The ResolveOME platform: integrated whole genome and whole transcriptome profiling from a single cell to unlock drug resistance mechanisms

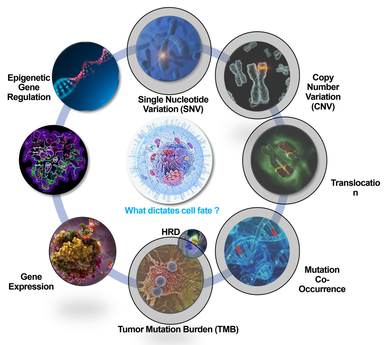
Tatiana V. Morozova, Jeff G. Blackinton, Isai Salas-González, Viren R. Amin, Victor J. Weigman, Jon S. Zawistowski, Gary L. Harton and Jay A. A. West
BioSkryb Genomics, Inc., Durham, NC.

Abstract

To elucidate drug resistance mechanisms in cancer and to deconvolve the molecular basis of variability in treatment responses, it is paramount to apply approaches that integrate multi-omics data. Existing methodologies aiming to couple genomic and transcriptomic information are limited in that the genomic information is based on a targeted approach, providing insight into only a small fraction of the genome.

We have developed a novel, true multi-omic platform, ResolveOME, that combines whole genome amplification with transcriptome profiling from the same single cell. The platform unifies template-switching single-cell RNAseq chemistry with modified ResolveDNA™ whole genome amplification (WGA) technology based on Primary Template-directed Amplification (PTA). The workflow then relies on affinity purification of first-strand cDNA and subsequent separation of the RNA/DNA fractions to allow for independent library preparation of the fractions followed by ResolveDNA™ library preparation and sequencing.

To demonstrate the validity of this platform we generated GM12878 cell data using the ResolveOME chemistry and compared it to data from the ResolveDNA™ WGA kit and to bulk RNA sequencing. Products from the DNA and RNA arms of the protocol demonstrated comparable product sizes and consistent yields to the bulk RNAseq and standard ResolveDNA™ chemistry. Low-pass sequencing of the DNA arm of the ResolveOME workflow showed robust performance, including high genomic coverage and high library diversity. Gene expression analysis of the RNA arm of ResolveOME revealed the ability to detect ~ 10K expressed genes in bulk RNA samples and ~ 8K expressed genes in the single cells, concordant with published data for GM12878 cells.



Our next focus is to extend the ResolveOME platform to established cancer drug resistant models in cell lines to identify SNVs, differentially expressed transcripts, and CNV contributing to drug resistance, as well as to primary patient samples.

We have demonstrated here that the ResolveOME platform provides unprecedented, and non-targeted multi-omics data from individual single cells. The utility of ResolveOME extends to all applications which require both whole genome and transcriptome data from a single cell and/or samples with limited biological material.

ResolveOME Workflow

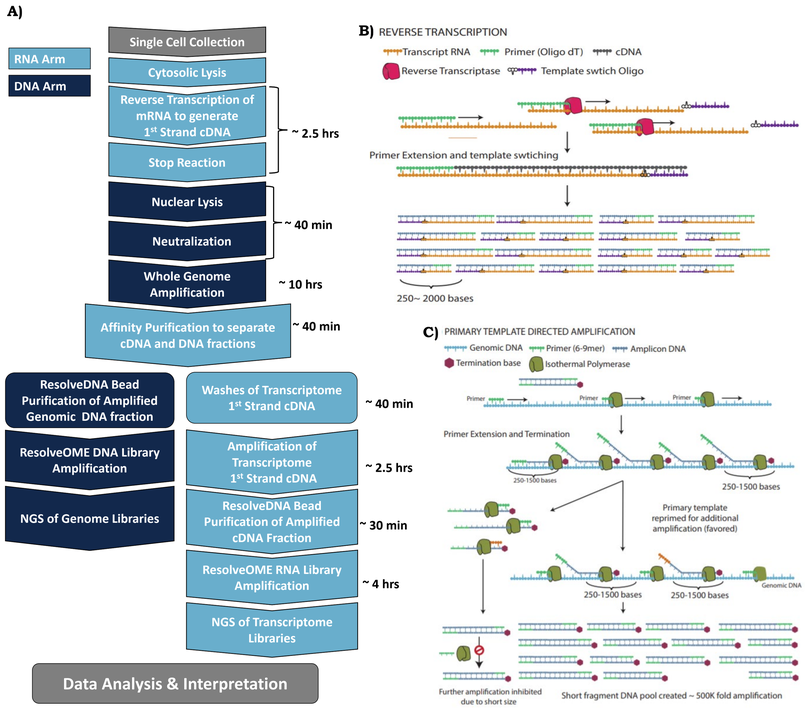
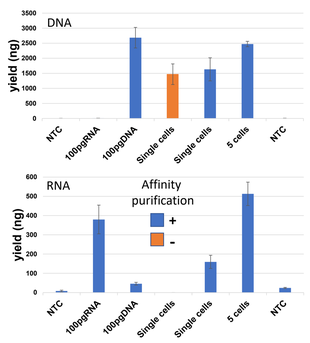


