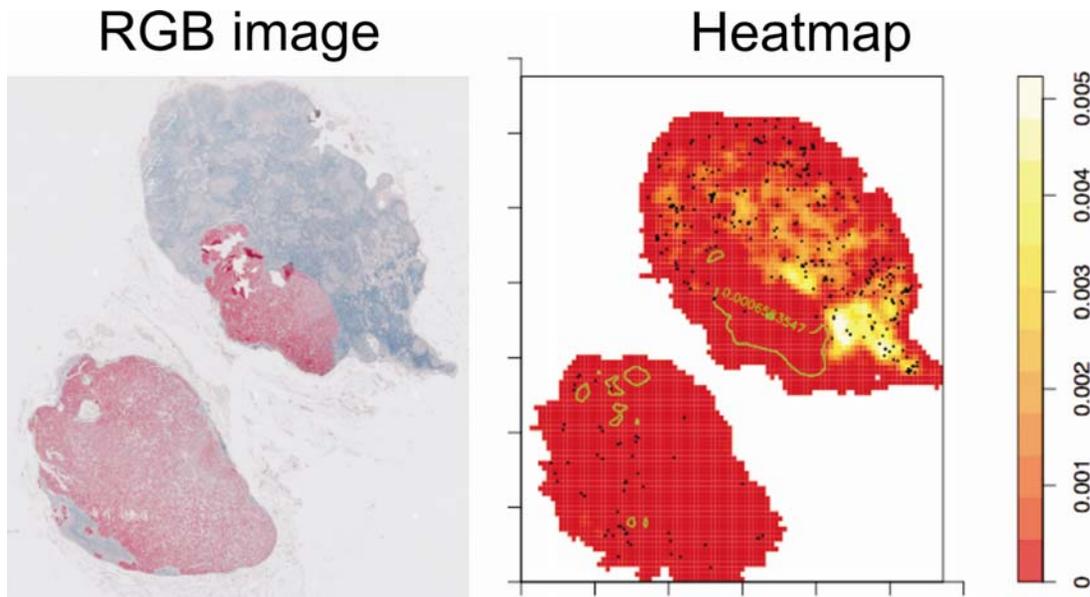
Our data showed that our novel technology is a powerful tool to study cell population sizes and spatial patterns between various cell types in tissue sections, and allows the investigation of the relationship between immunological profiles and clinical outcomes. In the near future, we plan to put together all the results obtained from our examination of tumor and immune cell populations in TDLNs to further investigate the mechanistic insights and propose novel mechanisms of immune dysfunction in cancer development. We also plan to explore any potential applications of specific immune cell profiles in TDLNs as novel prognostic tools for breast cancer.

**Figure C1.**

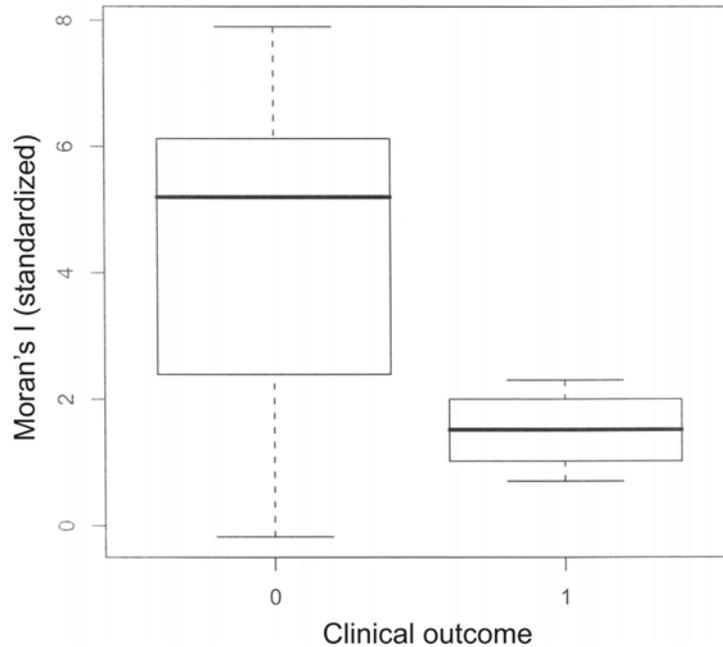


**Figure C1. Images of a TDLN cross section taken by Nuance™ at 200x magnification.**

A, The RGB image. Chromogens used were Vulcan Fast Red (Ki67 (cell proliferation marker), red), DAB (CD20(+)-B cells, brown), Ferangi Blue (CD3(+)-T cells, dark blue) and Bajoran Purple (cytokeratin (tumor), purple). Cellular nuclei were counterstained with hematoxylin (light blue). B, The reconstructed image with pseudo-colors that allowed a greater distinction of the cell populations as compared to the original image.

