AWARD NUMBER: W81XWH-11-1-0814

TITLE: Development of Technologies for Early Detection and Stratification of Breast Cancer

PRINCIPAL INVESTIGATOR: David R. Walt, PhD

CONTRACTING ORGANIZATION: Tufts University
Boston, MA 02111

REPORT DATE: December 2016

TYPE OF REPORT: Final

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

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
December 2016	Final	09/29/2011 – 09/28/2016

4. TITLE AND SUBTITLE
Development of Technologies for Early Detection and Stratification of Breast Cancer

5a. CONTRACT NUMBER
W81XWH-11-1-0814

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHOR(S)
Dr. David Walt, david.walt@tufts.edu;
Dr. Daniel Chiu, chiu@uw.edu;
Dr. Charlotte Kuperwasser, charlotte.kuperwasser@tufts.edu;
Dr. Gail Sonenshein, gail.sonenshein@tufts.edu;
Dr. Rachel Buchsbaum, rbuchsbaum@tuftsmedicalcenter.org

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Tufts University	20 Professors Row Medford, MA 02155
Tufts University School of Medicine	136 Harrison Ave. Boston, MA 02111
University of Washington	4333 Brooklyn Ave. NE Seattle, WA 98195
Tufts Medical Center	800 Washington St. Boston, MA 02111

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

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

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The overall goal of this work was to develop ultra-sensitive detection techniques to identify a panel of new biomarkers and indicators with diagnostic and predictive value in breast cancer. We successfully created several Single Molecule Array (Simoa) assays which were utilized to determine biomarker concentrations in breast cancer patients. We screened multiple biomarkers in breast cancer patient and healthy control samples and used additional markers, such as mtDNA, to improve the sensitivity and specificity of these assays. We used PSA to demonstrate that biomarkers can be measured at ultrasensitive levels within serum using Simoa prior to tumor formation in a mouse model. We successfully achieved full quantification of protein molecules within single cancer cells and discovered that highly cultured cells exhibited different protein expression levels than low-passage cells. Integration of ensemble Decision Aliquot Ranking (eDAR) with Simoa was accomplished to increase the throughput of the overall system.

15. SUBJECT TERMS
Single molecule detection, cancer biomarkers, ultra-sensitive protein assays, single cells, human-in-mouse model, miRNA, circulating tumor cells

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	58	19b. TELEPHONE NUMBER *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

Page

Table of Contents……..4

Statement of Work……...5

Introduction……..……..5

Body…….……...6

Key Research Accomplishments……..50

Reportable Outcomes……52

Conclusion……..54

References…….……54

Appendix 1: Publications…….……56

Appendix 2: Project Personnel…….……...……57

Statement of Work

The overall goal of this work is to develop ultra-sensitive detection techniques in order to identify a panel of new biomarkers and indicators with diagnostic and predictive value in breast cancer. During years 1 and 2 we will develop ultra-sensitive detection techniques and apply them to identifying prospective biomarkers using the Human-in-Mouse (HIM) model of breast cancer. During years 3-5 we will extend the findings in the HIM model to validate prospective biomarkers in human subjects with breast cancer. This work is broken down into specific tasks by investigator as follows:

INTRODUCTION

Despite the recent advances made in breast cancer diagnostics and treatment, in 2013 the United States is estimated to diagnose approximately 230,000 new cases of invasive breast cancer, resulting in approximately 40,000 deaths. Mammography is a powerful imaging technique for tumor detection; however, it lacks the ability to decipher benign from cancerous tumors, is unable to detect tumors smaller than 1 mm, misses approximately 20% of breast cancers potentially present at the time of screening, and has an 8-10% false positive rate. These drawbacks lead to inaccurate patient diagnosis, which can allow potentially fatal disease progression, or in the cases of over-treatment, unnecessary physical and emotional trauma. ELISA, the most common immunoassay for measuring proteins from breast tumors, excised samples, and serum, has a lower detection limit of ~1-10 pM, which is not sensitive enough to measure low abundance proteins, RNAs, and other biomarkers that could aid in the early and reliable diagnosis of cancer. There is a strikingly clear need to develop techniques capable of detecting biomarkers specific for breast cancer that will enable earlier diagnosis of disease, prediction of patient outcome, and improve therapeutic efficacy in a non-invasive manner. Our goals are to utilize ultrasensitive single molecule techniques developed in our laboratory to discover new biomarkers that meet these requirements within serum so that a simple blood test can be implemented. We are also working to characterize breast cancer biopsy samples with single cell resolution to discover the nature of the underlying heterogeneity in complex cell populations with the goal of correlating disease outcome with genotypes and phenotypes of individual cells.

David R. Walt, PhD, *Tufts University, Department of Chemistry, 62 Talbot Ave, Medford, MA 02155*

Task 1. Develop single molecule diagnostics using HIM (Human-in-Mouse) model (Years 1-2).

1a. Select approximately ten candidate markers (mt DNA, proteins, stem cell markers, etc.) in conjunction with collaborators (Months 1-6).
This task is complete. See Task 1b for the list of developed assays for protein and nucleic acid biomarkers.

1b. Develop single molecule assays for the candidate markers selected in 1a. Assays will be developed on microspheres and tested first on standard samples. Assay performance characteristics will be determined to ascertain that they will address concentration ranges of interest and requisite precision. For general protocols see: Nature Biotechnology, 2010, 28, 595-599 (Months 3-15)
In conjunction with collaborators, and based on recent literature reports, we have selected a list of candidate biomarkers. Overall progress on protein biomarker assay development is summarized in **Table 1**. Table 1 compares the assay sensitivity of commercially available ELISA kits and Simoa assays for each marker. Additionally, PSA is used as a model biomarker for mouse model and single cell studies since PSA has an established Simoa.

Table 1. Overall progress summary of breast cancer biomarker assay development. (LOD: Limit of Detection)

Biomarker	ELISA LOD (pg/mL)	SiMoA LOD (pg/mL)	Fold Improvement
BORIS	78	10	8x
FOXC2	156	13.8	11x
CRYAB	150	11	13x
ADAM8	62.5	3.2	20x
SNAIL	156	5.9	26x
SLUG	300	8.4	36x
ADAM12	125	3	40x
PR	12.5	0.3	42x
ERα	15.6	0.3	50x
HER2	14.8	0.3	50x
NEDD9	1000	16.1	63x
LCN2	40	0.42	100x
CA 15-3	1 U/mL	0.01 U/mL	100x
CDKN2D	10	0.1	100x
HSP70	125	0.8	156x
CYR61	3.8	0.02	190x
NRP1	31.2	0.136	229x
TBX3	116	0.24	480x
HIF1α	50	0.015	1000x
CA 19-9	5 U/mL	0.005 U/mL	1000X
VEGF	31.2	0.03	1000x

Multiplexing assays
Multiplex assays allow for in-depth analyses of blood samples without using larger sample volume, facilitating a more high-throughput approach with minimal sample material. Dye-encoded magnetic capture beads are used to differentiate up to 10 different bead types, representing 10 different targets. Several breast cancer protein biomarkers have been selected and tested for multiplex assay development. Overall progress on multiplex assay development is summarized in **Figure 1**.

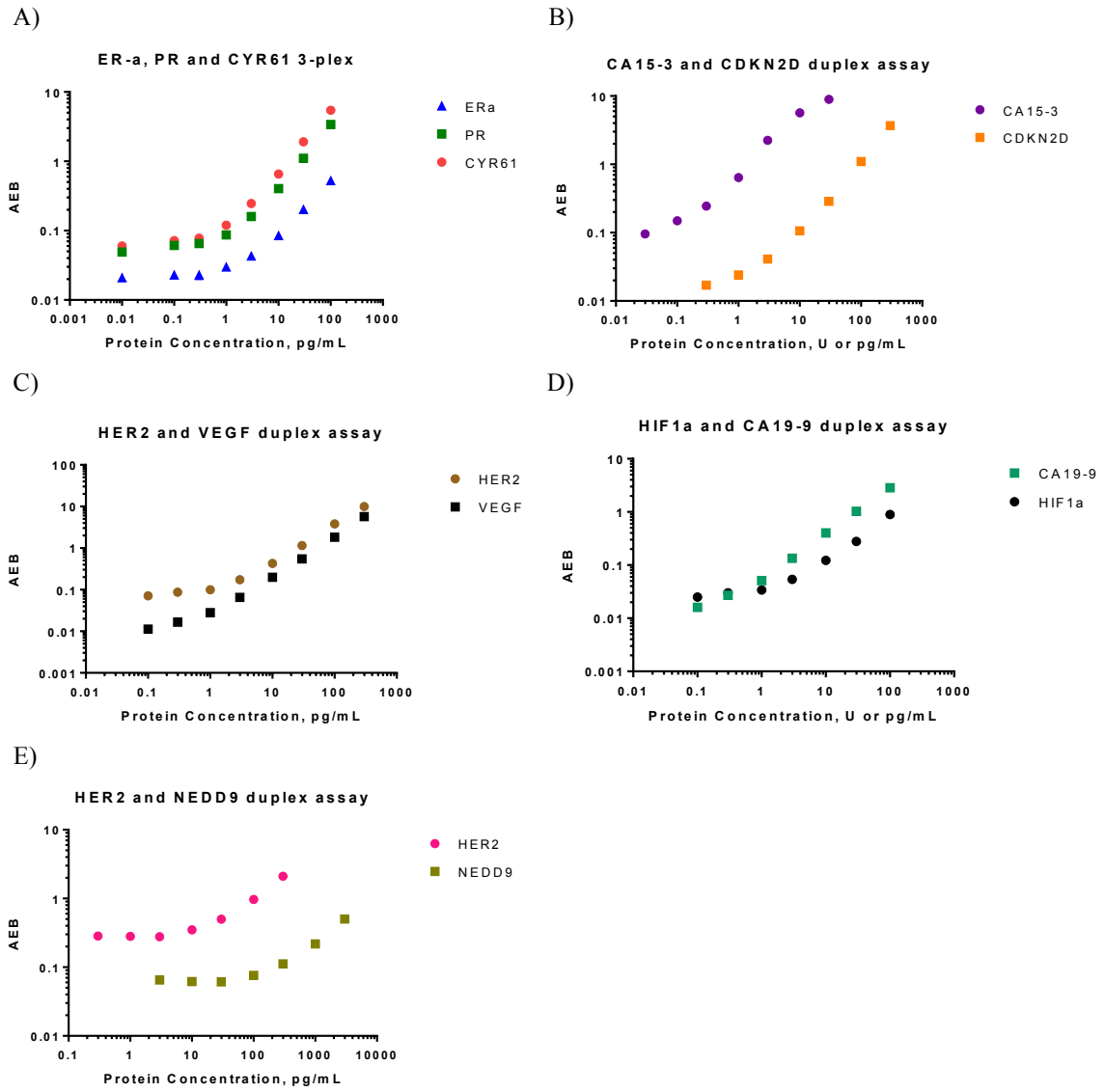


Figure 1. Plots of average number of enzyme per bead (AEB) against concentration of protein spiked into bovine serum **A**. Panel #1: ER-α, PR and CYR61 3-plex assay with LODs of 0.343 pg/mL, 0.231pg/mL, and 0.019pg/mL, respectively. **B**. Panel #2: CA15-3 and CDKN2D duplex assay with LODs of 0.030 U/mL and 0.39 pg/mL, respectively. **C**. Panel #3: HER2 and VEGF duplex assay with LODs of 0.030 U/mL and 0.39 pg/mL, respectively. **D.** Panel #4: HIF1 α and CA19-9 duplex assay with LODs of 0.208 pg/mL and 0.051 U/mL, respectively. **E.** Duplexed detection of HER2 and NEDD9. The LOD for HER2 was 0.59 pg/mL, while the NEDD9 plex had an LOD of 37 pg/mL.

Detection of microRNA as a biomarker for early stage breast cancer

MicroRNAs are involved in virtually every biological pathway, and thus represent a very rich source of biological information. However, detection of miRNA is traditionally challenging to achieve via conventional detection methods like quantitative reverse transcription PCR. We have used the Simoa platform to detect microRNA with high sensitivity and specificity.

As previously reported, we demonstrated microRNA detection was effective using a sandwich assay approach. In this approach, we used capture and detector probes that are partly composed of locked

nucleic acids (LNA) in place of natural DNA to provide higher stability and specificity. More recently, we have expanded this approach to demonstrate multiplex assays for microRNAs as well. First, we tested duplex assays with hsa-miR-21 and hsa-let-7a microRNAs or hsa-miR-141 microRNAs.

To study the expression levels of multiple miRNAs simultaneously, we developed multiplex assays using optimized conditions identical to those of the singleplex assays. To enable multiplexing, we used paramagnetic beads labeled with fluorescent dyes with different wavelengths and degrees of labeling to produce distinct subpopulations of beads. These encoded beads were then further labeled with capture probes for specific miRNAs (hsa-miR-21 and hsa-let-7a). By incubating with specific biotinylated detectors, we measured two miRNAs simultaneously with an average limit of detection of approximately 1 fM.

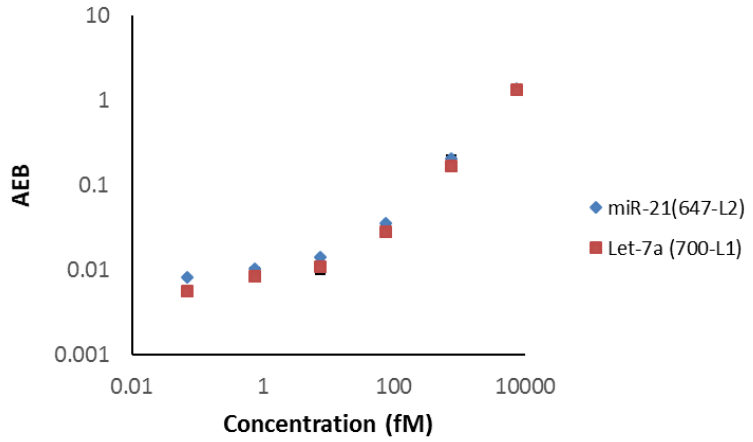
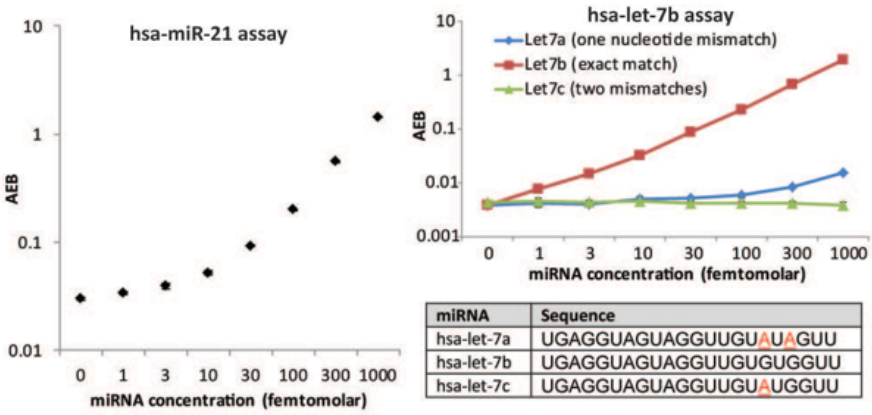


Figure 2. Multiplex assay for hsa-miR-21 and hsa-let-7a. Each measurement was conducted in the presence of equimolar amounts of the miRNA species. AEB=average enzyme per bead.

To confirm the reliability of our previously developed Simoa microRNA assays, we tested the detection of several microRNA using freshly-prepared reagents. Our results confirmed that the detection of miR-21 could be carried out with a limit of detection in the range 1 – 3 fM. By using the let7 family of miRNA to detect RNA targets that differed by only one or two single nucleotide mismatches, we confirmed that the assay was highly selective for the correct target. These results are well-matched with our previously reported assay performance, suggesting that the miRNA assays are reproducible from batch to batch of reagents. Although our previous results were obtained using a two-temperature hybridization protocol, the similar sensitivity and specificity reported here were obtained using one-temperature hybridization at 45 °C, suggesting that the two-temperature protocol is not strictly necessary for optimum assay performance.



miRNA	Sequence
hsa-let-7a	UGAGGUAGUAGGUUGUAUAGUU
hsa-let-7b	UGAGGUAGUAGGUUGUGUGGUU
hsa-let-7c	UGAGGUAGUAGGUUGUAUGGUU

Figure 3. Left: The hsa-miR-21 assay reliably demonstrates sensitivity to concentrations down to 1-3 fM. Right: Specificity of the hsa-let-7b assay. Signal was relatively low even for mismatch targets that differed by only one or two nucleotides from the correct target.

Using these conditions, we successfully tested assays for hsa-let-7a-5p, hsa-let-7b-5p, hsa-let-7c-5p, hsa-miR-16-5p, hsa-miR-21-5p, hsa-miR-25-3p, hsa-miR-126-3p, and hsa-miR-141-3p. All assays had LODs in the range 1 – 10 fM.

Detection of mutations in mitochondrial DNA

The Sonenshein lab has identified multiple mutations in mitochondrial DNA (mtDNA), but PCR does not have the sensitivity required to detect some of these mutations in small volumes of murine serum (provided by the Kuperwasser lab). The Simoa platform was applied to detect low abundance mutations of the target mtDNA sequences. A hybridization and ligation-based approach is employed as follows: (1) the target mtDNA sequence binds to complementary DNA capture and biotinylated detection sequences, forming a sandwich complex. If the mutation of interest is present, the hybridization is a perfect match. (2) On-bead ligation is performed; only the perfectly matched sequences are ligated. (3) Target mtDNA is washed away, leaving a labeled DNA sequence for subsequent detection via enzymatic readout. This process is illustrated in Figure 4. Target, capture, and detection sequences are shown in Table 2.

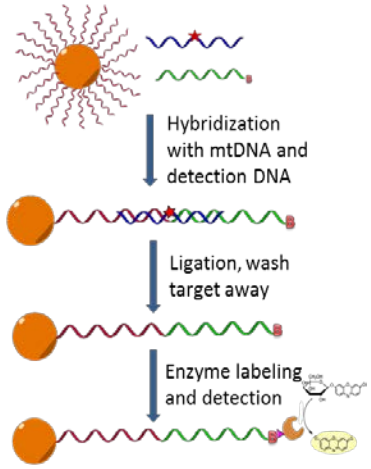


Figure 4. Hybridization and ligation- based detection of SNP-mtDNA.

Table 2. Sequences used in the hybridization-based mtDNA assay*

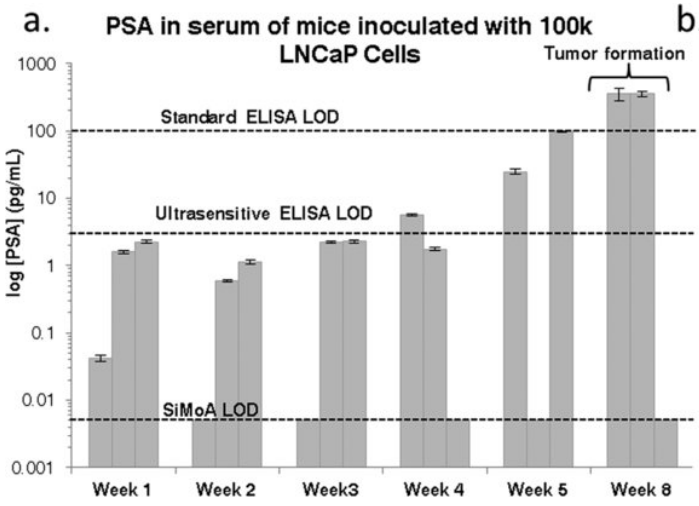
Sequence Name	Nucleotide Sequence
MT-CYB 14766 C>T Target	5'-ATGACCCCAATACGCAAAATTAACCCCCTAATAAAATTA-3'
MT-CYB SNP 14766 C>T Capture	5'-/AmMC12/TAATTTTATTAGGGGGTTAA-3'
MT-CBB 14766 C>T Detection	5'-/5Phos/TTTTGCGTATTGGGGTCAT/3Bio/-3'

* AmMC12 indicates amine modification with a C12 chain for attachment chemistry, and Phos refers to a phosphate modification to facilitate ligation between the phosphate group and a neighboring hydroxyl group.

These sequences were tested in a sandwich hybridization protocol without the ligation, which yielded a LOD of 16 fM. This approach afforded femtomolar sensitivity, but no specificity to distinguish between the wild type and SNP sequences. Results did not improve with the incorporation of a ligase.

1c. Screen blood samples from 'Human-In-Mouse' (HIM) Model obtained from Kuperwasser (Months 12-24). HIM tumors will be created from human breast epithelial cells collected from discarded tissues of women who have undergone reduction mammoplasty surgeries. For more information/details on model: Nature Protocols, 2006, 1, 595-599.

We successfully created a mouse model using low inoculums of LNCaP cells to study PSA in mouse serum over time. **Figure 5** and **Figure 6** shows results from mice injected subcutaneously with 100k and 10k LNCaP cells. The bar graph shows an increase in PSA concentration over time from week one to week five. The mice were sacrificed after eight weeks, which was when palpable tumors were observed in two of the three mice. PSA levels at week eight were significantly higher than previous measurements in the mice with tumors.

