AWARD NUMBER: W81XWH-14-1-0474

TITLE: Genomic Diversity and the Microenvironment as Drivers of Progression in DCIS

PRINCIPAL INVESTIGATOR: Carlo Maley

CONTRACTING ORGANIZATION: Arizona State University

REPORT DATE: October 2020

TYPE OF REPORT: Annual

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

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
October 2020	Annual	30Sep2019-29Sep2020

4. TITLE AND SUBTITLE
Genomic Diversity and the Microenvironment as Drivers of Progression in DCIS

5a. CONTRACT NUMBER
W81XWH-14-1-0474

5b. GRANT NUMBER
BC132057

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

CARLO MALEY

E-Mail: shelley.hwang@dm.duke.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

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

Arizona State University
PO Box 876011
Tempe, AZ 85287

8. PERFORMING ORGANIZATION REPORT

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

U.S. Army Medical Research and Development 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 project is designed to test whether genetic and/or tumor environmental heterogeneity is a driving force in progression of breast DCIS. Our project, a collaboration between Duke and ASU, have met our 60 month milestones in all 4 aims. Primary achievements for 60 months are:

1) Completed Case and control identification of 57 Pure DCIS & 61 Synchronous DCIS (DCIS with adjacent invasion) through extensive database and searching at Duke 2) Completed deep whole exome sequencing (WES) for 100 cases from 30-160ng of DNA isolated from archival FFPE specimens, 3) Comparison of analytic methods to characterize somatic mutations from this full exome sequencing, 4) Application of sequencing data for copy number assessment 5) Completed dual immune-staining on DCIS lesions using 6 pairs and 3 single antibodies, 6) Completed Image analysis of these stains, including quantitative analysis, 7) Completed identification of upstaged DCIS cases for the radiology aim, 8) Development of image analysis methods for digital mammograms, 9) Completed the validation Aim 4, including collection of DCIS that either did not progress or progressed to DCIS or invasive cancer, 10) Completed Aim 4 WES and Whole genome sequencing for 110 validation cases from 30-160ng of DNA isolated from archival FFPE 11) Full integration of team members over the past year via frequent conferencing, face to face meetings, and constant communication. This multi-disciplinary approach allowed our group to fully implement and we reach our year 5 project goals.

15. SUBJECT TERMS
DCIS, intra-tumor heterogeneity, genetic diversity, phenotypic diversity, somatic evolution, microenvironment, mammographic biomarkers

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
			UU	41	USAMRDC
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U			

TABLE OF CONTENTS

	Page
1. Introduction	1
2. Keywords	1
3. Accomplishments	1
4. Impact	32
5. Changes/Problems	33
6. Products	34
7. Participants & Other Collaborating Organizations	40

1. INTRODUCTION

Ductal carcinoma in situ (DCIS) of the breast is an increasingly common diagnosis that is related to aggressive screening patterns (mammography). This "pre-invasive" lesion may progress to invasive cancer, but does so at a relatively low frequency. Nonetheless, it is commonly treated with extensive surgery, radiation, and hormonal therapy even though most of these lesions would never progress to invasive cancer. Thus, there is a pressing clinical need to stratify the risk of DCIS tumors into those in need of intervention and those that can be safely monitored without intervention. Our project is designed to address this need by characterizing the evolvability of DCIS, detecting those that have a high likelihood of evolving to malignancy versus those that are likely to remain indolent.

2. KEYWORDS

DCIS, cancer progression, intra-tumor heterogeneity, genetic diversity, phenotypic diversity, somatic evolution, microenvironment, mammographic biomarkers

3. ACCOMPLISHMENTS

What were the major goals of the project?

Aim 1. *Determine whether genetic diversity of DCIS is greater in DCIS with adjacent invasive disease compared to DCIS without progression.* Genetic divergence of the DCIS component of tumors were measured based on exome sequencing on two separate regions of the tumor, as well as normal tissue, in patients with DCIS only or Synchronous DCIS. This allowed us to confirm the association between increased genetic diversity and higher probability of progression to malignancy. Additionally, we found differences in which genes were mutating in the two tumor types by comparing the biological pathways differentially affected by somatic mutations.

60 Month Milestones:

- Protocol preparation, IRB submission and approval: (Duke eIRB Pro00054515, initial Duke approval, 5/27/2014 and renewed annually), DOD IRB approval in place – **completed**

- Case identification and tissue block selection: Through a variety of available databases, we identified a large number of cases and controls with tissue available in the Duke Pathology archives. Each potential case and control requires extensive chart and pathology review in order to determine final eligibility and usability. For example, there is sufficient amount of the DCIS lesion (>2mm size) for isolation and DCIS is not too close to invasive cancer (it extends outside the invasive component). There must be two blocks with DCIS present that are >0.8cm apart. We identified **120** cases, with pathology review- **Completed**

- Sectioning of tissue blocks: New sections from candidate paraffin blocks are cut, one H&E was stained at the beginning and end of each set prior to pathology review. Remaining sections from candidate blocks (containing a sufficient amount of the DCIS lesion) are

used for macro-dissection and subsequent DNA extraction. Additional sections were also stored for immunohistochemical (IHC) analysis of key measures of tumor and micro-environmental heterogeneity. These slides are scanned for analytic and archival purposes. We completed both cases and controls in this manner- **Completed**.

- DNA extraction of test cases: **Completed**.

- Exome sequencing of test cases: **Completed.** We chose the Genome Center at Washington University where cutting-edge methods for producing high quality data from these FFPE specimens have been developed and refined. Over the past four years, Wash U. sequenced 30-160ng from 300 individual DNA samples derived from 100 subjects (germ line sample plus 2 DCIS containing samples). They were able to derive interpretable sequence data (minimum of 40X depth at 50% coverage) from 30-160ng of FFPE DNA with qualities summarized in Figures 2, 3, 4 and 5.

Case Selection

Breast tumors were classified using the World Health Organization (WHO) criteria [25]. Following pathology review, Pure DCIS, not associated with invasion and synchronous DCIS with Invasive Ductal Carcinoma (IDC), were included in this study IDC and DCIS were graded according to the Nottingham grading system [26] or recommendations from the Consensus conference on DCIS classification [27], respectively. Each case had matched normal tissues or normal node, confirmed by pathology review, to be devoid of neoplastic cells.

Table 1: Characteristics of Aim 1/2 Study Cohort

	Pure DCIS (n=57)	Adjacent DCIS (n=61)	
Age (mean)	56.9	57.5	
Race			
White	32	42	
Black	20	15	
Other	6	4	
Tumor Size (mean, cm)	3.5	4.6	
Nodal Status			
Positive	0	27	
Negative	57	34	
Grade		*DCIS*	*Invasive*
1	1	1	11
2	24	22	27
3	30	37	22
Surgery			
Lumpectomy	24	38	

Mastectomy	19	23
Estrogen Receptor		
Positive	39	42
Negative	9	19
Equivocal	0	0
Progesterone Receptor		
Positive	33	38
Negative	12	23
Equivocal	3	0
HER2 Status		
Positive	0	14
Negative	0	45

Analytic Pipeline Development for identification of somatic genetic alterations: Completed.

FFPE tumor samples has numerous advantages in cancer research, but they tend to generate artifacts when used for genomic analyses. In order to minimize this limitation, we developed and validated a new bioinformatic tool (Figure 1) to robustly identify somatic nucleotide variants and estimate intratumor heterogeneity. We optimized this strategy using 28 pairs of sequencing technical replicates—the same DNA sample sequenced twice. We validated our results on the same DNA using targeted primers using the AmpliSeq™ technology. After optimization, the mean similarity between replicates increased 5-fold, reaching 88.3% (range 66.7-100%), with a mean of 19.9 SNVs (range 1-61) per sample (Figure 2). This new algorithm provides a crucial improvement in detecting SNVs in FFPE samples. This work enabled us to achieve the main goal of this aim (see below), and has led to the development of an open-source tool that will benefit the cancer-research community (https://github.com/adamallo/ITHE). **This was published in *Briefings in Bioinformatics* (2021).**

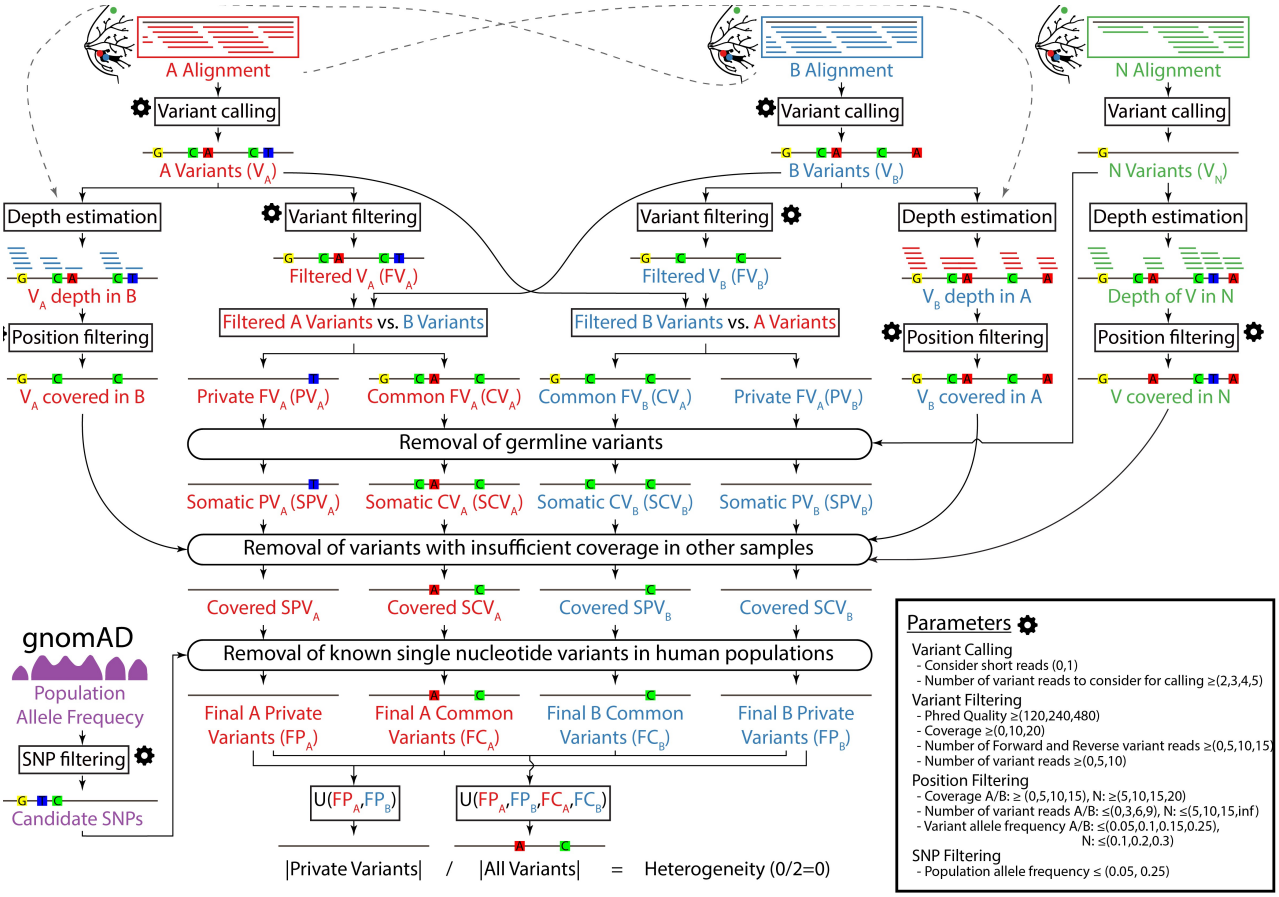

Figure 1: Flowchart of our novel algorithm to call SNV variants and estimate the genetic heterogeneity between two samples. Inputs: aligned sequences (BAM files) of the two samples (A, in red; and B, in blue) and their healthy tissue control (N, in green), population allele frequency data from the gnomAD database (single nucleotide polymorphisms, SNPs, in purple), and user-specified configuration parameters (gear icon). Outputs: estimate of the genetic heterogeneity between samples A and B, and set of variants (level of detail user-specified). The key step of this algorithm is the generation of two sets of private and common variants by comparing the variants in the two samples twice, alternatively filtering one of the sets and using all variants from the other. The parameters that control this pipeline, and the values assayed in our bioinformatic optimization, are detailed in the Parameters box.

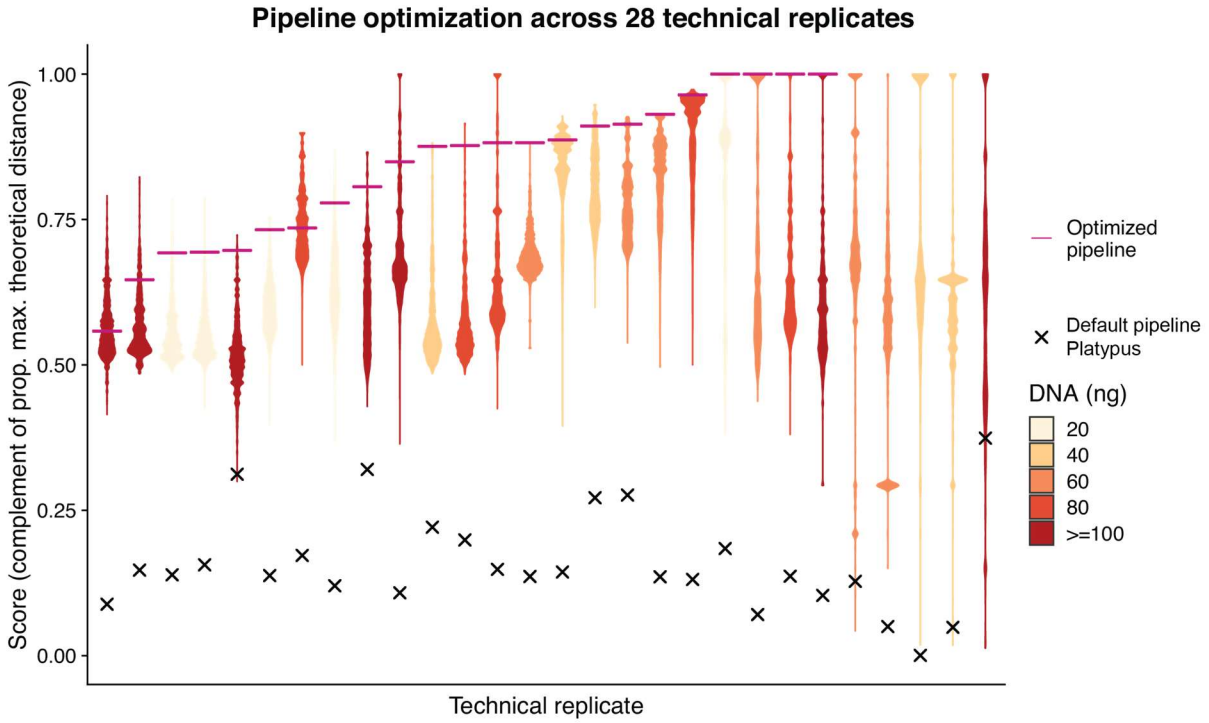

Figure 2: **Empirical optimization of the variant post-processing algorithm using technical replicates.** Each violin plot summarizes the distribution of optimization scores of the similarity obtained with 5,308,416 combinations of values of the 13 parameters that control the pipeline for one of the 28 technical replicates. (x= score before optimization using the default parameters of Platypus; — = similarity using the optimized parameter values for the entire data set; colors indicate the amount (ng) of DNA used as template)

Exome sequencing analysis.