**Figure C2.****Figure C2. Heatmap representation of the density of different cell types on lymph nodes.**

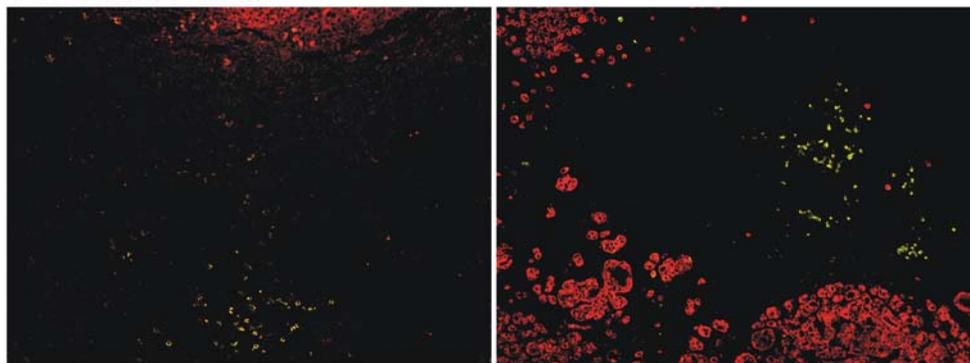
High tumor density areas are represented by green contours, high T cell density areas are represented by yellow and white contours, and the dendritic cells are represented by black dots.

**Figure C3.****Figure C3. Moran's I spatial autocorrelation analysis of dendritic cells.**

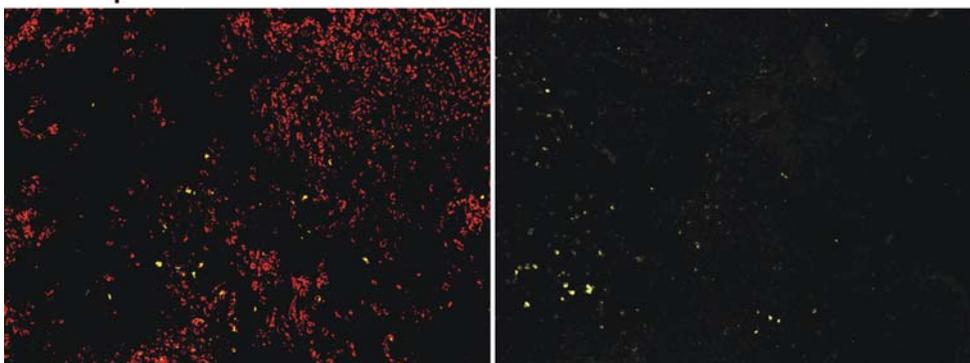
Dendritic cells in TDLNs from disease-free patients are more clustered compared to those from relapsed patients. Standardized Moran's I (y-axis) indicates the clustering index. Disease-free (n=7) and relapse (n=4) groups are indicated by clinical outcome "0 = disease free" and "1 = relapsed", respectively.  $p = 0.02$  between the two groups, as determined by Student's t-test.

