AWARD NUMBER: W81XWH-14-1-0474

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

PRINCIPAL INVESTIGATOR: Carlo C. Maley

CONTRACTING ORGANIZATION: University of California San Francisco
San Francisco, CA 94103

REPORT DATE: October 2015

TYPE OF REPORT: Annual

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

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

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

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

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

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
October 2015	Annual	30 Sep 2014 - 29 Sep 2015

4. TITLE AND SUBTITLE

Genomic Diversity and the Microenvironment as Drivers of Progression in DCIS

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-14-1-0474

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Carlo C. Maley

E-Mail: maley@asu.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

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

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO
1855 Folsom St. Suite 425
San Francisco, CA 94103-4249

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 goals of this project are to test if measures of genetic diversity, microenvironmental diversity, and/or mammographic biomarkers can be used to predict which DCIS tumors are most likely to progress to invasive breast cancer. We have applied for and received ethical approval to carry out the study at our primary site, and have completed a series of pilot experiments to determine the best resource (Washington University) that we will use to perform the genomic sequencing of our tumors. We are currently evaluating and optimizing antibodies to measure microenvironmental diversity as well as algorithms for analyzing mammograpies. We have also filtered through our tissue banks to identify the appropriate cases for the full studies, and have even begun work to secure samples for our final validation study. Finally, we have recently published three papers supported by this study in Nature Reviews Clinical Oncology and Philosophical Transactions of the Royal Society. We are making rapid progress and anticipate no problems in meeting our 24 month milestones.

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
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	11	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

1. INTRODUCTION

Both overdiagonosis and underdiagnosis have emerged as the primary problems bedeviling cancer screening and prevention. This is particularly true in breast cancer where ~75% of positive mammography results detect the pre-cancerous tumor ductal carcinoma *in situ* (DCIS). Many cases of DCIS will never progress to life-threatening cancer, and so treating all cases of DCIS as if they are cancer would expose women to unnecessary toxicity and morbidity (overdiagnosis and overtreatment). 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, 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. Diversity measures must be derived from geographically distinct areas of tumor. Genetic divergence of the DCIS component of tumors will be measured based on exome sequencing and SNP arrays run on two separate regions of the tumor, as well as normal tissue, in patients with DCIS either with or without invasion to determine the association between genetic diversity and progression to malignancy. Genetic diversity will be measured by the genetic divergence between the tumor samples, that is, the proportion of the genome that differs between the two samples from the same tumor.

24 Month Milestones:

- Protocol preparation, IRB submission and approval: **complete**[*]
- Case identification and tissue block selection: Through a variety of available databases, we have identified a large number of potential 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. We are now performing these reviews with newly created case report forms and databases to capture the information.

[*] IRB approval has been obtained at the Duke site where all the tissues are stored and processed. With the Maley lab moving from UCSF to ASU, IRB approval at ASU is pending for permission to analyze the genomic and phenotypic data produced by the Duke investigators.

- Sectioning and coring of tissue blocks: New sections from candidate paraffin blocks are made, stained with H&E, reviewed by the study pathologist, and these slides are scanned for analytic and archival purposes. Additional slides from useful blocks (containing a sufficient amount of the DCIS lesion of interest) are obtained and macro-dissected for DNA extraction. Additional sections (every other one) are also stored for immunohistochemical (IHC) analysis of key measures of heterogeneity. This process has been fully implemented and we are moving through both cases and controls in this manner.

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

- SNP and Exome sequencing of test cases: we have investigated a number of platforms and collaborators for the DNA sequencing and SNP analysis. Since we are working with small amounts of FFPE DNA, standard methodologies do not readily apply. Based on a pilot set of 14 DNA samples, we have settled on the Genome Center at Washington University run by Elaine Mardis. Dr. Mardis is working with us closely and her group has developed cutting-edge methods for producing high quality data from these specimens. In addition to full-exome capture, the method employs additional enrichment for a panel of 83 high value breast cancer genes to ensure high coverage of the most commonly altered driver genes. In the last month, Wash U. sequenced 20ng from 14 individual DNA samples derived from 4 subjects (germ line sample plus 2 DCIS containing samples with two duplicates=14) and returned the data to us for analysis. In addition, we also asked the Wash U. group to perform a basic analysis of the data for comparison to our informatics pipeline. Most important, they were able to derive interpretable sequence data from 20ng of FFPE DNA with average coverage ranging from 10-80X. Our group (Maley and Graham) analyzed these data and found numerous candidate mutations with estimated allele frequencies. The Wash U. group recently returned their analysis and we are now in the process of comparing the results.