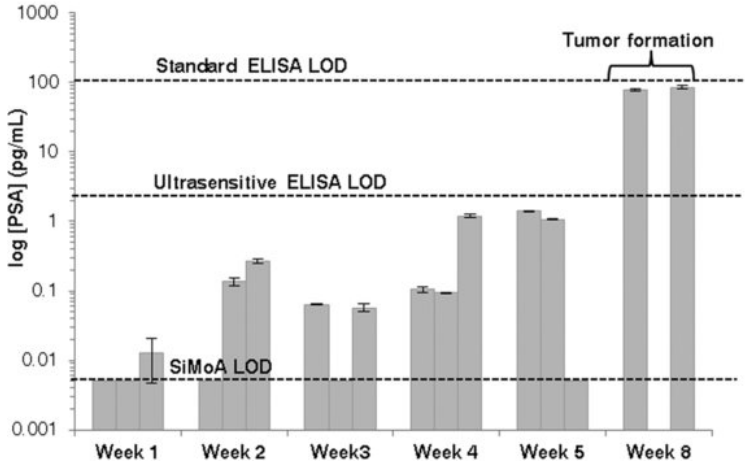


b. PSA in Serum of Mice Inoculated with 100k LNCaP Cells

Average Mouse Serum [PSA] (pg/mL)	Sample 1	Sample 2	Sample 3
Week 1	0.04 ± 0.01	1.59 ± 0.08	2.25 ± 0.11
Week 2	Below LOD	0.59 ± 0.02	1.12 ± 0.08
Week 3	Below LOD	2.21 ± 0.06	2.27± 0.10
Week 4	5.67 ± 0.14	1.75 ± 0.08	Below LOD
Week 5	24.9 ± 2.0	Below LOD	98.2 ± 2.1
Week 8	354 ± 70	357 ± 30	Below LOD

Figure 5. A. Bar graph of measured PSA concentrations in the serum of mice inoculated with 100k LNCaP cells. All measurements represent serum from terminal bleeds of individual mice. An average increase in PSA concentration is displayed over time. Three mice were examined for each time point, but several samples were below the 0.0109 pg/mL detection limit for the assay. Values for each measurement are shown in **B**. Three healthy control mice were also tested and were all below the assay LOD.

a. PSA in serum of mice inoculated with 10k LNCaP Cells



b. PSA in Serum of Mice Inoculated with 10k LNCaP Cells

Average Mouse Serum [PSA] (pg/mL)	Sample 1	Sample 2	Sample 3
Week 1	Below LOD	Below LOD	0.01 ± 0.01
Week 2	Below LOD	0.14 ± 0.02	0.27 ± 0.02
Week 3	0.06 ± 0.01	Below LOD	0.06 ± 0.01
Week 4	0.10 ± 0.01	0.09 ± 0.00	1.20 ± 0.06
Week 5	1.40 ± 0.01	1.06 ± 0.01	Below LOD
Week 8	79.2 ± 3.5	No Sample	85.8 ± 5.9

Figure 6. A. Bar graph of measured PSA concentrations in the serum of mice inoculated with 10k LNCaP cells. All measurements represent serum from terminal bleeds of individual mice. A general increase in PSA concentration is displayed over time. Three mice were examined for each time point, but several samples were below the 0.0109 pg/mL detection limit for the assay. Values for each measurement are shown in **B**. Three healthy control mice were also tested and were all below the assay LOD.

In general, PSA concentration increased over time, with a large increase at week eight, where tumor formation occurred. Several samples were below the LOD, but measurements for at least two mice were recorded for each time point. All measured PSA concentrations, including those with tumor formation, were well below the detection limit of standard ELISA. Through this study, we demonstrated that Simoa can be used to measure biomarkers in serum prior to tumor formation. This work has been published.[1] Since this study proved that our technique is valid for the early detection of tumors, we chose to pursue the study of clinical samples in lieu of developing more mouse models since this will yield more meaningful data as an end result.

Similar to the PSA mouse model study, the Kuperwasser lab created a ductal carcinoma in situ (DCIS) mouse model using the MCF10DCIS.COM cell line to determine if tumor progression could be monitored by measuring breast cancer biomarker candidates. Preliminarily, eight mice were injected with 100k cells and terminal bleeds were taken from a mouse once a week for eight weeks. We received the mouse serum samples from the Kuperwasser lab and tested Simoa protein assays with the samples. Tumors were palpable at the beginning of week 2 and were surgically removed at week 5. Three mice were monitored for recurrence after tumor removal but tumor formation was not observed. In conclusion, Simoa protein assays were able to detect the most protein markers in the samples. However, due to the mouse-to-mouse variation, it was hard to draw a conclusion about the pattern of the protein levels (**Figure 7**).

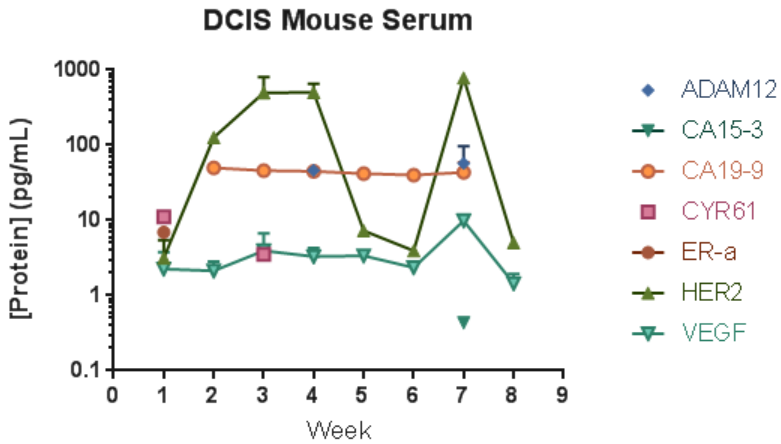


Figure 7. Levels of seven breast cancer biomarkers in DCIS mouse serum samples utilizing established Simoa protein assays. All measurements represent serum from terminal bleeds of individual mice.

Task 2. Apply single molecule diagnostic technique (developed in Task 1) to analysis of human serum samples (Years 3-5).

2a. Screen clinical samples obtained from Buchsbaum for presence of the markers that could be detected in HIM model (Months 25-30).

Dr. Buchsbaum consented patients and collected clinical samples. We received a total of 312 breast cancer patient serum samples including 66 pretreated breast cancer patient serum samples. These samples were tested against validated multiplex Simoa assays.

2b. Iterate 2a-2d until a sufficient number of markers has been identified for further investigation. Some candidate markers are expected to either not be found in the clinical samples or will not be predictive of disease state. Other markers may be difficult to measure due to lack of suitable binding reagents. As these markers drop out, others will be added to ensure there is a sufficient number when sample set is expanded (2c) (Months 25-48).

Preliminary experiments were performed testing human samples using commercially available sources. Breast cancer serum samples and healthy controls were obtained from the commercial supplier BioreclamationIVT. These samples were tested with the Simoa protein assays we developed to quantify the levels of candidate biomarkers in clinical samples and validate the usefulness of each biomarker. Through screening, the Simoa assays developed in our lab were able to detect most of the serum samples. **Figure 8** shows a graph of different markers that have been measured in healthy serum (n=37) and breast cancer serum (n=64) samples.

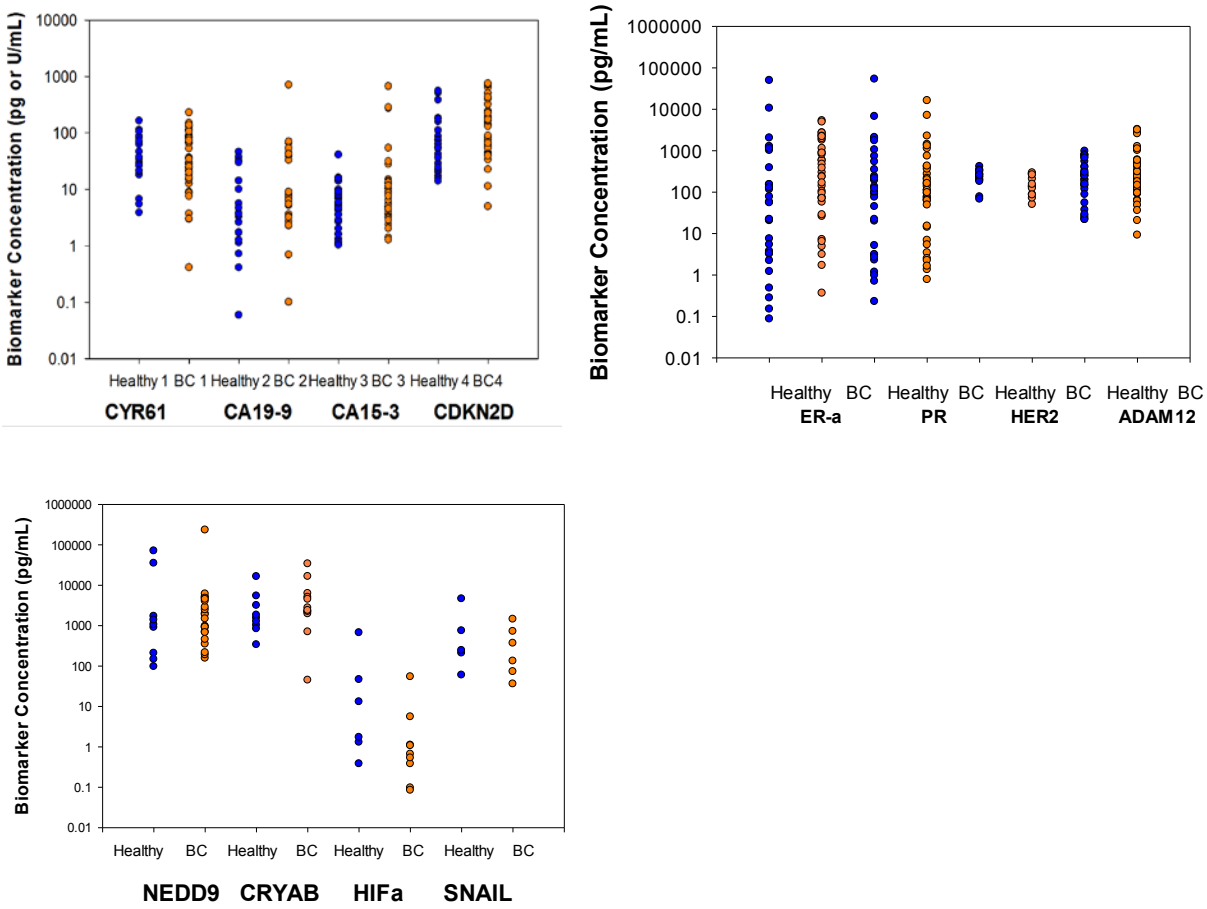
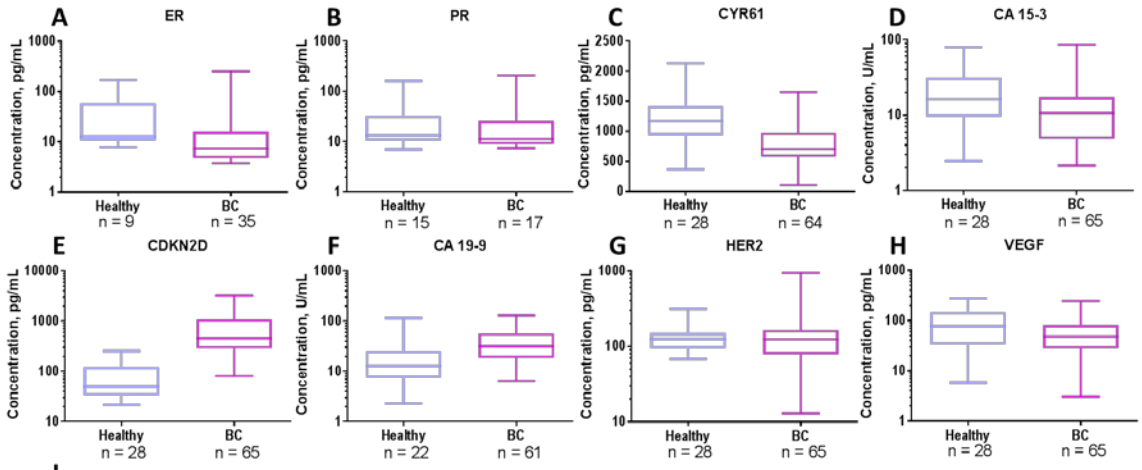


Figure 8. Measurements of protein biomarkers in commercially available healthy and breast cancer serum.

A second set of serum samples was tested with the multiplex assays we developed. The 28 healthy serum samples tested in this study were obtained from BioreclamationIVT from women who self-reported as healthy. Subject age ranged between 32 and 53 years of age. For the breast cancer serum samples, a prospective study was initiated at Tufts Medical Center, where patients were screened and diagnosed with breast cancer. These women consented to have their blood drawn for the study, and the serum was stored for later protein analysis. This population represents 66 women who had not yet undergone any therapeutic intervention for breast cancer at the time of sampling. **Figure 9** shows the results of these measurements, with a table describing the results of Mann-Whitney statistics comparing different groups of healthy and breast cancer patients. Based on these univariate statistics, it appears that CYR61, CDKN2D, CA 15-3, and CA 19-9 may have utility in distinguishing healthy and pretreated breast cancer serum samples received from Dr. Buchsbaum.



I

Sample Groups	Protein							
	ER	PR	CYR61	CA 15-3	CDKN2D	CA 19-9	HER2	VEGF
Healthy vs. All Cancer	*0.0373*	0.0892	*<0.0001*	*0.0068*	*<0.0001*	*0.0003*	0.820	*0.0418*
Healthy vs. Early Stage	*0.0282*	0.149	*<0.0001*	*0.0018*	*<0.0001*	*0.0012*	0.968	0.0580
HR+ vs. TNBC	0.651	0.171	0.307	0.8323	*0.0181*	0.634	0.593	0.394

Figure 9. Concentrations of different protein biomarkers in healthy and breast cancer serum. Each graph reflects measurements above the limit of detection, with the sample size listed below each group. Listed concentrations account for the assay's dilution factor. The markers shown are (a) ER, (b) PR, (c) CYR61, (d) CA 15-3, (e) CDKN2D, (f) CA 19-9, (g) HER2, and (h) VEGF. The table in (i) shows the p value results from Mann-Whitney statistical analysis of serum protein concentrations in healthy, all breast cancer, early stage (Stage 0, I, II), hormone receptor positive (HR+), and triple negative breast cancer samples. Values in red indicate a significant difference between the two groups, with a p value <0.05.

Detection of exosomes and exosomal proteins in breast cancer patient serum samples

There have been many reports showing that extracellular vesicles carry molecular information, which can affect neighboring cell signaling pathways as they are released. Non-denaturing lysis buffer was added in the serum samples and levels of certain biomarker candidates were then subsequently evaluated with hopes of finding these proteins within extracellular vesicles such as exosomes in the serum.

Interestingly, we found substantial NEDD9 and CRYAB levels in the serum samples treated with lysis buffer in both healthy and breast cancer samples from which both protein levels were undetectable in the serum alone (**Figure 10**). The Simoa calibration curve was not affected upon treatment with lysis buffer, confirming the lysis buffer did not impact nonspecific binding. These data is a proof of concept study that demonstrates that secreted candidate biomarkers can be detected using Simoa after adding lysis buffer to the serum samples.

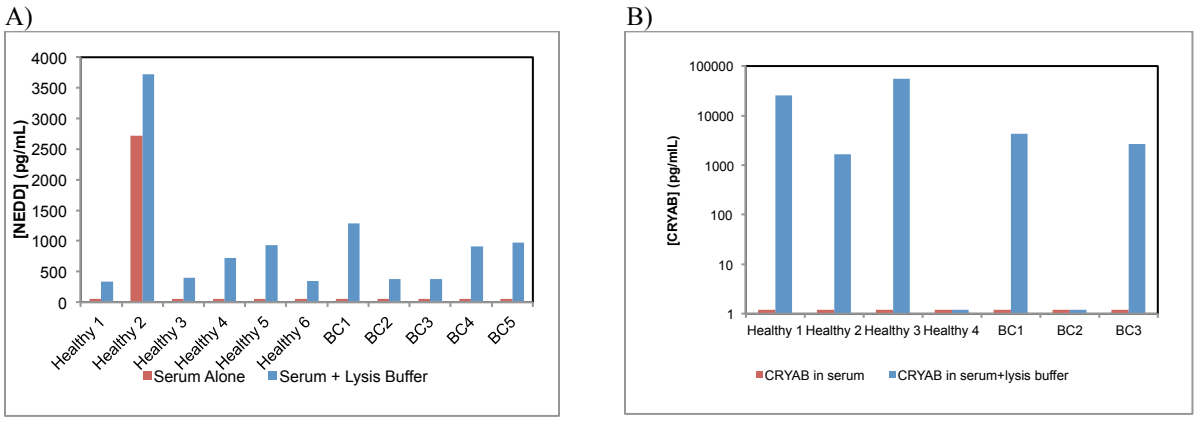


Figure 10. **A.** NEDD9 levels, **B.** CRYAB in healthy and breast cancer serum samples treated with the lysis buffer.

For evaluating exosomal proteins as breast cancer biomarkers, exosomes were collected from serum samples with an average size of approximately 110 nm determined by a Nanosight particle tracking system. **Figure 11** shows a comparison of ADAM12 and NEDD9 levels in serum versus the exosome lysate. We were able to find ADAM12 in the exosome lysate; however, the amount of protein released from the exosomes was much lower than in the original serum in all but one of eight samples. Unlike ADAM12, NEDD9 is not accessible for detection in serum. In fact, NEDD9 was only measured in one serum sample. However, NEDD9 was measured in serum samples following a lysis protocol for the serum, as described above. This indicates that NEDD9 may be contained in exosomes in the serum and is released upon lysis. We were able to detect NEDD9 in the exosome lysate in 3 out of 8 samples, but the amount of NEDD9 detected from the exosome lysate does not correlate well with the amount measured in serum or serum lysate. This result suggests that exosomes are not the only kind of vesicles that may contain NEDD9. In comparison to ADAM12 and NEDD9, we could not detect HER2 in the exosome lysate, even though the serum levels of HER2 are high for all the samples tested (average 585 pg/mL).

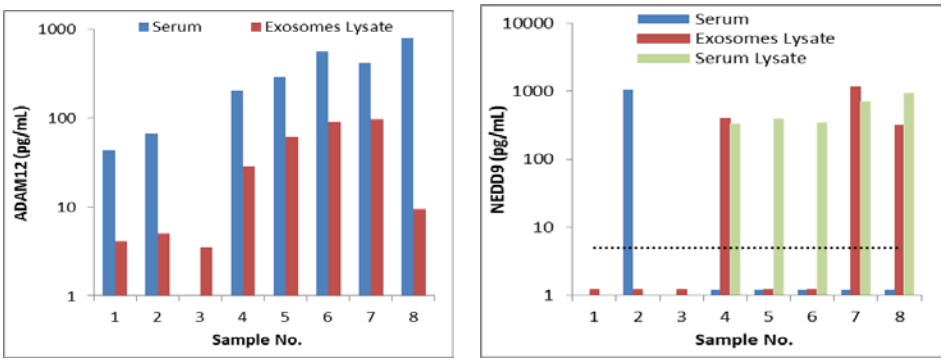


Figure 11. ADAM12 and NEDD9 protein level comparison between exosome lysate and original serum.

Based on the results above, we determined that proteins inside exosomes are not good biomarker candidates for use with the ultrasensitive Simoa platform. The proteins of interest were either measurable directly in serum (e.g., HER2 & ADAM12) at higher levels than in exosomes, or could be detected by lysing the serum directly (e.g., NEDD9). Overall, the results here demonstrate that exosomes are not superior to serum (or serum lysate) as a clinical sample for ultrasensitive protein marker detection. Further investigation of exosomal content as biomarkers may shift from proteins to nucleic acids like microRNAs.

2c. Expand sample set and run clinical samples with all assays developed for all markers (Year 5).

We received 325 breast cancer patient serum samples from Dr. Buchsbaum as described above in task 2a. These samples were tested against validated multiplex Simoa assays for eight protein markers. Expanded sample collection was not done in the grant period.

2d. Determine which markers in blood correlate with disease using data processing and computational methods available in the laboratory (Year 5).

Preliminary multivariate data analysis has been carried out to assess the predictive capabilities of several protein biomarkers. The supervised multivariate statistical method of partial least squares-discriminant analysis (PLS-DA) was employed to assess PR, ER α, and CDKN2D. Results were promising, with one model yielding 84% accuracy in distinguishing healthy individuals from early stage (I-II) breast cancer patients (**Table 3**).

	Description	Precision	True Positives	AUC	Number of samples	Overall Accuracy
Model 1	Healthy	72%	90%	0.88	31	81%
	BC Stage I-IV	92%	75%	0.88	44	
Model 2	Healthy	82%	87%	0.91	31	84%
	BC Stage I-II	86%	80%	0.91	30	
Model 3	Healthy	86%	81	0.88	31	45%
	BC Stage I-II	62%	70%	0.78	30	
	BC Stage III-IV	0%	0%	0.63	14	
Model 4	TNBC	50%	78%	0.73	14	68%
	Luminal	86%	63%	0.73	30	

Table 3. PLS-DA analysis results for different populations of breast cancer serum samples.

Similar analysis was performed with serum samples collected at Tufts Medical Center. All of the data obtained for these serum samples were used as inputs to evaluate the combined utility of eight protein markers and patient age toward a breast cancer diagnostic signature. Four models were tested comparing healthy samples to all breast cancer samples, healthy and Stage 0-II samples, Stage 0-II and Stage III-IV samples, and HR+ and TNBC samples, results of which are summarized in **Table 4**.

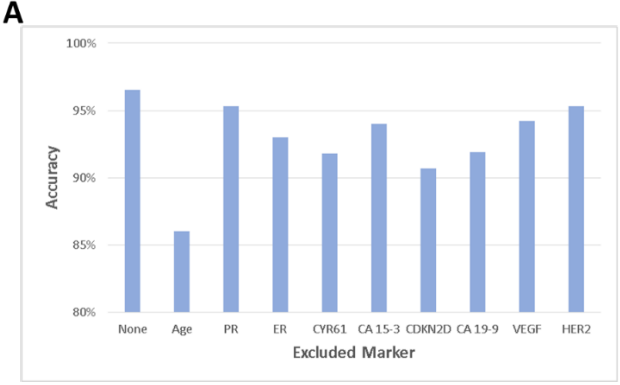
	Class Groups	Precision	True Positives	AUC	# Samples	Accuracy
Model 1	Healthy	90%	96%	0.98	28	96%
	All BC	98%	95%		66	
Model 2	Healthy	93%	96%	0.99	28	97%
	BC Stage 0-II	98%	97%		58	
Model 3	BC Stage 0-II	91%	88%	0.78	58	82%
	BC Stage III-IV	30%	38%		8	
Model 4	HR+ BC	88%	70%	0.56	54	66%
	TNBC	84%	38%		8	

Table 4. Description of each PLS-DA model, with precision, true positive rate, AUC value, the number of samples in each group, and the overall accuracy.