**Figure C4.**

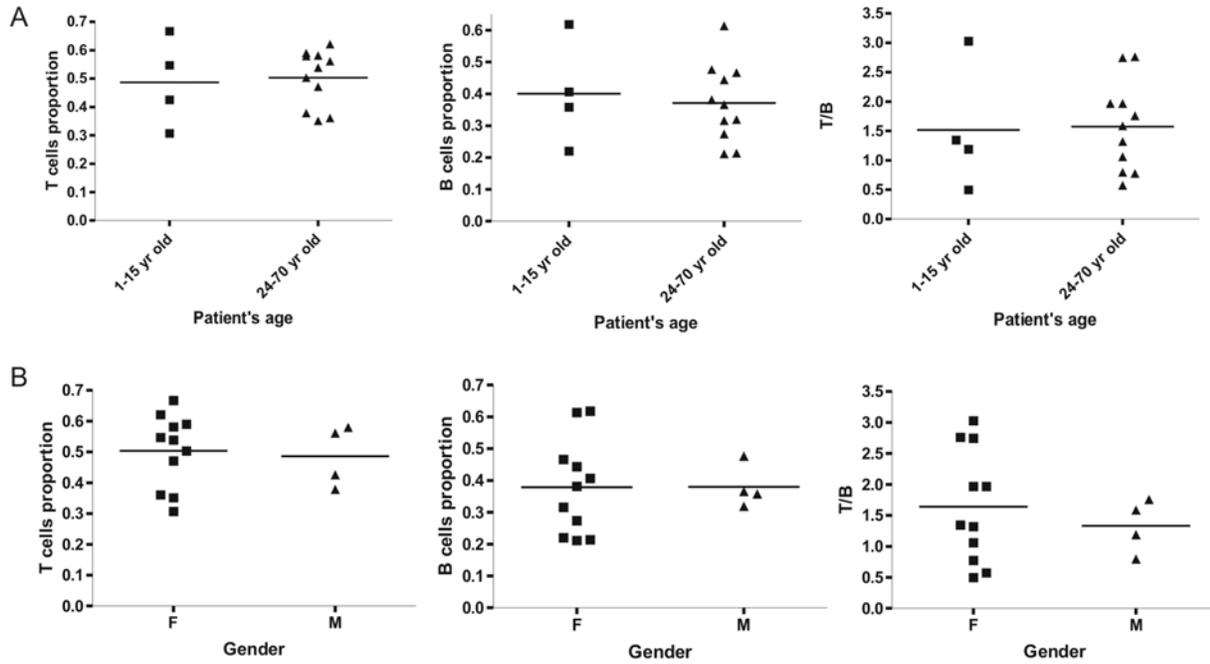
Disease-free



Relapsed



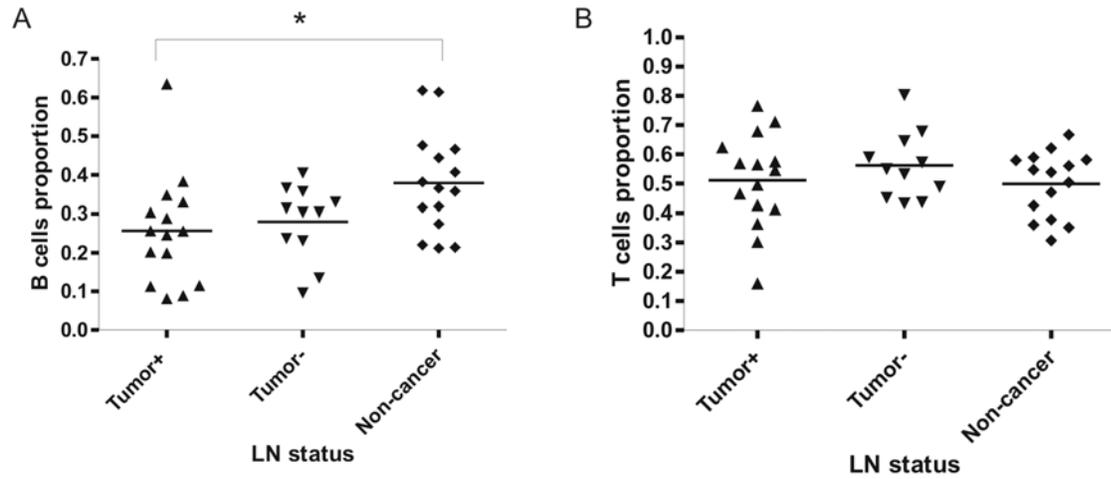
**Figure C4. Representative images of high versus low DCs clustering in TDLNs.** Images were taken at 100x magnification from mid-tumor density regions of TDLNs from disease-free and relapsed patients, respectively. DCs: bright yellow; tumor cells: red.