Minimum coverage for inclusion in this study was 40X over at least 50% of the exome. First, reads were aligned to the reference human genome GRCh37 and variants were detected using the variant caller Platypus [26]. We annotated the variants using Annovar software [27]. Our algorithm compares two cancer samples between each other and versus the control sample. The first cancer sample analyzed with higher stringent filtering steps is compared with the second sample with low stringency and vice versa. The variants selected with the more stringent filter are retained if present also in the sample with lower stringent filter. Somatic mutations are validated on the original sequences and germ line polymorphism removed. We used the optimized pipeline to identify single nucleotide variants (SNVs), based on our 28 technical replicates and controls, as described above.

Mutational burden. There is no statistically significant difference between Pure DCIS, synchronous DCIS and IDC samples (Fig. 3). This suggested that mutational burden would likely not be a good biomarker for risk stratification in DCIS. However, our validation study (Aim 4 below) found that patients with pure DCIS that had a recurrence later (of either DCIS or IDC), had a higher mutational burden in their pure DCIS than patients that did not recur (see below).

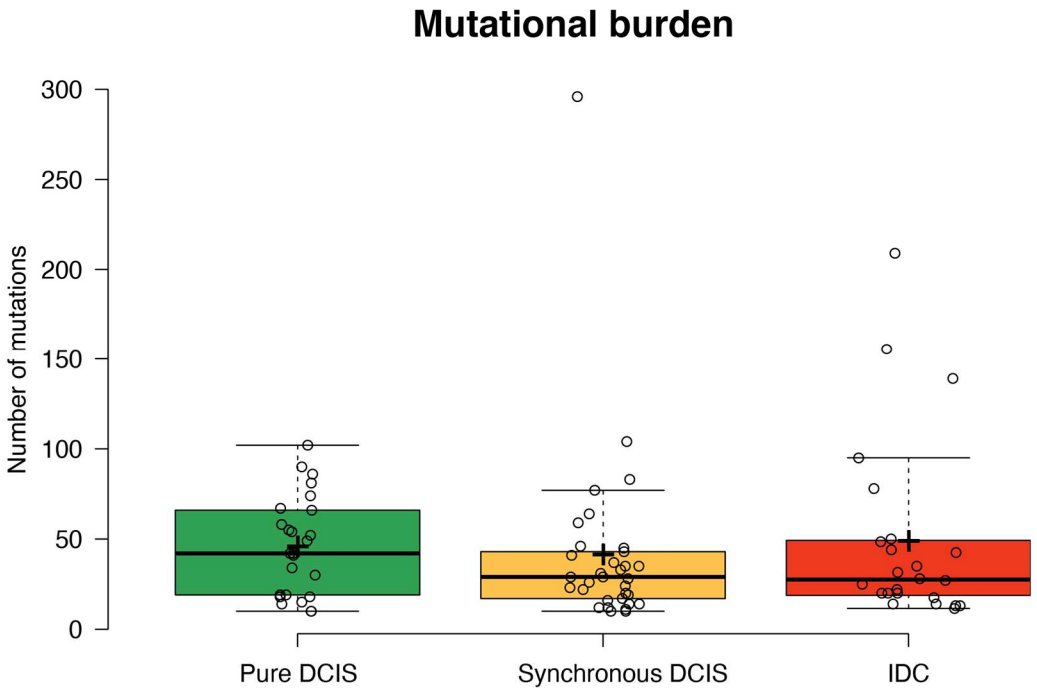

Figure 3: Mutational burden between pure and synchronous DCIS. There is no statistically significant difference between pure DCIS (n=25), synchronous DCIS (n=33), and IDC samples (n=24). Mutational burden is not diagnostic of cancer progression.

With more mutations in a tumor, we would expect that more differences would accumulate between regions. So, our null model would predict a positive correlation between mutational burden and divergence. However, there is a negative correlation between the number of mutations and the divergence in pure DCIS (Kendall's tau-b, correlation coefficient=-0.323, p=0.025), suggesting a surprising degree of genetic homogeneity in pure DCIS. In contrast, there is a positive, non-statistically significant correlation in synchronous DCIS and IDC samples themselves.

Divergence. We estimated the percent of mutations that are different between two regions of the same tumor (called genetic divergence). This can only reliably be estimated when there are sufficient mutations in the first place, to detect divergence. We limited our analysis to DCIS that had at least 10 SNVs detected by our pipeline. **We found statistically significant genetic divergence between Pure DCIS and synchronous DCIS** (adjacent to IDC) and, Pure DCIS and IDC (one-way ANOVA (F (2,75) = 8.959, p< 0.001, post-hoc Tukey test: Pure DCIS vs synchronous DCIS, p=0.005; Pure DCIS vs IDC, p<0.001, synchronous DCIS vs IDC=NS), Figure 4.

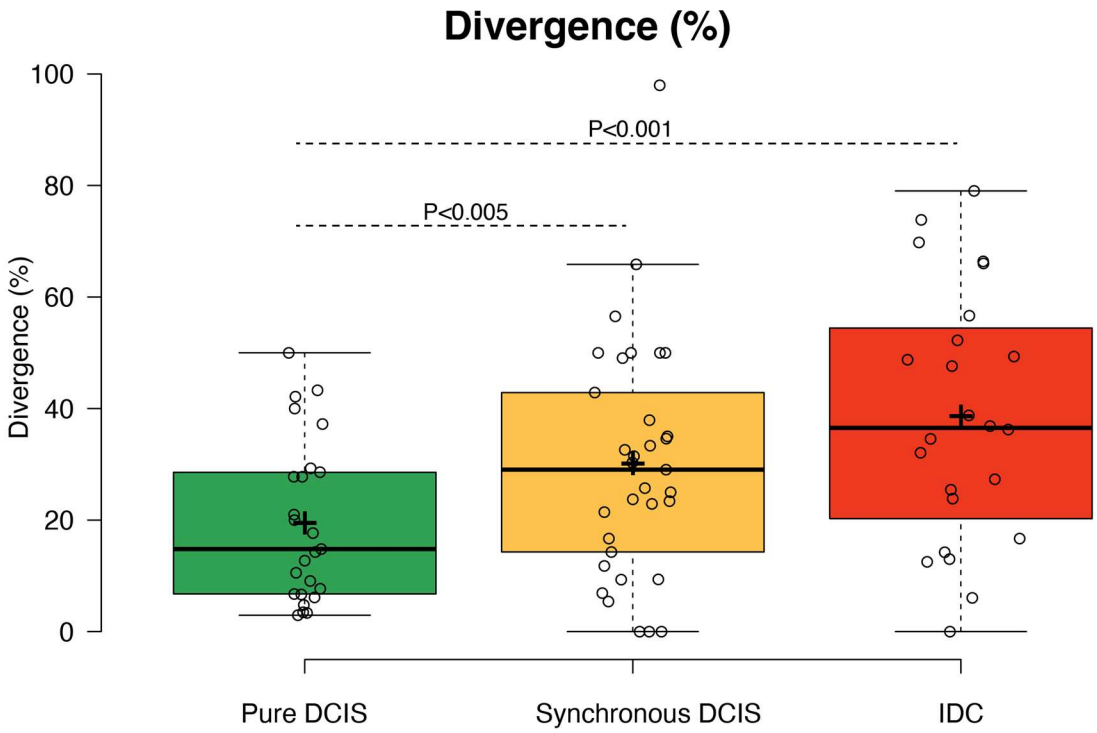

Figure 4: Genetic divergence between pure and synchronous DCIS. The genetic divergence between two regions of the same tumor is higher in synchronous DCIS and IDC samples than in Pure DCIS samples (Anova, p<0.001). Here divergence in the IDC column is calculated as the differences in the mutations carried in the one IDC sample compared to the adjacent DCIS sample.

Functional analysis. We further analyzed the mutated genes to evaluate the molecular processes or signaling pathways that are deregulated based on DAVID (https://david.ncifcrf.gov) gene functional analysis. Pure DCIS and synchronous DCIS have a distinct mutational profile. Cell adhesion genes (e.g. cadherins) are statistically significantly enriched in synchronous DCIS patients (Fold enrichment=6.11, p<0.001 FDR, DAVID). Pure DCIS patients have an enrichment of immune-related genes (Fold enrichment=11.02, p=0.047 FDR, DAVID). There is a marginal overlap between pure DCIS and synchronous DCIS mutated genes (8.6%). Functional gene analyses show a difference between pure DCIS and synchronous DCIS patients compatible with the acquired invasive capacity of synchronous DCIS cells. It is unclear why immune related genes would be more often mutated in pure DCIS than in DCIS adjacent to IDC. The enrichment of the cell adhesion genes mutated in the synchronous DCIS may predict their evolution of invasion into the surrounding tissues.

Conclusion

Our results for Aim 1 support our hypothesis that the genetic diversity (intra-tumor heterogeneity) of DCIS is associated with progression to IDC. Genetic divergence is higher in DCIS that is adjacent to IDC than in pure DCIS. This is the first time that genetic divergence has been tested as a biomarker in DCIS. The major caveat to this result is that we are using DCIS adjacent to IDC as a model of high risk DCIS. Aim 4 tests the prognostic value of mutational burden and genetic divergence by comparing pure DCIS that did versus did not recur (as either DCIS or IDC). The work from Aim 1 also generated a widely applicable bioinformatics tool for the large community of cancer biologists who wish to sequence tumor from small amounts of FFPE material.

Aim 2. *Determine whether phenotypic diversity of DCIS and the tumor microenvironment (TME) is greater in DCIS with adjacent IDC compared to DCIS without IDC.* Since genomics is not the sole driver of tumor behavior, we phenotypically characterized DCIS and its microenvironment including markers of hypoxia, migration, proliferation, matrix organization, and immune signaling in the same samples used in Aim 1. We computed measures of microenvironmental heterogeneity to determine if the specific components of the TME, or the differences between TMEs from the same tumor, distinguishes between DCIS with and DCIS without adjacent IDC.

60 Month Milestones:

- IHC staining of candidate markers on all cases – **completed**
- Expert scoring of all markers on all cases – **completed**
- Data analysis using distance metrics to determine which markers demonstrate significant heterogeneity that distinguishes pure DCIS from mixed DCIS/invasive cases – **completed**
- Stain Aim 4 cases with the most promising markers of tumor and microenvironmental heterogeneity – **completed**
- Scan IHC and H&E stained slides for Automated image analysis (AIA) – **completed**
- Training and validation of AIA for the identification and enumeration of cell types (epithelial, stromal, lymphocytes, blood vessels). Computer algorithms are trained by expert identification of cell types (study pathologist, Allison Hall). Accuracy of the computer identification is evaluated by comparison back to the expert scoring. Apply methods for quantitative image analysis – **completed**
- Test computer vision methods for measuring nuclear size as a surrogate for tumor grade – **completed**

Biomarker Staining and Scoring

We have analyzed our phenotypic diversity markers on a total of 118 cases (57 pure DCIS, 61 mixed DCIS/ invasive, Table 1). To evaluate these elements, we have used a detailed expert scoring that captures the distribution of intensity of staining. This allows us to fully evaluate heterogeneity between regions of the cancer following the original study design and the genetic analyses.

Table 2: Aim 2 IHC Markers

Scoring Method	Marker	Pure (n=57)	Invasive (n=61)
	ALDH_DCIS	57	56
	ALDH_stroma	57	56
	CA9	56	58
	ER_DCIS	48	56
	FASN_DCIS	39	44
	GLUT_DCIS	56	57
	COX2	57	54
Intensity (class 1-4)	HER2_DCIS	55	56
	PFAK_DCIS	56	57
	PR_DCIS	57	55
	RANK_DCIS	57	55
	COL15	55	56
Hotspot (class 1-4)	CD31	48	50
	CD68_macrophage index	56	57
Percentage	Ki67	57	56
(Class 1-2: pos/neg)	FOXP3	51	58
	P63	54	55

For each IHC marker, the percent of cells in each staining category (Table 2) were used to define a phenotypic profile of a region. Each region also has a genetic profile generated using the two DCIS regions that were sequenced in Aim 1. We calculated three types of statistics for each marker:

1. Mean intensity score = a weighted sum of the scores, weighted by staining intensity, normalized by the maximum possible staining intensity, which is a measure of the average amount of protein expression per cell, in the sample.
2. The mean Shannon entropy across the four intensity levels of staining, which is a measure of the within sample heterogeneity among cells for the protein.
3. The Earth Mover's Distance, which is a measure of between sample heterogeneity.

Mean Intensity

Figure 5a shows that only the mean of GLUT1 staining is statistically significantly different between the pure DCIS and the DCIS that is adjacent to invasive disease (unadjusted $p = 0.004$). CA9 may also help distinguish the two groups (unadjusted $p = 0.07$).