Model 1 compared the healthy cohort to the entirety of the breast cancer cohort, a sensitivity and specificity of 95% and 96%, respectively. Model 1 displayed an overall accuracy of 96% and an AUC of 0.98, which describes a successful classifier. Model 2, which compared healthy samples to Stage 0-II breast cancer, performed slightly better with an overall of 97%, with sensitivity and specificity values at

97% and 96%, respectively. Additionally, the variables with the highest impact on Model 2 were age, CDKN2D, CYR61, and CA 19-9 based on the exclusion of these markers from the models, shown in Figure 11. These findings agree with the Mann-Whitney statistical evaluation of individual markers, which supports the use of such univariate statistics to help assess the utility of individual markers.

A



B

Excluded Marker	Total Sample Number	Correct Assignments	Accuracy
None	86	83	97%
Age	86	74	86%
PR	86	82	95%
ER	86	80	93%
CYR61	86	79	92%
CA 15-3	86	81	94%
CDKN2D	86	78	91%
CA 19-9	86	79	92%
VEGF	86	81	94%
HER2	86	82	95%

Figure 12. A. Overall accuracy of Model 2 (Healthy vs Stage 0-II) is plotted against the marker excluded from the model. **B.** The number of correctly assigned samples, the total number of samples, and the resulting accuracy is listed for each marker exclusion scenario**.**

Model 3 evaluated the same biomarker signature to differentiate Stage 0-II and Stage III-IV breast cancer samples—this model demonstrated proficiency in classifying Stage 0-II samples with a sensitivity of 88%, but was less successful in identifying Stage III-IV, which only had a sensitivity of 38%. The AUC for this model was 0.78, with 82% accuracy. Although Model 3 is not as successful as Models 1 and 2, the signature shows some promise for use in tracking disease progression. The most influential variables in this model were CA 15-3 and CA 19-9, which supports the idea that different markers in a signature could be used for different purposes (i.e. screening, therapeutic efficacy, recurrence monitoring). Model 4 compared the hormone positive (HR+) population to the TNBC cohort, which performed poorly. This model had an accuracy of 66% and an AUC of 0.56, which places this model slightly above the discrimination of a random guess. This signature does not appear to be appropriate for differentiating breast cancer subtypes. Out of the four models, the protein signature combined with patient age displayed the most discrimination between Stage 0-II breast cancer and healthy samples.

Task 3. Develop single cell analysis methods to determine composition of a primary tumor.

3a. Select approximately ten candidate markers (SNPs, proteins, etc.) of value for single breast cancer cell analysis in conjunction with collaborators (Months 1-9).

We chose to utilize the pre-established PSA assay to validate Simoa at the single cell level. This assay has proven to be extremely sensitive and our goal was to take advantage of the sensitivity of this established assay for single cell resolution.[2] In order to test this assay within cells, we have cultured the prostate cancer cell line LNCaP (cell lines were purchased from ATCC and donated by the Kuperwasser lab). We have developed single cell assays using Simoa assays established in Task 1. Breast cancer cell lines BT-474, MDA MB231 and MCF7 were used to quantify these markers in single cells.

3b. Detect presence of mtDNA by screening for many sequences using cultured cell lines and HIM tissue samples to select which markers to use. Will require development of mtDNA microarrays and sample screening (Months 6-18).

The selection of mtDNA markers was performed in conjunction with the Sonenshein laboratory. The DNA assay development was put on hold to prioritize other tasks.

3c. Develop assays for the selected markers (Months 6-24).
The Simoa assay for PSA has already been developed. PSA was used as a proof of concept assay because it is well established. Simoa assays described in Task 1 were utilized for single cell studies of breast cancer cell lines such as BT-474, MDA MB231 and MCF7.

3d. Develop single cell assays for selected markers (Months 18-36).
We spent considerable time developing a Simoa single cell assay for detecting PSA in LNCaP cells as a proof-of-concept model since the PSA assay is well studied, reliable, and has a very low Simoa detection limit (~0.01 pg/mL). Previously, all assays were carried out manually using optical fiber bundle arrays, which require several hours of hands-on preparation. Additionally, cells were previously isolated by diluting a stock cell suspension. However, this preparation is inconsistent and challenging to verify accurate cell counts. Using our previous protocols, we could detect PSA levels in down to 5 cells. To improve our assay for single cell resolution, we adapted our protocols for the automated HD-1 analyzer developed by Quanterix Corporation and collaborated with the Chiu lab to isolate and verify single cells.

The HD-1 analyzer fully automates all assay incubation, washing, and imaging steps, yielding consistent sample handling and maintaining ultrasensitive detection. Single cells were isolated by pipetting 1 µL cell suspension into PCR tube caps the number of cells present in the droplet was confirmed under a microscope. Isolated cells were lysed in 64 µL lysis buffer, diluted in an additional 75 µL of sample diluent (total volume = 140 µL), and analyzed by HD-1.

Interestingly, we found that the LNCaP cell line purchased directly from ATCC exhibited significantly higher PSA expression than the cell line donated by the Kuperwasser lab. Cell line verification by short tandem repeat (STR) profiling indicated that the cell line donated by the Kuperwasser lab had an 88% match to LNCaP. Cell lines with ≥80% match in their STR loci are considered to be related. Cancer cells are genetically unstable so it is expected that protein expression will change in highly cultured cell lines. The cell line donated by the Kuperwasser lab exhibited an average of 6.97 $x10^4$ PSA molecules/cell compared to an average of 2.06 x 10^6 PSA molecules/cell found in the LNCaP cell line obtained from ATCC.

Figure 13 shows the distribution of PSA molecules detected in single LNCaP cells both from ATCC (lower passage) and from the Kuperwasser lab (highly passaged). From the results, we were able to measure the impact of genetic drift on protein expression at the single cell level and show that $LNCaP_B$ cells have significantly depressed PSA expression compared to $LNCaP_A$. These results indicate that genetic instability in cancer cells affects protein expression, and by extension, cancer progression, and also highlights the necessity of cell culture authentication. Importantly, we establish Simoa as a unique and facile approach to count protein molecules with single cell resolution to reveal unbiased phenotypic information. These data have been published.[3]

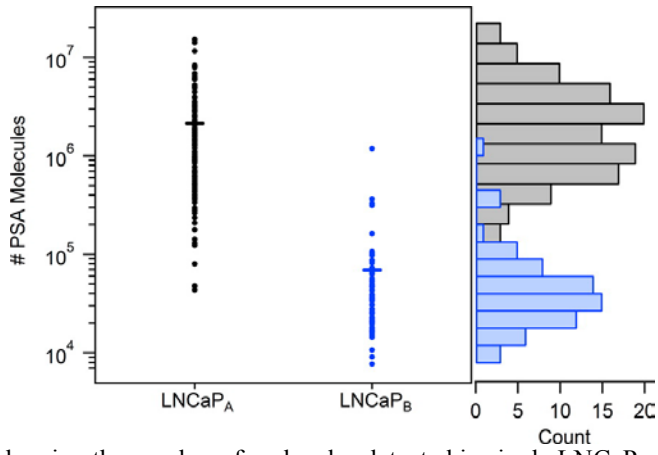


Figure 13. (Left) Dot plot showing the number of molecules detected in single LNCaP$_A$ and LNCaP$_B$ cells. Each dot represents a single cell. The average of each population is represented with a bar. (Right) Histograms illustrating the log-normal distribution of PSA molecules in individual LNCaP$_A$ (gray) and LNCaP$_B$ (blue) cells.

We demonstrated the ability to detect PSA in single LNCaP cells, and expanded that work to characterize breast cancer cell lines of different subtypes. The goal of this work is to use single cell Simoa to classify breast cancer cell subtypes and identify rare cells in a heterogeneous tumor population with single cell resolution. We cultured three different breast cancer cell lines to study the distribution of protein expression based on breast cancer subtype. MDA-MD-231, BT-474, and MCF7 cell lines were selected to represent triple negative breast cancer (TNBC), luminal B, and luminal A subtypes, respectively. In the past year, we have measured the levels of eight different protein biomarkers in bulk cell populations containing roughly 100 cells for each cell line (**Table 5**). These experiments will inform how to proceed to multiplexing markers for single cell studies.

The results in Table 5 clearly demonstrate the differences in protein expression observed between different cell lines. BT-474 overexpresses HER2 and PR, and shows measurable expression of ERα, CYR61, CA15-3, CA19-9, and ADAM12. MCF7 cells show considerably lower expression of HER2, which is consistent with the luminal A subtype. Likewise, MDA-MB-231 cells show higher concentrations of CYR61 and no detectable levels of HER2, PR, or ERα, as expected for TNBC.

Marker	MCF7 cells (pg/mL)	BT-474 cells (pg/mL)	MDA-MB-231 cells (pg/mL)	Assay LOD (pg/mL)	Assay LOD (# of molecules)
ER-a	1.7	3.7	Below LOD	0.2	200,466
PR	Below LOD	9.2	Below LOD	0.2	109,564
HER2	1.2	77	Below LOD	0.3	13,244
CA15-3 (U/mL)	0.21	Below LOD	Below LOD	0.1	
CYR61	0.54	2.5	7.2	0.03	18,060
HIF1a	Below LOD	0.48	0.04	0.015	188,125
VEGF	Below LOD	1.95	0.28	0.03	40,133
ADAM12	3.8	9.2	4.8	3	3,311,000
Subtype	**Luminal A**	**Luminal B**	**Triple Negative**		

Table 5. Comparison of Breast Cancer Protein Expression in ~100 Cells

We have been utilizing fluorescence-activated cell sorting (FACS) to obtain single cell samples. We quantified and analyzed the distribution patterns of protein biomarkers mentioned above in single MCF7, BT-474 and MDA MB231 cells, with results shown in **Figure 14**.

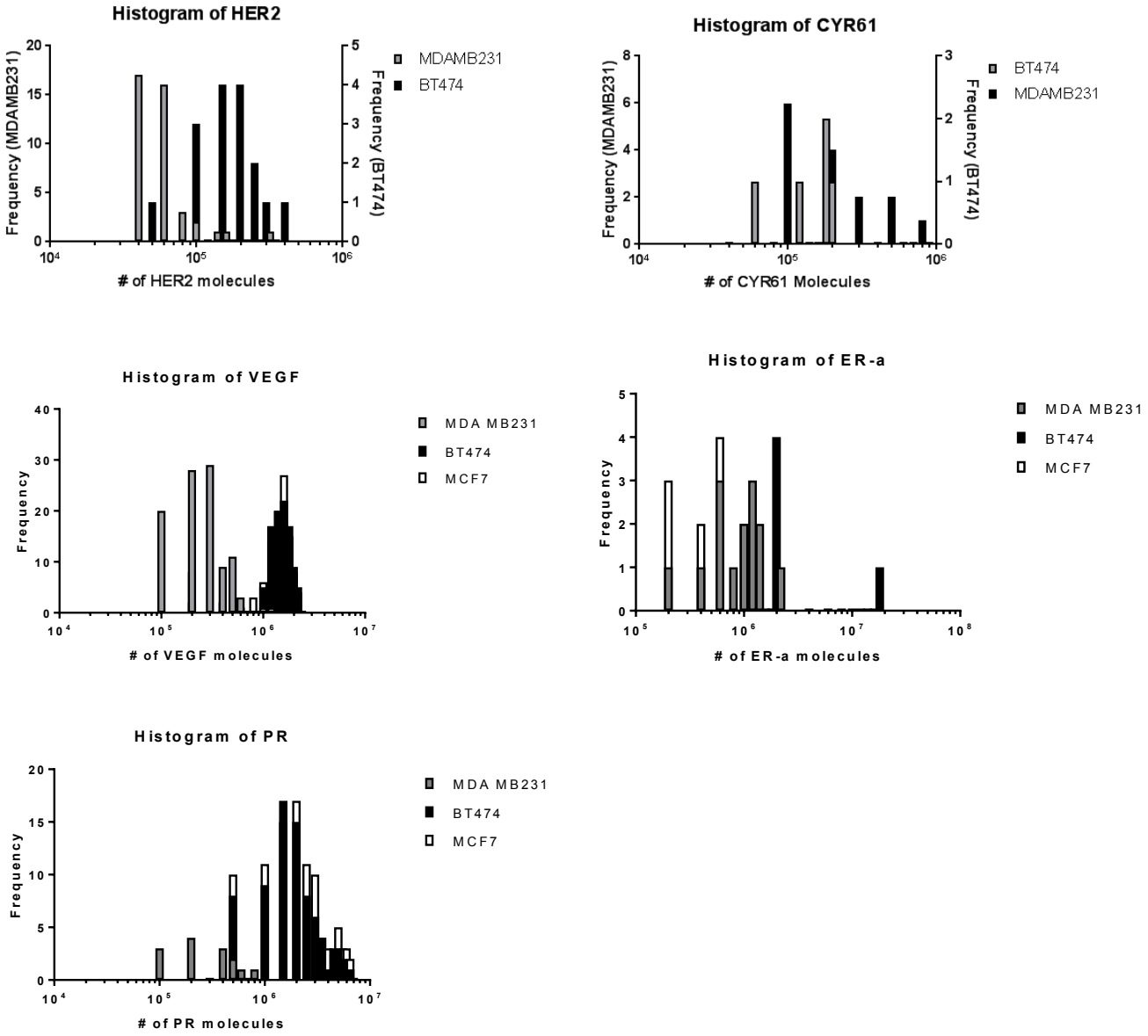


Figure 14. Histograms illustrating the log-normal distribution of ER-α, PR, HER2, VEGF and CYR61 molecules in individual breast cancer cells.

Additional work, including obtaining tissue samples from HIM to perform single cell analysis, determining if rare cells with more aggressive genotypes or protein levels in HIM samples can be observed, and confirming the presence of single cells in human breast cancer biopsy samples, was not undertaken in the current project period in order to prioritize other tasks.

Daniel T. Chiu, PhD, *University of Washington, Seattle Campus Box 351700 Seattle, WA 98195-1700*

Our overall goal was to develop and deploy new technologies for the early diagnosis of breast cancer and detection of recurrence, better prognosis monitoring, and stratifying this highly heterogeneous disease. To achieve these objectives, we developed new ultrasensitive methods for monitoring biomarkers from peripheral blood. Rather than focusing on one type of biomarker, we pursued a multi-pronged approach that used DNA, protein, and cell biomarkers to provide a more comprehensive and robust set of information about disease status.

Task 1. Work with Walt lab to develop/refine single-molecule and single-cell techniques for analyzing protein and cell biomarkers (Months 1-36).

For Task 1, we successfully developed and applied a new technology for the detection and analysis of cell biomarkers, in particular, circulating tumor cells (CTCs). CTCs have emerged as an important and valuable biomarker for the prognosis of breast cancer, and the sensitivity we demonstrated makes CTCs a robust biomarker for prognosis and as an indicator of treatment efficacy. The high sensitivity we achieved also makes CTCs a potential diagnostic tool. Below, we highlight the accomplishments under each subtask in more detail.

***1a. Develop microfluidic devices for sample preparation and optical manipulation. Initial devices will contain only sample preparation module for use by Walt lab (Years 1-2*).**

We successfully completed this task, where we now have a new technology called eDAR (ensemble Decision Aliquot Ranking) with which we can isolate circulating breast cancer cells from the peripheral blood of patients with ~95% recovery efficiency in a background of over billions of blood cells in less than 20 mins. **Figure 15** shows the eDAR microfluidic chip technology, which combines the following components: multi-color line-confocal fluorescence detection with a high sensitivity, a hydrodynamic switching mechanism, a cell trapping and subsequent purification process, and an identification and downstream analysis section. **Figure 16** shows the analytical performance of eDAR (see caption for detail); we were able to isolate one CTC from a background of ~5 billion blood cells with ~95% recovery efficiency (**Fig 16e**).

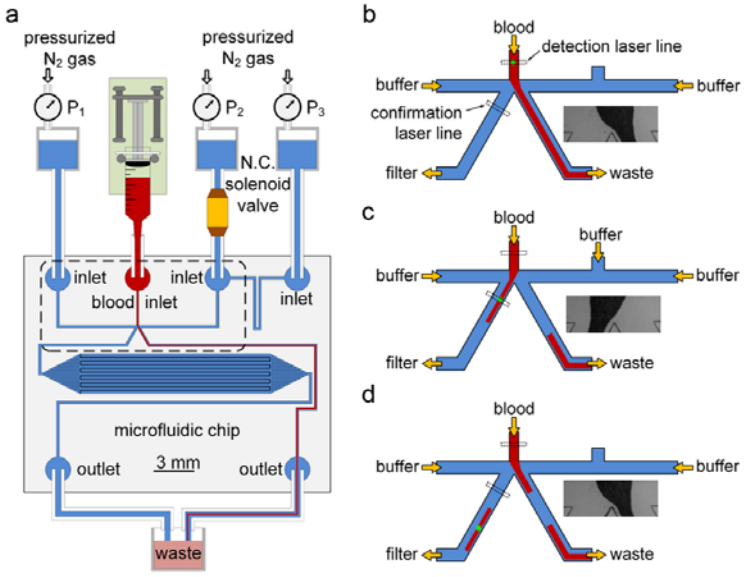


Figure 15. Microfluidic chip and hydrodynamic switching scheme of eDAR. a) General structure of the microfluidic chip and the configuration of the eDAR platform. The bottom left channel was used to collect sorted aliquots and transfer them to the subsequent purification area, which had 20,000 microslits. The area marked with a dashed box is further explained in b-d. b) The flow condition when no positive aliquot was ranked. c) The blood flow was switched to the CTC collection channel, and the sorted aliquot was confirmed by the second detector. d) The blood flow was switched back after the aliquot was sorted.

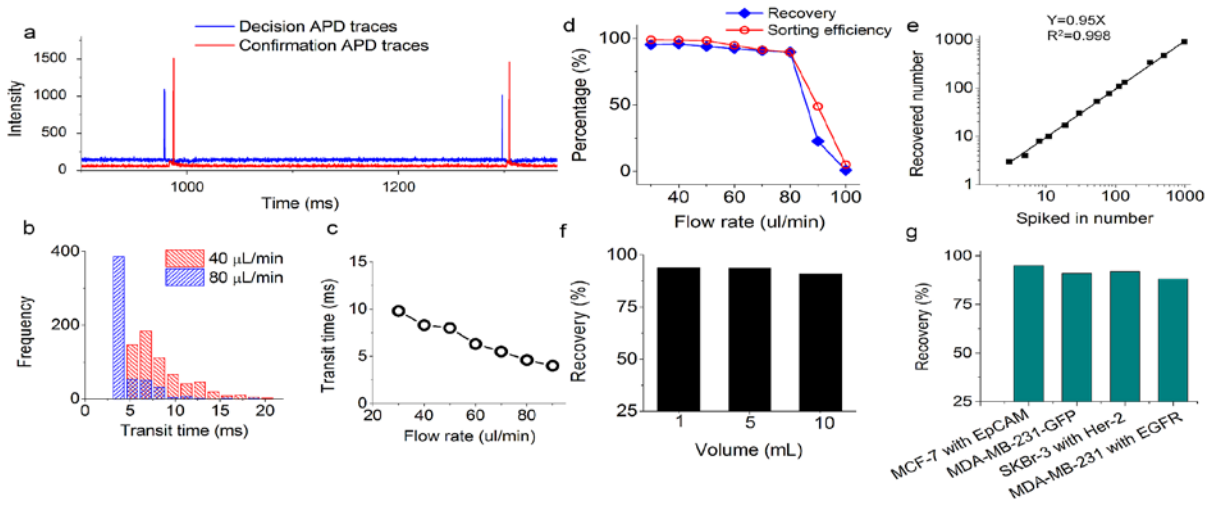


Figure 16. eDAR characterization and analytical performances. a) The segment of the detector (APD) data trace from a cancer sample that shows two events triggered the sorting, which then were confirmed by the second detection window. b) The distribution of transit time at flow rates of 40 and 80 μL/min, respectively. c) A plot showing the fastest average transit time was about 4 ms when the flow rate was 90 μL/min. d) The recovery and sorting efficiency value versus different flow rate. e) The recovery ratio of MCF-7 breast-cancer cells spiked into whole blood. f) The recovery ratio of 300 MCF-7 cells spiked into 1, 5 and 10 mL of whole blood aliquots. g) The recovery ratio of four selection schemes of four breast cancer cell lines spiked into whole blood.

We have also made a number of improvements and refinements to this new eDAR technology. One example is the use of better fluorescent probes for targeting breast cancer cells so we do not miss sub-types of cancer cells from patients, given the highly heterogeneous nature of these cells. **Figure 17** shows a yellow emitting Pdot (semiconducting polymer dot), which is a new family of ultra-bright fluorescent probes developed by our lab. These Pdots are much brighter than comparable fluorescent probes, such as quantum dots (Qdots). In fact, we routinely collect 10^9 photons from single Pdots, which is 1,000× to 100,000× more than many single dyes and green fluorescent proteins (GFPs), and 10-100 times more than Qdots. These new ultra-bright probes enable us to isolate CTCs with low-expression cell-surface biomarkers.

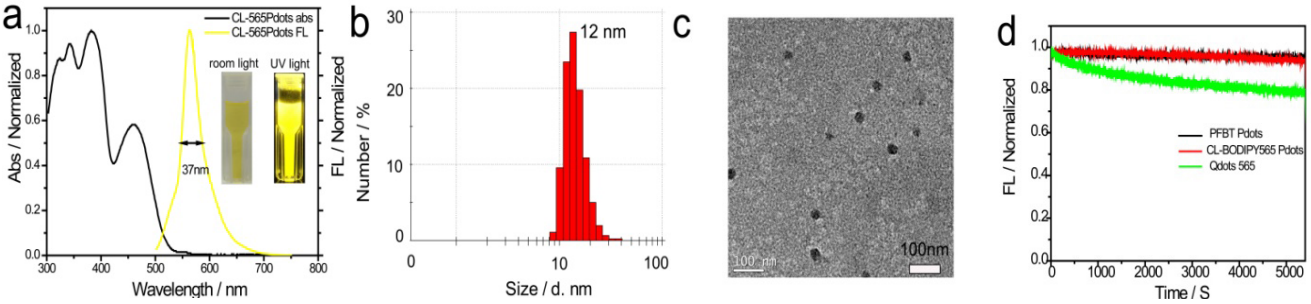
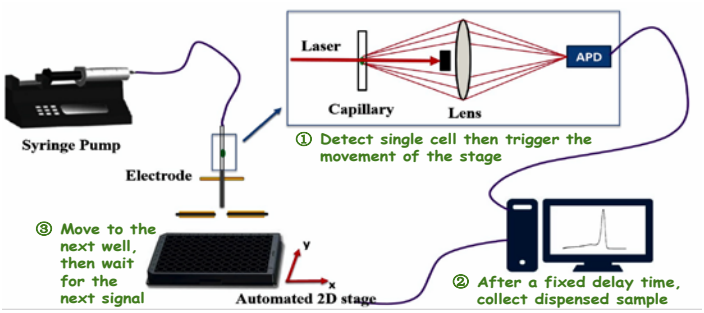


Figure 17. (a) Absorption and fluorescence spectra of CL-BODIPY 565 Pdots in water. (b) The histograms of the distribution of the sizes of CL-BODIPY 565 Pdots, measured by DLS (the mean size is 12 nm) (c) TEM images of CL-BODIPY 565 Pdots, (d) Photobleaching curves of PFBT Pdots (black line), CL-BODIPY 565 Pdots (red line), and Qdots-565 (green line).