Figure 1: ResolveOME workflow providing simultaneous RNA and DNA information from a single cell. Schematic diagram of the ResolveOME Workflow (**Panel A**). Isolated single cells first undergo 1st strand cDNA synthesis of cytosolic mRNA molecules using a template switching-based reverse transcription process (**Panel B**) on the entire polyA+ transcriptome. The synthesized cDNA molecule of each transcript then remains in the sample during the nuclear lysis and subsequent steps used to amplify the genomes of each cell (**Panel A**). ResolveOME is built upon the foundation of Primary Template-directed Amplification (PTA, **Panel C**)[1]. Similar to the ResolveDNA™ product, the genome of each cell is then denatured in preparation for random priming based genome amplification (**Panel A**). PTA utilizes isothermal amplification and proprietary termination chemistry to restrict amplicon size, preferentially redirecting random primers to the primary template. Affinity purification of cDNA parses the combined pool of amplified genomic DNA and the complete diversity of polyA+ derived cDNAs. After enrichment of the transcriptome, distinct libraries from both the transcriptome and genome fraction are prepared using ResolveOME DNA or RNA library kits for downstream next generation sequencing and analysis with BaseJumper software.

ResolveOME comparative genomic and transcriptomic single-cell performance

Figure 2. ResolveOME amplification yield from individual NA12878 cells. We obtained approximately 1-1.5 µg of DNA amplification product from single cells (top panel) and approximately 100-200 ng of cDNA product representing the single cell transcriptome (bottom panel). The presence of affinity purification does not significantly affect the PTA yield.



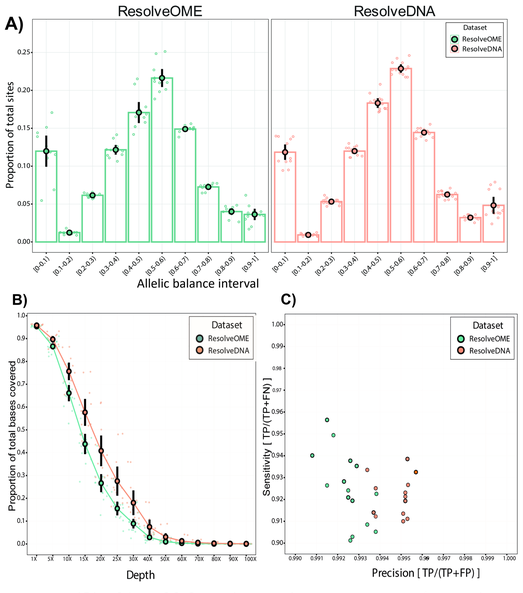
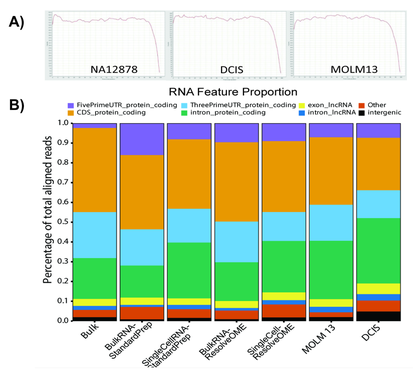
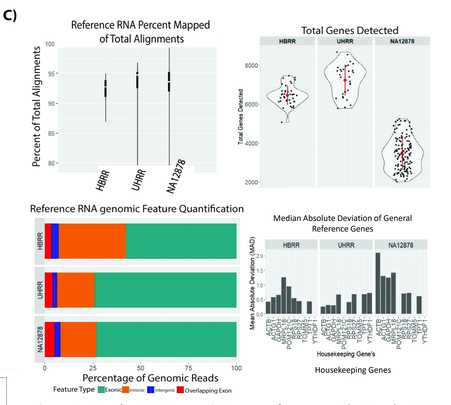