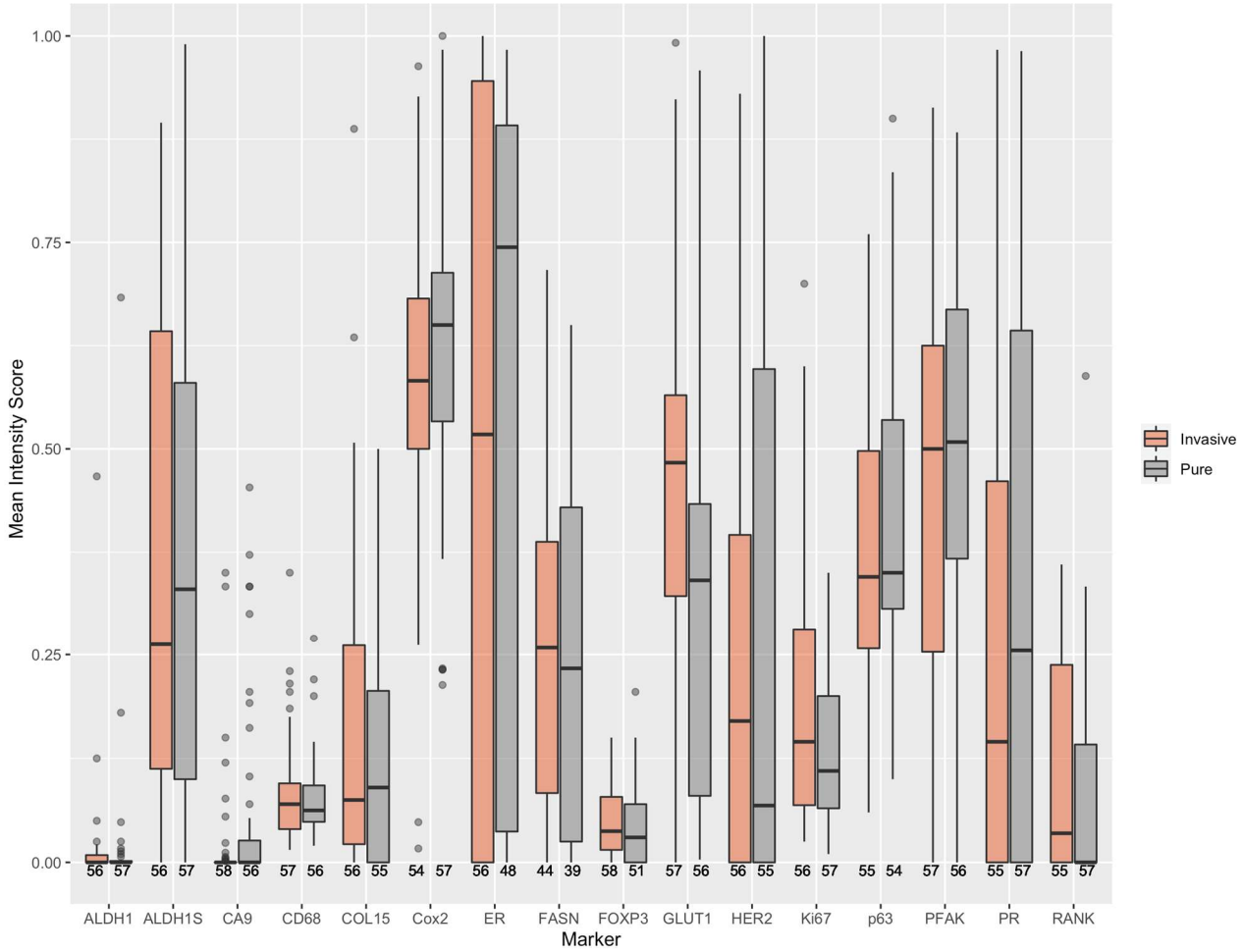

Figure 5a. Distributions of mean intensity for all Aim 2 markers. Only GLUT1 expression significantly differs between the groups, with greater expression in the DCIS adjacent to invasive cancer.

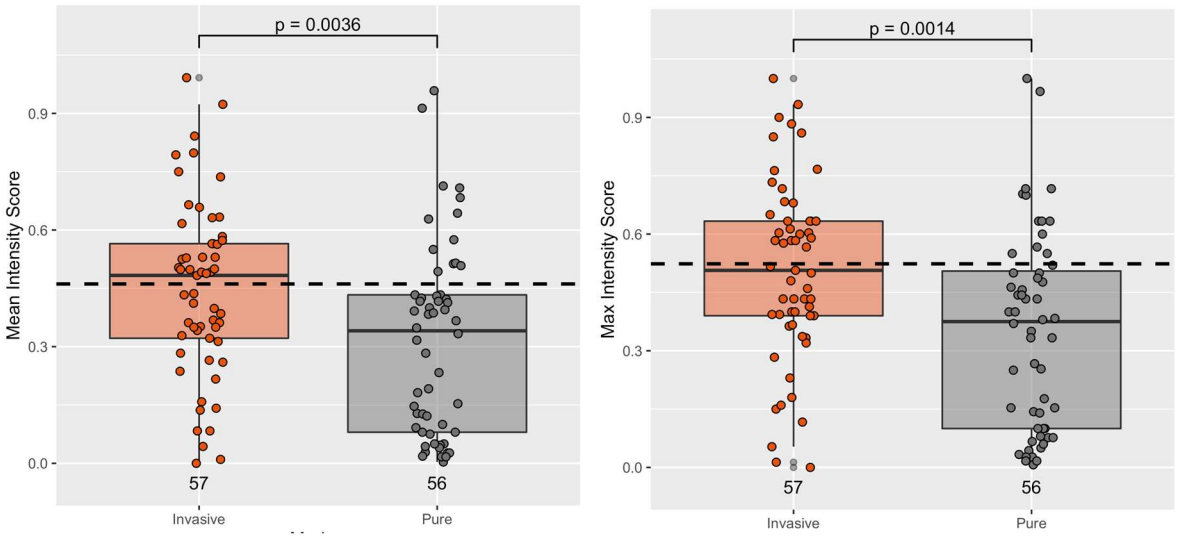

Figure 5b. Distributions of mean and maximum intensity score for GLUT1. There is significance in both these comparisons, with pure DCIS samples showing lower staining intensity. Horizontal lines represent the optimal cutoff for a logistic model classifier.

Within Sample Heterogeneity

Figure 5c shows the distributions of within sample heterogeneities as measured by the Shannon entropy statistic and averaged over the two samples from each patient. The only statistically significant difference we found was in ER staining (unadjusted $p = 0.02$), which is more heterogeneous in the pure DCIS samples. Within sample heterogeneity of PFAK may also be relevant for distinguishing the groups (unadjusted $p = 0.06$).

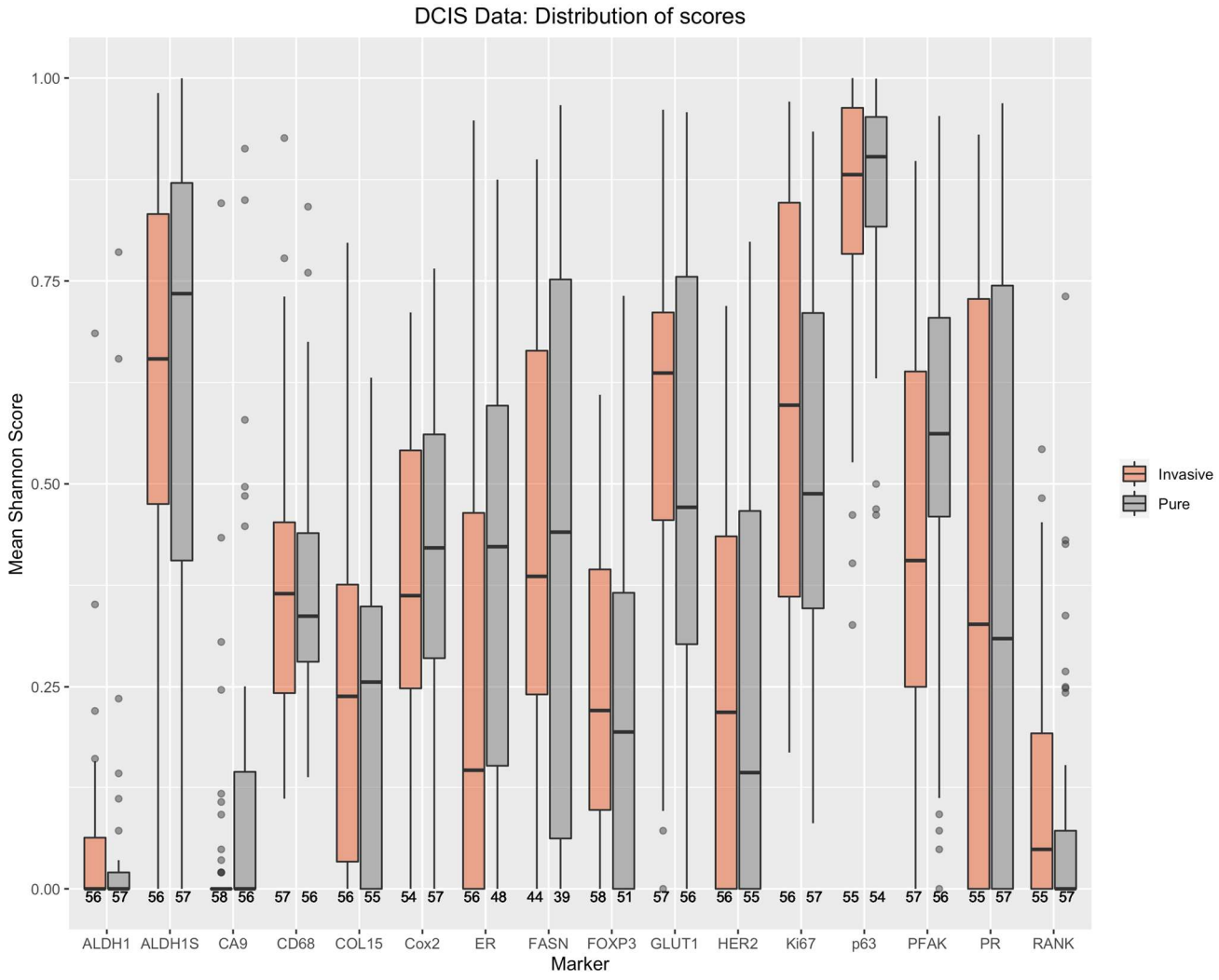

Figure 5c. Mean Shannon diversity score of marker staining's, averaged over the two samples for each patient. Only the Shannon diversity of ER is statistically significantly different between the two groups. Staining of ER is more homogeneous in the DCIS that is adjacent to invasive disease and more heterogeneous in the pure DCIS patients.

Between Sample Heterogeneity

We tested a number of measures of differences between samples. These are essentially metrics for how different are the distributions of staining intensities between the two samples from a tumor. We found that the Earth Mover's Distance was the best. This measures how much of one distribution you would have to move to transform it into the other distribution. It also takes into account how "far" you would have to move the different parts of the distribution. So the maximum Earth Mover's Distance would involve transforming a distribution of all cells not staining into a distribution of all cells staining with high intensity (or vice versa). Figure 6d shows the Earth Mover's distance between the two samples from each patient. We found a statistically significant difference in the Earth Mover's Distance between pure DCIS patients and DCIS adjacent to invasive disease for GLUT1 (unadjusted $p = 0.007$) and HER2 (unadjusted $p = 0.01$).

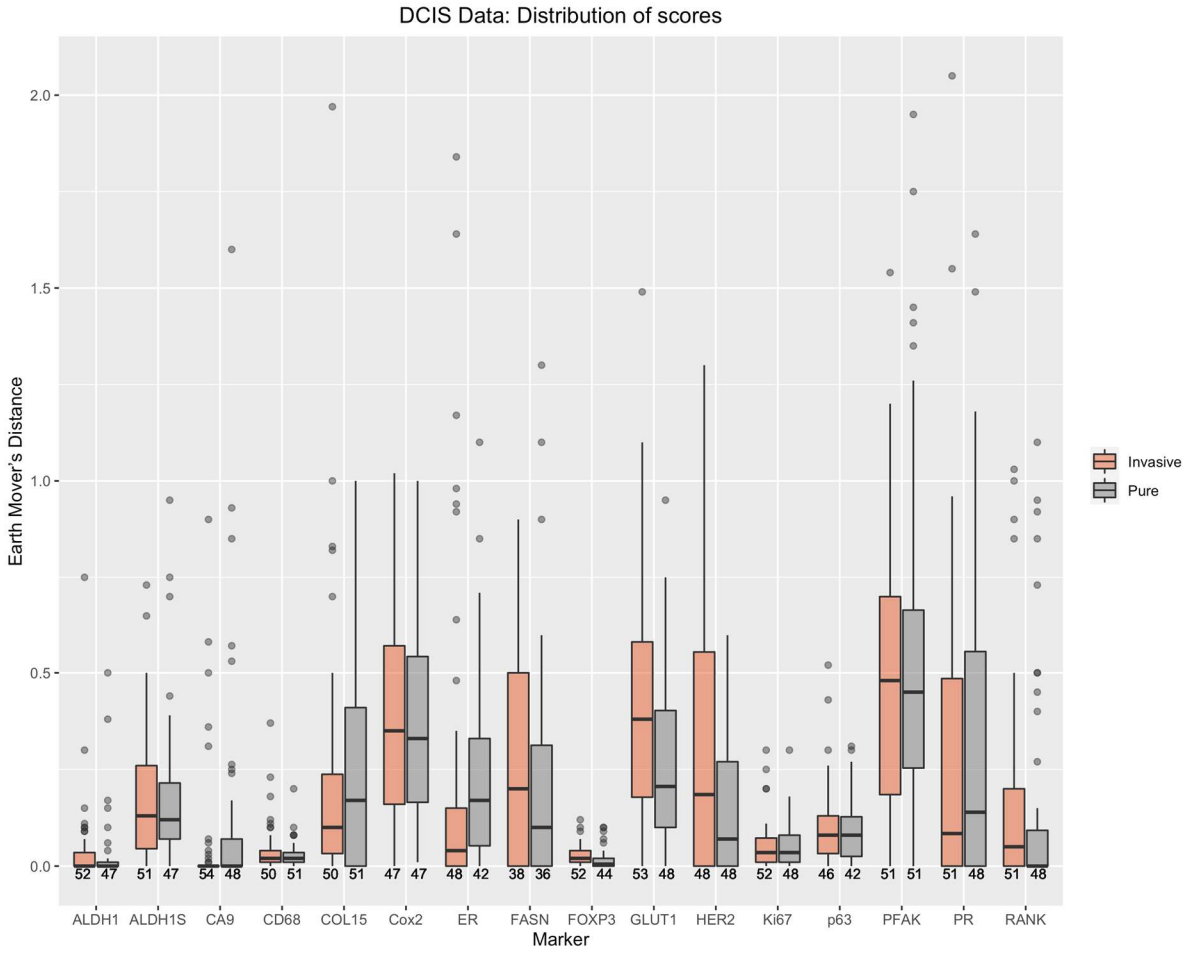

Figure 5d. Distributions of the Earth Mover's Distance for all markers and groups. This is a measure of the difference between two distributions. DCIS that is adjacent to invasive disease shows more differences between samples than pure DCIS for GLUT1 (p=0.007) and HER2 (p=0.01).