The individual CTCs isolated from breast cancer patients can be analyzed for protein expression and biomarker detection using the sensitive single-molecule technologies from the Walt lab.

1b. Integrate assays and optical methods developed in Walt lab into a microfluidic device for the analysis and profiling of biomarkers (Years 1-2).

We have completed this task and now have a method for the isolation and analysis of individual CTCs, including performing digital PCR and Simoa on the isolated single cells. **Figure 18** shows the single-cell dispensing platform that we have developed, in which individual cancer cells are detected by the optical system (inset) and then ejected into each well of a 96 well plate for high-throughput single-cell analysis.



Delay time = longest transition time + droplet formation time + droplet falling down time
= 0.45 s + 0.072 s (for 50nL droplets at 2.5mL/h) + 0.08 s (based on gravity)
= 0.602 s

Figure 18. Schematic showing the integrated setup for single-cell detection and dispensing. With a 1.5-3 cm long distance between cell detection and dispensing, the delay time is on average 0.45s. The droplet formation and dispensing time is about 0.15s.

1c. Develop integrated microfluidic and optical techniques for single cell analysis using model cell lines (Years 1-3).

We successfully achieved several types of single-cell analysis, such as digital PCR assay of the isolated cancer cells. **Figure 19** shows an example; here, we measured the number of RNA transcripts in individual cancer cells using both single-molecule RNA FISH (fluorescence in situ hybridization) and digital single-cell PCR.

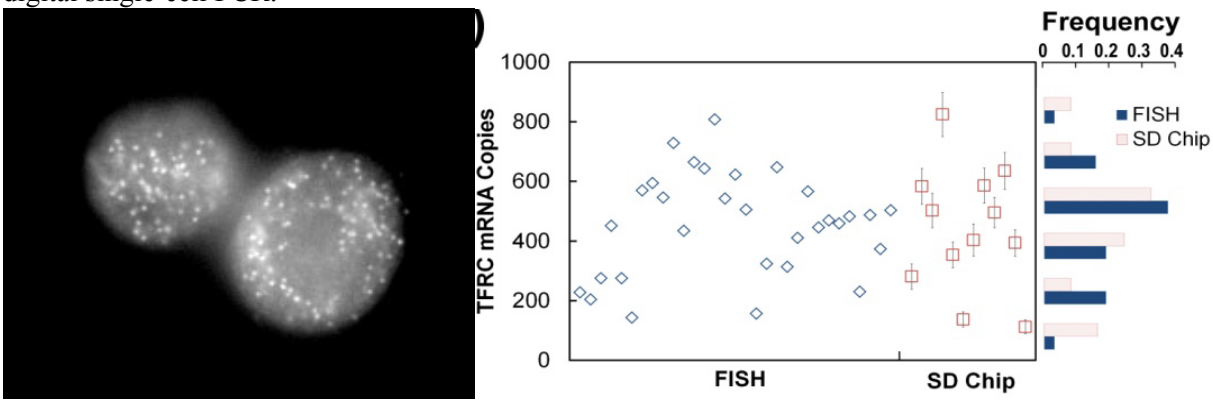


Figure 19. Digital assays showing the number of RNA transcripts in individual cells with single-molecule FISH and digital PCR. Left image is an example of single-molecule FISH we observed for the TFRC gene, and the right plot shows the variability in the copy number of the TFRC RNA in single cells as measured with FISH and digital PCR.

To further improve the throughput of analysis, we have also reported the parallel manipulation of single cells using bipolar electrode (BPE) based dielectrophoresis (DEP). Briefly, for the BPE-DEP based manipulation of cells, we have developed the theory that describes this technique and applied it to single-cell manipulation. **Figure 20** shows an example, in which BPE-DEP was used to trap a single cell in a microchamber (left panel), and by changing the frequency of the applied AC field, the trapped single cell was subsequently released from the chamber (right panel).

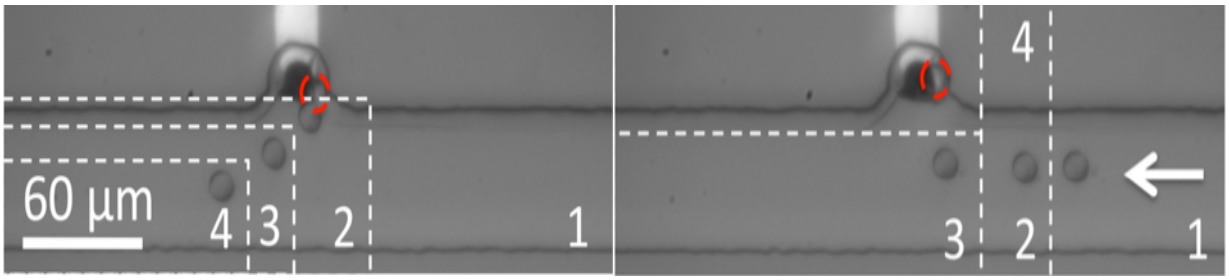


Figure 20. Left panel: nDEP attraction of a B-cell toward the BPE cathode in phosphate DEP buffer (4 s/slice). $E_{DC,avg}$ = 1.5 kV/m, $E_{RMS,avg}$ increased from 5.7 kV/m to 28.3 kV/m from t = 0 s (slice 1) to t = 8 s (slice 3). Right panel: Release of the trapped cell (2 s/slice) upon subsequent decrease of $E_{RMS,avg}$ to 5.7 kV/m (from slice 1 to slice 2).

Task 2. Work with the Walt lab to apply these methods to breast-cancer patient samples for the detection and validation of protein and cell biomarkers *(Months 25-60).*

We successfully completed this task and showed the high-sensitivity recovery of CTCs from breast cancer patients, even from patients with early stage cancer. We also developed and applied single-cell analysis techniques to study the isolated cancer cells. Below, we highlight the accomplishments under each subtask in more detail.

2a. Develop/submit amended proposal to University of Washington IRB to permit secondary use of currently archived patient samples (Months 1-12).

This task was completed in year 1, which gave us permission to use the archived clinical samples from our lab for this project.

2b. Apply sensitive techniques (1a and 1b) for the retrospective analysis and validation of biomarkers from archived patient samples (Months 25-60).

We completed this task and sent our archived samples to the Walt lab for Simoa analysis.

2c. Apply sensitive techniques (1a and 1b) for the prospective analysis and validation of biomarkers from patient samples (Months 36-60).

We successfully completed this task and **Table 6** below summarizes some of our results. These results show CTCs isolated from patients are highly heterogeneous, both in terms of the amount of cell-surface proteins that the CTCs express and the types of cell surface proteins present on the CTCs.

Table 6. Analysis of biomarkers in patient samples.

Sample Code	Labeling strategy	CTC in 2 ml	Subtype	Stage
B-140000007	1 step EpCAM	8	Triple -	Stage 4
B-140000053	1 step EpCAM	1.5	ER+, PR- Her2+	Stage 2
B-140000089	1 step EpCAM	0	ER/PR+ Her2-	Stage 2
B-140000122	1 step EpCAM	4	ER/PR - Her2+	Stage 4
B-140000169	Cocktail (EpCAM, HER2, EGFR, N-cadherin)	32	ER/PR+ Her2-	Stage 3
B-140000169	1 step EpCAM	8	ER/PR+ Her2-	Stage 3
B-140000477	Cocktail (EpCAM, HER2, EGFR, N-cadherin)	15	ER/PR + Her2 -	Stage 2
B-140000865	1 step	0	ER/PR- Her2-	Stage 4
B-140000931	1 step	29*	ER/PR+ Her2-	Stage 4
B-140002443	Cocktail(EpCAM, HER2, EGFR, N-cadherin)	19	ER/PR+ Her2-	Stage 4
B-140001155	Cocktail (EpCAM, HER2, EGFR, N-cadherin)	27*	ER+PR-Her2-	Stage 1A (T1N0)
B-140001155	1 step	5*	ER+PR-Her2-	Stage 1A (T1N0)
B-140001165	1 step	4*	ER/PR+ Her2-	Stage 2
B-140000068	1 step	48*	ER/PR+ Her2-	Stage 2
B-140001108	1 step	1, 4*	ER/PR+ Her2-	Stage 2
B-140001146	1 step	1, 5*	ER/PR- Her2-	Stage 2
B-140001146	Cocktail (EpCAM, HER2, EGFR, N-cadherin)	1	ER/PR- Her2-	Stage 2
B-140001196	1 step EpCAM	44	ER+ PR- Her2+	Stage 4
B-140001196	Cocktail (EpCAM, HER2, EGFR, N-cadherin)	75	ER+ PR- Her2+	Stage 4
B-140001222	Cocktail (EpCAM, HER2, EGFR, N-cadherin)	57	ER+ PR+ Her2-	Stage 2
B-140001222	1 step EpCAM	31	ER+ PR+ Her2-	Stage 2
B-140001362	1 step EpCAM	3	ER- PR- Her2+	Stage 4
B-140001385	1 step EpCAM	0	ER+PR+ Her2-	Stage 2
B-150001464	1 step EpCAM	2	ER/PR+ Her2-	Stage 2
B-150001464	Cocktail (EpCAM, HER2, EGFR, N-cadherin)	18	ER/PR+ Her2-	Stage 2

*These CTCs also expressed CD45.

To increase the number of CTCs isolated from patient samples, especially from early stage patients, we have also carried out the following three studies: (1) A systematic study to determine the minimum number of proteins needed on CTCs in order for eDAR to capture them efficiently; (2) develop ultra-bright probes based on Pdots for the detection of cells with low biomarker expression; (3) improve the ability of eDAR to capture CTCs with variable patterns of biomarker expression by developing a cocktail of antibodies to label multiple proteins on CTCs. We have already discussed Point 2 (Pdots) in a previous section, and below, we describe the outcomes of Point 1 and Point 3.

(1) Limit of detection (LOD) of eDAR: eDAR requires the use of fluorescent antibodies to bind to the biomarkers on the cancer cells to "light up" the cancer cells so they can be located and isolated. If there is a low level of biomarkers on the CTC, a correspondingly small number of fluorescent antibodies are bound to the CTC and thus the CTC may not be sufficiently bright for eDAR to pick up. To understand the LOD of eDAR, we characterized the brightness of the different breast cancer cell lines known to

express various amounts of the cell surface protein EpCAM (currently the cell surface marker used for isolation of CTCs) and then determined the number of EpCAMs present on each cell based on the number of fluorescent antibodies bound to EpCAM. Similarly, using MDA-MB-231 cells (which mimic triple-negative breast cancer), we have determined the number of several cell-surface biomarkers specific to breast-cancer cells. **Figure 21** shows the results; here, the number of cell-surface biomarkers was measured based on PE (phycoerythrin) fluorescence where there was a one-to-one correlation between the number of PEs and the number of cell-surface biomarkers.

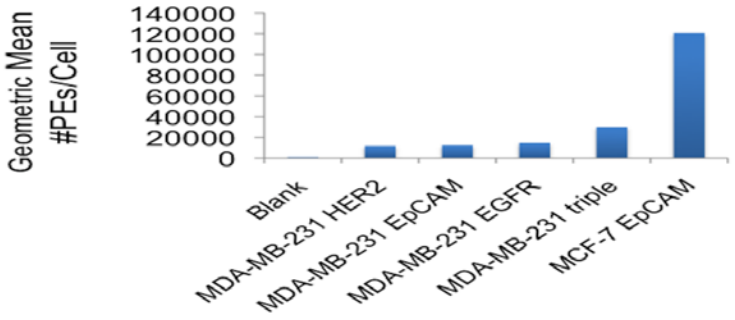


Figure 21. MDA-MB-231 cells labeled with three antibodies (EpCAM, HER2, EGFR) and a cocktail of all three, demonstrating that the antibody brightness is additive.

From these experiments, we found our current eDAR LOD is 5800 EpCAMs and LOQ (limit of quantitation) is 23500 EpCAMs per cell. While LOD is defined as three times the standard deviation (σ) of the background, LOQ is defined as 10σ, which is the threshold eDAR uses for robust sorting and to minimize any potential false sorting caused by fluctuations in the background signal.

(3) Improve the ability of eDAR to capture CTCs with variable patterns of biomarker expression by developing a cocktail of antibodies to label multiple proteins on CTCs. Many CTC detection methods, including ours, take advantage of the epithelial origin of CTCs, which provides the cells with surface markers that are distinct from those on other cells in the blood. However, for a tumor cell to migrate into the bloodstream, it is reported to undergo an epithelial-to-mesenchymal transition (EMT), losing some of its epithelial characteristics, such as structural rigidity, cell adhesion and epithelial markers, such as EpCAM and cytokeratin, and taking on a more mesenchymal phenotype. The phenotypic changes undergone by CTCs during the metastatic process mean that these cells may not express high levels of the epithelial markers, such as EpCAM, that are targeted by current CTC detection techniques. These "EpCAM^{low}" CTCs recently have received attention because they potentially can evade detection and have been linked with enhanced invasiveness and migration. Accurate population statistics describing CTC expression of epithelial markers is unavailable. More importantly, immunoaffinity techniques traditionally quantify recovery rate (percentage of cells captured) using cultured cancer cells that are EpCAM^{high} (e.g. the breast cancer cell line MCF-7) thus leading to an *upper estimate* of CTC recovery.

To address this issue, we developed a cocktail of CTC-specific primary antibodies targeting both mesenchymal and epithelial cell surface markers. We then used fluorescently labeled secondary antibodies to target all the primary antibodies bound to the CTCs. The use of a single fluorescently labeled secondary antibody to target the primary antibody cocktail provides important practical advantages, because this way, the relative amount of fluorescent antibodies present on the cell surface is additive without also seeing an additive increase in the background signal as would happen with fluorescently conjugated primary antibodies. We demonstrated the detection of CTCs that otherwise would be missed by targeting EpCAM alone with a dye-linked primary antibody. This immunolabeling strategy is an important improvement over our previously reported results as it addresses the heterogeneous nature of CTCs that makes them difficult to detect. The strategy also expands the limit of detection in a quantifiable manner, which is important for accurate CTC counts in patient samples.

The relevance and success of the labeling schemes described above needed to be evaluated with clinical samples. To quantify the improvement, we compared the new labeling scheme to our former labeling strategy of PE-anti-EpCAM alone. **Figure 22** shows a bar graph with the number of CTCs isolated with each labeling scheme (cocktail and PE-anti-EpCAM) for each of the samples obtained from four breast cancer patients. Patient 1 was stage 3, ER+ PR+ HER2-; patient 2 was stage 4, ER+ PR- HER2+; patient 3 was stage 2, ER+ PR+ HER2-; and patient 4 was stage 2, ER+PR+HER2-. In this experiment, each sample was divided into two 2-mL portions. The samples were analyzed by eDAR sequentially using each of the two different labeling schemes. Samples were processed within two days after the blood was drawn. The data in Figure 21 demonstrate that a significant improvement, **averaging 6×**, was observed with the new labeling strategy. Figures 21b and 21c show representative cells that were isolated and further labeled against CD45, cytokeratin, and DAPI (post-capture) to confirm CTC identity.

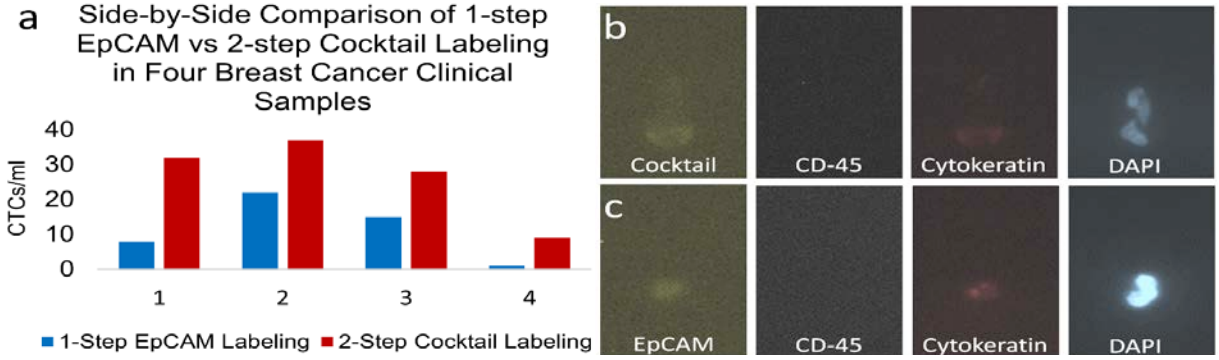


Figure 22. Cells from breast cancer clinical samples were sorted on an eDAR chip. Following sorting, the cells were fixed, permeabilized and labeled with CD45-alexa700, Pancytokeratin-alexa647 and DAPI to verify whether the cells were CTCs. a) Chart showing the number of CTCs recovered from each of three clinical breast samples, where 2 ml of blood was run with each 1-step EpCAM and 2-step cocktail. Patient 1 was stage 3, ER/PR+HER2-, patient 2 was stage 4, ER+PR-HER2+, and patient 3 was stage 2, ER/PR+HER2-. b) Fluorescence images of a cell captured by eDAR using the cocktail labeling scheme. c) A cell detected using the 1-step PE-anti-EpCAM labeling scheme.

These results demonstrate a significant improvement in CTC recovery with our new labeling scheme. They also validate the assertion that there are $EpCAM^{low}$ CTCs in the blood of breast cancer patients. We increased the capability to isolate and characterize these $EpCAM^{low}$ cells that represent a potentially more invasive population of CTCs. This information will be critical in creating a better diagnostic and prognostic tool for breast cancer as well as unraveling the mechanisms of metastasis and developing anti-metastatic drugs.

2d. Apply techniques developed for the prospective analysis of cancer cells (1a and 1c) from patient samples obtained in 2c with single-cell resolution (Months 36-60).

We have completed this task and have developed several single-cell analysis techniques for interrogating the isolated CTCs, especially for the quantification of gene and protein expression. **Figure 23** shows an example, which is a single-cell imaging method with which we were able to detect a large panel of protein biomarkers using cycles of sequential protein labeling followed by fluorescence imaging and photobleaching (see caption for detail).

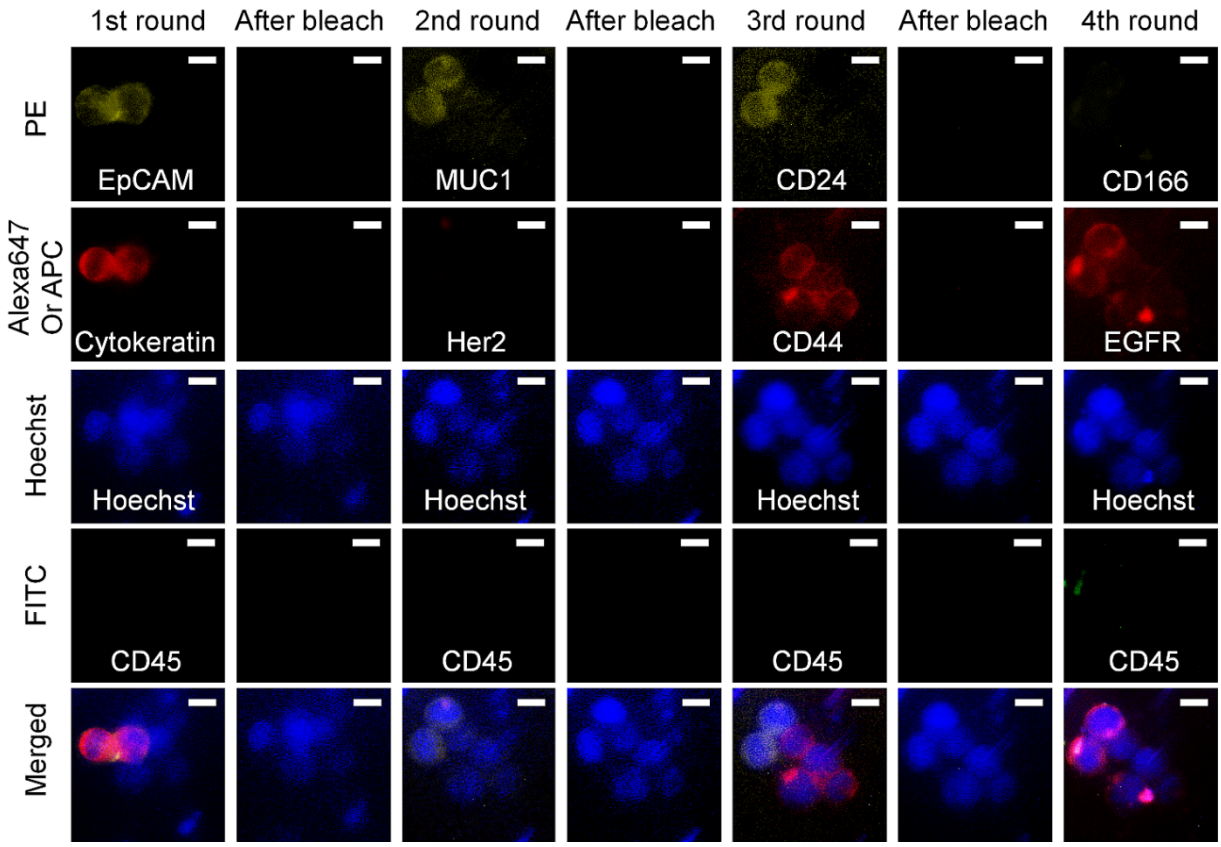


Figure 23. Sequential immunostaining and photobleaching results for six cancer cells trapped on an eDAR chip. The Hoechst nuclear stain was used as a positive control marker and CD45 was used to exclude the potential interference from WBCs. Eight protein markers were studied, including EpCAM/Cytokeratin, MUC1/Her2, CD44/CD24 and CD166/EGFR.