**Figure C5. Comparison of the distribution and T:B ratio in different age and genders**

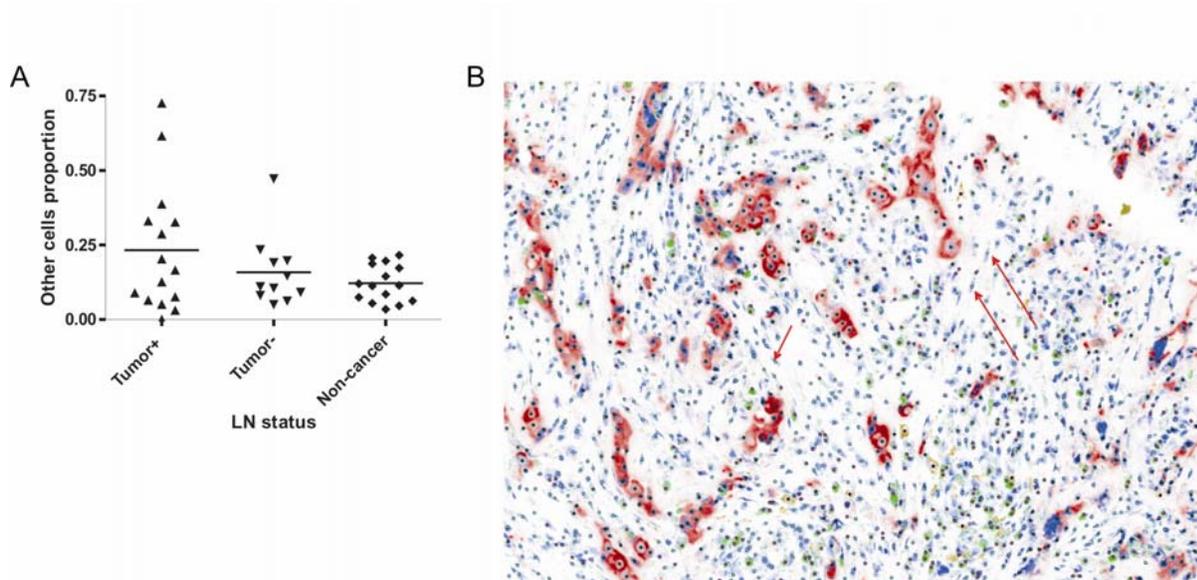
Different age and gender did not affect the distribution of T and B cells proportions, and T:B ratios in lymph node samples from non-cancer patients.

**Figure C6.**



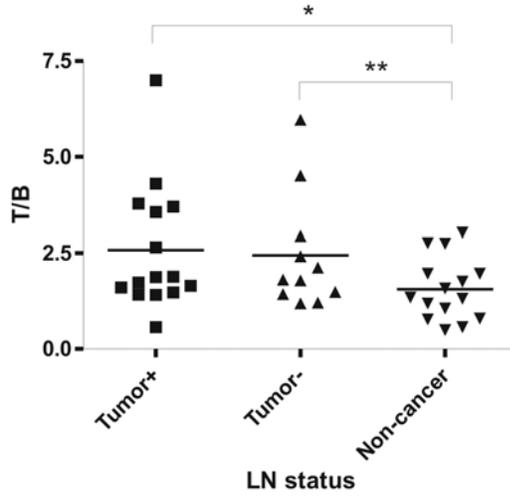
**Figure 6. Comparisons of proportions of T and B cells.**

Comparisons of proportions of T and B cells in tumor-positive and tumor-negative NSLNs from breast cancer patients, and in lymph nodes from non-cancer patients. *A*, B cell proportions. \* $p=0.01$ , as determined by Mann-Whitney test. *B*, T cell proportions.