Multivariate Discrimination of Pure from Adjacent DCIS

We took the 6 variables that showed statistical significance, or near statistical significance, for distinguishing pure DCIS from DCIS that is adjacent to invasive disease, and put them together in a multivariate logistic regression to determine which measures have independent predictive value. Those variables included mean intensity scores for GLUT1 and CA9, within sample heterogeneity (Shannon) scores for ER and PFAK, and between sample heterogeneity (Earth Mover's Distance) scores for GLUT1 and HER2. When combined, only within sample heterogeneity of ER (unadjusted p = 0.002) and between sample heterogeneity of GLUT1 (unadjusted p = 0.008) remain statistically significant. This two variable model has an area under the curve (AUC) of 0.72 for the receiver-operator characteristic (ROC) curve for distinguishing pure DCIS from DCIS adjacent to invasive disease. When we added genomic divergence from Aim 1 to the model, our sample size (and statistical power) shrinks to 55 patients, because we can only reliably measure divergence in tumors with at least 10 mutations. In this case, all three measures provide independent predictive value: Within sample heterogeneity of ER (unadjusted p = 0.004), between sample heterogeneity of GLUT1 (unadjusted p = 0.02), and divergence (unadjusted p = 0.02). That three variable linear model provides an AUC=0.85 for the ROC curve. None of these p-values remain significant after the most conservative Bonferroni multiple testing adjustment, but these

analyses were applied to our training/discovery cohorts (Aims 1&2). We dealt with the potential of false positives in this analysis by testing just GLUT1 (between sample heterogeneity), ER (within sample heterogeneity) and genomic divergence in our validation cohort (Aim 4, see below).

Tumor Infiltrating Lymphocytes (TIL)

Typically, cancer researchers measure the absolute number of lymphocytes that infiltrate a tumor. However, the exact spatial relationships between lymphocytes and neoplastic cells may indicate important biological interactions between those cells. The Yuan lab has developed deep neural network methods to automatically identify ducts and classify epithelial, lymphocyte and stromal cells in H&E images. This allows us to calculate spatial statistics of the relationships between those cell types in DCIS ducts. We found that in the DCIS that is adjacent to invasive disease, lymphocytes colocalized with DCIS cells more than in the pure DCIS cases (Figure 5e), even though there were fewer lymphocytes infiltrating those ducts (Figure 5e). **These results have been published in NPJ Breast Cancer [3].**

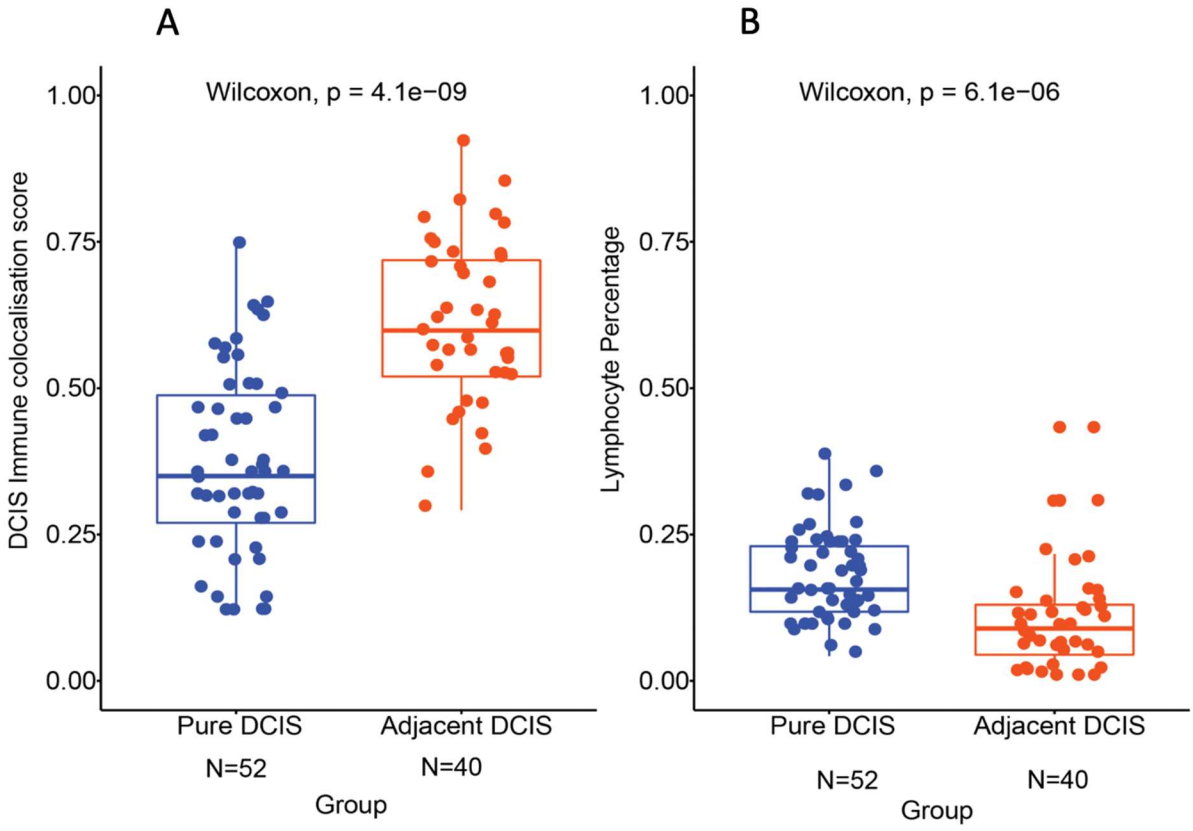

Figure 5e. ***Comparison of colocalization (A) and lymphocyte count (B)*** **in pure and synchronous DCIS.** If we only quantify the total number of lymphocytes that have infiltrated the DCIS ducts, we find that there are fewer infiltrating lymphocytes in DCIS that is adjacent to invasive disease ($p < 10^{-5}$), however they tend to be more spatially closely associated with the DCIS cells than in the pure DCIS cases. Using the Morisita-Horn index for measuring the intermixing or colocalization of DCIS cells and lymphocytes, we found that there is more colocalization of those cell types in DCIS that is adjacent to invasive disease than in pure DCIS ($p < 10^{-8}$).

Spatial Analysis of the Hypoxic DCIS Microenvironment.

To enable spatial mapping of hypoxia and T cell regulation in DCIS, we designed and reported in our last year's progression report, an end-to-end deep learning framework for single cell identification in histology images. Following that, we hypothesize, and provide preliminary data, that T-reg recruitment is spatially dependent on hypoxic microenvironment and these two spatial factors underpinned evolution of DCIS to invasive cancers. To this end, we have achieved two aims: 1.1 to develop and validate a computational pipeline that accurately segment DCIS ducts in IHC images; 1.2 to test spatial dependency of T-cell regulation on hypoxia and to compute microenvironmental divergence to determine if specific components of the TME, or the divergence between TMEs from the same tumor, differs between DCIS with and DCIS without adjacent IDC.

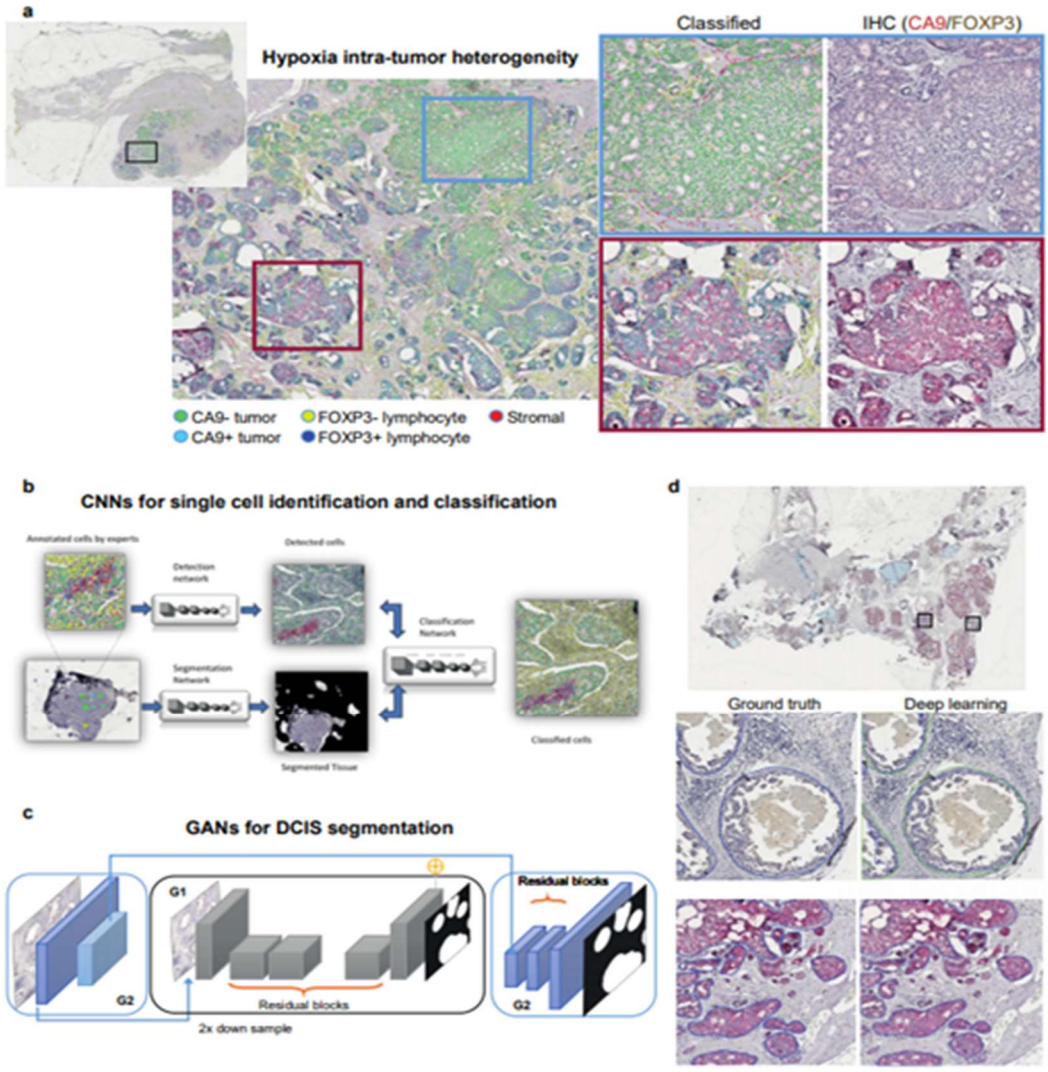

Figure 6 - Studying intra-tumor heterogeneity of hypoxia in DCIS using deep learning and digital pathology. a. An illustrative example of a DCIS tumor with high spatial intra-tumor heterogeneity of hypoxia. Shown are images of IHC dual-staining with CA9 and FOXP3, cells were classified into 5 types based on their expression of CA9 and FOXP3 and morphological features. b. The deep learning pipeline using Convolutional Neural Networks (CNNs) for single-cell analysis. c. Generative adversarial networks (GANs) for semantic segmentation of individual DCIS ducts. d. An example of DCIS tumor where individual DCIS ducts have been segmented using GANs. Two high resolution examples show ground truth obtained from annotations by pathologists and output from GANs.

Develop and validate a computational pipeline that accurately segments IHC images DCIS ducts

Following single cell identification (Figure 6a-b), we employed generative adversarial networks (GANs) for the detection and segmentation of DCIS ducts in CA9/FOXP3 IHC images (Figure 6c-d). Given the large size of whole slide images, we used an extended version of GANs (HD-GAN) for analyzing high-resolution histology images and generating semantic label maps corresponding to the target regions on the whole slide images. This method is based on conditional GANs that uses a robust adversarial learning objective together with new multi-scale generator and discriminator architectures. This model has enabled us to analyze our images at high image resolutions (i.e. 1024×1024 resolution) and predict large and variable size and shape objects on the whole slide images. The network was trained on 18 whole slide images and the performance of the model was tested on 8 unseen slides. HD-GAN achieved average Dice score of 0.90 for the segmentation performance. This model outperformed other approaches tested owing to the inclusion of paired image supervision and wider field of view that was considered for preparing the training and the testing datasets. These AI developments enabled us to quantify the spatial relationship between T cell regulation and hypoxia in DCIS.

This work has been submitted to the MICCAI workshop on Computational Pathology; COMPAY 2021.

Spatial colocalization of hypoxic tumor cells and T-regulatory cells

The deep learning-predicted cell classes and their location was used to compute the abundance of cell types and the degree of colocalization between pairs of cell types, in pure DCIS and DCIS adjacent to IDC samples, as well as the IDC itself.

Cellular colocalization*:* The Morisita Horn Index was used to quantify colocalization of FOXP3+, FOXP3- lymphocytes with CA9+ and CA9- epithelial cells. Briefly, The Morisita Horn Index computes the statistical significance in spatial co-occurrence of a pair of cell types within a spatial region defined by Voronoi tessellation. The Voronoi tessellation normalizes the number of cells in each tile of the tessellation. The number of seeds needed to perform Voronoi tessellation (prior to computing colocalization) was set as the cube root of all cells within the region of interest.

For the pure DCIS samples which are composed exclusively of DCIS duct regions, cellular abundance and colocalization were computed across whole slide images. This produced a single abundance and colocalization index (per pair of cell types), per whole-tumor section. DCIS/IDC samples are composed of DCIS ducts (segmented by deep learning DCIS segmentation) as well as invasive components. Given our interest in analyzing each component individually, we adapted the following approach: the entire tissue area was divided into non-overlapping Voronoi polygons which were classified into 3 types (Figure 6a): 1) **Synchronous DCIS** polygons composed of only DCIS ducts, 2) **IDC** polygons composed predominantly of invasive cancer cells and 3) **Mixed** polygons composed of both ducts and invasive regions, denoting the interface between DCIS and invasive regions. Cellular abundance and colocalization was computed separately for synchronous DCIS and IDC regions, producing two distinct abundance/colocalization indices per whole section tumor images. The discrepancy in number of synchronous DCIS regions (n=27, rather than 56 to

match the IDC invasive regions) is due to a proportion of IDC samples (n=29) which did not contain synchronous DCIS regions i.e., 'pure invasive' samples. These samples were not used for comparisons described above, since they produce null synchronous DCIS colocalization index (owing to lack of synchronous DCIS regions).