Gail E. Sonenshein, Ph.D., *Tufts University School of Medicine*, *150 Harrison Ave., Boston, MA 02111*

Task 1. Perform mitochondrial DNA mutational analysis. (Months 1 to 16).

1a. Isolate mitochondria from cultured mouse and human mammary breast epithelial or cancer cells that exist in our laboratory and extract DNA and subject it to massively parallel sequencing and bioinformatics to identify mutations. (Months 1-8).

To begin the project, we decided to first determine whether we could specifically amplify mitochondrial DNA (mtDNA) from a total DNA preparation rather than to isolate mitochondria from the various cells and tissues. For amplification, we needed to identify primers that are specific for mtDNA and that are unable to amplify their pseudogenes that exist in the nucleus. After consultation with a number of groups, we chose to use the mtDNA specific primers developed by the laboratories of Taylor and Turnbull, which allow for amplification of mtDNA from a single cell without apparent amplification of the nuclear pseudogenes.[4] To confirm that these primers are unable to amplify nuclear pseudogene DNA, we obtained rho zero cells from Eric Schon (Columbia University). The rho zero cells have almost no mtDNA but retain the nuclear pseudogenes. Thus, they serve as an excellent negative control. We also received matching control 143B cells. DNA amplification with various primer sets was profoundly reduced with the rho zero cell DNA, as reported by Taylor and Turnball, with only an occasional non-specific band seen. In contrast, effective amplification of the control DNA was readily detected (**Figure 24**). Notably these primers could amplify human and mouse mtDNA specifically from total DNA preparations and hence we did not require mitochondria isolation. Thus, in all of our subsequent work, we used the mitochondrial primers of Taylor and Turnbull.

- PCR protocol was optimized
- 58°C annealing temp and 20 cycles
- DNA conc (50 ng/µl)

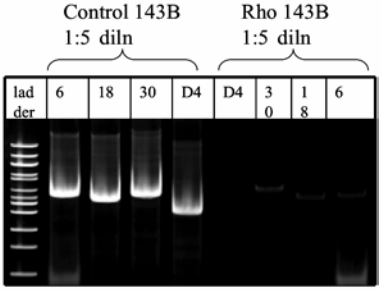


Figure 24. Mitochondrial specific primers of Taylor and Turnbull amplify mitochondrial DNA but not nuclear pseudogenes.

As a reference for subsequent discussion of specific mutations in the mitochondrial genome, the linearized view of mtDNA is given in **Figure 25** with positions of identified genes, tRNAs, rRNAs, and control D-loop DNA indicated.

Using the Taylor and Turnbull primers, we have characterized the mtDNA sequences in an untransformed human breast epithelial cell line (MCF10A), and in wide spectrum of breast cancer cell lines, including carcinogen-transformed derivative lines of MCF10A, HER2+, estrogen receptor α (ER) positive (ER+) and ER negative (ER-) breast cancer cell lines, inflammatory breast cancer (IBC) cells, and also in tumor tissue derived from these lines. The cell lines analyzed are listed in **Table 7**. We compared the mtDNA sequence information obtained to the Revised Cambridge Reference Sequence (rCRS) for the human mitochondrial genome (in the MITOMAP human mitochondrial genome database). Our analysis of the cell lines led to the identification of distinct mtDNA mutations. The findings on the number of mutations found in each cell line are shown in **Table 8**. All of the breast cancer cell lines have multiple mtDNA mutations. The SUM149 inflammatory breast cancer (IBC) line has the highest number of mutations at 99, which did not vary depending upon whether the population was stem-like ($ALDH^{hi}$) or not ($ALDH^{lo}$). The transformed BP1 and D3 lines were derived from MCF10A cells following treatment with benzo[α]pyrene (BαP) and 7,12-dimethylbenz[α]anthracene (DMBA), respectively.[5] These lines have nine additional mtDNA mutations compared to the MCF10A parental cells (and are missing one mtDNA mutation present in the parent line). The IBC SUM190 line, MDA-MB-231, Hs578T and MDA-MB-468 cell lines all contained 30-40 mutations. In Task 2 below, we assessed whether the two ER+ lines listed in Table 7 display altered mtDNA mutations when grown in mouse models. To complete this analysis, we

first determined the mtDNA mutations in the cultured cells (Table 8). We found that the two ER+ cell lines contained 16-19 mtDNA mutations. The lines with the lowest numbers of mtDNA mutations were BT549 and the ER-, PR- and HER2+ IBC cell line (MDA-IBC-3) which was generated from breast cancer cells isolated from the pleural effusion fluid collected from an IBC breast cancer patient; these lines contained 13 and 14 mtDNA mutations, respectively (Table 8). The sequence information data for each of these lines are given below.

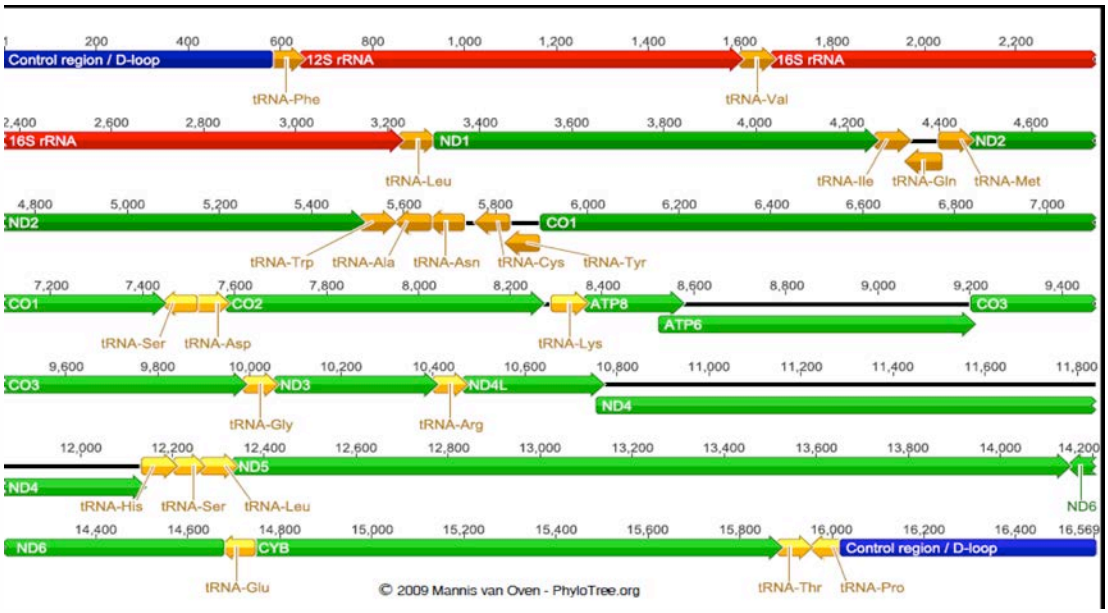


Figure 25. Linearized view of the mtDNA with positions of the 13 genes, 22 tRNAs, 2 rRNAs and the control D-loop indicated.

Table 7: Table listing the panel of cell lines, including both untransformed breast epithelial and breast cancer cell lines, analyzed for mitochondrial DNA mutations.

Untransformed	**ER+**	**ER-**	**Inflammatory breast cancer (IBC)**	**Carcinogen transformed**
HMEC	MCF7	BT549	SUM149	BP1
MCF10A	T47D	MDA-MB-231	SUM190	D3
		Hs578T	MDA-IBC-3	
		MDA-MB-468	SUM149-ALDH+	
		MDA-MB-231-LM1	SUM149-ALDH-	
		MDA-MB-231-LM2		
		SUM1315		

Type	Cell lines	No. of mutations
ER+	MCF7	16
ER+	T47D	19
ER-	BT549	13
ER-	MDA-MB-231 parental cells and isolates that preferentially metastasize to the lung LM1/LM2	31
ER-	Hs578T	32
ER-	MDA-MB-468	33
ER-	SUM1315	30
IBC	SUM149 and isolates that are either $ALDH^{hi}$ or $ALDH^{lo}$	99
IBC	SUM190	41
IBC	MDA-IBC-3	14
Carcinogen treated	BP1	33
Carcinogen treated	D3	33
Untransformed	MCF10A	25

Table 8. Number of mitochondrial DNA mutations observed in untransformed breast epithelial and breast cancer cell lines in comparison to the reference human mitochondrial genome.

A list of the 99 mtDNA mutations detected in the SUM149 IBC cell line is given in **Table 9**. Ten of these mutations are commonly seen in breast cancer patients. One is within the tRNA-asp gene, which likely has profound on protein synthesis within the mitochondria. Compilations of the mutations in the ER- and HER2+ cell lines are given in **Table 10** and **Table 11**, respectively. The mutations in the BP1 and D3 lines, which arise as a result of carcinogen treatment of MCF10A, are giving in **Table 12**. All of these data were used below in Task 1c to identify a panel of common mtDNA mutations.

We also studied the effect of hypoxia (1% oxygen) on mtDNA mutations. This was done since hypoxia typically precedes angiogenesis and tumor growth and it has been reported to cause mutations in nuclear genes. We found that even after prolonged exposure of MCF10A, BP1 and D3 cells to Hypoxia (72 h, 1 week and 2 weeks), there were no additional mtDNA mutations observed (data not shown). Recently another lab has published a similar finding.[6]

TN-IBC SUM149 breast cancer cell line

Gene	n.p.	mutation	Gene	n.p.	mutation	Gene	n.p.	mutation
D-Loop	16184-1618	DEL C	16S rRNA	2394	DEL A	COX III	9301	C-T
D-Loop	16192-1619	INS C	16S rRNA	2706	A-G	COX III	9540	T-C
D-Loop	16223	C-T	16S rRNA	2758	G-A	COX III	9722	T-C
D-Loop	16265	A-C	16S rRNA	2885	T-C	ND3	10274	T-C
D-Loop	16278	C-T	16S rRNA	3107	DEL N	ND3	10321	T-C
D-Loop	16286	C-G	ND1	3594	C-T	ND3	10398	A-G
D-Loop	16294	C-T	ND1	3666	G-A	ND4L	10586	G-A
D-Loop	16311	T-C	ND1	4104	A-G	ND4L	10688	G-A
D-Loop	16360	C-T	ND2	4769	A-G	ND4	10792	A-G
D-Loop	16519	T-C		5899	Insert-C	ND4	10793	C-T
D-Loop	16527	C-T	COX I	5951	A-G	ND4	10810	T-C
D-Loop	73	A-G	COX I	6071	T-C	ND4	10873	T-C
D-Loop	151	C-T	COX I	6150	g-a	ND4	11329	A-G
D-Loop	152	T-C	COX I	6253	T-C	ND4	11654	A-G
D-Loop	182	C-T	COX I	6723	G-A	ND4	11719	G-A
D-Loop	186	C-A	COX I	7028	C-T	ND4	12049	C-T
D-Loop	189	A-C	COX I	7055	A-G	ND5	12705	C-T
D-Loop	195	T-C	COX I	7076	A-G	ND5	12810	A-G
D-Loop	198	C-T	COX I	7142	T-C	ND5	13105	A-G
D-Loop	247	G-A	COX I	7146	A-G	ND5	13149	A-G
D-Loop	263	A-G	COX I	7256	C-T	ND5	13485	A-G
D-Loop	297	A-G	COX I	7300	T-C	ND5	13506	C-T
D-Loop	315	Insert-C	COX I	7337	G-A	ND5	13650	C-T
D-Loop	316	G-A	COX I	7389	T-C	ND5	13789	T-C
D-Loop	471	T-C	tRNA-Asp	7521	G-A	ND5	14000	T-A
D-Loop	523-524	DEL AC	COX II	8027	G-A	ND5	14178	T-C
12S rRNA	750	A-G	ATPASE 8	8468	C-T	ND6	14560	G-A
12S rRNA	769	G-A	ATPASE 6	8655	C-T	CYB	14766	C-T
12S rRNA	825	T-A	ATPASE 6	8701	A-G	CYB	14861	G-A
12S rRNA	1018	G-A	ATPASE 6	8784	A-G	CYB	14911	C-T
12S rRNA	1420	T-C	ATPASE 6	8860	A-G	CYB	15016	C-T
12S rRNA	1438	A-G	ATPASE 6	8877	T-C	CYB	15326	A-G
16S rRNA	2156	INS A	ATPASE 6	9072	A-G	CYB	15784	T-C

Table 9. Mitochondrial DNA mutations observed in Triple-negative IBC SUM149 breast cancer cell line in comparison to the reference human mitochondrial genome. Mutations highlighted in yellow are commonly observed in breast cancer cell lines and patient samples.

ER- breast cancer cell lines

ER -Ve							ER -Ve						
Gene	NP	Mutation	BT549	MDA-MB-231	Hs578T	MDA-MB-468	Gene	NP	Mutation	BT549	MDA-231	Hs578T	MDA-468
D-Loop	73	A>G	+	+	+								
D-Loop	150	C>T				+	COX II	9824	T>C				
D-Loop	152	T>C	+				tRNA-GLY	10005	A>T				
D-Loop	153	A>G		+			ND3	10398	A>G			+	+
D-Loop	195	T>C		+		+	ND3	10400	C>T				
D-Loop	228	G>A			+		ND4	10819	A>G				+
D-Loop	241	A>G			+		ND4	10873	T>C				+
D-Loop	263	A>G	+	+	+	+	ND4	11017	T>C				
D-Loop	291	A>T			+		ND4	11084	A>G				
D-Loop	303	INS C					ND4	11168	A>G				
D-Loop	309-311	INS C	+	+			ND4	11251	A>G			+	
D-Loop	309	C>T				+	ND4	11314	A>G				+
D-Loop	310	T>C				+	ND4	11719	G>A		+	+	+
D-Loop	462	C>T			+		ND4	11722	T>C				
D-Loop	482	T>C			+		ND4	12084	C>T		+		
D-Loop	489	T>C			+		ND5	12612	A>G			+	
D-Loop	523-524	INS C		+			ND5	12705	C>T		+		+
12sRNA	709	G>A		+			ND5	12771	G>A				
12S rRNA	750	A>G	+	+	+		ND5	13105	A>G				+
12S rRNA	1415	G>A					ND5	13708	G>A			+	
12S rRNA	1420	T>C					ND5	13966	A>G		+		
12S rRNA	1438	A>G	+	+	+	+	ND6	14212	T>C				+
16S rRNA	1719	G>A		+			ND6	14364	G>A				
16S rRNA	2352	T>C				+	ND6	14470	T>C		+		
16S rRNA	2626	T>C					CYB	14766	C>T		+	+	+
16S rRNA	2706	A>G		+	+	+	CYB	14783	T>C				
16S rRNA	2772	C>T					CYB	14798	T>C			+	
16S rRNA	3010	G>A			+		CYB	14869	G>A				+
16S rRNA	3200	T>C				+	CYB	14905	G>A				+
ND1	3394	T>C			+		CYB	15043	G>A				
ND1	3550	G>A	+				CYB	15119	G>A			+	
ND1	4216	T>C			+		CYB	15200	G>A	+			
ND1	4257	A>G	+				CYB	15301	G>A				+
ND1	4386	T>C					CYB	15310	T>C		+		
ND2	4769	A>G		+	+	+	CYB	15326	A>G	+	+	+	+
ND2	4823	T>C				+	CYB	15452	C>A			+	
ND2	4958	A>G					tRNA-PRO	15968	T>C			+	
ND2	5263	C>T				+	D-Loop	16021	C>T				+
ND2	5894	A>C					D-Loop	16069	C>T			+	
COX1	6221	T>C		+			D-Loop	16093	T>C		+		
COX1	6227	T>C				+	D-Loop	16126	T>C			+	
COX1	6371	C>T		+			D-Loop	16189	T>C		+		
COX1	6455	C>T					D-Loop	16192	C>T			+	
COX1	7028	C>T		+		+	D-Loop	16209	T>C				
COXII	8222	T>C			+		D-loop	16223	C>T		+		+
ATPASE8	8506	T>C		+			D-Loop	16248	C>T	+			
ATPASE6	8701	A>G				+	D-Loop	16265	A>G		+		
ATPASE6	8860	A>G	+	+	+		D-Loop	16278	C>T		+		
COX II	9299	A>G					D-Loop	16320	C>T				+
COX II	9540	T>C				+	D-Loop	16324	T>C				
COX II	9662	A>G					D-Loop	16519	T>C	+	+	+	+

Table 10. Mitochondrial DNA mutations observed in ER negative breast cancer cell lines in comparison to the reference human mitochondrial genome. +, indicates the presence of a specific mutation. Mutations marked in **red** are high frequency mutations.

Her2+ breast cancer cell lines

			Her2 +ve					Her2 +ve	
Gene	NP	Mutation	SUM190	MDA-IBC-3	Gene	NP	Mutation	SUM190	MDA-IBC-3
D-Loop	73	A>G	+	+					
D-Loop	150	C>T			COX II	9824	T>C	+	
D-Loop	152	T>C			tRNA-GLY	10005	A>T	+	
D-Loop	153	A>G			ND3	10398	A>G	+	
D-Loop	195	T>C			ND3	10400	C>T	+	
D-Loop	228	G>A			ND4	10819	A>G		
D-Loop	241	A>G			ND4	10873	T>C		
D-Loop	263	A>G	+	+	ND4	11017	T>C	+	
D-Loop	291	A>T			ND4	11084	A>G	+	
D-Loop	303	INS C	+		ND4	11168	A>G		+
D-Loop	309-311	INS C	+		ND4	11251	A>G		
D-Loop	309	C>T			ND4	11314	A>G		
D-Loop	310	T>C			ND4	11719	G>A	+	
D-Loop	462	C>T			ND4	11722	T>C	+	
D-Loop	482	T>C			ND4	12084	C>T		
D-Loop	489	T>C	+		ND5	12612	A>G		
D-Loop	523-524	INS C			ND5	12705	C>T	+	
12sRNA	709	G>A		+	ND5	12771	G>A	+	
12S rRNA	750	A>G	+	+	ND5	13105	A>G		+
12S rRNA	1415	G>A		+	ND5	13708	G>A		
12S rRNA	1420	T>C		+	ND5	13966	A>G		
12S rRNA	1438	A>G	+	+	ND6	14212	T>C		
16S rRNA	1719	G>A			ND6	14364	G>A	+	
16S rRNA	2352	T>C			ND6	14470	T>C		
16S rRNA	2626	T>C	+		CYB	14766	C>T	+	
16S rRNA	2706	A>G	+		CYB	14783	T>C	+	
16S rRNA	2772	C>T	+		CYB	14798	T>C		
16S rRNA	3010	G>A		+	CYB	14869	G>A		
16S rRNA	3200	T>C			CYB	14905	G>A		
ND1	3394	T>C			CYB	15043	G>A	+	
ND1	3550	G>A			CYB	15119	G>A		
ND1	4216	T>C			CYB	15200	G>A		
ND1	4257	A>G			CYB	15301	G>A	+	
ND1	4386	T>C	+		CYB	15310	T>C		
ND2	4769	A>G	+		CYB	15326	A>G	+	+
ND2	4823	T>C			CYB	15452	C>A		
ND2	4958	A>G	+		tRNA-PRO	15968	T>C		
ND2	5263	C>T			D-Loop	16021	C>T		
ND2	5894	A>C	+		D-Loop	16069	C>T		
COX1	6221	T>C			D-Loop	16093	T>C		
COX1	6227	T>C			D-Loop	16126	T>C		
COX1	6371	C>T			D-Loop	16189	T>C		
COX1	6455	C>T	+		D-Loop	16192	C>T		
COX1	7028	C>T	+		D-Loop	16209	T>C	+	
COXII	8222	T>C			D-loop	16223	C>T	+	
ATPASE8	8506	T>C			D-Loop	16248	C>T		
ATPASE6	8701	A>G			D-Loop	16265	A>G		
ATPASE6	8860	A>G	+	+	D-Loop	16278	C>T		
COX II	9299	A>G	+		D-Loop	16320	C>T		
COX II	9540	T>C	+		D-Loop	16324	T>C	+	
COX II	9662	A>G	+		D-Loop	16519	T>C		+

Table 11. Mitochondrial DNA mutations observed in Her2+ breast cancer cell lines in comparison to the reference human mitochondrial genome. +, indicates the presence of a specific mutation. Mutations marked in **red** are high frequency mutations.

Gene	n.p.	Mutation	MCF10A	BP1	D3
D-Loop	73	A-G	+	+	+
D-Loop	185	G-A		+	+
D-Loop	228	G-A		+	+
D-Loop	263	A-G		+	+
D-Loop	295	C-T		+	+
D-Loop	307	insert C		+	+
D-Loop	315	insert C		+	+
D-Loop	388	A-T	+		
D-Loop	462	C-T	+	+	+
D-Loop	489	T-C	+	+	+
D-Loop	538	A-G	+		
12S rRNA	750	A-G	+	+	+
12S rRNA	1438	A-G	+	+	+
16S rRNA	2706	A-G	+	+	+
16S rRNA	3010	G-A	+	+	+
ND1	4216	T-C	+	+	+
ND2	4769	A-G		+	+
COX1	6249	G-A	+	+	+
COX1	7028	C-T		+	+
ATPase-6	8860	A-G	+	+	+
ND3	10084	T-C	+	+	+
ND4	11251	A-G	+	+	+
ND4	11719	G-A	+	+	+
ND5	12612	A-G	+	+	+
ND5	13063	G-A		+	+
ND5	13708	G-A	+	+	+
CYB	14766	C-T	+	+	+
CYB	14798	T-C	+	+	+
CYB	15326	A-G	+	+	+
CYB	15452	C-A	+	+	+
CYB	15766	A-G	+	+	+
D-Loop	16069	C-T	+	+	+
D-Loop	16126	T-C	+	+	+
D-Loop	16319	G-A	+	+	+

Table 12. Mutations observed in carcinogen treated cancer cell lines in comparison to parent untransformed line (MCF10A). All mutations highlighted in **red** are mutations unique to MCF10A. Mutations highlighted in **blue** are new mutations observed in the carcinogen-treated line.