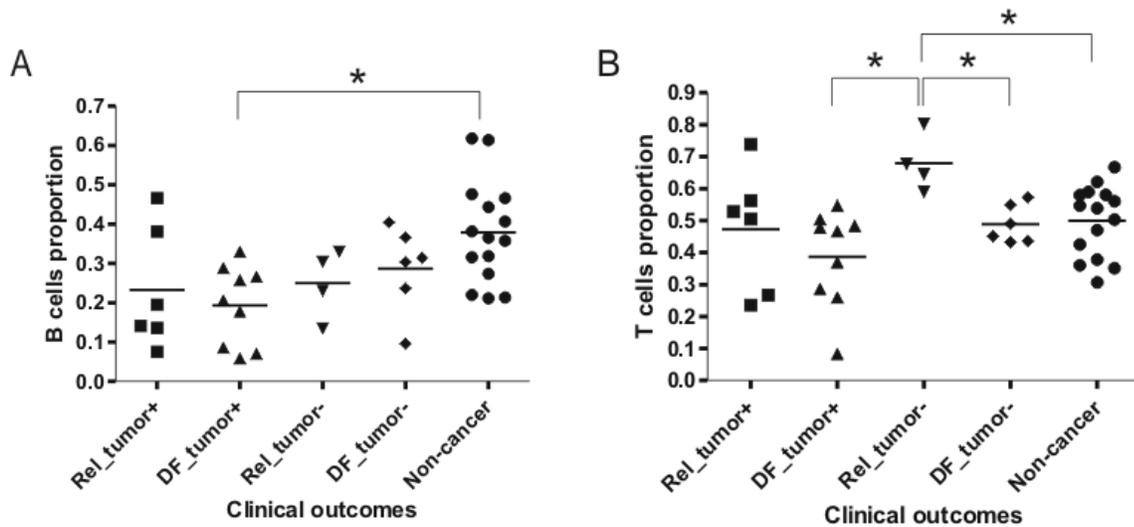
**Figure C7.**

**Figure C7. Proportion of “other cells” population increased in tumor-positive NSLNs.**

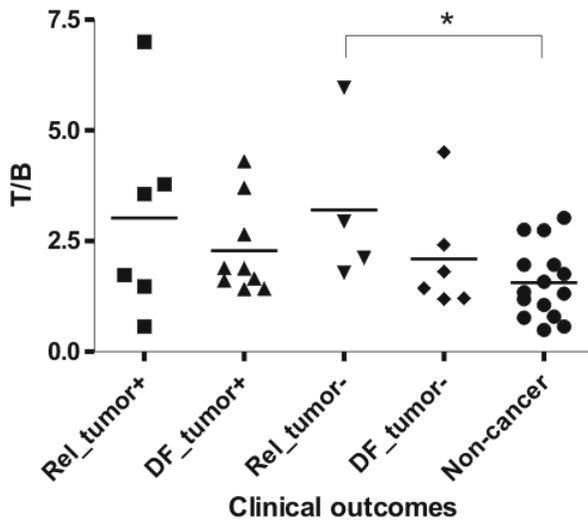
*A*, Comparisons of proportions of other cells in tumor-positive and tumor-negative NSLNs from breast cancer patients, and in lymph nodes from non-cancer patients. *B*, Desmoplasia in tumor-invaded lymph nodes. Arrows indicate fibrous-like cells commonly found between clusters of tumor cells (red membrane stains) in the lymph nodes.



**Figure C8. The ratio of T:B cells in breast cancer patients and healthy controls**  
 The ratio of T:B cells was higher in NSNs from breast cancer patients compared to the ratio in non-cancer patients. \*p=0.05; \*\*p=0.09, as determined by Mann-Whitney test.



**Figure C9. Comparisons of proportions of T and B cells**  
 Comparisons of proportions of T and B cells in tumor-positive and tumor-negative NSLNs from breast cancer patients with different clinical outcomes, and in lymph nodes from non-cancer patients. Rel=relapsed; DF=disease free. A, B cell proportions. \*p=0.002. B, T cell proportions, \*p<0.02 as determined by Mann-Whitney test.



**Figure C10. The ratio of T:B**

The ratio of T:B cells was higher in NSLNs from relapsed cancer patients compared to the ratio in disease-free and non-cancer patients. Rel=relapsed; DF=disease free. \*p=0.04, as determined by Mann-Whitney test.

Personnel: Lee, Holmes, Schwartz, Setiadi, Holmes.

D. Synthesizing a useful model of breast cancer through mathematical and computational modeling

Originally proposed in SOW:

To integrate our experimental data and observations into a mathematical model to address the dynamics of cancer cells and the immune response in the tumor and lymph node. This will ultimately enable us to perform *in silico* experiments to quickly test novel therapeutic strategies for breast cancer.