Spatial colocalization of CA9 and FOXP3 positive cells in pure DCIS and IDCIS samples

Colocalization between FOXP3+ and FOXP3- lymphocytes with CA9+ and CA9- epithelial cells was compared across the aforementioned sample groups (Table 3, Figure 7b). Among the pairwise comparisons (tested using pairwise Wilcoxon rank sum tests), IDC regions had significantly higher FOXP3+ and CA9+ colocalization, compared to synchronous DCIS regions (p=0.0004) and pure DCIS samples (p=0.0007) (Table 3). Similarly, we report significantly higher FOXP3+ lymphocytes abundance (number of FOXP3+ lymphocytes/total number of lymphocytes) in IDC than in synchronous DCIS regions of the IDC samples (p=3.8e-6), indicating a preferential localization of FOXP3+ lymphocytes to the invasive components in IDC (Figure 7c). To ensure that cellular abundance and clinical covariates do not confound colocalization, we report that higher FOXP3+ CA9+ colocalization in IDC regions compared to pure DCIS samples is independent of abundance of FOXP3+ lymphocytes and CA9+ epithelial cells (which do not vary significantly between these groups p=0.44 and p=0.48 respectively, Figure 7c), ER status (adjusted p=0.004), and Grade (DCIS grades, adjusted p=0.01).

We explored the differential ductal microenvironments between pure DCIS and synchronous DCIS regions (i.e., ducts adjacent to IDC). To this effect, pure DCIS samples had significantly higher FOXP3+ CA9+ colocalization (p=0.0007, Figure 7a) as well as abundance of CA9+ (p=2.3e-9) and FOXP3+ cells (p=4.4e-6, Figure 7c), compared to synchronous DCIS regions. The difference in this spatial pattern between pure DCIS and synchronous DCIS remained significant after adjusting for FOXP3+ lymphocyte (adjusted p=0.0005) and CA9+ epithelial cell abundance (adjusted p=0.0012).

In comparison, there was no significant difference in colocalization of FOXP3- lymphocytes with CA9+ epithelial cells between the 3 groups (Table 3).

Taken together, we found that the spatial microenvironmental phenotype of IDC regions differ from pure DCIS samples, with increased colocalization of FOXP3+ lymphocytes and CA9+ epithelial cells in IDC regions, independent of cellular abundance and ER status. Notably, our study revealed that the ductal microenvironment of pure DCIS and synchronous DCIS vary, with significant differences in spatial organization of hypoxic CA9+ epithelial cells and FOXP3+ lymphocytes.

Table 3: Comparison of colocalization between FOXP3+/- lymphocytes with CA9+/- epithelial cells across IDCIS-invasive, synchronous DCIS or pure DCIS sample groups. P-values are reported from Wilcoxon Ranksum test, adjustment method: Holm

Colocalization between	IDC vs Synchronous DCIS	IDC vs Pure DCIS	Synchronous DCIS vs Pure DCIS
	p-values from pairwise comparisons using Wilcoxon rank sum test (adjustment method: Holm)		
FOXP3+ and CA9+	0.0004	0.0007	0.0007
FOXP3- and CA9+	0.15	0.60	0.15
FOXP3+ and CA9-	0.04	0.74	0.05
FOXP3- and CA9-	0.04	0.45	0.04

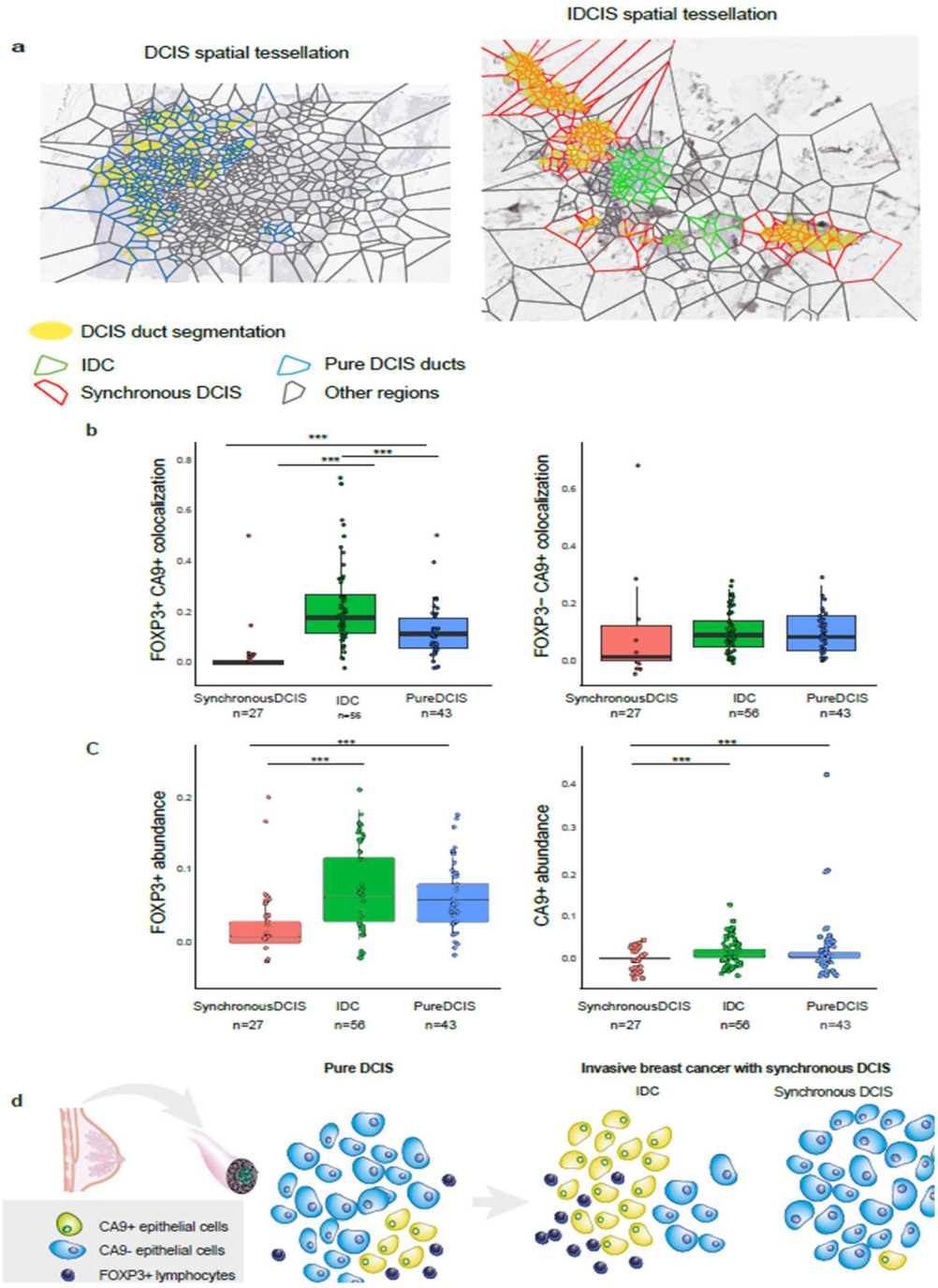

Figure 7. Applying the Morisita Horn index to measuring spatial colocalization of cells in DCIS. a. Voronoi tessellation of tumor region in a representative pure DCIS samples (left pane) and IDC (labeled IDCIS-invasive) samples (right pane), denoting normal DCIS ducts (blue polygons), IDC component (green polygons) and synchronous DCIS (red polygons). Grey polygons (in both DCIS and IDC samples) represent other cell types (fibroblasts, normal epithelium, adipose tissue and artefacts) excluded for analysis. Yellow-shaded regions represent ducts segmented by automated duct segmentation b. Boxplots comparing colocalization between FOXP3+ and FOXP3- lymphocytes with CA9+ epithelial cells in pure DCIS samples (n=44), IDC invasive (n=56) and synchronous DCIS (n=27) regions. c. Boxplots comparing FOXP3+ lymphocytes, CA9+ epithelial cells in pure DCIS samples (n=44), IDC (n=56) and synchronous DCIS (n=27) regions. FOXP3+/- lymphocyte abundance = number of FOXP3+/- lymphocytes/total number of lymphocytes. CA9+/- epithelial cell abundance = number of CA9+/- epithelial cells/total number of epithelial cells. d. Proposed model for hypoxic epithelial cells promoting T-regs recruitment selected during DCIS progression. ***: p<0.001. P values from pairwise Wilcoxon ranksum test, adjusted using Holm method for multiple testing.

Our analysis integrating deep learning, computational pathology, and spatial statistics on customized immunohistochemistry panel revealed a novel spatial pattern of preferential colocalization between FOXP3+ lymphocytes and CA9+ epithelial cells in DCIS. Compared with pure DCIS samples, the degree of CA9+ epithelial cell and FOXP3+ lymphocyte colocalization was significantly higher in the invasive compartment of invasive breast cancer (IDCIS-invasive), but significantly lower in the synchronous DCIS compartment. These differences were independent of the abundance of these cell types. Therefore, our study reiterates differential microenvironments between pure-DCIS and IDC compartments [1–4] . However, we also present evidence that the ductal microenvironment of pure DCIS and synchronous DCIS vary, with significant differences in spatial organization of hypoxic CA9+ epithelial cells and FOXP3+ lymphocytes [1–4]. Based on these data, our proposed model is that hypoxic epithelial cells promoting T-regs recruitment are selected during DCIS progression, resulting in stronger, preferential colocalization of these cells within the invasive compartment (Figure 7d).

This work was published NPJ Breast Cancer (*Narayanan, P. L et al,* 2021)

Aim 3. *Create and test a computational learning algorithm to compare mammographic characteristics and diversity measures in pure DCIS compared to DCIS with IDC.* A weighted computational algorithm using mammographic features of lesional and stromal characteristics as well as heterogeneity measures derived from Aims 1 and 2 will be constructed. The tool will be designed to allow for radiologic discrimination between good and poor prognosis DCIS, and will be evaluated in a validation set.

60 Month Milestones:

- Collection of 700 Duke DCIS mammography cases – completed.
- Development of mammography radiomics model on 400 cases to predict upstaging – **completed.**
- Testing of mammography radiomics model to predict upstaging on 300 independent cases – **completed.**
- First stage of radiologist observer study to predict upstaging from mammography – completed.
- Second stage of radiologist observer study to predict upstaging from mammography – **completed.**
- Detection of microcalcification clusters in mammography using anomaly detection approach – **completed.**

Extension Milestones:

- Submitting radiomics classification model study to journal – **completed.**
- Expansion of anomaly detection model and submission to journal – **in progress**.

In previous years, we completed two major studies. Much of the effort of this extension year has been to submit and revise these for publication.

First, we developed a mammography radiomics classification model to distinguish between pure DCIS and upstaged cases with occult IDC. The logistic regression model was trained (with cross-validation) using 400 cases and then applied to an independent, internal test set of 300 cases. Performances generalized very well, and provided different operating thresholds that may influence active surveillance or advanced surgery. The manuscript was revised and recently resubmitted to *Radiology*.

Second, we created a deep learning anomaly detector for the semi-supervised detection of calcifications in mammography. This work was published as a full-length conference proceedings paper (Peng *et al.* 2020). Based on feedback from a journal submission, we expanded the scope of this project to include comparison to other methods and more rigorous analysis. The manuscript recently came back from initial review and will be resubmitted to *IEEE Transactions in Biomedical Engineering*.

Highlighting just one of the many new analyses in the second study, we compared the effects of training with varying numbers of cases, all of which were negative mammograms that did not contain suspicious calcifications. The free-response receiver operating characteristic (FROC) curves are shown below. Note the plateau effect such that performance for the three largest datasets were all the same.

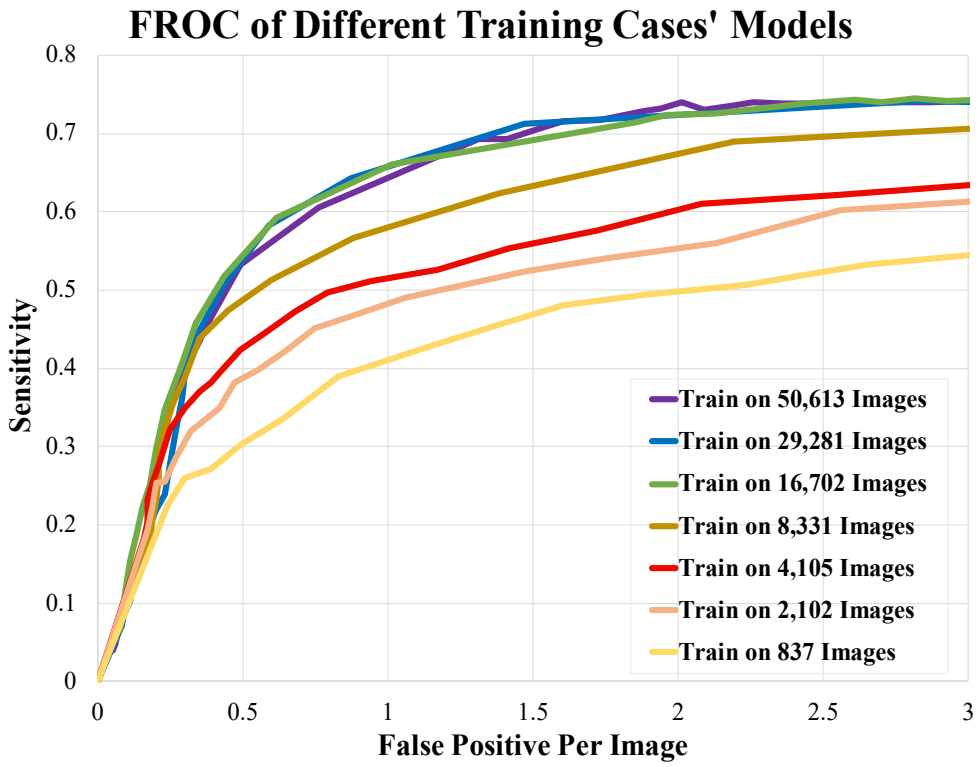

Figure 8. FROC on test set with different number of autoencoder training images. Performance improves with more cases, until reaching a plateau at 16,702 images.

Aim 4. *Test the predictive performance of the best diversity measures in an independent validation set of pure DCIS with and without subsequent invasive recurrence.* Genotypic and phenotypic measures of diversity derived from Aims 1-2 will be applied to an independent case-control, longitudinal, tissue bank of DCIS with and without invasive or DCIS recurrence to validate their utility. The Duke IRB approved protocol was approved at 12 sites. In the last two years, we accrued cases of pure DCIS with no disease recurrence or cases that recurred with DCIS or invasive cancer. Slides were shipped to Duke for macro-dissection followed by DNA analysis and immunodetection of phenotypic heterogeneity.