Figure 3: Performance of the genomic arm of ResolveOME relative to ResolveDNA. We performed high-depth sequencing to estimate the allelic balance, genomic coverage, and single nucleotide variant calling precision and sensitivity across 10 NA12878 single cell replicates from ResolveDNA and ResolveOME workflows. ResolveOME (orange) allelic balance ~ 86% was comparable to the ~88% obtained for ResolveDNA (green, **Panel A**). Here the allelic balance is shown in deciles of observed allele frequency (AF) across known heterozygous positions. Allelic dropouts are called when AF is < 0.1 or > 0.9. Cumulative genomic coverage at a range of depths did not significantly differ (**Panel B**) between the workflows. Each dot represents a cell replicate within a dataset and error plots denote the variability of coverage at a given depth. The allelic balance and coverage obtained from the ResolveOME workflow afforded the ability to call genome-wide SNVs with accuracy and confidence: a SNV calling sensitivity range of 0.90-0.95 and with precision >0.99, similar in performance to ResolveDNA data (**Panel C**).





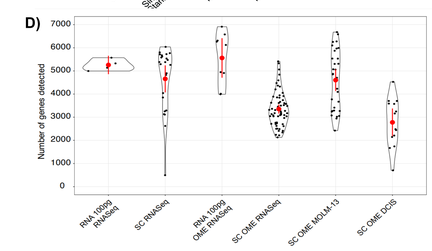


Figure 4: ResolveOME transcriptome performance. The ResolveOME transcriptome workflow captures full-transcript (vs. end-counting) information to enable the detection of isoforms and gene fusions. This chemistry enables even gene body coverage across the summation of all transcripts in NA12878, DCIS and MOLM-13 datasets (**Panel A**). RNA feature quantifications are shown in **Panel B**, highlighting the increased ability to identify 5' and intronic features relative to end counting and averaged bulk datasets. Features and proportions were derived from Qualimap summarizations of our transcriptome definition file. Performance patterns of common metrics with well characterized Human Brain Reference RNA (HBRR), Universal Human Reference RNA (UHRR) standards and the NA12878 cell line are shown in **Panel C**.

Clockwise from the top left, the distribution of reads assigned to transcriptome, coding region features, unique genes detected and the median absolute deviation (MAD) of common housekeeping genes. We also examined the dynamic range, computed various markers of DNA contamination, sample degradation, and/or bias as a percentage of exonic and intergenic mapping (less than 5 %) as characteristics of the ResolveOME RNA fraction. **Panel D** shows expressed genes detected with ResolveOME chemistry compared to RNAseq datasets employing the same chemistry yet in the absence of the overlayed PTA reaction or affinity purification.

ResolveOME in acute myeloid leukemia (AML) model of quizartinib drug resistance and in primary ductal carcinoma in situ/invasive ductal carcinoma (DCIS/IDC) single cells