1b. Isolate mitochondria from normal tissue and tumors, extract DNA and subject it to massively parallel sequencing (Months 8-12).

We developed methods to extract DNA from both frozen and formalin fixed paraffin embedded (FFPE) tumor tissue. For frozen tumor samples, the tumor tissue was ground in a chilled mortar pestle and DNA extracted using a genomic DNA kit (Macherey-Nagel). To extract DNA from FFPE tissue samples, the EX-WAX DNA extraction kit (EMD Millipore) was employed to avoid use of xylene for paraffin dissolution as it affects DNA quality. An average of 2000 ng of total DNA was obtained from 4-5 slices of FFPE tissue ~5 microns thick. DNA was amplified using a double primer method and the Q5 DNA polymerase, which worked better than the Taq and AmpliTaq gold polymerases.

These procedures were used to isolate DNA from orthotopic tumors grown in mice using various breast cancer cell lines, including SUM149, MDA-MB-231 and its clones and to analyze the mtDNA. The mitochondrial mutations present in the tumors were identical to the cell lines grown in culture (data not shown). Thus, no new mutations were detected in the tumor tissue during its growth in the mammary fat pad.

These results confirmed that we could identify mtDNA mutations in tumor samples obtained from mice. The data also suggested that during the relatively rapid time for tumor growth in this model that the tumor microenvironment in the NOD/SCID mouse model did not induce novel mtDNA mutations in the cancer cells. We also used this technology to analyze tumor tissue obtained from the Human in Mouse (HIM) breast tumor model, which was provided by the Kuperwasser lab (Task 3a).

1c. Compare mtDNA from normal mammary tissue and from tumor tissue using bioinformatics to identify mutations. (Months 10-16).

We used the MITOMAP human mitochondrial genome database and identified a total of **40** mtDNA mutations in the breast cancer cell lines listed above that were in common with those observed in breast cancer patient tissue (Table 13). The highest concentrations of mutations were observed in the D-loop and Complex I. Several of the D-loop mutations were not specific to breast cancer. There are no characteristic mutations that distinguish ER+ vs ER- cell lines. These data along with the results obtained in Aims 2 and 3 have been used to generate an initial panel of mtDNA mutations to be tested in patient samples.

Mutation	Gene	Mutation	Gene
73 A>G	D-LOOP	7028 C>T	CO1
152 T>C	D-LOOP	7521 G>A	CO1
185 G>A	D-LOOP	8764 G>A	ATPASE6
189 A>G	D-LOOP	8701 A>G	ATPASE6
195 T>C	D-LOOP	8860 A>G	ATPASE6
198 C>T	D-LOOP	9540 T>C	COIII
228 G>A	D-LOOP	10873 T>C	ND4
263 A>G	D-LOOP	10398 A>G	ND4
295 C>T	D-LOOP	11251 A>G	ND4
311 INSERT C	D-LOOP	11719 G>A	ND4
489 T>C	D-LOOP	12612 A>G	ND5
523 A-DEL	D-LOOP	12705 C>T	ND5
750 A>G	12s	13650 C>T	ND5
769 G>A	12S	13263 G>A	D-LOOP
1438 A>G	12S	15326 A>G	D-LOOP
2706 A>G	16S	16278 C>T	D-LOOP
3010 G>A	16S	16294 C>T	D-LOOP
4216 T>C	ND1	16311 T>C	D-LOOP
4769 A>G	ND2	16319 G>A	D-LOOP
3594 C>T	ND1	16519 T>C	D-LOOP

Table 13. Common mutations observed in breast cancer cell lines and patient tissue. All mutations highlighted in **red** are high frequency mutations not specific to breast cancer and in **black** are low frequency mutations more specific to breast cancer.

Task 2. Characterize mtDNA mutations resulting from oncogene expression in breast tumors (Months 16-24).

2a. Compare mtDNA from mammary tissue vs tumors or derived cell lines of transgenic mice driven by oncogenes implicated in breast cancer. Samples obtained from Kuperwasser (months 16-24)

We obtained unmatched serum and tumor tissue samples from HIM tumors with SV40/*K-Ras* mutations from the Kuperwasser lab. We identified 32 mtDNA mutations induced by the K-*Ras* oncogene in these tumors (**Table 14**). Interestingly, K-Ras oncogene seems to induce a novel tRNA mutation in these spontaneously occurring tumors. Other known tRNA-cys mutations are associated with a variety of diseases thus suggesting that this novel finding could be important in identifying K-Ras induced tumors.

We also looked at HIM tumor samples with a mutation in the *BRCA1* oncogene. Breast tumors from patients with *BRCA1* gene mutation are very difficult to detect by mammography and these patients are more susceptible to irradiation. DNA was analyzed from mice carrying HIM tumors with *BRCA1*

mutations. We identified 31 mutations in the BRCA1 tumor sample (**Table 15**). Interestingly we found two tRNA mutations in this tumor tissue, which is very rare. Of note, one of these mutations A12308G (tRNA-leu) has been reported to correlate with breast cancer risk and colorectal cancer risk.[7-8] It is interesting to note that SUM149 cell line, which carries a *BRCA1* mutation, also has an unusual tRNA mutation G7521A (tRNA-asp), as discussed above.

Overall, our analysis indicates that oncogenes may lead to specific mtDNA mutations in breast cancer that may prove to be good biomarkers for disease onset and recurrence.

Gene	n.p.	Mutation
D-Loop	146	T>C
D-Loop	195	T>C
D-Loop	263	A>G
D-Loop	310	insert C
D-Loop	317	insert C
12S rRNA	709	G>A
12S rRNA	750	A>G
12S rRNA	775	C>A
12S rRNA	794	insert C
12S rRNA	1438	A>G
16S rRNA	1793	G>A
16S rRNA	3248	insert G
ND1	3437	G>C
ND1	3692	G>C
ND1	4083	G>A
ND2	4700	insert A
ND2	4769	A>G
ND2	5095	T>C
tRNA-cys	5773	G>T
COX1	6218	insert T
COX1	6572	insert G
COX2	8107	A>G
ATPASE6	8764	G>A
ATPASE6	8860	A>G
COX3	9812	C>T
ND5	13102	A>C
ND5	14122	C>T
CYB	15016	insert T
CYB	15326	A>G
D-Loop	16068	T>C
D-Loop	16288	T>C
D-Loop	16362	T>C

Table 14. mtDNA mutations seen in HIM mammary tumors driven by the SV40/*K-Ras* oncogene. Mutations highlighted in yellow are commonly observed in breast cancer cell lines and patient samples.

Gene	n.p	Mutation
D-loop	16327	C>T
D-Loop	16519	T>C
D-Loop	73	A>G
D-Loop	146	T>C
D-Loop	263	A>G
D-Loop	285	T>C
D-Loop	315	ins c
12S rRNA	1438	A>G
16S rRNA	2387	T>C
16S rRNA	2706	A>G
ND1	3425	A>G
ND2	4769	A>G
ND2	5460	G>A
COX1	6365	T>C
COX1	7028	C>T
ATPASE8	8395	C>T
ATPASE6	8860	A>G
tRNA-Arg	10469	insert G
ND4	10886	T>C
ND4	11467	A>G
ND4	11566	A>G
ND4	11719	G>A
tRNA-Leu	12308	A>G
ND5	12372	G>A
ND5	12879	T>C
ND5	13104	A>G
CYB	14766	C>T
CYB	15110	G>A
CYB	15148	G>A
CYB	15172	G>A
CYB	15326	A>G

Table 15. mtDNA mutations observed in BRCA1 Mutant HIM tumor tissue. Mutations highlighted in yellow are commonly observed in breast cancer cell lines and patient samples. The mutation highlighted in blue is in a mitochondrial tRNA gene reported to be associated with breast cancer risk.

2b. Characterize mtDNA alterations in tumor lines that will be utilized by the Kuperwasser laboratory in their 'HIM' analysis. DNA samples will be analyzed as in Task 1 above (Months 18-24).
The Kuperwasser laboratory has proposed using ER+ tumor lines MCF7 and T47D in their HIM model. Thus, we characterized the mtDNA alterations in these lines. Our data are summarized in **Table 16**.

ER+ breast cancer cell lines

MCF7		T47D	
263	**A>G**	**263**	**A>G**
315	**INS C**	309	INS C
750	A>G	**315**	**INS C**
1438	**A>G**	**1438**	**A>G**
2591	A>G	2355	A>G
4769	**A>G**	2442	T>C
5899	INS C	2706	A>G
6776	T>C	3847	T>C
8860	**A>G**	4767	A>G
9966	G>A	**4769**	**A>G**
13260	T>C	7028	C>T
14319	T>C	**8860**	**A>G**
15326	**A>G**	13188	C>T
15380	A>G	13481	G>A
16148	C>T	14766	C>T
16519	**T>C**	**15326**	**A>G**
		15674	T>C
		16362	T>C
		16519	**T>C**

Table 16. Mitochondrial DNA mutations observed in MCF and T47D cell lines. **Bold** font indicates the mutations found in common between the two cell lines.

We also obtained tumor tissue from the HIM tumors derived from the SUM1315 breast cancer cell line. DNA from these tissues was amplified and found to have no new mutations when compared to SUM1315 cell line. These data, along with others obtained from our lab, confirmed that the tumor microenvironment is not inducing additional mtDNA mutations during tumor progression under the conditions being used.

Task 3. Recurrence prediction: Determine whether mtDNA mutations can be used to test for breast tumor regression in mice (Months 18-60).

We used the technology developed in our laboratory to assess whether mtDNA in the blood can be used to evaluate tumor regrowth and recurrence. Briefly, NOD/SCID mice were injected with 2.5 x 10^6 MDA-MB-231 cells in the mammary fat pad. Mice were bled a day prior to injection and at day seven and day 16 after cancer cell inoculation. The tumors were resected when they reached an average volume of 0.2 cm^3. Mouse 3 died before resection. Blood samples were collected at 10, 15, 17, 25, 32 and 33 days after resection. As shown in **Table 17** below, tumor resection was successfully performed on 11 mice, which were then followed for a total of 33 days. Out of these 11 mice, six mice had primary recurrence over the five week period. All of the mice developed macro or micro metastases to the lung, including the ones

that had no primary recurrence. Four of the mice developed brain metastases (Table 17). We analyzed the whole blood and tissue samples for mtDNA mutations.

Table 17. Summary of the primary recurrences, or brain or lung metastases in specific mice that have undergone tumor resection. Most of the metastases observed in the brain were smaller micro metastases and those observed in the lung were macro metastases.

Mouse No#	# 1	# 2	#3 Died	# 4	# 5	# 6	#7	#8	#9	#10	#11	# 12
Primary recurrence	+ 3/20/14 17 days after resection	+ 4/03/14 33 days after resection		+ 3/7/14 4 days after resection	+ 3/7/14 4 days after resection	None	None	+ 3/14/14 11 days after resection	+ 3/7/14 4 days after resection	None	None	None
Brain metastasis	+m	+m		ND	+m	–	+	–	ND	–	–	–
Lung metastasis	+m	+		ND	+	+m	+	+	ND	+	+	+

+m: multiple metastases, +: single metastases, –: no metastases, ND: not determined.

One question addressed in this study was whether mtDNA can be observed in mice with early recurrence. We were able to identify mtDNA mutations in blood obtained from mice 10 days after resection. Not only did we observe mutations in mice that had early recurrence at Day 4 after resection (mice 4, 5, and 9), we also observed mutations in mice that showed primary recurrence at a slightly later date (mouse 8) (**Table 18**). We thus showed that mitochondrial DNA mutations were observed in mice with both early and late recurrence.

Additionally, we examined whether mtDNA mutations can be observed in mice with no primary recurrence but lung and brain metastases. As mentioned above, our data showed that we can detect mtDNA mutations in blood of mice that show recurrence (Table 18). We then analyzed the blood of mice that did not show recurrence, like mice 6, 7, and 12, to determine if we can predict tumor recurrence before it happens. We were able to detect mtDNA mutations in mice with no primary recurrence as early as 10 days after resection. To confirm that the presence of mtDNA mutation was from the metastasis and not from the residual cells left from the tumor, we analyzed blood obtained from these mice 17 days after resection (Table 18). Our results showed that we can identify mtDNA mutations in the blood of all mice tested, thus providing evidence that specific mtDNA mutations may hold potential as markers of disease for early detection of breast cancer recurrence in patients.

Table 18. Mutations observed in blood samples of mice with recurrence collected 10 or 17 days after resection. All mutations highlighted in red are observed in blood samples of mice 4 and 9, ten days after resection. Some of these mutations were also observed in 5, 6, 7, 8, and 12. This was the earliest time point at which we collected blood after resection. Mutations marked in **red** were identified in the blood, while mutations in **black** have not been observed in the blood.

MDA_MB-231			Observed in blood of mice 4 and 9	Observed in blood of mice 5, 8 and 7 on 3/13/14 10 days after resection (AFR)	Observed in blood of mice 7, 6, 12 on 3/13/14 10 days AFR	Observed in blood of mice 7, 12, 6 3/20/14 17 days AFR
gene	np	mutation				
D-Loop	16093	T>C	yes			
D-Loop	16189	T>C	yes			
D-Loop	16265	A>G				
D-Loop	16278	C>T				
D-Loop	16519	T>C				
D-Loop	73	A>G	yes	yes		
D-Loop	153	A>G	yes	yes		
D-Loop	195	T>C	yes	yes		
D-Loop	263	A>G	yes	yes		
D-Loop	309-311	ins c	yes			
D-Loop	523-524	INS C				
12sRNA	709	G>A				
12sRNA	750	A>G	yes			
12sRNA	1438	A>G				
16sRNA	1719	G>A				
16sRNA	2706	A>G				
ND2	4769	A>G	yes			
COX1	6221	T>C	yes	yes	yes (7,6)	
COX1	6371	C>T	yes	yes	yes (7,6)	
COX1	7028	C>T				
ATPASE8	8506	T>C	yes			
ATPASE6	8860	A>G	yes			
ND4	11719	G>A	yes			
ND4	12084	C>T	yes			
ND5	12705	C>T	yes			yes (12, 7,6)
ND5	13966	A>G				
ND6	14470	T>C	yes			
CYB	14766	C>T	yes	yes		
CYB	15310	T>C	yes	yes	yes (6)	yes (12)
CYB	15326	A>G	yes	yes	yes (6)	yes (12)
D-loop	16223	C>T				

We also sought to establish a panel of mtDNA mutations for early detection and recurrence. To date, we have obtained patient tissue samples for panel validation. We are currently identifying the total number of mtDNA mutations observed in each TNBC patient sample and the copy number for each mutation.

3b. DNA will be isolated from serum collected before resection and weekly after resection that will be provided by the Kuperwasser laboratory (Months 18-30).

We have demonstrated that we can detect mtDNA in the blood of mice at a stage the tumor is barely palpable, when circulating tumor cells (CTCs) are detectable. Together these data suggest that mtDNA mutations can be used as a marker for detection of early stage breast cancer recurrence in patients. We tested blood samples obtained from Kuperwasser lab from SUM1315 derived HIM tumors and were also looking for an increase in the frequency of mutations with tumor progression. Our studies showed that serum volume size is less than 20 µl per time point and is insufficient for DNA extraction using our protocol. We tried to pool some samples together from the same time point, but the DNA obtained was still insufficient for PCR. Thus, we propose in the future to focus on whole blood samples, which contain more mtDNA.

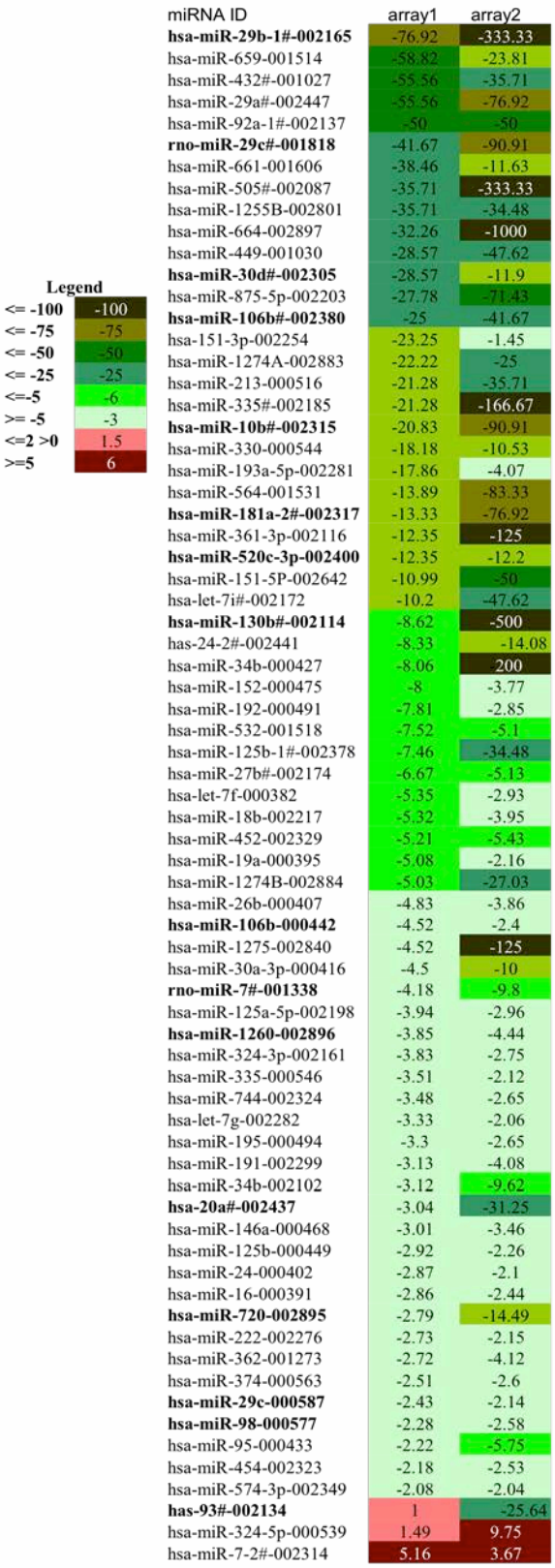
In summary, in the last five years we have identified common mutations observed in breast cancer that can be used as potential markers for early detection and recurrence. We identified novel mtDNA mutations induced by KRAS and BRCA1 mutant proteins leading to breast cancer. Specifically we have identified unusual tRNA mutations in a cell lines and a tumor sample with a BRCA1 mutant, which has been linked to increased breast cancer risk in patients. Taken together, these data justify further study into the role and importance of mtDNA tRNA mutations in BRCA1 carriers as markers for increased risk and early detection. We have also successfully completed our mouse experiment and have answered important questions about the detection of mtDNA mutations during tumor progression and recurrence. We have shown that we can detect mtDNA mutation as early as 10 days after resection when there is no primary recurrence. Most importantly we can detect mtDNA mutation in blood of mice that have no primary recurrence but have lung and brain metastases.

Task 4.

As the Walt lab made a significant attempt to assess the use of miRNAs as breast cancer biomarkers, we focused some of our efforts on these small non-coding RNAs. These miRNAs represent potential biomarkers for breast cancer detection as many are released from tumor cells and can be present stably in the blood stream. Our initial strategy was to test whether miRNAs were regulated by a protein ADAM8, which is present on 50% of all breast cancer metastases. Elucidation of specific miRNAs could provide new important breast cancer biomarkers for the Walt Lab to develop a detection assay. As background, we have previously demonstrated that ADAM8 (**A D**isintegrin **a**nd **M**etalloprotease protein 8) mediates cell adhesion, migration, signaling, and proteolysis of extracellular matrix. Notably, high ADAM8 mRNA levels correlate with metastatic relapse in patients and almost half of all breast metastases are positive for ADAM8 staining.[9] While substantial progress was made in accomplishing our goals in this task, we terminated our efforts when the Walt lab decided not to pursue miRNAs in their project. Our work, including the mouse and patient studies, which were not part of the original SOW, were completed with funding from other sources. The published findings are described below. The support from this grant was gratefully acknowledged in our publication (Das et al., 2016).