60 Month Milestones:

- We obtained approval to obtain specimens through the Translational Breast Cancer Research Consortium (TBCRC) and Duke IRB approval. We identified 12 high volume academic medical center consortium members who obtained regulatory approval, DOD approval and MTA's- – **completed.**

- The REDCap database is used for data entry online and slide inventory control. To date we have obtained cases from all 12 sites. We received 106 cohort 0, 75 cohort 1 and 61 cohort 2 cases thus achieving final project accrual and analysis goals – **completed.**

- Aim 4 WES data analysis – **completed** (a manuscript is in process)

- Nanostring was completed on a subset (n=20) of these cases. Additional Nanostring cases (n=150) are currently under way **– in process.**

Table 4: Aim 4 Clinical summary Table

	Pure DCIS (N=106) Cohort 0	DCIS with DCIS recurrence (N=75) Cohort 1	DCIS with Invasive recurrence (N=61) Cohort 2
Year of Diagnosis			
Median	2009	2008	2006
Age at Diagnosis			
Median	54.7	57	50.6
Mean (±SD)	55.5 (± 9.3)	55.8 (± 9.6)	52.9 (±9.6)
Grade			
1	6 [5.7%]	7 [9.3%]	5 [8.2%]
2	38 [35.8%]	33 [44.0%]	20 [32.8%]
3	62 [58.5%]	35 [46.7%]	36 [59.0%]
Treatment			
Lumpectomy w Radiation	61 [57.6%]	45 [60.0%]	28 [46.0%]
Lumpectomy no Radiation	5 [4.7%]	20 [26.7%]	13 [21.3%]
Lumpectomy Radiation Unknown	1 [0.9%]	1 [1.3%]	1 [1.6%]
Mastectomy	39 [36.8%]	9 [12.0%]	19 [31.1%]
Time to Recurrence/Disease Free follow-up (months)			
Mean (±SD)	108.4 (±39.0)	53.0 (±39.9)	72.3 (±45.2)

Aim 4 was designed to test the best markers that distinguish pure from synchronous DCIS found in Aims 1 and 2 in DCIS that either did or did not recur as invasive cancer. From the results described above, we identified three promising markers to validate: 1) genomic divergence, 2) Within region heterogeneity of ER staining, and 3) between region heterogeneity of GLUT1 staining. We used the clinical samples shown in Table 3 to test if any of these markers can predict progression from pure DCIS to recurrence of DCIS or progression to IDC.

Because this longitudinal sample set is the most relevant for clinical impact, we are also using it as a discovery cohort to test if whole genome copy number profiling, using low pass whole genome sequencing, and expression profiling, by Nanostring as well as RNA SMART3-Seq, can predict recurrence or progression. These data sets, using technology that was not available when our DOD grant was written, will greatly increase our ability to discover markers related to disease progression which is the primary goal of this project.

Mutational burden and genomic divergence

In our cross-sectional study in Aim 1, we found that mutational burden did not distinguish pure DCIS from DCIS adjacent to IDC, but that the degree of genomic divergence did (excluding cases with too few mutations to detect divergence). In our longitudinal study, we found the opposite. We found more mutations in the DCIS that progressed to IDC or recurred as DCIS than we found in the DCIS that did not recur (Figure 9). However, we found no difference in the genomic divergence between the two sequenced regions for any of the cohorts (Figure 10).

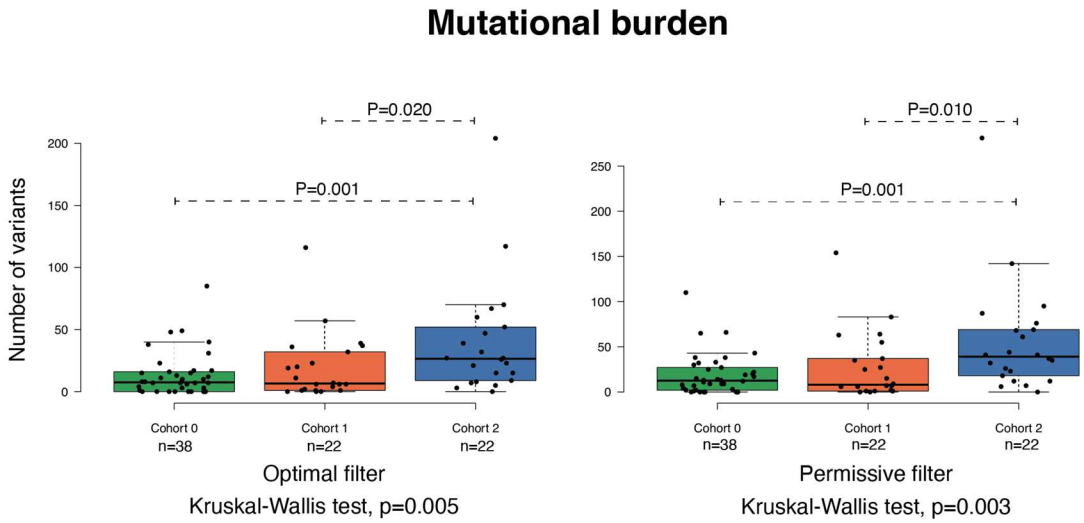

Figure 9. The number of mutations we detected in the pure DCIS at baseline was significantly different between our two cohorts, no matter whether we used our optimized filter (left panel) or a more permissive filter (right panel). Cohort 0 are the tumors that did not recur; Cohort 1 are the tumors that recurred as DCIS and Cohort 2 are the tumors that progressed to IDC.

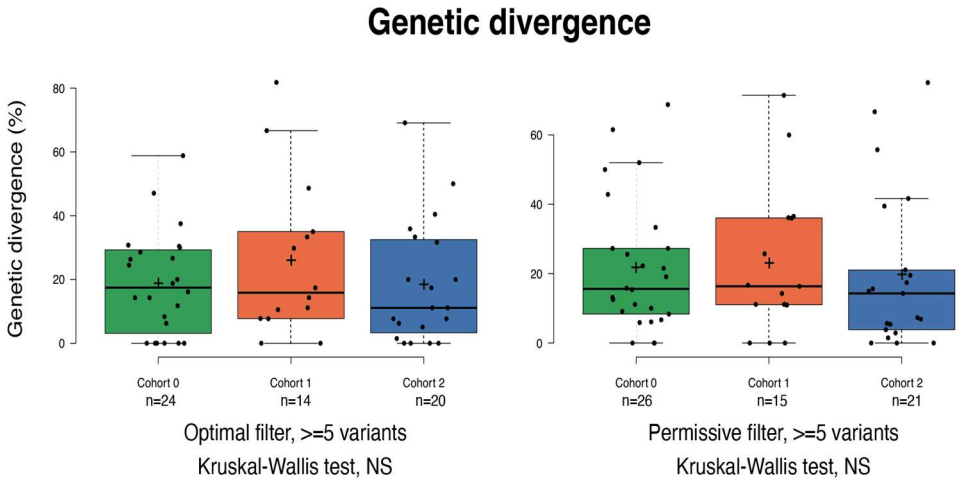

Figure 10. The genetic divergence is the percent of mutations that differ between the two regions that we sequenced in a tumor. There were no significant differences between the cohorts regardless of whether we used our optimized filter (left panel) or a more permissive filter (right panel). Cohort 0 are the tumors that did not recur; Cohort 1 are the tumors that recurred as DCIS and Cohort 2 are the tumors that progressed to IDC.

Between region heterogeneity of GLUT1

Our previous GLUT1 results from Aim 1 were not validated in our longitudinal study (using a Kruskal-Wallis test, see Figure 11). Using the Earth Movers Distance to characterize the difference between the two regions we assayed for each DCIS, we found no significant differences between any of the three cohorts (pure DCIS that did not recur, pure DCIS that recurred as DCIS, and pure DCIS that progressed to IDC).

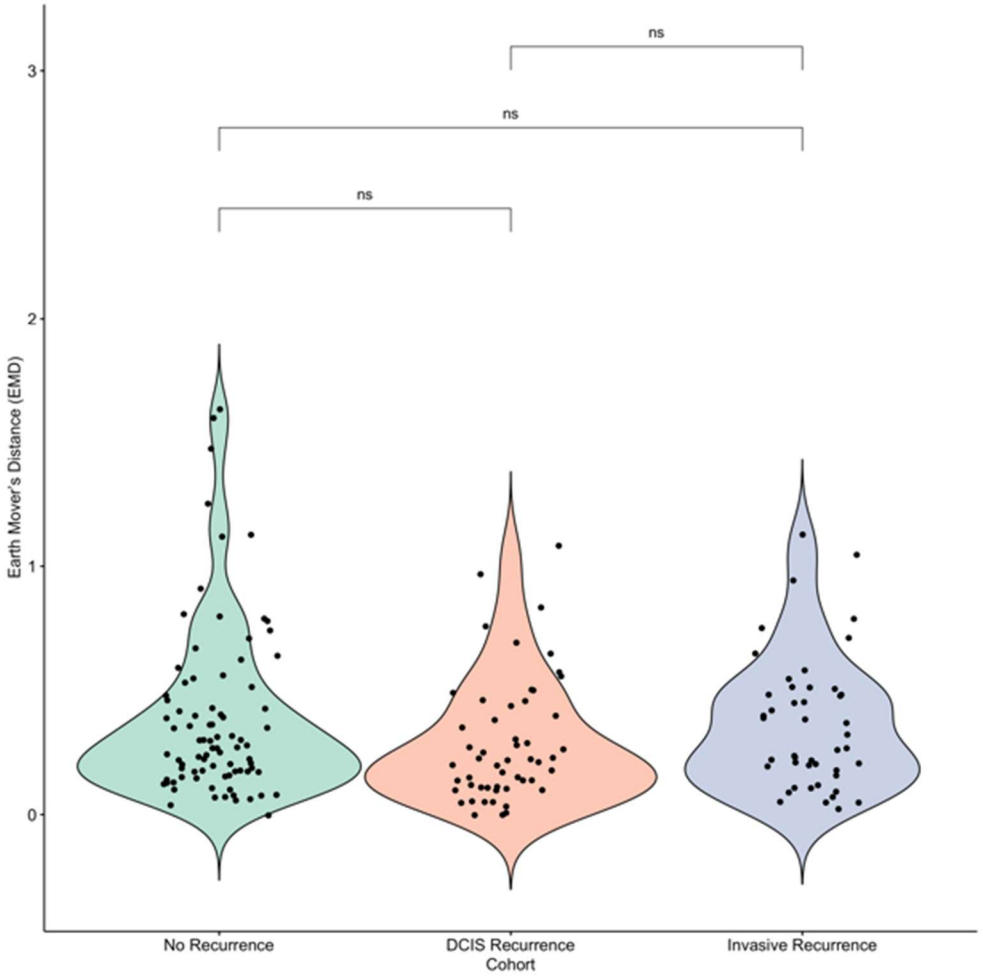

Figure 11. Unlike the cross-sectional study in Aim 1, we found in the longitudinal study that where were no significant differences in the between region heterogeneity (Earth Movers Distance) of GLUT1 staining for any of the three groups: Pure DCIS with no recurrence, pure DCIS that recurred as DCIS, and pure DCIS that progressed to IDC.

Within region heterogeneity of ER

In our ER staining, we noticed that the highest risk DCIS was often associated with either high intensity staining, across most of the DCIS cells, or a complete lack of staining. While, low risk DCIS more often had a lot of variation in ER staining across the cells. The Shannon index captures this distinction imperfectly, as it would score a moderate level of staining across all the cells the same as a high level of staining across all the cells. So we developed a variation on the Shannon index to better distinguish the extreme cases of high or no staining from variation in staining. Our new statistic is the absolute value of the Cumulative Density Index (CDI). CDI lumps the two extremes together and distinguishes them from cases where the staining profile was more diverse.

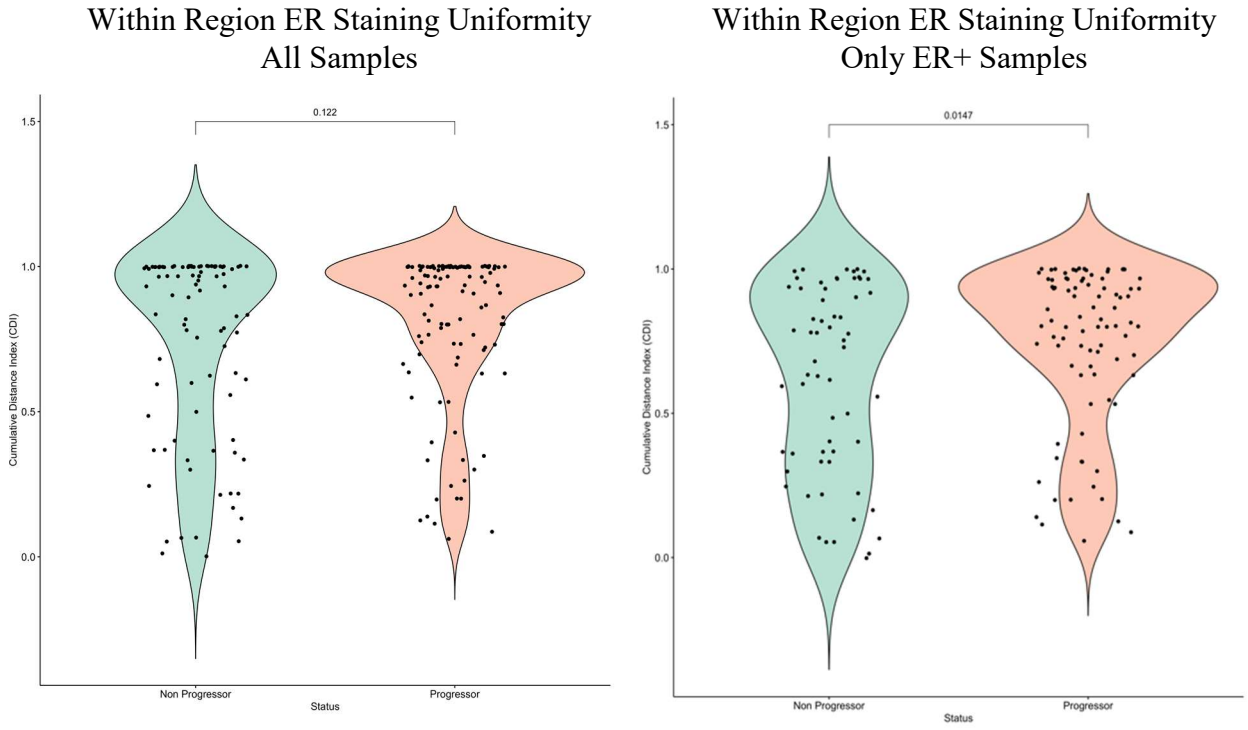

Figure 12. When we include all the samples (left panel), there is no clear distinction between the DCIS that did not recur or progress (green) compared to those that recurred as DCIS or as IDC (orange). However, if we exclude DCIS that were classified as ER- by both the clinical assay for that patient's tumor, and by our RNA expression data (right panel), then there were significantly more DCIS cases with heterogeneous staining (low CDI scores) that did not recur or progress (p=0.015).

