Cell Communication in Antiestrogen Resistance

Robert Clarke & Yue Wang

E-Mail: clarker@georgetown.edu & yuewang@vt.edu

VIRGINIA POLYTECHNIC INST & STATE UNIVERSITY

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

Approved for Public Release; Distribution Unlimited

More women die from the estrogen receptor positive (ER+) breast cancer subtype than from any other. The proportion of early ER+ recurrences (=5 years since diagnosis) approaches that for all triple-negative breast cancers alone. Late recurrences (>5 years after diagnosis), the result of dormancy, are most common in ER+ disease and can arise decades after the initial diagnosis. Since recurrent breast cancers have escaped the effects of endocrine therapies, and are lethal, we will study endocrine resistance (Tamoxifen; Fulvestrant). Our primary objective is to identify what drives breast cancer growth and determine how to stop it. We will learn about why some breast cancers are aggressive and others are indolent, and why/how some breast cancers lay dormant for years and then re-emerge.
TABLE OF CONTENTS

1. Introduction
2. Keywords
3. Accomplishments
4. Impact
5. Changes/Problems
6. Products
7. Participants & Other Collaborating Organizations
8. Special Reporting Requirements
9. Appendices
1. Introduction

~70% of newly diagnosed breast cancers are ER+ [1]. Many of these women die because metastatic ER+ disease becomes treatment resistant. Resistance is multiscale, i.e., evident at many levels, with genetic, cellular, and phenotypic features (including intratumor heterogeneity; ITH), all are molecularly manifested, and functionally realized, as networked changes in the transcriptome and proteome. We will take a systems biology approach to portray the proteome and transcriptome topology of treatment-induced adaptive remodeling of cell admixtures in vitro and in vivo. Overarching goals are to understand the principles of this remodeling and uncover the mechanisms that confer endocrine resistance in breast cancer, leading to new treatment strategies.

2. Keywords

Drug resistance, admixing, ecology, multiscale modeling
3. Accomplishments

A) Major goals (and related subtasks) of the project from approved SOW:

<table>
<thead>
<tr>
<th>SPECIFIC AIM 1 (specified in proposal)</th>
<th>Timeline</th>
<th>Site 1</th>
<th>Site 2</th>
<th>Percent complete</th>
<th>Date completed (if 100%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Major Task 1 (Aim 1a)</strong></td>
<td>Months</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Subtask 1:</strong> Determine prevalence of R and/or P cells in S populations</td>
<td>1-6</td>
<td>Dr. Clarke</td>
<td>Dr. Sengupta</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Subtask 2:</strong> Determine the effects of different S:R ratios on response to TAM and ICI in MCF-7, LCC, T47D, and ZR-75-1 S and R matched cell models <em>in vitro</em></td>
<td>1-18</td>
<td>Dr. Clarke</td>
<td>Dr. Sengupta</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Milestone(s) Achieved:** Identified effects of S:R ratio on responsiveness to TAM and ICI *in vitro* in multiple breast cancer cell models and identified optimal admix ratios for *in vivo* studies.

- Local IACUC approval (annual renewal required only – approval for studies already in place)
  - 1 Dr. Clarke

- Local IRB approval (add this award as an exemption for use of existing data – no new clinical data will be generated in this BT#2)
  - 1 Dr. Clarke

**Subtask 3:** Determine the effects of different S:R ratios on response to TAM and ICI in S:R matched cell models *in vivo* (models and admixes guided by the optimal* results in Major Task 1/Subtask 2)

It is difficult to provide direct numbers until the *in vitro* work is completed. A standard design for a single would include the following (n=15/group as in application):
- R cells alone ± ICI (15+15=30)
- S cells alone ± ICI (15+15=30)
- R+S cells at a single ration ± ICI (15+15=30)

Total = 90/experiment

We may do 3 such experiments over the 18-month period for 270 mice.

**Milestone(s) Achieved:** Identified effects of S:R ratio on responsiveness to TAM and ICI *in vivo*

<table>
<thead>
<tr>
<th><strong>Major Task 2 (Aim 1b)</strong></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subtask 1:</strong> Determine the role of GJIC in the ability of R to make S cells resistant to TAM and ICI in MCF-7, LCC, T47D and ZR-75-1 matched cell models <em>in vitro</em> (guided by the optimal experimental conditions from Aim 1a)</td>
<td>6-18</td>
<td>Dr. Clarke</td>
<td>Dr. Sengupta</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Subtask 2:</strong> Determine the role of microvesicles and protein secretion</td>
<td>6-18</td>
<td>Dr. Clarke</td>
<td>Dr. Sengupta</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(transwell) in the ability of R to make S cells resistant to TAM and ICI in MCF-7, LCC, T47D and ZR-75-1 matched cell models in vitro (informed by the optimal experimental conditions identified in Aim 1a)