The project was initiated to identify miRNAs regulated by ADAM8. To this end, we knocked down ADAM8 RNA levels in MDA-MB-231 cells using either of two ADAM8 siRNAs for 72 hours. RNA was isolated from two independent experiments (runs) and subjected to TaqMan low density miRNA array cards that could detect all miRNAs known at the time. We identified 68 miRNAs differentially regulated upon ADAM8 knockdown (**Figure 26**), including decreased levels of secreted miR-720, and confirmed changes using real-time PCR. We selected miR-720 for further study, as it has been detected in the serum

in breast cancer patients, and confirmed it is indeed regulated by ADAM8. Ectopic overexpression of wild-type ADAM8 or forms that lack metalloproteinase activity similarly induced miR-720 levels. The disintegrin and cysteine-rich domains of ADAM8 were shown to induce miR-720 via activation of a β1-integrin to ERK signaling cascade. Knockdown of miR-720 led to a significant decrease in migratory and invasive abilities of TNBC cells. Conversely, miR-720 overexpression rescued these properties. A profound increase in plasma levels of miR-720 was detected 7 days after TNBC cell inoculation into mouse mammary fat pads when tumors were barely palpable (**Figure 27**). Concordantly, miR-720 levels were found to be significantly higher in serum samples of TNBC patients with high ADAM8 expression. Thus, miR-720 in a panel with other miRNAs downstream of ADAM8 hold promise as biomarkers for early detection or of treatment response of ADAM8-positive TNBCs.



miRNA ID	array1	array2
hsa-miR-29b-1#-002165	-76.92	-333.33
hsa-miR-659-001514	-58.82	-23.81
hsa-miR-432#-001027	-55.56	-35.71
hsa-miR-29a#-002447	-55.56	-76.92
hsa-miR-92a-1#-002137	-50	-50
rno-miR-29c#-001818	-41.67	-90.91
hsa-miR-661-001606	-38.46	-11.63
hsa-miR-505#-002087	-35.71	-333.33
hsa-miR-1255B-002801	-35.71	-34.48
hsa-miR-664-002897	-32.26	-1000
hsa-miR-449-001030	-28.57	-47.62
hsa-miR-30d#-002305	-28.57	-11.9
hsa-miR-875-5p-002203	-27.78	-71.43
hsa-miR-106b#-002380	-25	-41.67
hsa-151-3p-002254	-23.25	-1.45
hsa-miR-1274A-002883	-22.22	-25
hsa-miR-213-000516	-21.28	-35.71
hsa-miR-335#-002185	-21.28	-166.67
hsa-miR-10b#-002315	-20.83	-90.91
hsa-miR-330-000544	-18.18	-10.53
hsa-miR-193a-5p-002281	-17.86	-4.07
hsa-miR-564-001531	-13.89	-83.33
hsa-miR-181a-2#-002317	-13.33	-76.92
hsa-miR-361-3p-002116	-12.35	-125
hsa-miR-520c-3p-002400	-12.35	-12.2
hsa-miR-151-5P-002642	-10.99	-50
hsa-let-7i#-002172	-10.2	-47.62
hsa-miR-130b#-002114	-8.62	-500
has-24-2#-002441	-8.33	-14.08
hsa-miR-34b-000427	-8.06	200
hsa-miR-152-000475	-8	-3.77
hsa-miR-192-000491	-7.81	-2.85
hsa-miR-532-001518	-7.52	-5.1
hsa-miR-125b-1#-002378	-7.46	-34.48
hsa-miR-27b#-002174	-6.67	-5.13
hsa-let-7f-000382	-5.35	-2.93
hsa-miR-18b-002217	-5.32	-3.95
hsa-miR-452-002329	-5.21	-5.43
hsa-miR-19a-000395	-5.08	-2.16
hsa-miR-1274B-002884	-5.03	-27.03
hsa-miR-26b-000407	-4.83	-3.86
hsa-miR-106b-000442	-4.52	-2.4
hsa-miR-1275-002840	-4.52	-125
hsa-miR-30a-3p-000416	-4.5	-10
rno-miR-7#-001338	-4.18	-9.8
hsa-miR-125a-5p-002198	-3.94	-2.96
hsa-miR-1260-002896	-3.85	-4.44
hsa-miR-324-3p-002161	-3.83	-2.75
hsa-miR-335-000546	-3.51	-2.12
hsa-miR-744-002324	-3.48	-2.65
hsa-let-7g-002282	-3.33	-2.06
hsa-miR-195-000494	-3.3	-2.65
hsa-miR-191-002299	-3.13	-4.08
hsa-miR-34b-002102	-3.12	-9.62
hsa-20a#-002437	-3.04	-31.25
hsa-miR-146a-000468	-3.01	-3.46
hsa-miR-125b-000449	-2.92	-2.26
hsa-miR-24-000402	-2.87	-2.1
hsa-miR-16-000391	-2.86	-2.44
hsa-miR-720-002895	-2.79	-14.49
hsa-miR-222-002276	-2.73	-2.15
hsa-miR-362-001273	-2.72	-4.12
hsa-miR-374-000563	-2.51	-2.6
hsa-miR-29c-000587	-2.43	-2.14
hsa-miR-98-000577	-2.28	-2.58
hsa-miR-95-000433	-2.22	-5.75
hsa-miR-454-002323	-2.18	-2.53
hsa-miR-574-3p-002349	-2.08	-2.04
has-93#-002134	1	-25.64
hsa-miR-324-5p-000539	1.49	9.75
hsa-miR-7-2#-002314	5.16	3.67

Figure 26. Heat map representation of miRNA expression displaying a >2-fold change presented as fold changes in *siADAM8* compared to *siCtrl* in two independent experiments (Arrays 1 and 2). The legend for the fold changes in the heat map is given to the left.

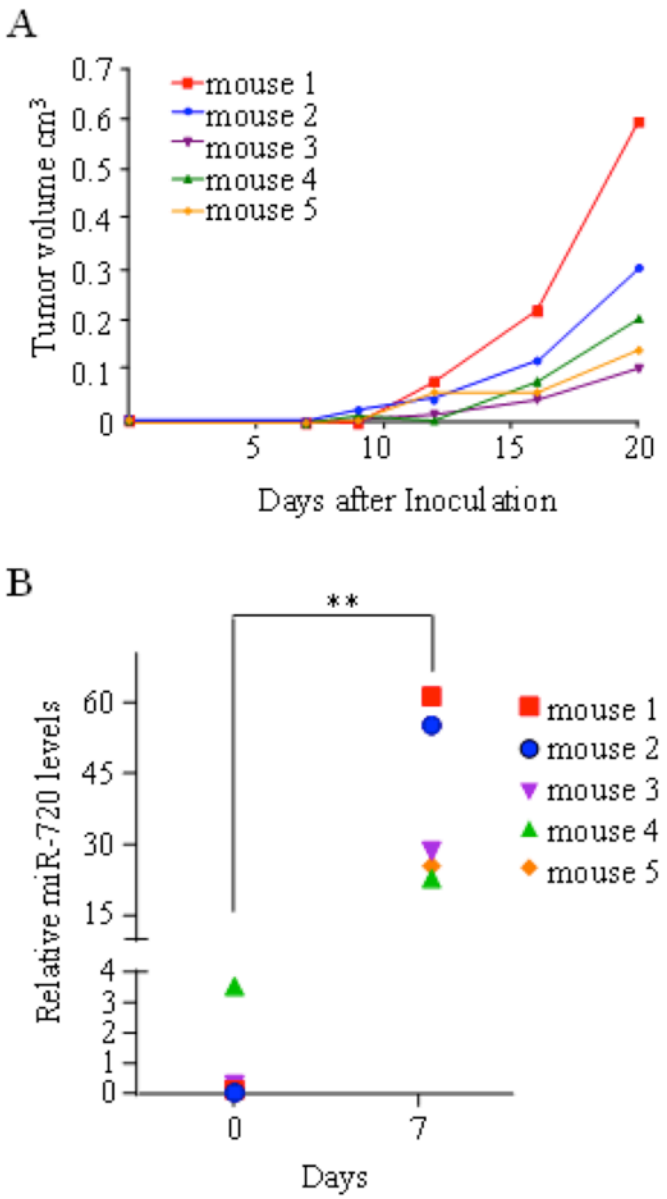


Figure 27. miR-720 is detected in plasma of mice before tumors become palpable.
MDA-MB-231 shCtrl-3 cells (2.5 x 10^6 in 40 µL 50% Matrigel) were implanted on Day 0 into the mammary fat pad of female mice (n = 5). Tumor volume was measured on the indicated days. Blood was collected on the day prior to implantation and on Day 7 by submandibular bleeding and centrifuged to obtain clear plasma. Total RNA was collected from plasma samples and miR-720 levels were determined using Real-Time PCR. A) Tumor growth in individual mice from day 0 to day 21. Tumor volume (cm^3) is presented. B) Relative miR-720 levels in plasma of the individual mice at day 0 and day 7. **P < 0.005, Student's t-test.

Rachel Buchsbaum, Tufts Medical Center, 75 Kneeland St., Boston, MA 02111

The Buchsbaum research effort includes both laboratory and clinical research aims. The purpose of the laboratory effort is to determine the effect of the tumor microenvironment on serum markers of metastasis in the HIM model. The purpose of the clinical research effort is to validate the serum markers of metastasis derived from the Walt and Chiu pre-clinical experiments using human serum samples.

Task 1. Determine effect of tumor microenvironment on serum markers of metastasis in the HIM model (Months 1-36).

1a. Develop mammary fibroblast lines with range of Tiam1 (both down- and up-regulated). (Months 1-6)

Completed as detailed in prior reports.

1b. Utilize HIM breast tumor model in collaboration with Dr. Kuperwasser.
In earlier Annual Progress Reports, we summarized our analysis of the effects of Tiam1-manipulated fibroblasts in the 3D in vitro breast cancer microenvironment, as well as two mouse models of human breast cancer metastasis. We also showed initial results using a chemical inhibitor (here called inhibitor 1) to the Tiam1-deficiency pathway and its effect in blocking metastasis in the second mouse model. In this model, the mice receive xenografts with breast cancer cells isolated from 3D co-cultures after exposure to engineered mammary fibroblasts. In the original protocol, we used inhibitor 1 to treat 3D tissue cultures, prior to xenograft implantation in the mice.

During the prior year, we have initiated a modification to the second mouse model of human breast cancer metastasis. In the new modified model, mice are treated directly with the inhibitor, rather than treating the co-cultures with the inhibitor. To do this, we worked with our supplying collaborator to derive a suitable supply of the inhibitor. We applied for and received approval from our institutional DLAM to amend our current animal protocol to permit this treatment. At the end of the prior year, and continuing in the current reporting year, we initiated the 3D tissue cultures necessary for the implantation experiments, and treated sixteen additional mice on this modified protocol. All mice developed primary tumors at the site of implantation, and received direct dosing with inhibitor 1, consisting of daily i.p. injections beginning approximately two weeks after implantation (once staples were removed). Mice tolerated the treatment for only 2-3 weeks before developing weight loss and ruffled fur, requiring sacrifice one-two weeks earlier than usual. Terminal serum specimens were obtained on all mice. Pathology studies were obtained on tumors and lungs, confirming primary tumors and showing no clear lung metastases. This is likely due to the early sacrifice required by developing toxicity to inhibitor 1 treatment.

Thus, we have focused on the development of more potent, less toxic inhibitors, together with our collaborator. From an earlier screen of six additional related compounds testing potential inhibitory ability against the Tiam1-deficiency pathway, we have now identified two potential novel inhibitors. During the earlier reporting period, these screening tests were complicated by varying issues affecting the reproducibility of results. These have now been largely resolved. During the current reporting period we have focused on one of the two novel inhibitors (here termed inhibitor 2). We have now reproducibly demonstrated several-fold increased potency of inhibitor 2 in both the screening studies as well as more complex first and second-phase functional assays of migration, invasion, and cancer stem cell populations. We have now also gone on to test 15 derivative compounds synthesized by our collaborator based on the findings with inhibitors 1 and 2. We have identified active inhibition of the Tiam1-deficiency pathway in three of these derivatives, which include derivatives of both inhibitors 1 and 2. We are currently conducting expanded dose titration experiments in vitro to establish optimal conditions for testing this new compound in functional studies based on the 3D culture model. The goal is to establish an optimized

inhibitor compound for testing in the mouse xenograft models. As before, serum samples will be obtained and banked during all mouse experiments for potential marker assays.

Finally, during this reporting period we also initiated experiments testing the effect of our inhibitor 1 in an expanded phenotypic range of breast cancer cell lines. Our original tests were performed using triple-negative breast cancer lines. We have found that two hormone-receptor positive cell lines formed spheroids in 3D co-culture with our engineered fibroblasts, and that exposure to Tiam1-deficiency fibroblasts enhanced cancer cell invasion into matrix. Experimental results with these two new lines suggest that incorporation of inhibitor 1 blocks the effect of Tiam1-deficient fibroblasts on cancer cell invasion and mammosphere formation. This result suggests that the Tiam1-deficiency pathway in breast cancer-associated fibroblasts can affect multiple different sub-types of breast cancers, thus increasing the potential clinical impact of these studies.

Task 2. Perform retrospective clinical trial to validate candidate markers from Walt single molecule studies using banked repository samples. (Months 1-60)

During the previous reporting periods, we worked with statisticians in the Tufts Research Design Center within the Tufts Clinical and Translational Sciences Institute to define needed numbers of human serum samples for testing of the initial biomarker panel, based on prevalence estimates derived from reports of circulating tumor cells.

We also identified potential samples within the NIH PLCO (Prostate, Lung, Colorectal and Ovarian) Cancer screening trial. In particular, this trial includes samples from women prior to development of breast cancer, which are otherwise very difficult to find in suitable number and sample volume. A Concept Application was submitted to the PLCO as the initial step in the process for access to the samples, based on recent data from the Walt and Kuperwasser labs derived from proof-of-concept experiments demonstrating the ability to detect circulating PSA in mice implanted with human prostate cancer cells prior to the development of measurable tumors.

During this reporting period, we received notification that our application to the EEMS-PLCO trial was denied for the winter cycle. The primary reason was the volume of serum sample required. Secondary concerns were the justification of the request for post-diagnosis serum samples and a request to see more details of the preliminary data. At the current time, the EEMS-PLCO is not accepting new applications due to internal restructuring. We anticipate resubmission of a revised application once the application process is again open. The Walt lab is working on reducing the sample volumes required.

During this reporting period, we also continued our search for other suitable biorepositories for provision of serum samples. This search has included extensive review of many of the biorepository resources listed through the NCI Specimen Resource Locator (https://specimens.cancer.gov/) as well as other biorepositories. To date, we have identified two other potential sources of specimens suitable for this project: the Baltimore Longitudinal Study of Aging (http://www.blsa.nih.gov/) and the Gundersen Foundation BioBank (http://www.gundersenhealth.org/biobank). The defined criteria for sample suitability include the ability to obtain accompanying curated clinical information regarding tumor status and date of sample acquisition with respect to treatment.

Multiple phone calls to the BLSA requesting further information about access to samples have not been returned. However, our initial application to the Gundersen Foundation has identified over 130 suitable samples for this project in a search of their biorepository. Access to these samples will require significant additional cost, so these samples have not yet been obtained.

We have initiated the request for verification to our institutional IRB regarding our judgment that working with de-identified serum samples from public biobanks such as the EEMS-PLCO and the Gundersen

Foundation BioBank does not constitute Human Subjects Research. Our IRB has confirmed our judgment that working with de-identified serum samples from public biobanks such as these does not constitute Human Subjects Research. Our IRB has requested that we obtain confirmation from the biobanks that samples will be de-identified, prior to accessing any such samples.

Task 3. Perform prospective clinical trial for predictive and prognostic markers in women with newly diagnosed breast cancer.

Progress on developing a bank of human serum samples for validation of candidate markers is continuing. The goal is to establish a fully curated bank of serum samples from all women with newly diagnosed breast cancer, women who are being followed after completing treatment for localized breast cancer, and women with active metastatic breast cancer being treated at Tufts Medical Center. Particular focus has been on banking serum from women with newly diagnosed breast cancer, prior to definitive surgery, as these samples are not generally available from outside serum collections. Recruitment of consenting subjects and serum samples to our Tissue Repository is currently ongoing. During this period, we have also successfully recruited other clinicians to this effort, most importantly breast cancer surgeons and breast center personnel. The resulting significant increase in consenting subjects and requests for sample collections exposed a number of logistical issues in the process of handling specimens in the main laboratory prior to delivery to the Tissue Repository. During the current reporting period, substantial progress was made in identifying and addressing these process issues. At this time, we have successfully collected over 170 samples. The success rate of obtaining and banking ordered samples from consenting subjects has improved to 96% during this reporting year (135/140 samples) from 64% previously (41/64 samples).

Charlotte Kuperwasser, PhD, *Tufts University School of Medicine, 136 Harrison Avenue, Boston, MA 02111*

Task 1 Early Detection (Months 1 to 36)

1a. Utilize the 'Human-In-Mouse' breast tumor model to determine if one can identify biomarkers of early tumor progression and tumor growth. Collect blood and breast tissues at various stages of progression: Normal, hyperplasia, DCIS, invasive cancer (Months 1-18). HIM tumors will be created from human breast epithelial cells collected from discarded tissues of women who have undergone reduction mammoplasty surgeries. For more information/details on model: Nature Protocols, 2006, 1, 595-599.

As described in the quarterly reports, Task 1 has been completed.

We have generated nearly 20 tumors and have provided the Walt lab with all the samples they have requested for their studies.

In addition, while the animal studies and generation of HIM tumors have been valuable for the studies by the Walt and Sonenshein laboratories, the latter portion of the project was focused on testing human samples for the markers identified and validated from HIM samples.

Task 2. Response to chemotherapy and hormone therapy (Months 24-48).

2a. Inject traditional breast cancer cell-line based xenografts into mice using hormone receptor positive lines and hormone receptor negative and Her2+ lines (MCF7, T47D, HCC1418, SUM225, BT20, SUM149, SUM159)(Months 24-36). Human breast cancer cell lines are all commercially available and previously characterized. For more information/details on these cell line-based xenografts see: Breast Cancer Research, 2010;12(5):1-17.

Cell-line based xenografts have been initiated and terminal blood collections have been made with MCF7, T47D and SUM1315 lines (N=10, 10, 12, respectively).

2b. Allow tumors to form and reach 5mm and assess serum levels of EMT/CSC markers. (Months 24-36)

2c. Treat animals with chemotherapy (Paclitaxel/Taxol) or anti-estrogen (Tamoxifen) and measure serum levels of EMT/CSC markers and mtDNAs during tumor regression and after cessation of therapy.

2d. Continue to monitor tumor growth. (Months 36-48)

2e. Determine whether the levels of EMT/CSC markers correlate to response to therapy or to tumor recurrence (Months 36-48). Predicted Total mice =(30x2replicates)=60

These studies were not undertaken due to the shift in focus on human samples during the course of the project evolution.

Task 3. Recurrence prediction (Months 36-60)

3a. Create tumors using the 'Human-In-Mouse' breast tumor model. Allow tumors to form and reach 5-8mm in diameter and assess serum levels of oncogenes used to create tumors (Months 36-50).

3b. Surgically resect the tumors and measure blood serum levels of EMT/CSC markers as well as mitochondrial DNAs during tumor regression and after cessation of therapy after surgery (Months 36-50).

3c. Monitor mice and measure the levels of EMT/CSC markers weekly to determine if they increase or correlate with regrowth of tumor (Months 48-60). Predicted Total mice=(80x2replicates)=160;Predicted total number of reduction mammoplasty tissues= 10

As described in the quarterly report, we have completed a study to measure multiplexed biomarkers in the serum of mice harboring a human model of DCIS to invasive cancer progression. In this pilot study, we generated 10 MCF10.DCIS.COM xenografts with $1.0x10^6$ cells and provided serum from these animals taken once a week for 10 weeks. Each week we examined the stage of cancer progression from no tumor, to DCIS, to invasive cancer. Since tumors developed so rapidly with this model, tumors were resected from half of the animals and serum was collected to assess recurrence. This experiment determined: (1) when during tumor formation are the various biomarkers detected, (2) whether the levels of biomarker detection correlate with DCIS to invasive cancer progression, and (3) whether this model progress to metastatic cancer and if so, do biomarker levels change as disease progresses.

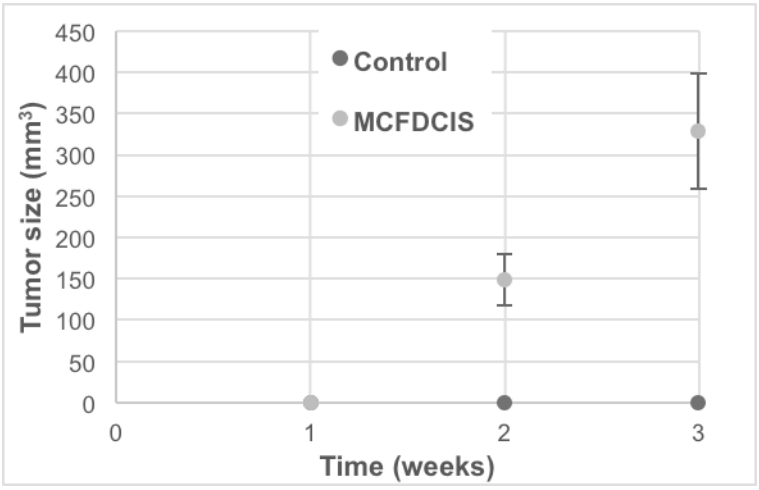


Figure 28: Mammary tumor growth in NOD-SCID control or experimental animals 1-3 weeks after injection of MCF10ADCIS.COM (MCFDCIS) cells.

We injected mouse mammary glands with luciferase-expressing MCF10ADCIS.COM cells. These cells form comedo DCIS-like lesions that spontaneously progress to invasive

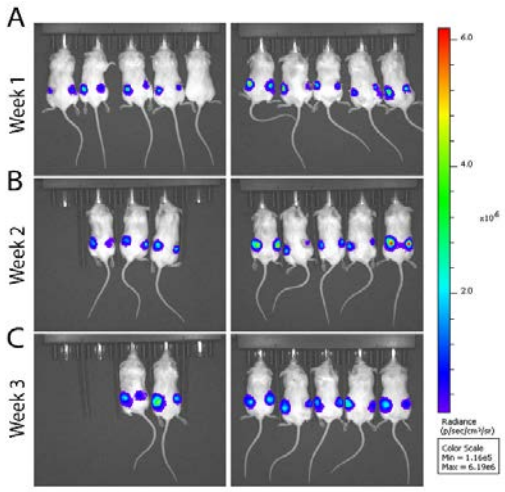


Figure 29. *In vivo* bioluminescence imaging of luciferase expression in NOD-SCID animals A) 1 week, B) 2 weeks, and C) 3 weeks after injection of luciferase expressing MCF10ADCIS.COM cells.

tumors. For this study, $1x10^6$ MCF10ADCIS.COM cells were injected orthotopically into the mammary glands of immunodeficient NOD-SCID mice (n=9, 2 glands/mouse). One control mouse did not receive mammary gland injections. Tumor formation was being assessed by palpitation once a week and tumor growth curves calculated from weekly caliper measurements after 9 weeks. We observed an increase in tumor size each week (**Figure 28**). In addition, we have been imaging tumor cells in vivo using the IVIS 2000 system in order to monitor tumor growth on a weekly basis. Bioluminescence can be seen one week after mammary gland injections (**Figure 29**). We observe that the tumors are getting progressively larger (Figure 29 B & C). We collected mammary glands and serum from each mouse that was be used to detect breast cancer biomarkers over the course of disease progression (**Table 19**).

Table 19.

Week	Experimental Details	Tumor (Y/N)	Serum Collection Date	Tumor Removal Date
Control	Control Mouse-No Cells Injected	N	6/22/2016	6/22/2016
1	Week 1-1 week after injection	Tumors not palpable	6/22/2016	6/22/2016
2	Week 2-2 weeks after injection	Y	6/28/2016	6/28/2016
3	Week 3-3 weeks after injection	Y	7/5/2016	7/5/2016
4	Week 4-4 weeks after injection	Y	7/12/2016	7/12/2016
5	Week 5-5 weeks after injection	Y	7/19/2016	7/19/2016
6	Week 6-3 days after tumor removal	N	7/22/2016	7/19/2016
7	Week 7-1 week after tumor removal	N	7/26/2016	7/19/2016
8	Week 8-2 weeks after tumor removal	N	8/3/2016	7/19/2016
Note: JRC010: Week 5 - 5 weeks after injection; This mouse died during surgery. Janel tried to get as much serum as possible. Used 1:30 dilution for ER-a, PR and CYR61 Simoa runs but data was not retrieved due to the instrument mechanical dysfunction.				
Note: MCFDCIS cells were injected on 6/14/2016				
Note: Tumors were surgically removed on 7/19/2016				

As described above in the Walt Lab section, we provided 10 serum samples and Simoa assays were performed for ADAM12, CA15-3, CA19-9, CYR61, ER, HER2, HIF1a, PR and VEGF (**Figure 7**). Simoa could detect the most of protein markers in the samples. Because of mouse-to-mouse variation, it is hard to make a conclusion regarding the pattern of the protein levels.