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 will phenotypically characterize DCIS and its microenvironment including markers of hypoxia, migration, proliferation, matrix organization, and immune signaling in the same samples used in Aim 1. We will employ automated image analysis 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.

We are pleased to report that we have brought a new collaborator into the team, Dr. Yinyin Yuan from the Center for Evolution and Cancer at the Institute for Cancer Research in London. Dr. Yuan is an expert in computational image analysis of histological sections of breast cancer, and the application of ecological and other spatial statistics to those images [1-4].

24 Month Milestones:

- IHC staining of candidate markers (test cases): We have obtained a series of antibodies representing our initial targets including ER, PR, KI-67, COL15A1, RHOA, RAC, CA9, HIF1a, FOXP3, and cleaved Caspase 3. We have piloted dual staining for sets of these antibodies on other breast specimens and will soon be staining for these antigens on cases and controls. Dual staining conditions must be optimized in collaboration with Dr. Yinyin Yuan's lab who will be doing the automated, quantitative scoring and analysis of the stained tissues.

- Scan IHC results for Automated image analysis (AIA): Not started yet.

- Automated image analysis (AIA) of tumor and stromal markers of heterogeneity: Dr. Yuan's team is adapting their algorithms for dual staining. They already have successfully analyzed both clustering of cell types [2, 3], and co-localization (interleaving) among different cell types (manuscript under review).

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.

24 Month Milestones:

- Define permissible values for each input class: For automated identification of lesions representing DCIS on mammography, we created preliminary algorithms for the detection, segmentation, and clustering of microcalcifications. The multi-step process is based upon median filtering and global as well as local thresholding, with several false positive rejection steps using clustering and morphology rules. Using images from 12 randomly selected subjects, we performed a grid search to optimize initial algorithm parameters. We are in the process of implementing initial algorithms to automatically extract imaging features from the resulting microcalcification clusters. We have also developed graphical user interfaces to facilitate radiologists providing ground truth for lesion size and location. By the end of year 1, we will have preliminary but fully functional algorithms for both cluster identification and feature extraction.

- Identify test set and validation set: We are identifying the cohort of subjects to be used for the main study. Based on our inclusion and exclusion criteria, we have conducted several searches into our electronic medical records to identify qualifying DCIS cases from our institution. From over 1300 initial candidates, we have so far identified 161 potential subjects, from which we will verify availability of imaging and other required clinical data.

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 recurrence to validate their utility.

24 Month Milestones: This aim will be carried out after aims 1-3 are complete. However, we have already initiated the process for obtaining the validation specimens. In order to obtain these specimens, we presented the concept to the TBCRC and it was approved in principle pending submission and review of the formal protocol.

What was accomplished under these goals?

See above for the major activities undertaken to meet our goals. As we are in the preliminary stage of our project, we do not have significant results to report as yet.

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

Nothing to report.

How were the results disseminated to communities of interest?

Nothing to report.

What do you plan to do during the next reporting period to accomplish the goals?

Aim 1: While we plan to derive copy number variation (CNV) from the sequencing data, our goal was to use SNP arrays as the primary source of data for CNV assignment. Wash. U. has piloted the use of next generation sequencing libraries to probe SNP arrays. This work is now underway and we expect the first data set to arrive in next 2 months. In the next budget period, we anticipate sequencing exomes of up to 100 specimens as per the original goals of the grant.

Aim 2: We will begin to analyze cases and controls using a series of antibody stains described in the proposal. Scanned images of these stained slides will be shared with Dr. Yuan for image analysis and quantification. Dr. Yuan's team will adapt their algorithms to quantify dual stained slides.

Aim 3: We anticipate creating a preliminary database of at least 50 cases, which will be sufficient to drive the continued development of the mammography lesion identification and feature extraction algorithms. These images will be analyzed with the algorithms to derive measures of tumor heterogeneity as per the original goals of the proposal.

Aim 4: The TBCRC protocol is in final draft form and will be submitted to the TBCRC for review within the next month. Once it is reviewed and approved, we will begin to accrue these cases and controls for validation from participating institutions. This will begin in the second year of the budget period.