To better understand the impact of the immune response to breast cancer, we are first developing a mathematical model for the dynamics of the primary T cell response. The main focus of the model is on the T cell-mediated response and particularly on the function of adaptive regulatory T cells (iTregs). In our model, we consider two paradigms for the expansion and contraction of

the T cell response: 1) T cells expand and contract based on a proliferation program that is determined early on during initial T cell stimulation and 2) T cell expansion proceeds in an antigen-dependent manner, but contraction is induced by the appearance of iTregs, which suppress activated T cells.

We formulated mathematical models of both paradigms as systems of delay differential equations (DDEs), in which the time delays correspond to the duration of T cell divisions. Each equation corresponds to a population of cells that contribute to the development of the overall T cell response. In the models, we consider antigen-presenting cells (APCs), two types of T cells (conventional and regulatory), and antigen stimulation. In addition, we separate the dynamics into two separate compartments, the lymph node and the tissue. The motivation of this paper is to understand the mechanisms of immune regulation, whether primarily by intrinsic T cell programs or by external regulation via iTregs.

In the paper, we compared the two modeling paradigms and concluded that the mechanism of regulation by T cell programs alone is not sufficient to explain the robustness of the immune response to highly variable initial conditions, in particular the potential variability of the precursor frequencies of antigen-specific T cells from patient to patient or even from day to day. In addition, we concluded that T cell dynamics regulated by the appearance of iTregs later during the course of the immune response leads to more robust and realistic dynamics. Hence, we hypothesize that regulatory T cells, specifically iTregs that differentiate from stimulated effector cells, serve to control the duration and magnitude of the T cell response.

This hypothesis implies that T cell responses are greatly limited in size and duration by the appearance of iTregs. For this reason, even the addition of high antigenic stimulus or stimulation over an extended period of time does not significantly increase the efficacy of the T cell response. These results imply that attempts to induce or enhance an anti-cancer T cell response in a patient must consider the potential impact of regulatory T cells on the efficacy of the vaccination strategy.

Hence, we propose that T cell vaccines, particularly against cancer, must either be given at optimal time points to minimize the negative impact of regulatory T cells or that the differentiation and proliferation of regulatory T cells must be inhibited temporarily during the vaccination treatment.

Continuing on a previous paper investigating the dynamics of naturally-occurring regulatory T cells, our current model involving iTregs also shows that T cell dynamics proceed in two phases. In the first phase, CD8+ cells remain sequestered in the lymph node during a period of rapid proliferation. During the second phase, the CD8+ population emigrates to the tissue, where it quickly destroys the target population. Furthermore, the transition between the two phases is mediated by iTregs as also concluded in the previous study.

During the emigration phase, regulatory cells suppress activity in the lymph node by suppressing activated T cells. This change in environment causes the majority of T cells to stop dividing and start emigrating to the tissue, where they begin to destroy target cells. Since iTregs appear later in the immune response, there is an adequate delay between the initiation of the conventional T cell response and the regulatory T cell response that allows conventional T cells to proliferate sufficiently before emigrating to the tissue. Much later in the immune response, regulatory T cells in the lymph node also emigrate to the tissue and suppress the remaining T cells that are lingering after target elimination. The two-phase process results in an effective immune response if the transition between phases occurs within an appropriate time window when the conventional T cells have expanded enough to be effective but not too much to potentially induce a collateral impact on normal tissue cells.

Although normally functioning regulatory cells appropriately control the duration and magnitude of an immune response, in tumor microenvironments, the chronic presence of regulatory cells serves to weaken existing immune responses against tumor cells. As a net step, we plan to incorporate tumor cells and the tumor microenvironment into our models of T cell dynamics.