**Milestone(s) Achieved:** Identified role GJIC, microvesicles and protein secretion (transwell) in the ability of R to make S cells resistant to drug and how this is affected by different S:R ratios. Identified conditions to allow design and execution of in vitro studies with guggulsterone and/or GW4869 (experiments will be done if supported by data and if time permits)

<table>
<thead>
<tr>
<th>Subtask 3: Collect and store materials (e.g., cell lysates) from optimal conditions for omics studies in Aim 2</th>
<th>Timeline</th>
<th>Site 1</th>
<th>Site 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dr. Clarke Dr. Sengupta</td>
<td>1-24</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SPECIFIC AIM 2 (specified in proposal)**

<table>
<thead>
<tr>
<th>Major Task 3 (Aim 2a)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subtask 1:</strong> Collect RNA and protein from the materials stored from Aim 1a (this will be done as the optimal experiments are identified above)</td>
</tr>
<tr>
<td><strong>Subtask 2:</strong> Perform array and proteome data collection, processing of raw data from Major Task 3/Subtask 1 (above), and initial analyses (e.g., CAM, kDDN)</td>
</tr>
</tbody>
</table>

**Milestone(s) Achieved:** Create initial signaling maps of what is communicated by R to S to confer resistance and how this is affected by different S:R ratios

**Major Task 4 (Aim 2b)**

| Subtask 1: Build initial mathematical models of cell population remodeling dynamics (in vitro and in vivo data) | 4-24 | Dr. Bansal |
| Subtask 2: Build final mathematical models of cell population remodeling dynamics (in vitro and in vivo data) | 24-36 | Dr. Bansal |

**Milestone(s) Achieved:** Identified how endocrine therapies and the starting ratios of S:R cells affects population responses to treatment

**Major Task 5 (Aim 2c)**

| Subtask 1: Use the data from Aims 1 and 2 to design and execute novel drug combination and scheduling studies in silico (mathematical modeling), e.g., ICI+DNMTi | 18-36 | Dr. Clarke Dr. Sengupta Dr. Bansal | Dr. Wang |
| Subtask 2: Design and execute novel drug combination and scheduling studies in vitro using the predictions in Major Task 5/Subtask 1 | 18-36 | Dr. Clarke Dr. Sengupta Dr. Bansal | Dr. Wang |

**Milestone(s) Achieved:** Identified novel optimized (activity vs. toxicity) combination regimens in vitro.

| Subtask 3: A small number of predictions from the in vitro modeling in Major Task 5/Subtask 2 will be tested in vivo (we anticipate completing ~5 such animal studies) | 18-36 | Dr. Clarke Dr. Sengupta Dr. Bansal | Dr. Wang |

**Milestone(s) Achieved:** Identified novel optimized (activity vs. toxicity) combination regimens in vitro.
It is difficult to provide direct numbers until the *in vitro* work is completed. A standard design for a single would include the following (n=15/group as in application):

- R cells alone + Vehicle (15)
- S cells alone + Vehicle (15)
- R cells alone + Drug A and + Drug B (15+15=30)
- S cells alone + Drug A and + Drug B (15+15=30)
- R+S cells at a single ratio with Vehicle, + Drug A and + Drug B (15+15+15=45)

Total = 135/experiment

We may do 4 such experiments over the funding period (n=540 maximum number mice).

**Milestone(s) Achieved:** Identified novel optimized (activity vs. toxicity) combination regimens *in vivo.*

<table>
<thead>
<tr>
<th>SPECIFIC AIM 3 (specified in proposal)</th>
<th>Timeline</th>
<th>Site 1</th>
<th>Site 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Major Task 6 (Aim 3a)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Subtask 1:</strong> Initial CAM and kDDN modeling of microarray data from human tumors (public and in-house datasets); data will be fed back to Aim 2 to increase clinical relevance</td>
<td>1-12</td>
<td>Dr. Wang</td>
<td>100% Site 2</td>
</tr>
<tr>
<td><strong>Subtask 2:</strong> Update models using outcomes from Aim 2 and study if candidate molecules from Aim 2 are associated with clinical outcome (univariate and multivariate)</td>
<td>12-36</td>
<td>Dr. Clarke, Dr. Sengupta, Dr. Bansal</td>
<td>10% Site 2</td>
</tr>
</tbody>
</table>