KEY RESEARCH ACCOMPLISHMENTS

- Developed Simoa assays for several breast cancer protein biomarkers.
- Established a method to detect and distinguish miRNAs with high sensitivity and specificity.
- Performed preliminary studies to quantify candidate biomarker levels in healthy and diseased human serum and evaluate usefulness of each marker.
- Characterized multiple breast cancer cell lines with established Simoa assays.
- Measured PSA content of two different populations of single LNCaP cells and observed differences based on protein expression.
- Developed multiplex assays for the detection of multiple breast cancer proteins simultaneously.
- Multivariate analysis (PLS-DA) of eight proteins and age performed well as a classifier, with an overall accuracy of 97% and 0.99 AUC for distinguishing between healthy and early (Stages 0-II) breast cancer.
- Isolated CTCs from breast cancer patients using eDAR.
- Characterized the performance of eDAR.
- Developed ultra-bright probes for the isolation of CTCs with low biomarker expression.
- Improved the ability of eDAR to capture CTCs with variable patterns of biomarker expression.
- Demonstrated a 6-fold increase in recovery rate of CTCs from patient-derived samples.
- Developed a new eDAR device for the simultaneous and selective isolation of multiple subpopulations of CTCs.
- Explored high-throughput methods for single-cell dispensing for downstream analysis of CTCs using digital nucleic acid and protein analysis.

- Developed a method for storing whole blood for over 3 days so blood samples from mouse models and from patients could be shipped from Tufts or other hospital sites for analysis of cell biomarkers.
- Developed and validated a method for the digital PCR analysis of single cancer cells.
- Developed a technique based on sequential imaging and bleaching to study a large panel of biomarkers present on cancer cells.
- Developed a new technique for the high-throughput parallel manipulations of single cells using bipolar electrode based dielectrophoresis.
- Identified primers and methods to sequence the human and mouse mtDNA genome within a population of total DNA.
- Established techniques to sequence mtDNA from blood and FFPE tumor tissue samples
- Identified 99 mtDNA mutations in SUM149 IBC cells
- Identified the mtDNA mutations in four ER- negative breast cancer cell lines
- Identified the mtDNA mutations in two HER2+ IBC lines (SUM190 and MDA-IBC-3 cells)
- Identified 40 specific mtDNA mutations that are also prevalent in patient samples.
- Determined that tRNA mutations are frequently present in IBC breast tumors
- Determined that rare tRNA mutations are present in both a BRCA1 mutant breast tumor and cell line. Work is ongoing to evaluate the prognostic efficacy of using these mtDNA mutations in liquid biopsies.
- Developed methods to extract DNA from frozen and formalin fixed paraffin embedded (FFPE) tumor tissue.
- Determined that no additional mtDNA mutations are induced *in vivo* when tumor cells are injected in NOD/SCID mice or in HIM mouse models.
- Identified mtDNA mutations in ER+ breast cancer line lines
- Showed that we can detect mtDNA mutations as early as 7 days after tumor cell orthotopic implantation in mice when no tumor can be palpated.
- Detected mtDNA, as judged by characteristic mutations, from tumors after resection (surgery) in mice with no primary recurrence but with lung and brain metastases, suggesting the DNA is a marker for tumor metastasis.
- Detected tumor mtDNA mutations following tumor surgery in mice early on before recurrence of the primary tumor.
- Developed an initial panel of common mtDNA mutation in breast cancer that is being validated in patient samples.
- Demonstrated ADAM8, which is highly expressed in almost 50% of all breast cancer metastases, regulates over 60 miRNAs, including the secreted miR-720
- Established a method to detect miRNAs in mouse serum
- Identified a new compound with increased potency against metastasis and decreased toxicity based on in vitro screening and functional assays.
- Identified new derivative compounds with potential activity against metastasis based on in vitro screening.
- Established maximal tolerated dose of inhibitor 1 in mice for testing against breast cancer metastasis.
- Established expanded range of breast cancer sub-types with sensitivity to our initial inhibitor.
- Identified available human serum banks for biomarker testing, including the NIH PLCO screening trial and the Harvard Cohorts Biorepository.
- Established effective process for serum banking from women with breast cancer at Tufts Medical Center enabling development of initial breast cancer biomarker panel; active monitoring ongoing.

REPORTABLE OUTCOMES

- Anand, R.K., Johnson, E.S., Chiu, D.T. "Negative dielectrophoretic capture and repulsion of single cells at a bipolar electrode: the impact of faradaic ion en-richment and depletion" *Journal of the American Chemical Society*, **2015**, 137, 776-783.

- Baig, S.; Schubert, S. M.; Walter, S. R.; Usmani, K.; Walt, D.R. "Ultrasensitive Assays for Early Breast Cancer Detection". Poster presentation. The Pittsburgh Conference on Analytical Chemistry and Applied Spectroscopy, 2015 March 8-12, New Orleans, LA.

- Baig, S.; Schubert, S. M.; Walter, S. R.; Walt, D. R. "Single Molecule Assay Development for Breast Cancer Detection". Poster presentation. 250th American Chemical Society National Meeting and Exposition, 2015 August 16-20, Boston, MA.

- Baig, S., Schubert, S. M., Walter, S.R., Walt, D. R. "Development of Ultra-Sensitive Assays for Early Breast Cancer Detection and Single Cell Analysis". Oral presentation for Tufts University Breast Cancer Working Group; 2014 April 28, Boston, MA.

- Baig, S., Schubert, S. M., Walter, S.R., Walt, D. R. "Development of Ultrasensitive Protein Assays for the Early Detection of Breast Cancer". Poster presented at: Bioanalytical Sensors GRC; 2014 June 21; Newport, RI.

- Cohen, L., Walter, S. R., Schubert, S. M., Manesse, M., Duan, B. D., Lee, M., Walt, D. "Single Molecule Arrays for Measuring Protein and mRNA Expression Levels in Single Cells". Poster presentation. Gordon Research Conference, 2016. Hong Kong.

- Das, S.G., Romagnoli, M., Mineva, N. D., Sonenshein, G.E. "MicroRNAs Downstream of ADAM8 as Therapeutic Targets and Non invasive Biomarkers for Triple Negative Breast Cancer". Poster presentation. AACR-NCI-EORTC Molecular Targets and Cancer Therapeutics Conference; October 19-23, 2013; Boston, MA

- Das, S. G., Romagnoli, M., Mineva, N. D., Barillé-Nion, S., Jézéquel, P., Capone, M., Sonenshein, G. E. "miR-720 is a downstream target of an ADAM8-induced ERK signaling cascade that promotes the migratory and invasive phenotype of triple negative breast cancer cells", *Breast Cancer Research*, **2016**, 18:40.

- Ghatak, P.; Hartman, M. R.; Walt, D. R. "Detection of microRNAs as Potential Biomarkers of Early Stages in Breast Cancer". Poster presentation. Extracellular Biomarkers Summit, 2015 March 16, Cambridge, MA.

- Johnson, E.S., Anand, R.K., Chiu, D.T. "Improving detection of circulating tumor cells with low numbers of the targeted surface antigen by ensemble decision aliquot ranking (eDAR)" *Analytical Chemistry*, **2015**, 87, 9389-9395.

- Rong, Y., Yu, J., Zhang, X., Sun, W., Ye, F., Wu, I., Zhang, Y., Hayden, S., Zhang, Y., Wu, C., Chiu, D.T. "New yellow fluorescent semiconducting polymer dots with high brightness, small size and narrow emission for biological applications" *ACS Macro Letters*, **2014**, 3, 1051-1054.

- Schubert, S. M., Baig, S., Walt, D. R. "Single Molecule Assays for Early Breast Cancer Detection". Poster presentation. Tufts Cancer Center Symposium; 2013 June 25; North Grafton, MA.

- Schubert, S. M., Baig, S., Velasquez, E.F., Walt, D. R. "Single Molecule Assays for Early Breast Cancer Detection". Poster presentation. Tufts Research Day on the Tufts Innovation Institute; 2013 September 20; Boston, MA.

- Schubert, S. M., Baig, S., Walter, S.R., Walt, D. R. "The Development of Serum-Based Single Molecule Assays for Early Detection of Cancer". Poster presentation. Single Molecule Approaches to Biology GRC; 2014 July 15 & 17th; Lucca (Barga), Italy.

- Schubert, S. M., Baig, S., Walt, D. R. "Single Molecule Assays for Early Breast Cancer Detection". Poster presentation. Pittsburgh Conference; 2014 March 3; Chicago, IL.

- Schubert, S. M.; Arendt, L. M.; Zhou, W.; Baig, S.; Walter, S. R.; Buchsbaum, R. J.; Kuperwasser, C.; Walt, D. R., "Ultra-sensitive protein detection via Single Molecule Arrays towards early stage cancer monitoring", *Scientific Reports*, **2015,** 5, 11034.

- Schubert, S. M.; Walter, S. R.; Manesse, M.; Walt, D. R. "Protein Counting in Single Cancer Cells", *Analytical Chemistry*, **2016,** 88*,* 2952-2957.

- Schubert, S. M.; Baig, S.; Walter, S. R.; Palacios, M.; Walt, D. R. "Development of Serum-Based Single Molecule Assays for the Early Detection of Cancer." Oral Presentation. 250th American Chemical Society National Meeting and Exposition, August 16-20, 2015, Boston, MA.

- Sonenshein, G.E. "ADAM8, a Driver of Metastasis, is a Novel Target for Treatment of Triple-Negative Breast Cancer". Oral presentation. Northeastern Breast Cancer Research Conference, 8 May 2015, Lebanon, NH. Keynote Speaker.

- Thompson, A.M., Gansen, A., Paguirigan, A.L., Kreutz, J.E., Radich, J.P., Chiu, D.T. "The self-digitization microfluidic chip for the absolute quantification of mRNA in single cells" *Analytical Chemistry,* **2014,** 86, 12308-12314.

- Walter, S.R., Schubert, S. M., Baig, S., Walt, D. R. "Ultrasensitive Detection of Breast Cancer Biomarkers and Single Cell Protein Analysis". Oral talk and poster presentation. Bioanalytical Sensors GRC; June 26, 2014, Newport, RI.

- Walter, S. R.; Schubert, S. M.; Manesse, M.; Walt. D. R. "Quantifying Protein Expression in Single Cells." Oral Presentation. 250th American Chemical Society National Meeting and Exposition, August 16-20, 2015. Boston, MA.

- Zhang, Y., Yu, J., Gallina, M.E., Sun, W., Rong, Y., Chiu, D.T. "Highly fluorescent fluorinated semiconducting polymer dots for cellular imaging and analysis" *Chemical Communication,* **2013**, 49, 8256-8258.

- Zhao, M., Wei, B., Chiu, D.T. "Imaging multiple biomarkers in eDAR captured rare cells with sequential immunostaining and photobleaching" *Methods*, **2013**, 64, 108-113.

- Zhao, M., Nelson, W.C., Wei, B., Schiro, P.G., Hakimi, B.M., Johnson, E.S., Perdue, R.K., Gyurkey, G., White, L.M., Whiting, S., Chiu, D.T. "New generation of ensemble-decision aliquot ranking based on simplified microfluidic components for large-capacity trapping of circulating tumor cells" *Analytical Chemistry,* **2013**, 85, 9671-9677.

- Zhao, M., Schiro, P.G., Chiu, D.T. "Ensemble-decision aliquot ranking (eDAR) for CTC isolation and analysis" *Circulating Tumor Cells: Isolation and Analysis,* **2016**, 53-83.

- Zhao, M., Wei, B., Nelson, W.C., Chiu, D.T. "Simultaneous and selective isolation of multiple subpopulations of rare cells from peripheral blood using ensemble-decision aliquot ranking" *Lab Chip*, **2015**, 15, 3391-3396.

CONCLUSION

In conclusion, tremendous progress has been made toward the development of a blood test for early stage breast cancer. A total of 18 Simoa assays have now been successfully developed, most with LODs significantly lower than their commercial ELISA counterparts. While some of these assays can still be optimized, preliminary screening of eight protein biomarkers in breast cancer patients and healthy control serum samples has resulted in a classification model that is highly predictive of early stage breast cancer. Although additional samples must be screened to gain greater statistical information and further data processing is required, our preliminary results are quite promising. Additional protein biomarkers can still be added to the panel, as well as miRNA and mtDNA markers, to improve the sensitivity and specificity of the Simoa assays.

We have also successfully achieved full quantification of protein molecules within single cancer cells using Simoa. A cell labeling and selection system has been developed that enables cells to be identified and characterized by multiplexed antibody labeling. Cells matching user-defined selection criteria can be identified and isolated. After isolation, the cells can be subjected to lysis and analyzed to quantify protein copy numbers in single breast cancer cells using several of the established Simoa assays described above.

We have achieved all our goals and demonstrated that the ultrasensitivity of the Simoa platform is able to detect meaningful differences in the serum of healthy controls and breast cancer patients. These multiplexed protein biomarkers may one day enable early detection, further classification and stratification of patients, as well as the determination of progression and therapeutic efficacy.

REFERENCES

1. Schubert, S. M.; Arendt, L. M.; Zhou, W.; Baig, S.; Walter, S. R.; Buchsbaum, R. J.; Kuperwasser, C.; Walt, D. R., Ultra-sensitive protein detection via Single Molecule Arrays towards early stage cancer monitoring. *Scientific Reports* **2015,** *5*, 11034.
2. Wilson, D. H.; Hanlon, D. W.; Provuncher, G. K.; Chang, L.; Song, L.; Patel, P. P.; Ferrell, E. P.; Lepor, H.; Partin, A. W.; Chan, D. W.; Sokoll, L. J.; Cheli, C. D.; Thiel, R. P.; Fournier, D. R.; Duffy, D. C., Fifth-Generation Digital Immunoassay for Prostate-Specific Antigen by Single Molecule Array Technology. *Clinical Chemistry* **2011,** *57* (12), 1712-1721.
3. Schubert, S. M.; Walter, S. R.; Manesse, M.; Walt, D. R., Protein Counting in Single Cancer Cells. *Analytical Chemistry* **2016,** *88* (5), 2952-7.
4. He, L.; Chinnery, P. F.; Durham, S. E.; Blakely, E. L.; Wardell, T. M.; Borthwick, G. M.; Taylor, R. W.; Turnbull, D. M., Detection and quantification of mitochondrial DNA deletions in individual cells by real-time PCR. *Nucleic Acids Research* **2002,** *30* (14), e68-e68.
5. Calaf, G.; Russo, J., Transformation of human breast epithelial cells by chemical carcinogens. *Carcinogenesis* **1993,** *14* (3), 483-92.

6. Oliveira, P. H.; Boura, J. S.; Abecasis, M. M.; Gimble, J. M.; da Silva, C. L.; Cabral, J. M. S., Impact of hypoxia and long-term cultivation on the genomic stability and mitochondrial performance of ex vivo expanded human stem/stromal cells. *Stem Cell Research* **2012,** *9* (3), 225-236.
7. Covarrubias, D.; Bai, R. K.; Wong, L. J.; Leal, S. M., Mitochondrial DNA variant interactions modify breast cancer risk. *Journal of Human Genetics* **2008,** *53* (10), 924-8.
8. Mohammed, F.; Rezaee Khorasany, A. R.; Mosaieby, E.; Houshmand, M., Mitochondrial A12308G alteration in tRNA(Leu(CUN)) in colorectal cancer samples. *Diagnostic Pathology* **2015,** *10*, 115.
9. Romagnoli, M.; Mineva, N. D.; Polmear, M.; Conrad, C.; Srinivasan, S.; Loussouarn, D.; Barillé-Nion, S.; Georgakoudi, I.; Dagg, Á.; McDermott, E. W.; Duffy, M. J.; McGowan, P. M.; Schlomann, U.; Parsons, M.; Bartsch, J. W.; Sonenshein, G. E., ADAM8 expression in invasive breast cancer promotes tumor dissemination and metastasis. *EMBO Molecular Medicine* **2014,** *6* (2), 278-294.

APPENDIX 1: PUBLICATIONS

Anand, R.K., Johnson, E.S., Chiu, D.T. "Negative dielectrophoretic capture and repulsion of single cells at a bipolar electrode: the impact of faradaic ion en-richment and depletion" *Journal of the American Chemical Society*, **2015**, 137, 776-783.

Das, S. G., Romagnoli, M., Mineva, N. D., Barillé-Nion, S., Jézéquel, P., Capone, M., Sonenshein, G. E. "miR-720 is a downstream target of an ADAM8-induced ERK signaling cascade that promotes the migratory and invasive phenotype of triple negative breast cancer cells", *Breast Cancer Research*, **2016**, 18:40.

Johnson, E.S., Anand, R.K., Chiu, D.T. "Improving detection of circulating tumor cells with low numbers of the targeted surface antigen by ensemble decision aliquot ranking (eDAR)" *Analytical Chemistry*, **2015**, 87, 9389-9395.

Rong, Y., Yu, J., Zhang, X., Sun, W., Ye, F., Wu, I., Zhang, Y., Hayden, S., Zhang, Y., Wu, C., Chiu, D.T. "New yellow fluorescent semiconducting polymer dots with high brightness, small size and narrow emission for biological applications" *ACS Macro Letters*, **2014**, 3, 1051-1054.

Schubert, S. M.; Arendt, L. M.; Zhou, W.; Baig, S.; Walter, S. R.; Buchsbaum, R. J.; Kuperwasser, C.; Walt, D. R., "Ultra-sensitive protein detection via Single Molecule Arrays towards early stage cancer monitoring", *Scientific Reports*, **2015,** 5, 11034.

Schubert, S. M.; Walter, S. R.; Manesse, M.; Walt, D. R. "Protein Counting in Single Cancer Cells", *Analytical Chemistry*, **2016,** 88*,* 2952-2957.

Thompson, A.M., Gansen, A., Paguirigan, A.L., Kreutz, J.E., Radich, J.P., Chiu, D.T. "The self-digitization microfluidic chip for the absolute quantification of mRNA in single cells" *Analytical Chemistry,* **2014,** 86, 12308-12314.

Zhang, Y., Yu, J., Gallina, M.E., Sun, W., Rong, Y., Chiu, D.T. "Highly fluorescent fluorinated semiconducting polymer dots for cellular imaging and analysis" *Chemical Communication,* **2013**, 49, 8256-8258.

Zhao, M., Wei, B., Chiu, D.T. "Imaging multiple biomarkers in eDAR captured rare cells with sequential immunostaining and photobleaching" *Methods*, **2013**, 64, 108-113.

Zhao, M., Nelson, W.C., Wei, B., Schiro, P.G., Hakimi, B.M., Johnson, E.S., Perdue, R.K., Gyurkey, G., White, L.M., Whiting, S., Chiu, D.T. "New generation of ensemble-decision aliquot ranking based on simplified microfluidic components for large-capacity trapping of circulating tumor cells" *Analytical Chemistry,* **2013**, 85, 9671-9677.

Zhao, M., Schiro, P.G., Chiu, D.T. "Ensemble-decision aliquot ranking (eDAR) for CTC isolation and analysis" *Circulating Tumor Cells: Isolation and Analysis,* **2016**, 53-83.

Zhao, M., Wei, B., Nelson, W.C., Chiu, D.T. "Simultaneous and selective isolation of multiple subpopulations of rare cells from peripheral blood using ensemble-decision aliquot ranking" *Lab Chip*, **2015**, 15, 3391-3396.

APPENDIX 2: PROJECT PERSONNEL

Tufts University, Department of Chemistry

David Walt	PI
Mark Hartman	Postdoctoral Scholar
Bishnu Regmi	Postdoctoral Scholar
Stephanie Walter	Postdoctoral Scholar
Danlu Wu	Postdoctoral Scholar
Liangxia Xie	Postdoctoral Scholar
Shazia Baig	Graduate Student
Jennifer Geldart-Flashman	Graduate Student
Payel Ghatak	Graduate Student
Soyoon Hwang	Graduate Student
Carolyn Lipovsky	Graduate Student
Stephanie Schubert	Graduate Student
Thy Vu	Graduate Student
Aaron Amardey-Wellington	Undergraduate Research Associate
Sika Gadzanku	Undergraduate Research Associate
Kudret Usmani	Undergraduate Research Associate
Zhixin Xia	Undergraduate Research Associate

University of Washington

Daniel Chiu	PI
Bryant Fujimoto	Research Staff (with Ph.D)
Rui Han	Postdoctoral Scholar
Chun-Ting Kuo	Postdoctoral Scholar
Robert Lorenz	Postdoctoral Scholar
Wyatt Nelson	Postdoctoral Scholar
Robbyn Perdue	Postdoctoral Scholar
Thomas Schneider	Postdoctoral Scholar
Bingchuan Wei	Postdoctoral Scholar
Li Wu	Postdoctoral Scholar
Xu Wu	Postdoctoral Scholar
Fangmao Ye	Postdoctoral Scholar
Yong Zhang	Postdoctoral Scholar
Mengxia Zhao	Postdoctoral Scholar
Yue Zhang	Postdoctoral Scholar
Eleanor Johnson	Graduate Student
Jingyu Shen	Graduate Student
Alison Thompson	Graduate Student

Tufts University, School of Medicine

Charlotte Kuperwasser	PI
Gail Sonenshein	PI
Lisa Arendt	Senior Research Associate
Jessica McCready	Research Associate
Nora Mineva	Research Associate
Sonia Das	Postdoctoral Scholar & Research Associate
Patricia Keller	Postdoctoral Scholar
Devin Rosenthal	Postdoctoral Scholar
Anna Wronski	Postdoctoral Scholar

Benjamin Dake	Graduate Student
Tufts Medical Center	
Rachel Buchsbaum	PI
Kun Xu	Research Associate