In this sense, a low value of CDI indicates higher diversity (or higher entropy) and a high value indicates an extreme staining intensity case with high coverage (thus more uniformity of ER expression – it was rare for a sample to show uniform but intermediate values of ER staining). There is not a strong distinction in the heterogeneity (CDI scores) of ER staining between the DCIS that did or did not recur (Figure12 left panel). However, our colleagues gave us feedback that they felt we should not be including tumors that are ER- in this analysis because you can't get heterogeneity of ER staining in a tumor that simply does not express ER. So we analyzed the data excluding cases that were ER- by both the clinical assay that had been used to manage the patient, and by our RNA data from that tumor. Once those were excluded, we did find a significant difference in the within region ER staining heterogeneity by CDI (Figure12 right panel, unadjusted p=0.015). A Cox proportional hazards regression, controlling for age at diagnosis, and dichotomizing the CDI scores based on the training set best threshold (CDI=0.567), shows strong statistical significance for ER staining heterogeneity (HR=3, p=0.007, Figure 13 KM).

Patients with CDI scores below 0.0567 had a median of 30 months before they experienced a recurrence of their DCIS or progression to IDC, compared to patients with high CDI scores. Note that the predictive value of the ER CDI score is particularly high for the first 80 months of follow-up, after which the confidence intervals on the Kaplan-Meier curves start to overlap (Figure 13 KM). We suspect this is because the late recurrences are a mixture of tumors that derive from the primary tumor as well as tumors that evolved independently from the primary tumor. In the latter case, assays of the primary tumor are unlikely to be predictive of the independent tumors that developed later. Even if a late recurrence did derive from the primary tumor, the long interval between the occurrences of those tumors implies that the two tumors may have evolved very different phenotypes, which would also weaken the predictive value of assays done on the primary tumor.

Figure 13 KM. Pure DCIS with ER CDI scores above 0.567 were assigned to the high (red) group, and tumors with lower scores were assigned to the low (green) group. There was a statistically significant difference in time to progression (recurrence of DCIS or progression to IDC) between the two groups (Median 30 months delay in progression for the low CDI group; Logistic regression p=0.011; Cox regression HR=3, p=0.007).

The discrepancies between our findings in Aim 1 and the results of our validation study in Aim 4 suggest that DCIS that is synchronous with IDC cannot be fully credentialed as a model for pure DCIS that will recur or progress. Unfortunately, it is difficult to collect DCIS samples with enough follow-up to do the valuable longitudinal studies, like we did in Aim 4, and so we saved those samples for our validation aim, rather than use half of them for Aim 1 and thereby be statistically under-powered for both aims.

TBCRC histology automated image analysis

To characterize tissue spatial architecture and the microenvironment of DCIS for the TBCRC (Aim 4) cohort, we designed an end-to-end deep learning framework for histology image analysis. Our primary objectives were: 1) to develop and validate a computational pipeline that accurately classify different cell types; 2) to automatically identified individual DCIS ducts; 3) to test spatial dependency of lymphocyte and tumor cells; 4) to compute microenvironmental divergence among different DCIS to determine if specific components of the TME, or the divergence between TMEs from the same tumor, differs between DCIS that does or does not progress. We have achieved objectives 1-3.

Deep-learning pipelines for Aim 4 TBCRC HE digital histology

To objectively and accurately score cancer cells and their spatial relationships with lymphocytes within DCIS samples, we developed a deep learning approach using convolutional neural networks (CNNs, Figure 14). The deep learning framework in this study consists of three parts:

Tissue segmentation - The fully automated tissue segmentation was performed to remove background and reduce noise and artefacts, allowing for computational efficiency and reduced processing time in subsequent image analyses steps. Tissue segmentation was performed using a pre-trained Micro-Net-512 [7, 8]. (Figure 14-a).

Single cell segmentation - To morphometrically analyze every single cell, we used the Micronet single cell segmentation trained model [6] (Figure 14-b).

Single cell classification - The SC-CNN (Spatially Constrained Convolutional Neural Network) model trained on the LUNG TRACERx cohort [7] was applied to individually classify each cell type. (Figure 14-c).

Performance of deep learning models for single cell detection and classification (92% accuracy, Table 5) was evaluated by comparison with ground truth i.e., pathologist's annotations of 10,486 single cell annotations.

Table 5- Evaluation of single cell classification

	fibroblast	lymphocyte	Tumor
Sensitivity	0.7716	0.918	0.981
Specificity	0.9801	0.982	0.925
Accuracy	0.876	0.95	0.953

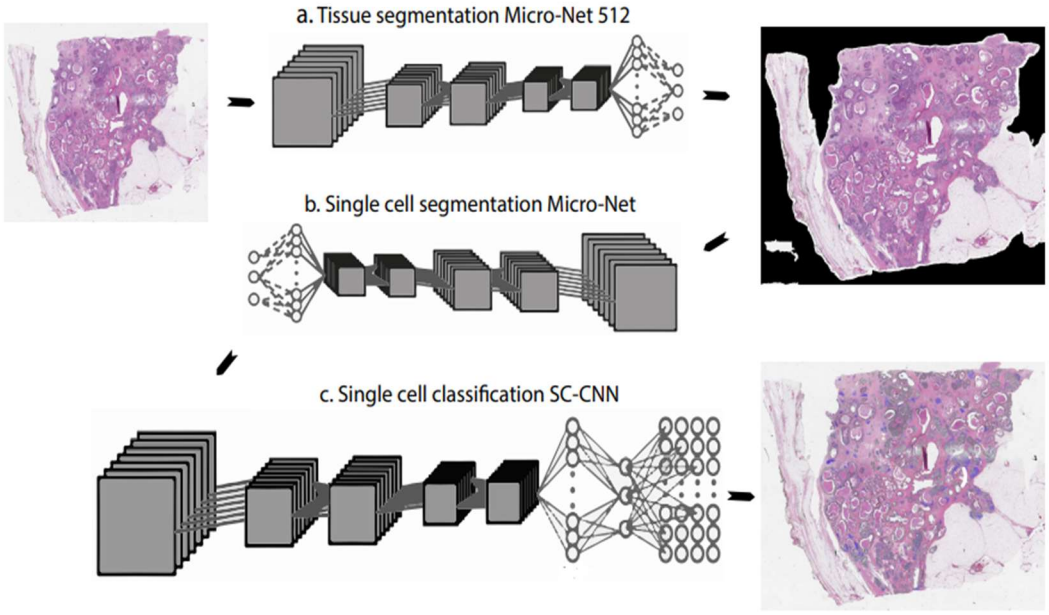

Figure 14. The deep learning pipeline using Convolutional Neural Networks (CNNs) for single-cell analysis. a. Tissue segmentation Micronet 512, b. Single cell segmentation Micro- Ner , c. SC-CNN single cell classification.

High concordance between histology deep learning results and RNAseq

We also evaluated the performance of the classification method through the comparison of the Deep learning pipeline results with the RNA seq immune estimate using Cibersort (Figure 2).

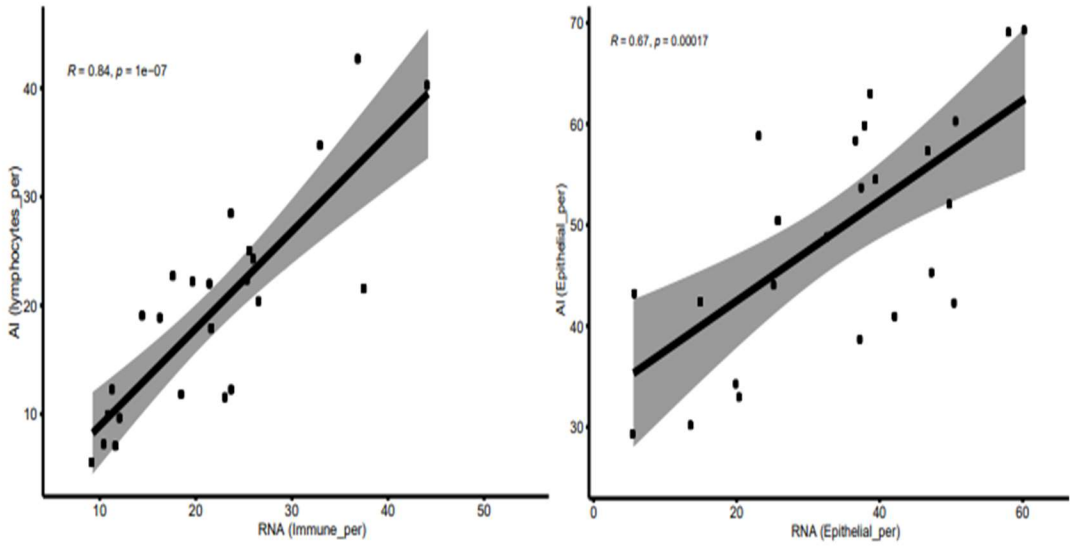

Figure 15. Comparison of AI method with the RNA seq immune estimate from Cibersort.

DCIS segmentation

We employed a generative adversarial network (GAN) for the detection and segmentation of DCIS ducts in H&E images. Given the large size of whole slide images, we used an extended version of GANs [2] for analyzing high-resolution histology images and generating semantic label maps corresponding to the target regions (DCIS ducts in our case). This model enabled us to analyze images at high resolutions and predict ducts of variable size and shape. The network was trained on hand drawn annotations of 30 whole slide H&E images (H.M.H pathologist) and the performance of the model was tested on annotations from 10 unseen slides. This model achieved the Dice score of 0.93 for the segmentation performance.

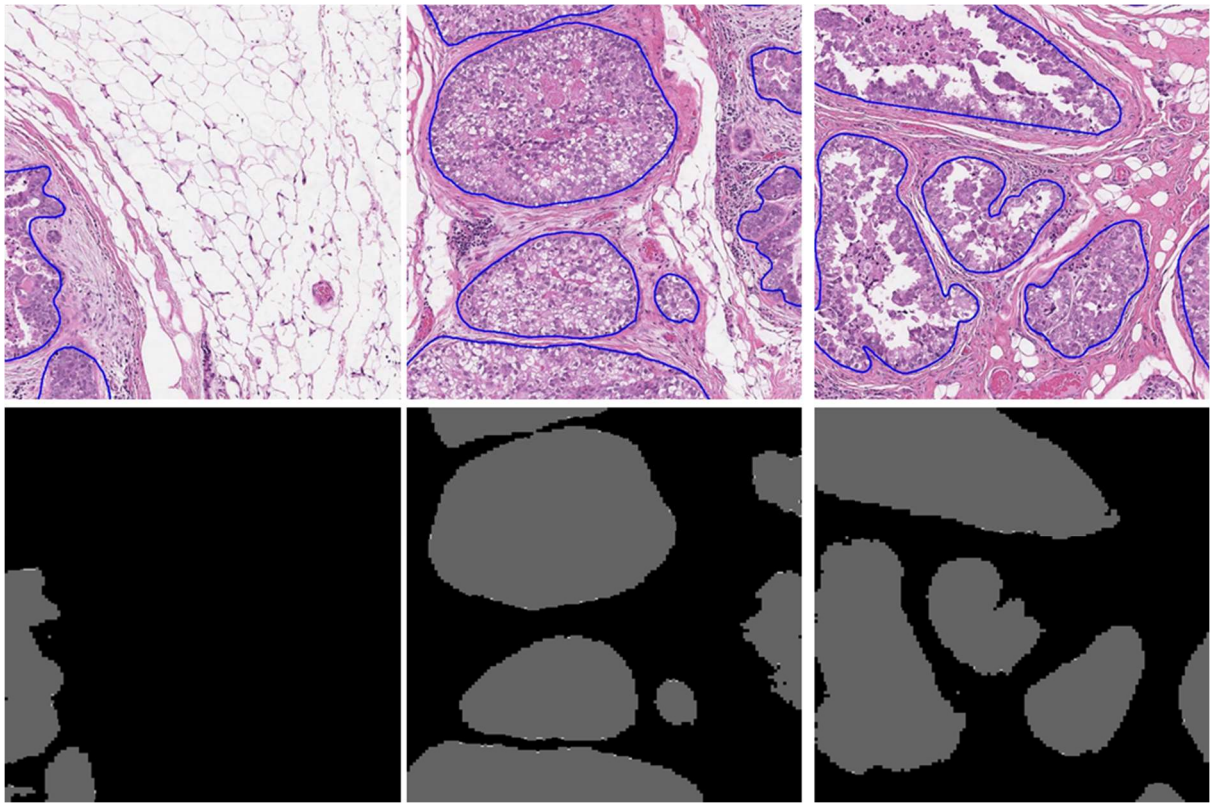

Figure 16 Example of Automated DCIS identification and segmentation. The top row is the sample of manual annotations and the masks (bottom row) are the predicted DCIS from the proposed network based on the manual annotations.

What was accomplished under these goals?

Our primary goals have been met including, testing and finding evidence that genetic diversity within DCIS is associated with progression to invasive ductal carcinoma (Aim 1).
This entailed development of a bioinformatics tool for reliably detecting mutations in small amounts of fixed DNA. We also identified immunohistochemical markers that reveal intra-tumor phenotypic heterogeneity that also distinguishes pure DCIS from DCIS that is adjacent to invasive disease (Aim 2). We have acquired more radiology imaging data sets and established the computer vision algorithms for their analysis. Further, we accrued sufficient cases and controls at Duke to fulfill the Aim 1 and 2 goals of the project. Overall, we completed the proposed work in the project period along the time line that was provided.

What opportunities for training and professional development has the project provided?

We hired a new post-doctoral fellow and technical staff in the previous year to continue expanding our analysis. Luis Cisneros joined the ASU data analysis team and attended a breast cancer meeting. Lunden Simpson joined the Duke team and has been trained in use and analysis of Nanostring. Rui Hou has been partly funded by this project for her entire PhD dissertation research, resulting in multiple conference presentations and manuscripts. Priya Narayanan has completed her PhD in the Yuan Lab.