4. IMPACT

Successful completion of this project will lead to a variety of biomarkers (genetic, IHC and radiographic) to distinguish high risk from low risk DCIS. This would reduce patient suffering and conserve clinical resources for the women with low risk DCIS, and focus management efforts and clinical resources on women with high risk disease, potentially justifying the risks of interventions. As the project is in its initial stages, these important impacts await in the future.

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

Nothing to report.

What was the impact on other disciplines?

Nothing to report.

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

So far the problems that have emerged have been primarily technical. Sequencing from small amounts of FFPE tissue is relatively new. An initial pilot experiment with BGI America essentially failed due to those challenges, probably due to poor DNA quality in some samples. However, with appropriate quality control testing of the DNA before submission for sequencing, the sequencers at Wash. U. have proven that they can deliver quality results for good prices.

We would also like to remove as much contaminating normal cells as possible from the tissues before extracting the DNA for sequencing. This makes the sequencing more sensitive to picking up mutations in the cancer cells. We evaluated laser capture microdissection, but found it prohibitively labor intensive with very low yield. We found a good compromise with macrodissection of the tissue blocks.

We are currently developing our automated imaging analyses of dual stained tissue sections with Dr. Yuan. Dual staining is challenging in and of itself, because one must find staining conditions that work well for both antibodies. We anticipate that there may be difficulties distinguishing the two colors in the same pixel when a cell is positive for both markers, but this has yet to be tested.

If that proves insurmountable, we will pair a nuclear stain with a cytoplasmic stain so that they do not overlap.

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 to report.

Significant changes in use or care of vertebrate animals.

Not applicable.

Significant changes in use of biohazards and/or select agents

None to report.

6. PRODUCTS

Publications

1. 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.
2. 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.
3. 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.
4. 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. Under review at Nature Medicine. Acknowledged federal support.
5. 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. Under review at Breast Cancer Research. Acknowledged federal support.

Website(s) or other Internet site(s)

None.

Technologies or techniques

Nothing to report.

Inventions, patent applications, and/or licenses

Nothing to report.

Other Products

We are working on developing both databases and specimen collections of DCIS, but they are not yet complete.

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 (no change)
Dr. Joseph Lo (Ph.D.): Duke University (no change)
Dr. Jay Baker (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 (no change)

New:

Name:	*Yinyin Yuan*
Project Role:	*Co-investigator*
Researcher Identifier (e.g. ORCID ID):	0000-0002-8556-4707
Nearest person month worked:	0 (but this will be approximately 1 for future years)
Contribution to Project:	*Dr. Yuan will lead the algorithmic development and automated quantification of the IHC imagery from the tumors*

Funding Support:	*The Institute for Cancer Research, London, supports Dr. Yuan's salary. This grant will support a postdoc in her lab (yet to be hired) to carry out the work.*

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

Nothing to report.

What other organizations were involved as partners?

Organization Name: Washington University
Location of Organization: *St. Louis, MO*
Partner's contribution to the project: *(*Facilities *&* Collaboration) We are contracting with Wash. U. to provide the exome sequencing for our project. We are also informally collaborating with Dr. Elaine Mardis and her breast cancer team on this project.

8. SPECIAL REPORTING REQUIREMENTS

This is a collaborative award with Dr. Shelley Hwang at Duke. This technical report is being submitted due to a move of Dr. Maley from UCSF to ASU. The first year technical report will be submitted by Dr. Hwang at the Duke site at the end of our first year.

9. APPENDICES

References:

1. Heindl, A., Nawaz, S. & Yuan, Y. Mapping spatial heterogeneity in the tumor microenvironment: a new era for digital pathology. *Lab Invest* **95**, 377-84 (2015).
2. Nawaz, S., Heindl, A., Koelble, K. & Yuan, Y. Beyond immune density: critical role of spatial heterogeneity in estrogen receptor-negative breast cancer. *Mod Pathol* **28**, 766-77 (2015).
3. Yuan, Y. Modelling the spatial heterogeneity and molecular correlates of lymphocytic infiltration in triple-negative breast cancer. *J R Soc Interface* **12** (2015).
4. Yuan, Y. et al. Quantitative image analysis of cellular heterogeneity in breast tumors complements genomic profiling. *Sci Transl Med* **4**, 157ra143 (2012).