**Milestone(s) Achieved:** Identified clinically relevant molecules associated with ITH and endocrine resistance

| **Subtask 4:** A small number of predictions from the *in vitro* modeling in Major Task 5/Subtask 3 will be tested *in vivo* (~5 such experiments will be done) | 18-36    | Dr. Clarke, Dr. Sengupta, Dr. Bansal | 10% Site 2 |
| A small number of predictions from the *in vitro* modeling in Major Task 5/Subtask 2 will be tested *in vivo* (we anticipate completing ~5 such animal studies) |
| It is difficult to provide direct numbers until the *in vitro* work is completed. A standard design for a single would include the following (n=15/group as in application):
- R cells alone ± Drug A (15+15=30)
- S cells alone ± Drug A (15+15=30)
- R cells alone ± Drug B (15+15=30)
- S cells alone ± Drug B (15+15=30) |
R+S cells at a single ratio + Drug A + Drug B (15+15=30)
Total = 150/experiment
We may do 3-5 such experiments over the funding period (n=750 maximum number mice).

**Milestone(s) Achieved:** Identified novel therapeutic strategies for ER+ breast cancer to prevent, delay or reverse resistance, and do so within minimized toxicity. These insights could be used to design clinical trials to be done outside this research program.

**Major Task 7 (Aim 3b)**

<table>
<thead>
<tr>
<th>Subtask 1: Test candidate molecules from the model predictions in Aims 2 and 3a. For example, as described in the narrative section, genes upregulated in resistant cells relative to sensitive cells will be overexpressed (cDNA; regulatable and/or constitutive promoters) in sensitive cells and knocked down in resistant (RNAi) if their mRNA or protein is still present in sensitive cells. The gene will be knocked out (CRISPR) in resistant cells if the gene is known to be lost or expression is undetectable in sensitive cells. The reverse experiments will be done where a gene is down regulated or lost in resistant cells relative to its expression/presence in sensitive cells.</th>
<th>12-36 months</th>
<th>Dr. Clarke Dr. Sengupta Dr. Bansal</th>
<th>Dr. Wang</th>
</tr>
</thead>
</table>

**Milestone(s) Achieved:** Identified mechanistically relevant molecules associated with ITH and endocrine resistance

**B) What was accomplished under these goals?**

Intratumor heterogeneity serves as both, an underexploited information source in characterizing complex tumor phenotypes, and a major confounding factor in studying individual subclones. Having analytic tools to define the molecular landscape of intratumor heterogeneity, and to determine how subclones are remodeled with phenotypic transitions, will be essential in the proposed research. We develop a fully unsupervised deconvolution method to dissect complex tumors into molecularly distinctive subclones based on bulk expression profiles. We implement an R package, deconvolution by Convex Analysis of Mixtures (debCAM) that can automatically detect subclone-specific markers, determine the number of constituent subclones, calculate subclone proportions in individual samples, and estimate subclone-specific expression profiles, Fig. 1A. We demonstrate the performance and biomedical utility of debCAM on gene expression, methylation, proteomics, and imaging data. With enhanced data preprocessing and prior knowledge incorporation, debCAM software tool will allow biologists to perform a more comprehensive and unbiased characterization of intratumor heterogeneity and remodeling in many biomedical contexts (e.g., cell-cell communications).
We have tested debCAM on biologically mixed gene expression profiles involving both cancer subclone and immune cell subtypes (GSE64385). The 12 samples represent the mixtures of five immune cell subtypes and one cancer cell line in various known proportions. The optimal scatter simplex blindly identified by debCAM is given in **Fig. 1B**, showing six distinctive vertices (**Fig. 1B**). To validate the accuracy of subtype-specific markers blindly detected by debCAM, we superimpose the color-coded known markers in the scatter simplex showing close proximity to the corresponding vertices (**Fig. 1C**, the cancer cell line specific markers are unavailable for comparison). More convincingly, the sample-wise subtype proportions estimated by debCAM match the ground truth almost perfectly, with \( r = 0.975–0.996 \).

We further performed radiogenomics feature based deconvolution of intratumor heterogeneity on breast cancer. The multiscale data are generated from TCGA cohort (n=87), covering DCE-MRI (TCGA-BRCA) and gene expressions (TCGA). The MRI tumour area was decomposed by debCAM-CM, resulting in three microvasculature compartments that represent plasma input, fast-flow kinetics and slow-flow kinetics. The prognostic MRI features were associated with the gene expression modules of cancer-related pathways. Furthermore, a multigene signature for each prognostic imaging feature was established (p=0.035), and the prognostic value for overall survival was confirmed in an additional cohort from TCGA (p=0.027).
Regulatory rewiring of molecular networks represents an important information source in characterizing complex diseases with phenotypic transitions. We have developed one of the first few data-driven and knowledge-fused differential network analysis methods, namely, Differential Dependency Network (DDN), to detect statistically significant coupling or uncoupling between pairs of molecular entities that are dependent on the specific phenotypes or experimental conditions. We have conducted some pilot DDN analyses on the multi-omics data on ovarian cancer. The DDN analysis identifies in a complex and often unknown overall molecular circuitry a network of differentially connected molecular entities. The subnetwork of hub molecular entities is used to assist in the inference of potential key pathways, and to generate novel hypotheses for further validation and investigations.