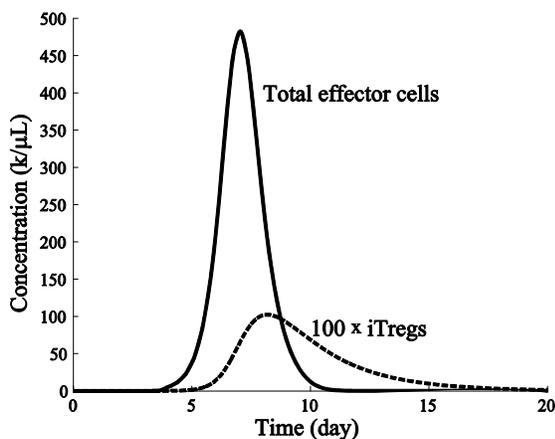


Figure D1. Time evolution of effector and iTreg populations over time. The peak of the iTreg response roughly coincides with the peak of the T cell response, but the iTreg response decays slower.

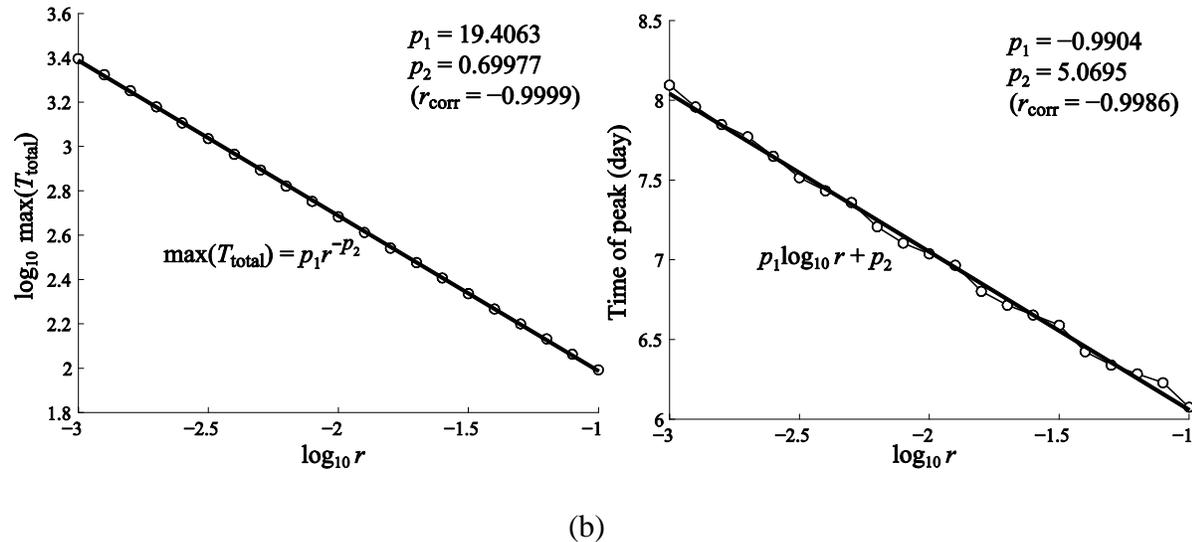


Figure D2. Log-log plots of the dependence of T cell dynamics on  $T_0(0)$ , the initial concentration of naïve T cells. (a) Maximum T cell expansion level versus  $T_0(0)$ . The linear regression shows that the maximum expansion level is roughly proportional to  $T_0(0)^{1/3}$ . The linear correlation  $r_{\text{corr}} = 0.9987$ . (b) The time of T cell peak versus  $T_0(0)$ . The linear correlation  $r_{\text{corr}} = -0.9984$ .

Personnel: Lee, Levy, \*Kim. \*Kim was a PhD graduate student in mathematics who was involved in this project. He has completed his PhD and has moved on to a postdoctoral position at University of Utah, where he still collaborates on this project at no cost to the award. **A third PhD postdoctoral fellow with expertise in bioinformatics, data integration and analysis would greatly enhance the success of this project.**

#### Overall Personnel

1. Peter P. Lee, MD – project PI (50% effort on EHSA).
2. Erich Schwartz, MD, PhD – Stanford Pathology (no salary requested on EHSA).
3. Denise Johnson, MD and Fred Dirbas, MD – Stanford Surgical Oncology (no salary requested on EHSA).
4. Susan Holmes, PhD – Stanford Statistics (1 month per year, as 33% of 3-month summer period).

