Single-cell identity and state coupled with genome-wide SNV and CNV in primary breast cancer with ResolveOME™ profiling

J.S. Zawistowski[1]; I. Salas-Gonzalez[1], T.V. Morozova[1], T. Tate[1], K.A. Kennedy[1], D.M. Arvapalli[1], S.D. Velivela[1], J.G. Blackinton[1], J.R. Marks[2], E.S. Hwang[2], G.L. Harton[1], V.J. Weigman[1], J.A.A. West[1]

[1]BioSkryb Genomics, Durham, NC, USA [2]Department of Surgery, Duke University Medical Center, Durham, NC, USA

Abstract

The molecular events governing the transition from ductal carcinoma in situ (DCIS) to invasive breast cancer are still being elucidated, whereby precise definition of these events has the potential provide a therapeutic window of intervention. To simultaneously expose both genomic and transcriptomic underpinnings in primary breast cancer samples and to ascertain intratumoral heterogeneity we utilized ResolveOME™ to profile single cells from tumor biopsies of DCIS/invasive ductal carcinoma (IDC). While earlier single-cell methods have importantly unified assessment of copy number variation (CNV) and transcriptomics, they do not yield complete and uniform genome-wide coverage for single nucleotide-level data, made possible with ResolveOME™'s employment of primary template-directed amplification. As input into ResolveOME™, we stratified singulated mastectomy samples by epithelial cell adhesion molecule (EpCAM) surface marker expression with FACS. The genomic arm of ResolveOME™ followed by analysis with BaseJumper software cataloged genome-wide single nucleotide variation (SNV) in 24 single cells expressing either high or low levels of EpCAM, including the identification of oncogenic *PIK3CA* N345K, while identifying cooccurring DCIS/IDC prototypical chromosomal loss of 11q, 13q and 16q/17p harboring tumor suppressor loci. Concurrently, the transcriptomic arm of ResolveOME™ enabled the calling of cell identity with the Human Cell Atlas transcriptional database, revealing monocytic and fibroblastic infiltration in the biopsy samples in addition to the expected cells of epithelial identity. The coupling of SNV data and transcriptome data critically unveiled phenotypic cell state heterogeneity, whereby an epithelial cell with both *PIK3CA* N345K and associated chromosomal losses manifested with a stemness/fibroblastic gene expression signature characteristic of the EpCAM low clade of patient cells. We are currently expanding ResolveOME™ profiling to additional DCIS/IDC patient tumors to comprehensively define driver and regulatory SNV while simultaneously distinguishing between infiltration of non-epithelial cell types from instances of epithelial morphing of physiological cell state within a biopsy. Furthermore, we are defining cell lineage at both the CNV and SNV level as additional single cells are sequenced for each patient sample. These data collectively highlight the molecular complexity and heterogeneity even among a small number of biopsied cells, and underscore the criticality of the interplay of DNA/RNA tiers of information when positing oncogenic mechanisms.

Introduction & Background

ResolveDNA®
- Single nucleotide variation (SNV): coding + non-coding
- Copy number variation (CNV)
- Small indels
- Large deletion and amplification detection
- Tumor mutation burden

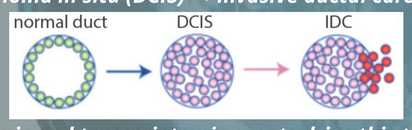
ResolveOME™
- Regulatory variant discovery
- Cell identity and cell state coupled to genomic lesion
- Neoantigen discovery
- Ploidy : gene expression correlation
- T cell / B cell receptor

scRNA-seq
- Differential gene expression
- Cell type identification
- RNA editing
- Non-coding RNA
- Transcript isoform utilization

Unique capabilities of *unified* genomic and transcriptomic data not possible in isolation

For understanding cancer heterogeneity and clonal evolution, ResolveOME™ (Ref. 1) uniquely enables interrogation of crucial DNA:RNA relationships (purple) in addition to the mainstay outputs of single cell RNA sequencing (red) or whole-genome sequencing through ResolveDNA® (blue).

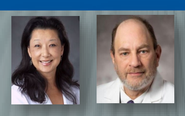
Ductal carcinoma in situ (DCIS) → invasive ductal carcinoma (IDC)



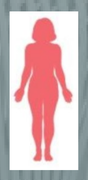
What genomic and transcriptomic events drive this transition?