For selected pathways, DDN analysis on MAPK signaling pathway also highlighted AKT gene as one of the hub genes pivotal in the network rewiring events. AKT is modulated by PTEN mutation which is one of the key definers of HRD status. DDN method is used as an integrated analysis tool to detect rewiring events in KSR between HRD and non-HRD groups on the LC-MS/MS quantitated expression of kinases and HuProt array-based activity quantitation of substrates. Two kinases PTK2 and PTK2B linked with significant rewiring events were found dysregulated in ovarian cancer in cell line models.

C) What opportunities for training and professional development has the project provided?
Two doctoral graduate research assistants are supported in part by this project. In return, they have also contributed to developing data analytics tools and pipelines.

D) How were the results disseminated to communities of interest?
debCAM: a Bioconductor R package for fully unsupervised deconvolution of intratumor heterogeneity, has been disseminated. Availability and implementation: [http://bioconductor.org/packages/debCAM](http://bioconductor.org/packages/debCAM)

DDN availability and Implementation: Open-source DDN Java code with compiled package are freely available for download at [http://apps.cytoscape.org/apps/kddn](http://apps.cytoscape.org/apps/kddn)

E) What do you plan to do during the next reporting period to accomplish the goals?
We will apply both debCAM and DDN tools to analyze the data generated in-house at Georgetown University Medical Center site. We will jointly interpret the results and develop sensible hypotheses. We will revise the analytics experimental designs toward more focused testing on the hypotheses. We will further improve the data analytics tools by incorporating the feedback from GUMC team.

4. Impact
Describe distinctive contributions, major accomplishments, innovations, successes, or any change in practice or behavior that has come about as a result of the project relative to:

What was the impact on the development of the principal discipline(s) of the project?
The debCAM R package implements and tests the latest functionalities of the debCAM algorithm pipeline in the literature. We have demonstrated real biomedical utilities of debCAM tool on gene expression, proteomics, methylation, and imaging data. These applications have led to novel findings and hypotheses. The present software also enhances data preprocessing, accelerates robust scatter simplex identification, and integrates supervising information such as known markers.

What was the impact on other disciplines?
Both debCAM and DDN tools can have broad applications in data sciences, engineering, and geosciences.

What was the impact on technology transfer?
Nothing to Report.

What was the impact on society beyond science and technology?
Nothing to Report.

Describe how results from the project made an impact, or are likely to make an impact, beyond the bounds of science, engineering, and the academic world on areas such as:

5. Changes/Problems
Nothing to Report.

Changes in approach and reasons for change
N/A.

Actual or anticipated problems or delays and actions or plans to resolve them
N/A.

Changes that had a significant impact on expenditures
N/A.

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

Significant changes in use or care of human subjects

Significant changes in use or care of vertebrate animals

Significant changes in use of biohazards and/or select agents

6. Products
Journal publications.
Robert Clarke, John J Tyson, Ming Tan, William T Baumann, Lu Jin, Jianhua Xuan, and Yue Wang, “Systems biology: perspectives on multiscale modeling in research on endocrine-related cancers,” *Endocrine-Related Cancer*, 2019. (accepted)

Books or other non-periodical, one-time publications.

Other publications, conference papers, and presentations.
N/A.

Website(s) or other Internet site(s)
N/A.

Technologies or techniques
N/A.

Inventions, patent applications, and/or licenses
N/A.

Other Products
Software: debCAM, DDN.

7. Participants & Other Collaborating Organizations

Name: Yue Wang
*Project Role*: Principal Investigator (Partnering PI)
*Research Identifier (ORCID ID)*:
*Nearest person months worked*: 1.0
*Contribution to project*: Designed data analytics methodology

Name: Lulu Chen
*Project Role*: Graduate Research Assistant
*Nearest person months worked*: 6.0
*Contribution to project*: Developed data deconvolution tool debCAM

Name: Minta Lu
*Project Role*: Graduate Research Assistant
*Nearest person months worked*: 6.0
*Contribution to project*: Developed integrated pathway prioritization tool IPP

Has there been a change in the active other support of the PI or senior personnel since the last reporting period?
Nothing to report.

What other organizations were involved as partners?
Nothing to report (aside from Partnering PI institution, Virginia Tech).

8. Special Reporting Requirements
The Initiating PI (Robert Clarke, PhD) has submitted an independent annual report for this period.

9. Appendices