How were the results disseminated to communities of interest?

Work based on aims 1 and 2 presented at the San Antonio Breast Cancer Symposium in December 2019, and for aim 3 at that conference in December 2017.
The work based on aim 3 resulted in several papers [5, 6, 8, 14, 15, and 16]. We also presented talks and posters at SPIE Medical Imaging in February 2017, 2018, 2019, and 2020.

What do you plan to do during the next reporting period to accomplish the goals? N/A this is the final reporting period

4. IMPACT

Successful completion of this project has led to a variety of biomarkers (genetic, IHC and radiographic) to distinguish high risk from low risk DCIS. These results will be further carried forward and validated as part of the Human Tumor Atlas Consortium (HTAN).

What was the impact on the development of the principal discipline(s) of the project?

We continue to advance the field's understanding of DCIS progression and the impact of tumor heterogeneity on the fate of DCIS. The final deliverables of this proposal will impact how DCIS is regarded both by the scientific and clinical communities.

What was the impact on other disciplines?

We have contributed to emerging knowledge regarding the digital radiographic characteristics of DCIS and continue to extend the applications for machine learning in breast cancer. We are one of the most active teams in the field, as evidenced by numerous publications and invited talks.

What was the impact on technology transfer?

Nothing to report.

What was the impact on society beyond science and technology?

Nothing to report.

5. CHANGES/PROBLEMS

Changes in approach and reasons for change

There have been no changes in approach.

Actual or anticipated problems or delays and actions or plans to resolve them

None in this reporting period. Prior challenges have been reported and resolved.

Changes that had a significant impact on expenditures

None

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

None

Significant changes in use or care of human subjects

None

Significant changes in use or care of vertebrate animals.

Not applicable.

Significant changes in use of biohazards and/or select agents

Not applicable

6. PRODUCTS

Publications

1. Fortunato, A, Mallo, D, Rupp, SM, King, LM, Hardman, T, Lo, JY, Hall, A, Marks, JR, Hwang, ES, Maley, CC. "A new method to accurately identify single nucleotide variants using small FFPE breast samples" Brief Bioinform. 2021 Jun 11. pii: 6296507. doi: 10.1093/bib/bbab221. Published. Acknowledged federal support

2. Narayanan, P. L., Raza, S. E. A., Hall, A. H., Marks, J. R., King, L., West, R. B., Hernandez, L., Guppy, N., Dowsett, M., Gusterson, B., Maley, C., Hwang, E. S., Yuan, Y., "Unmasking the immune microecology of ductal carcinoma in situ with deep learning" NPJ Breast Cancer, Epub 2021/03/03, doi: 10.1038/s41523-020-00205-5. . PubMed PMID: 33649333. Published. Acknowledged federal support.

3. Peng Y, Hou R, Ren Y, Grimm LJ, Marks JR, **Hwang ES, Lo JY**, "Microcalcification localization and cluster detection using unsupervised convolutional autoencoders and structural similarity index", Proc. SPIE 11314, Medical Imaging 2020: Computer-Aided Diagnosis, 1131403 (16 March 2020). Published. Acknowledged federal support

4. Grimm LJ, Neely B, Hou R, Selvakumaran V, Baker JA, Yoon SC, Ghate SV, Walsh R, Litton TP, Devalapalli A, Kim C, Soo MS, Hyslop T, **Hwang ES**, **Lo JY**. "Mixed-methods study to predict upstaging of DCIS to invasive disease on mammography". AJR Am J Roentgenol. Epub 2020/08/14. doi: 10.2214/AJR.20.23679. PMID: 32783550. Published. Acknowledged federal support.

5. Hou R, Grimm LJ, Mazurowski MA, Marks JR, King L, **Maley CC**, **Hwang ES, Lo JY**, "A multitask deep learning method in simultaneously predicting occult invasive disease in ductal carcinoma in-situ and segmenting microcalcifications in mammography," Proc. SPIE. 11314, Medical Imaging 2020: Computer-Aided Diagnosis, 1131405 (23 March 2020); doi.org/10.1117/12.2549669. Published. Acknowledged federal support.

6. Hou R, Mazurowski MA, Grimm LJ, Marks JR, King LM, **Maley CC, Hwang ES**, **Lo JY**. Prediction of Upstaged Ductal Carcinoma in situ Using Forced Labeling and Domain Adaptation. IEEE Trans Biomed Eng. 2019. Epub 2019/09/11. doi: 10.1109/TBME.2019.2940195. PubMed PMID: 31502960. Published. Acknowledged federal support.

7. Grimm LJ, Miller MM, Thomas SM, Liu Y, **Lo JY, Hwang ES**, Hyslop T, Ryser MD, "Growth Dynamics of Mammographic Calcifications: Differentiating Ductal Carcinoma in Situ from Benign Breast Disease," Radiology, 292:77-83 (2019). Epub 2019/05/22. doi: 10.1148/radiol.2019182599 PubMed PMID: 31112087. Published. Acknowledged federal support.

8. Hou R, Ren Y, Grimm LJ, Mazurowski MA, Marks JR, King L, **Maley CC, Hwang ES**, **Lo JY**, "Malignant microcalcification clusters detection using unsupervised deep autoencoders," Proc. SPIE 10950, Medical Imaging 2019: Computer-Aided Diagnosis, 109502Q (13 March 2019). Published. Acknowledged federal support.

9. Walther, V., Hiley, C.T., Shibata, D., Swanton, C., Turner, P.E., and **Maley, C.C.**: Can oncology recapitulate paleontology? Lessons from species extinctions. Nature Reviews Clinical Oncology, 12:273-285, 2015. doi:10.1038/nrclinonc.2015.12. Published. Acknowledged federal support.

10. Caulin, A.F., **Maley, C.C.:** Solutions to Peto's Paradox Revealed by Mathematical Modeling and Cross-Species Cancer Gene Analysis. Philosophical Transactions of the Royal Society of London B, 370 (1673):20140222. Published. Acknowledged federal support.

11. Aktipis, C.A., Boddy, A.M., Jansen, G., Hibner, U., Hochberg, M.E., **Maley, C.C.**, Wilkinson, G.S.: Cancer across the tree of life: Cooperation and cheating in multicellularity. Philosophical Transactions of the Royal Society of London B, 370 (1673):20140219. Published. Acknowledged federal support.

12. Noemi Andor, Trevor A. Graham, Marnix Jansen, Li C. Xia, C. Athena Aktipis, Claudia Petritsch, Hanlee P. Ji, **Carlo C. Maley**: Pan-cancer analysis of the extent and consequences of intra-tumor heterogeneity. Published. Nature Medicine 22:105-13, 2016. Acknowledged federal support.

13. **Carlo C. Maley**, Konrad Koelble, Rachael Natrajan, Athena Aktipis and Yinyin Yuan: An ecological measure of immune-cancer colocalization as a prognostic factor for breast cancer. Breast Cancer Research 17:1-13, 2015. Published. Acknowledged federal support.

14. Shi B, Grimm LJ, Mazurowski MA, Baker JA, Marks JR, King LM, **Maley CC, Hwang ES, Lo JY**, "Can Occult Invasive Disease in Ductal Carcinoma In Situ Be Predicted Using Computer-extracted Mammographic Features?" Academic Radiology, 24 (9), 1139-1147 (2017). PMC5557686. Published. Acknowledged federal support.

15. Shi B, Grimm LJ, Mazurowski MA, Baker JA, Marks JR, King LM, **Maley CC, Hwang ES, Lo JY**, Prediction of Occult Invasive Disease in Ductal Carcinoma in Situ Using Deep Learning Features, J Am Coll Radiol, 2018 Mar;15(3 Pt B):527-534. Epub 2018/02/06 Acknowledged federal support

16. Shi B, Grimm LJ, Mazurowski MA, Marks JR, King LM, **Maley CC, Hwang ES**, **Lo JY**, Prediction of occult invasive disease in ductal carcinoma in situ using computer-extracted mammographic features, Proc. SPIE 10134, Medical Imaging 2017: Computer-Aided Diagnosis, Armato SG, Petrick NA, Eds., 101341I (2017). Published. Acknowledged federal support.

17. Shi B, Grimm LJ, Mazurowski MA, Marks JR, King LM, Maley CC, **Hwang ES**, **Lo JY**, "Can upstaging of ductal carcinoma in situ be predicted at biopsy by histologic and mammographic features?" Proc. SPIE 10134, Medical Imaging 2017: Computer-Aided Diagnosis, Armato SG, Petrick NA, Eds., 101342X (2017). Published. Acknowledged federal support.

18. Abegglen, L.M., Caulin, A.F., Chan, A., Lee, K., Robinson, R., Campbell, M.S., Kiso, W.K., Schmitt, D.L., Waddell, P.J., Bhaskara, S., Jensen, S.T., **Maley, C.C**.†, Schiffman , J. D.†: Potential Mechanisms for Cancer Resistance in Elephants and Comparative Cellular Response to DNA Damage in Humans. JAMA, 314:1850-1860, 2015. Published. Acknowledged federal support.

19. Li, X., Paulson, T.P., Galipeau, P.C., Sanchez, C.A., Liu, K., Kuhner, M.K., **Maley, C.C**., Self, S.G., Vaughan, T.L., Reid, B.J., Blount, P.L.,: Assessment of esophageal adenocarcinoma risk using somatic chromosome alterations in longitudinal samples in Barrett's esophagus. Cancer Prevention Research, 8:845-56, 2015. Published. Acknowledged federal support.

20. Kostadinov, R., **Maley, C.C**., Kuhner, M.K.: Bulk genotyping of biopsies can create spurious evidence for hetereogeneity in mutation content. PLoS Computational Biology, 12:e1004413, 2016. Published. Acknowledged federal support.

21. Andor, N., **Maley, C.C**., Ji, H. P. Genomic Instability in Cancer: Teetering on the Limit of Tolerance. Cancer Research 77:2179-2185, 2017. Published. Acknowledged federal support.

22. Tollis, M., Boddy, A. M., **Maley, C.C ,**. Peto's Paradox: How has evolution solved the problem of cancer prevention? BMC Biology 15:60, 2017. Published. Acknowledged federal support.

23. Fortunato, A., Boddy, A., Mallo, D., Aktipis, A., **Maley, C.C**. †, & Pepper, J. W. †: Natural Selection in Cancer Biology: From Molecular Snowflakes to Trait Hallmarks. Cold Spring Harbor Perspectives in Medicine, a029652, 2016. († = co-senior authors) Published. Acknowledged federal support.

24. **Maley, C.C**., Aktipis, A., Graham, T.A., Sottoriva, A., Boddy, A.M., Janiszewska, M., Silva, A.S., Gerlinger, M., Yuan, Y., Pienta, K.J., Anderson, K.S., Gatenby, R., Swanton, C., Posada, D., Wu, C.-I., Schiffman, J.D., **Hwang, E.S**., Polyak, K., Anderson, A.R.A., Brown, J.S., Greaves, M., Shibata, D.: Classifying the Evolutionary and Ecological Features of Neoplasms. Nature Reviews Cancer, Sept. 15, 2017. Published. Acknowledged federal support.

Citations

25. Hoon Tan, P., et al., *The 2019 WHO classification of tumors of the breast.* Histopathology, 2020.

26. Elston, C.W. and I.O. Ellis, *Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up.* Histopathology, 1991. **19**(5): p. 403-10.
27. *Consensus Conference on the classification of ductal carcinoma in situ. The Consensus Conference Committee.* Cancer, 1997. **80**(9): p. 1798-802.
28. Rimmer, A., et al., *Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications.* Nat Genet, 2014. **46**(8): p. 912-918.
29. Wang, K., M. Li, and H. Hakonarson, *ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data.* Nucleic Acids Res, 2010. **38**(16): p. e164

Technologies or techniques

Nothing to report

Inventions, patent applications, and/or licenses

Nothing to report

Other Products

none

Grant Funding Resulting from Extension of these results

RFA-CA-17-035 Pre-Cancer Atlas (PCA) Research Centers (U2C) (PI: Hwang; coPIs West, Maley) NCI Human Tumor Atlas Network (HTAN) 2018/09/01-2024/08/30

PCORI 1505-30497 (Hwang) Comparison of Operative versus Medical Endocrine Therapy for low risk DCIS: The COMET Trial. 2016/07/01-2023/06

BCRF 19-074 (Hwang) 2019/10/02-2020/09/30**:** Molecular and Radiologic Predictors of Invasion in a DCIS Active Surveillance Cohort.

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

Co-PI: Dr. Shelley Hwang (M.D., M.P.H.): Duke University (no change)
Co-PI: Dr. Carlo C. Maley (Ph.D.): Arizona State University (no change)

Co-Investigators:
Dr. Jeffrey Marks (Ph.D.): Duke University (no change)
Dr. Joseph Geradts (M.D.): Duke University (departed during year one)
Dr. Allison Hall (M.D.): Duke University, replacing Dr. Geradts.

Dr. Joseph Lo (Ph.D.): Duke University (no change)
Dr. Jay Baker (M.D.): Duke University (no change)
Dr. Yin Yin Yuan (Ph.D): Institute for Cancer Research, UK (no change)
Dr. Lars Grimm (M.D.): Duke University (no change)
Dr. Trevor Graham (Ph.D.): Barts Cancer Institute, Queen Mary University of London (no change)
Dr. C. Athena Aktipis (Ph.D.): Arizona State University (no change)
Dr. Shane Jensen (Ph.D.): University of Pennsylvania (departed during year one)

Post-Docs:
Dr. Mengyu Wang (Ph.D): Duke University (departed during year one)
Dr. Violet Kovacheva (Ph.D): Institute for Cancer Research, UK (departed during year two)
Dr. Narayanan (Ph.D): Institute for Cancer Research, UK, replacing Dr. Kovacheva.
Dr. Sobhani (Ph.D): Institute for Cancer Research, UK.
Dr. Lorraine King (Ph.D): Duke University (no change)
Dr. Bibo Shi (Ph.D): Duke University (departed during year two)
Dr. Rui Hou, ECE (Ph.D): student, Duke University (no change)
Tushar Fakrul Islam, Associate in Research, Duke University
Dr. Yinhao Ren, BME Ph.D. student, Duke University (no change)
Dr. Angelo Fortunato (Ph.D): Arizona State University (no change)
Dr. Diego Mallo (Ph.D): Arizona State University (no change)
Dr. Luis Cisneros (Ph.D): Arizona State University