We initiated a collaboration with Duke University Medical Center to utilize ResolveOME™ to expose DNA:RNA relationships, as well as to define the variability of those relationships between single cells, that may be contributing specifically to the *transition* of DCIS to invasive cancer.

Duke University School of Medicine



E. Shelley Hwang, MD, MPH (Surgery)
Jeffrey Marks, PhD (Surgery, Pathology)



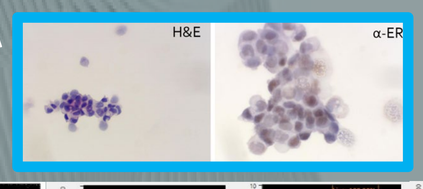
IRB: PRO00034242 "Biologic Characterization of the Breast Cancer Tumor Microenvironment"

Methodology

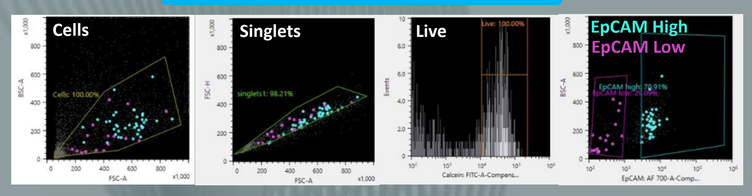
Scope: Perform ResolveOME™ unified genomic/transcriptomic workflow on DCIS/IDC single cells from a mastectomy specimen

Goal: Uncover genomic lesions and place in the context of cell identity and phenotypic cell state from transcriptomic data of the same cell

A



B



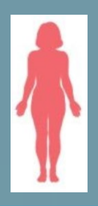
(A) **H&E staining and estrogen receptor immunostaining of DCIS/IDC singulated cells**. Prior to FACS, a slide smear of the patient sample was performed to assess morphology by H&E and estrogen receptor positivity (brown staining).

(B) **FACS scheme to enrich for ductal epithelial cells.** Shown is the gating hierarchy employed to enrich for viable single cells and to stratify those cells by EpCAM protein expression levels. Both EpCAM High and EpCAM low single cells were dispensed into PCR plate wells for ResolveOME™ reactions. The transcriptomic arm of ResolveOME in this study allowed cellular identity signatures to then be correlated with EpCAM High or Low status determined by FACS.

(C) **Manifest for the patient sample utilized in this study.** Highlighted are aspects pertinent to the current and previous DCIS/IDC presentations of this patient.

C

Patient manifest

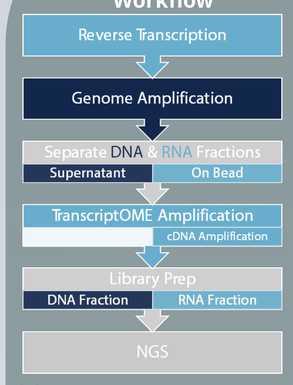


Current presentation
- Age 61
- African American
- 7 cm DCIS and 1.2 cm invasive cancer
- ER+ PR+ HER2 –
- DCIS grade II, invasive grade I
- Right breast mastectomy: normal and tumor digested to single cells
- H&E FFPE tissue for genomic DNA isolation

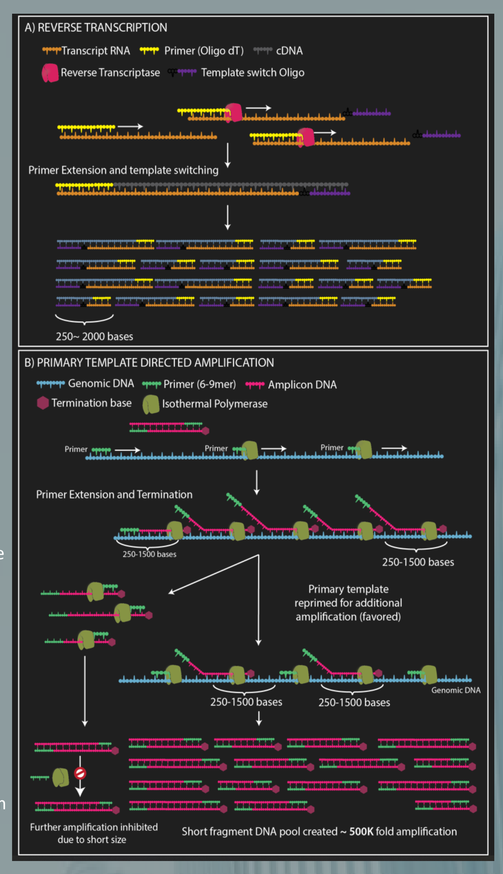
Previous presentation
- DCIS in left breast 14 years prior to current presentation
- radiotherapy and Arimidex
- Discontinued use of Letrozole