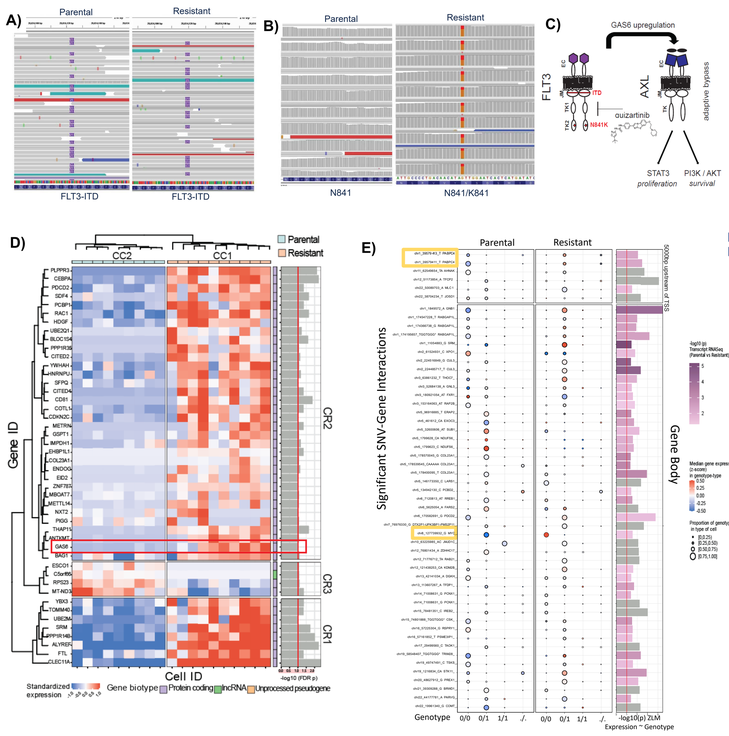
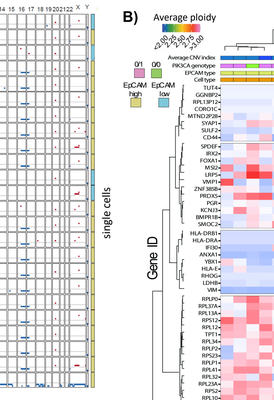


Figure 5: Genotypic and transcriptomic analysis in MOLM-13 cells. For generation of the quizartinib-resistant MOLM-13 line, cells were continually treated with 2 nM quizartinib or DMSO vehicle control for matched parental control line and drug replenished at each subculturing until emergence of resistant clones at 5 weeks duration in culture. Following Calcein AM, propidium iodide and DAPI staining, singlet and live cell (DAPI/PI negative, top 70% Calcein-AM positive) gating was established and ~2.0E6 MOLM-13 single cells were sorted for ResolveOME workflow. MOLM-13 cells are sensitive to the FLT3 inhibitor quizartinib, a selective type II kinase inhibitor targeting *FLT3,* due to the presence of an internal tandem duplication (ITD) mutation. ResolveOME verified the presence of the ITD mutation in both parental (P) and quizartinib-resistant (R) single cells (**Panel A**). In addition, a secondary *FLT3* mutation, N841K, was detected in all quizartinib resistant cells (**Panel B**). N841K has been also found in AML patient samples and resides in the activation loop of FLT3. Thus, N841K is likely contributing to quizartinib resistance in this model likely by preventing efficiency of drug binding (**Panel C**). Differential expression analysis revealed upregulation of *Gas6* (**Panel D, red box**), a ligand for the receptor tyrosine kinase AXL. The AXL pathway, specifically through downstream STAT3 cell proliferation and PI3K/AKT survival signaling (**Panel C**), has been shown to be a bypass pathway for FLT3 inhibition. Proximal SNVs with significant prevalence biases between "P" and "R" MOLM-13 cells are shown for differentially expressed transcripts (**Panel E**). These variants represent candidate regulatory SNVs and highlight genomic plasticity. A larger circle indicates a higher prevalence of the given genotype; color indicates the directionality of expression change. Yellow color boxes highlights *MYC*, as well as a putative promoter mutation in the mRNA stability factor *PABPC4*.

Figure 6. Patient information and FACS strategy. Tissue for DCIS/IDC studies was obtained in accordance with the Duke University Medical Center IRB for the clinical trial (left panel). Singulated cells (~4.2E5) derived from mastectomy tissue were thawed at 37°C and incubated with 2 µg/ml anti-human CD326 (EpCAM) antibody conjugated with AlexaFluor 700 at 4C in the dark for 1h to enrich for ductal epithelial cells by EpCAM epitope enrichment. Singlet (FSC-A / FSC-H, BSC-A / BSC-W) and live-cell (top 70% Calcein-AM positive) gating followed by gating on EpCAM high and EpCAM low signal (right panel).