5. HongXiang Yu, PhD - post-doc 1, 100% effort on EHSA – immunological and microarray studies.
6. Ning Yan, PhD - post-doc 2, 100% effort on EHSA – data analysis.
7. Diana Simons - research assistant 1, 100% effort on EHSA – to aid in immunological, histology, and microarray studies.

#### KEY RESEARCH ACCOMPLISHMENTS:

- Recruited over 170 breast cancer patients into this study – acquired tumor, TDLN, and blood samples for analyses.
- Optimized methods for analysis of fresh and archive samples by flow cytometry, function assays, and DNA microarray analysis to study immune and tumor cells within tumor and TDLN specimens.
- Demonstrated a defect in IFN signaling in peripheral blood lymphocytes from breast cancer patients, melanoma patients, and GI patients.
- Identified spatial patterns of breast cancer and immune (dendritic cells) within TDLNs which predict clinical outcome and may provide mechanistic insights.
- Identified gene expression patterns within TDLNs which show blunted proliferation of immune cells and may lead to mechanistic insights.

#### Outline of the project plan for the next 12 months

- Continue recruiting patients into study and acquiring samples.
- Continue functional assays of lymphocytes from tumor, TDLNs, and peripheral blood.
- Complete microarray analysis of patient sample sets. Each set includes tumor cells, tumor infiltrating immune cells, immune cells from TDLN, and immune cells from blood.
- Analyze IFN signaling in lymphocytes from TDLNs and tumor, as compared to those in peripheral blood. Determine extent of signaling abnormalities and potential mechanism of IFN signaling defect in breast cancer.

REPORTABLE OUTCOMES: Three manuscripts arising from this work have been recently submitted for publication.

1. Critchley-Thorne RJ, Simons DL, Yan N, Dirbas FM, Johnson DL, Swetter SM, Carlson RW, Fisher GA, Koong A, Holmes SP, Lee PP. Impaired Interferon Signaling is a Common Immune Defect in Human Cancer. Submitted.
2. Kim PS, Lee PP, Levy D. Regulatory T Cells Produce a More Robust Primary T Cell Response. Submitted.
3. Setiadi AF, Kohrt HB, Levic E, Johnson D, Schwartz E, Holmes SP, and Lee PP. Quantitative, Architectural Analysis of Tumor-Draining Lymph Nodes in Breast Cancer. Submitted.

## CONCLUSIONS:

In year 3, we built upon the foundation from the first two years of this award, and are making progress in multiple areas of this project. We now have an efficient system in place to recruit patients into this study and procure their samples. Limited numbers of subjects available and limited amounts of clinical materials available from each subject remain major challenges to the success of this project – we continually attempt to address and solve this issue by reducing the cell numbers that we need for each assay. We have developed a powerful set of immunological assays and molecular tools to study these samples in greater detail than previously possible. We continue to uncover dramatic changes in the immune cell populations within tumors, TDLNs, and peripheral blood from breast cancer patients. These findings are reported above, and have led to 3 manuscripts under review. We look forward in the coming year to build upon this momentum and specifically to elucidate insights into the immunobiology of breast cancer. In the coming year and beyond, we will begin to focus on translating our early findings into novel therapeutic strategies for the immunotherapy of breast cancer.

## REFERENCES:

Critchley-Thorne, R., *et al.* Down-Regulation of the Interferon Signaling Pathway in T Lymphocytes from Patients with Metastatic Melanoma. *PLoS Medicine* 2007; 4:0897

Murray, P.J., *et al.*, The JAK-STAT signaling pathway: input and output integration. *J Immunol.* 2007 Mar 1;178(5):2623-9.

Bivand, R., et al, More on Spatial Data Analysis. *R News* **1**, 13-17 (2001).

Ullah, A., Giles, DEA. *Handbook of Applied Economic Statistics*, (CRC Press, New York, 1998).

Moran, P.A. The interpretation of statistical maps. *Journal of the Royal Statistical Society* **10**, 243-251 (1948).

MAQC Consortium. The MicroArray Quality Control (MAQC) project shows inter- and in-trapplatform reproducibility of gene expression measurements. *NATURE BIOTECHNOLOGY*, 24(9):1151{1161, Sep 2006.

APPENDICES: None at this time.

SUPPORTING DATA: Tables and figures are integrated into the text above