Methodology

A **ResolveOME™ Workflow**



B **ResolveOME™ RNA and DNA reactions**



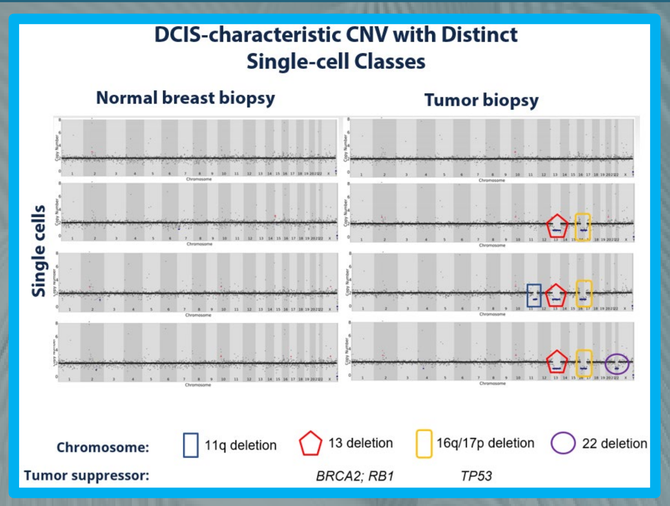
(A) **Generalized stages of ResolveOME™ joint transcriptomic and genomic single-cell workflow.** Following single cell isolation, reverse transcription occurs first, with subsequent nuclear lysis and genome amplification in the same tube. The RNA fraction is then separated from the DNA fraction by affinity purification, whereby two distinct library sets are created and sequenced together.

(B) **Molecular biology schematic of ResolveOME™.** Top: Template switching-based cDNA synthesis yielding full-length transcript information. The resultant first-strand cDNA is conjugated with an affinity tag to allow subsequent separation from the genomic amplification product.

Bottom: ResolveDNA® genomic amplification, driven by Primary Template-directed Amplification (Ref. 2), results in unprecedented genomic coverage, uniformity and allelic balance for accurate variant calling—through the fundamental principle of amplicon size attenuation and prevention of re-amplification of the amplicons.

Results

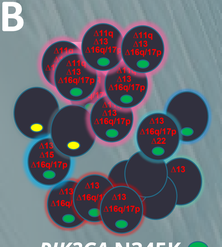
A



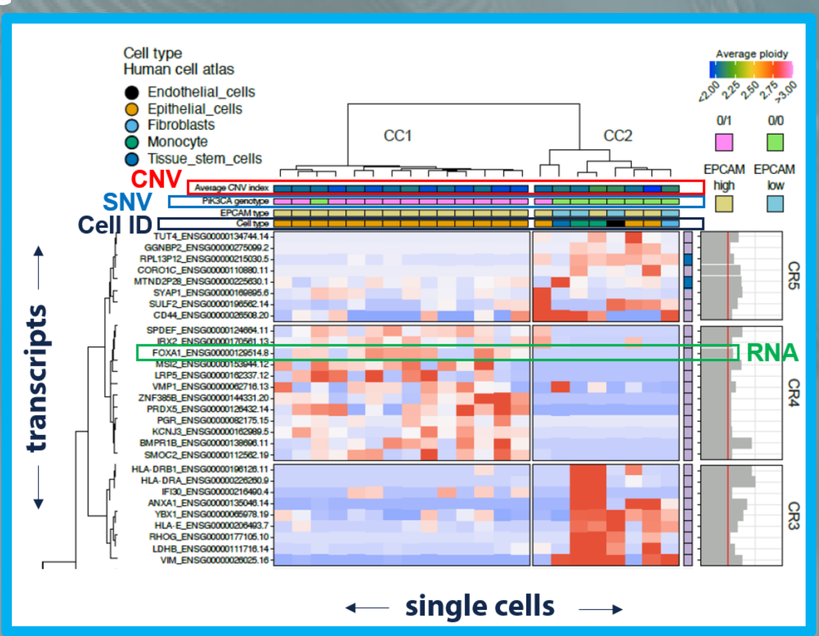
(A) **Copy number alterations detected in DCIS/IDC patient cells.** We identified distinct classes of chromosomal loss in single cells from tumor but not from normal ipsilateral biopsy control single cells. These specific chromosomal loci harbor prototypical tumor suppressor genes and have been observed to be frequently lost in various stages of DCIS advancement (Ref. 3).