- ❑ Age 61, Female
- ❑ 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 bulk gDNA isolation

Singulated patient cells



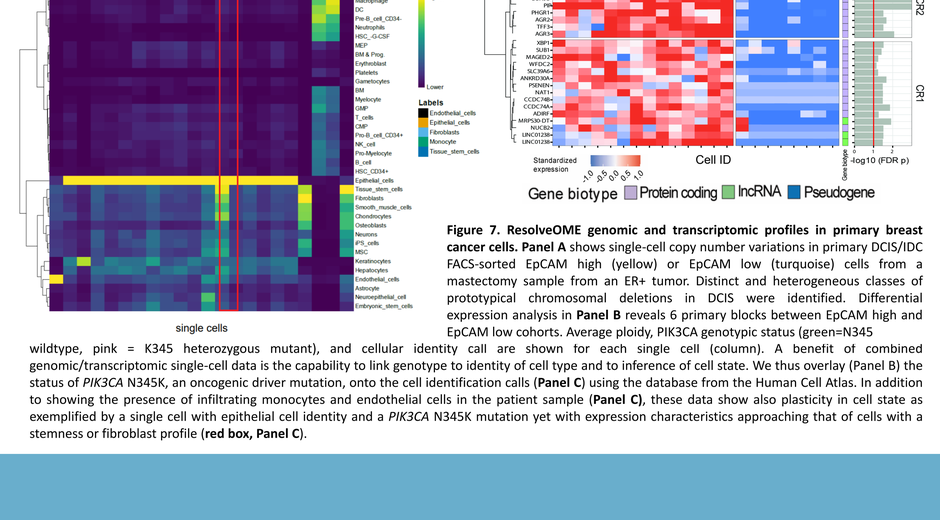


Figure 7. ResolveOME genomic and transcriptomic profiles in primary breast cancer cells. Panel A shows single-cell copy number variations in primary DCIS/IDC FACS-sorted EpCAM high (yellow) or EpCAM low (turquoise) cells from a mastectomy sample from an ER+ tumor. Distinct and heterogeneous classes of prototypical chromosomal deletions in DCIS were identified. Differential expression analysis in **Panel B** reveals 6 primary blocks between EpCAM high and EpCAM low cohorts. Average ploidy, PIK3CA genotypic status (green=N345 wildtype, pink = K345 heterozygous mutant), and cellular identity call are shown for each single cell (column). A benefit of combined genomic/transcriptomic single-cell data is the capability to link genotype to identity of cell type and to inference of cell state. We thus overlay (Panel B) the status of *PIK3CA* N345K, an oncogenic driver mutation, onto the cell identification calls (**Panel C**) using the database from the Human Cell Atlas. In addition to showing the presence of infiltrating monocytes and endothelial cells in the patient sample (**Panel C**), these data show also plasticity in cell state as exemplified by a single cell with epithelial cell identity and a *PIK3CA* N345K mutation yet with expression characteristics approaching that of cells with a stemness or fibroblast profile (**red box, Panel C**).

SUMMARY

- The ResolveOME multi-omic solution empowers simultaneous evaluation of the genome and transcriptome in single cells to determine cell identity and cell state, and to concurrently determine transcriptomic consequence of genomic variation/mutation.
- The attributes of Primary Template-directed Amplification (PTA)[1] enable accurate genome-wide calling of single nucleotide variation, not possible with existing methods for the unification of DNA+RNA information.
- In a model of acute myeloid leukemia drug resistance, ResolveOME exposed both DNA and RNA modes of resistance: a secondary mutation in *FLT3* encoding the drug target as well as transcriptional adaptive bypass through AXL signaling.
- In primary breast cancer, prototypical chromosomal loss and driver *PIK3CA* mutations were present in cells with epithelial identity while infiltrating monocytes were also identified.
- ResolveOME transcriptomics revealed attributes of cell state plasticity.
- Transcriptomes *and* genomes are dynamic; both tiers of data are required to define their interplay and their role in oncogenesis and in drug resistance.

Reference & Acknowledgments

1. 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).

We thank the patients involved in this study for their gracious gift of tissue samples that made the scientific insights in this manuscript possible. Tissue for single-cell DCIS/IDC studies was obtained in collaboration with Dr. E. Shelley Hwang and Dr. Jeffrey Marks, and in accordance with the Duke University Medical Center IRB for the clinical trial PRO00034242 "Biologic Characterization of the Breast Cancer Tumor Microenvironment."

Single cells

bioRxiv THE PREPRINT SERVER FOR BIOLOGY

Cells explored. Answers revealed.