(B) **Genotypic class diversity and CNV/SNV interplay in single DCIS/IDC cells.** Whole genome sequencing of single tumor cells revealed mutually exclusive oncogenic missense mutations N345K (green oval) and H1047R (yellow oval). Intriguingly, the chromosomal losses defined in (A) were only detected in the cells harboring *PIK3CA* N345K but not H1047R, suggesting distinct and heterogenous oncogenic mechanisms occurring in this individual tumor—*PIK3CA* H1047R may be sufficient to drive aberrant growth whereby N345K may require the synergism of tumor suppressor loss.

B



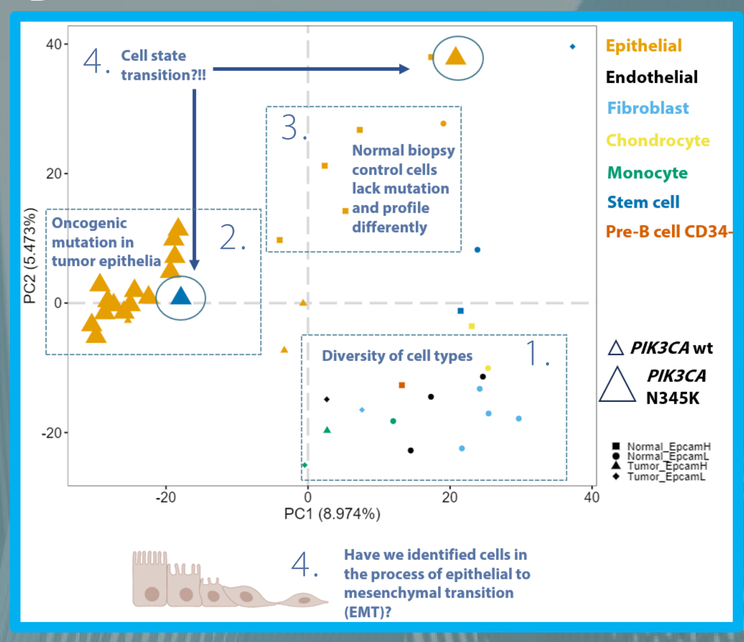
C



(C) Differential gene expression of DCIS/IDC single cells layered with additional omic tiers of ResolveOME™. Columns represent single cells and rows represent transcripts. Two predominant gene expression clades emerge (CC1 and CC2) from the clustering, which coincide with EpCAM status (tan or light blue) of the single cells. In addition to RNA-level information, such as *FOXA1* upregulation in EpCAM High cells (green box), ResolveOME™ enables the overlaying of CNV information (red box, A), *PIK3CA* N345K mutation status (blue box) as well as cell identification (dark blue box). Note the diversity of cell types from the tumor biopsy within the EpCAM Low-predominant expression clade.

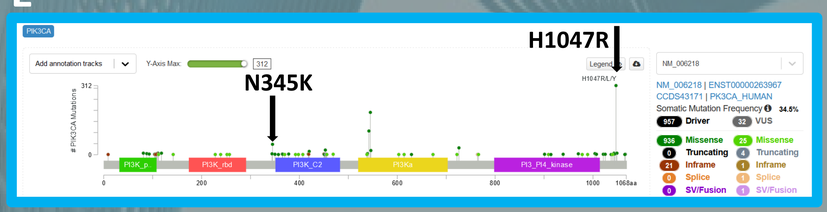
Results

D



(D) Principal component analysis of DCIS/IDC single cell gene expression overlayed with *PIK3CA* mutation status. Tumor biopsy cells or ipsilateral control biopsy single cells FACS-binned as either EpCAM High or Low were subjected to ResolveOME™ combined transcriptional and genomic profiling, where each point represents a single cell. SingleR and the Human Cell Atlas database was utilized in addition to proprietary algorithmic modifications for cell type identification (icon colors). The status of the oncogenic *PIK3CA* N345K mutation is overlayed (icon size). The EpCAM Low bin of cells exemplified the cell type diversity present in the biopsy, and this diversity of cell types expectedly lacked oncogenic *PIK3CA* changes (1). The EpCAM High bin of tumor biopsy cells profile epithelial and harbor *PIK3CA* N345K (2) while normal biopsy control cells separate from tumor cells by gene expression profile and lack *PIK3CA* N345K (3). Intriguingly, a single cell with stemness identity harbors *PIK3CA* N345K and a single cell with epithelial identity with oncogenic *PIK3CA* N345K clusters more closely with the expression profiles of normal biopsy cells—both indicative of phenotypic cell state plasticity (Ref. 4) potentially in the context of epithelial to mesenchymal transition (4).

E



(E) N345K and H1047R DCIS/IDC mutations within *PIK3CA* mutation spectrum. Shown is a cBioPortal representation of the secondary structure of PIK3CA and the location and frequency of the N345K and H1047R missense mutations in a cohort of 2473 invasive and metastatic breast cancer samples (MSK, Cancer Discovery 2022). H1047R is the most frequent *PIK3CA* missense mutation in breast cancer and N345K functionally disrupts the interaction between the p85 regulatory subunit and the p110α catalytic subunit of PI3 Kinase (Ref. 5).

(F) Extension of the current study to a larger patient cohort. The insights presented here are currently being extended to additional patient samples (IRB: PRO00034242 "Biologic Characterization of the Breast Cancer Tumor Microenvironment"), with the goals of defining inter-patient and intra-patient single-cell genomic and transcriptomic heterogeneity and defining oncogenic driver mechanisms informed by unified RNA:DNA information.

F **Extension of study: additional patient samples**

Sample number	Tumor / Normal	Dissociated cell number	Typing
1	Tumor	2.4 E5	IDC (ER+/PR+/HER2-)
2	Tumor	1.1 E5	IDC (ER+/PR+/HER2-)
3	Tumor	1.5 E5	IDC (ER+/PR+/HER2-)
4	Tumor	2.2 E5	ER+/PR+/HER2-
5	Tumor	3.2 E5	IDC
6	Tumor	7.7 E4	IDC
7	Normal	2.7 E4	N/A
8	Tumor	5.3 E5	IDC; DCIS present, > 8 cm
9	Normal	4.3 E4	N/A
10	Tumor	1.8 E5	IDC
11	Tumor	2.8 E5	IDC, DCIS present, non-extensive
12	Tumor	1.2 E6	TNBC

Summary

- ResolveOME™ provides, in single cells, the union of whole genome and full-length transcript RNAseq data, yielding oncogenic insights not possible with genomic or transcriptomic data in isolation
- In a primary DCIS sample, ResolveOME™ unveiled distinct and single-cell heterogenous classes of copy number loss which harbored tumor suppressor loci and associated exclusively with a specific *PIK3CA* missense mutation, N345K
- Through ResolveOME™ DNA/RNA unification, *PIK3CA* missense mutations predominated in tumor but not normal biopsy control single cells and biopsy cell type diversity was exposed in a non-EpCAM enriched population
- Transcriptional profiles coupled to genotypic *PIK3CA* mutation status were indicative of phenotypic cell state conversion or epithelial-mesenchymal transition in some single cells

References

1) Zawistowski, J.S. *et al.* Unifying genomics and transcriptomics in single cells with ResolveOME amplification chemistry to illuminate oncogenic and drug resistance mechanisms. bioRxiv preprint (2022) doi: https://doi.org/10.1101/2022.04.29.489440
2) Gonzalez-Pena, V. *et al.* Accurate genomic variant detection in single cells with primary template-directed amplification. *Proc Natl Acad Sci U S A* **118**, e2024176118 (2021).
3) Gorringe, K. L. *et al.* Copy number analysis of ductal carcinoma in situ with and without recurrence. *Mod Pathol* 28, 1174–1184 (2015).
4) Moris, N., Pina, C. & Arias, A. M. Transition states and cell fate decisions in epigenetic landscapes. *Nat Rev Genet* **17**, 693–703 (2016).
5) Gymnopoulos, M., Elsliger, M.-A. & Vogt, P. K. Rare cancer-specific mutations in PIK3CA show gain of function. *Proc Natl Acad Sci U S A* **104**, 5569–5574 (2007).

Acknowledgments

We thank the patients for their generous donation of tissue to enable the insights obtained in this study